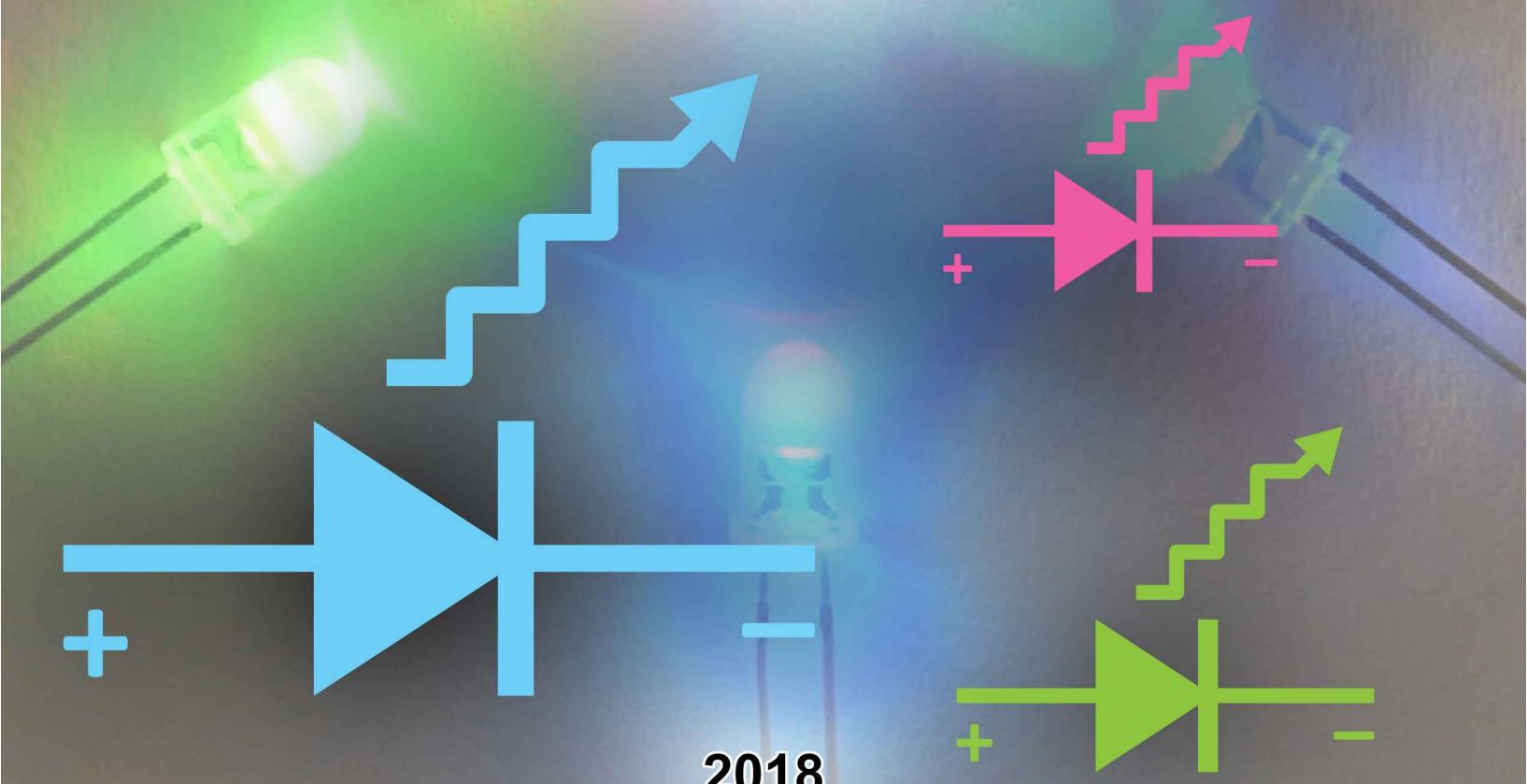


Light-Emitting Diodes

3rd Edition

E. Fred Schubert



Light-Emitting Diodes

3rd edition

E. Fred Schubert
Rensselaer Polytechnic Institute
Troy, New York, USA

ISBN: 978-0-9 863826-6-6

Publisher: E. Fred Schubert

First published: February 2018

Present updated and corrected version: August 2018

© E. Fred Schubert, all rights reserved

The 1st edition of the book “Light-Emitting Diodes” was published in 2003. The 2nd edition was published in 2006. The current 3rd edition of the book is a substantial expansion of the second edition and has 37 chapters. The book includes a thorough discussion of white light-emitting diodes (LEDs), phosphor materials used in white LEDs, packaging technology, and the various efficiencies and efficacies encountered in the context of LEDs. The background of light, color science, and human vision is provided as well. The fully colored illustrations of the 3rd edition are undoubtedly beneficial given the prominent role of light and color in the field of LEDs. The book is a comprehensive discussion of the LED, particularly its semiconductor physics, electrical, optical, material science, thermal, mechanical, and chemical foundations. The 3rd edition is published in electronic format in order to make the book affordable and easily accessible to a wide readership.

Table of Contents

1 – History of SiC light-emitting diodes.....	1-1
1.1 Silicon carbide (SiC).....	1-1
1.2 Henry Round's demonstration of the first LED.....	1-2
1.3 Oleg Lossev's research on SiC LEDs.....	1-5
1.4 Kurt Lehovec's explanation of electroluminescence from p-n junctions	1-8
1.5 Post 1950s developments in SiC LEDs	1-9
1.6 LED centennial in 2007.....	1-9
1.7 References	1-10
2 – History of III–V light-emitting diodes	2-1
2.1 Compound semiconductors.....	2-1
2.2 The first III-V compound semiconductor LED	2-3
2.3 History of $\text{GaAs}_y\text{P}_{1-y}$ alloy semiconductors and LEDs	2-4
2.4 History of GaAs and AlGaAs infrared and red LEDs	2-9
2.5 History of GaP and GaAsP LEDs doped with optically active impurities	2-12
2.6 History of AlGaNp visible LEDs	2-18
2.7 History of GaN metal–semiconductor emitters.....	2-19
2.8 History of GaInN p-n junction LEDs.....	2-22
2.9 References	2-26
3 – History of white light-emitting diodes	3-1
3.1 The advent of the white LEDs	3-1
3.2 Transition from LEDs to LED-based lighting.....	3-9
3.3 LEDs entering new fields of applications	3-14
3.4 LED replacement light bulbs	3-17
3.5 References	3-18
4 – Radiative and non-radiative recombination.....	4-1
4.1 Radiative electron–hole recombination	4-1
4.2 Radiative recombination for low-level excitation	4-2
4.3 Radiative recombination for high-level excitation	4-6
4.4 Bimolecular rate equations for quantum well structures	4-6
4.5 Luminescence decay	4-7
4.6 Shockley–Read–Hall non-radiative recombination	4-8
4.7 Auger non-radiative recombination	4-13
4.8 Non-radiative recombination at surfaces	4-16
4.9 Competition between radiative and non-radiative recombination	4-20
4.10 Recombination in direct-gap and indirect-gap semiconductors	4-22
4.11 References	4-23
5 – Theory of radiative recombination	5-1
5.1 Quantum mechanical model of recombination.....	5-1
5.2 The van Roosbroeck–Shockley model	5-3
5.3 Temperature and doping dependence of recombination	5-7
5.4 The Einstein model	5-9

5.5 References	5-11
6 – Recombination modeled by the ABCCD-model	6-1
6.1 Relation between optical and electrical carrier generation	6-1
6.2 Free carrier concentrations in the active region	6-1
6.3 Recombination expressed in terms of the A, B, and C coefficients and f(n) term	6-2
6.4 Determination of the A, B, and C coefficients in absence of carrier leakage	6-7
6.5 Carrier leakage and the ABCCD model	6-10
6.6 References	6-12
7 – Current–voltage characteristics	7-1
7.1 Diode current–voltage characteristic based on Shockley equation	7-1
7.2 Diode forward voltage	7-6
7.3 Diode ideality factor.....	7-9
7.4 Diode parasitic resistances	7-11
7.5 Measurement of diode parasitic resistances.....	7-12
7.6 Diagnosis of potential problems in the current-voltage characteristic.....	7-13
7.7 Ohmic contacts	7-17
7.8 Physics of ohmic contacts	7-19
7.9 Multilayer ohmic contacts	7-21
7.10 References	7-22
8 – Carrier transport effects in the active region	8-1
8.1 Carrier distribution in p-n homojunctions	8-1
8.2 Carrier distribution in p-n heterojunctions.....	8-2
8.3 Effect of heterojunctions on device resistance	8-3
8.4 Carrier loss in double heterostructures.....	8-7
8.5 Carrier overflow in double heterostructures.....	8-10
8.6 Electron-blocking layers.....	8-13
8.7 Effect of heterojunctions on diode voltage	8-15
8.8 References	8-17
9 – Design of current flow.....	9-1
9.1 Current-spreading layer	9-1
9.2 Theory of current spreading	9-7
9.3 Current crowding in LEDs on insulating substrates	9-10
9.4 Buried contact design	9-15
9.5 Lateral injection schemes	9-16
9.6 Current-blocking layers.....	9-17
9.7 Surface-leakage current.....	9-19
9.8 References	9-20
10 – Alternating-current (AC) and high-voltage (HV) LEDs.....	10-1
10.1 Simple circuits for AC-driven LEDs	10-1
10.2 LED bridge rectifier configurations for AC-driven LEDs	10-3
10.3 Analysis of pulsed drive current in AC-driven LEDs	10-4
10.4 Challenges associated with AC-driven LEDs	10-6
10.5 High-voltage (HV) LEDs	10-7

10.6 References	10-10
11 – Basic optical properties	11-1
11.1 Emission spectrum	11-1
11.2 The light escape cone	11-4
11.3 Emission pattern	11-7
11.4 The lambertian emission pattern	11-7
11.5 Encapsulants	11-10
11.6 Temperature dependence of emission intensity.....	11-12
11.7 References	11-13
12 – Reflectors.....	12-1
12.1 Metallic reflectors, reflective contacts, and transparent contacts	12-1
12.2 Total internal reflectors	12-6
12.3 Distributed Bragg reflectors.....	12-8
12.4 Omnidirectional reflectors.....	12-19
12.5 Specular, diffuse, and mixed reflectors	12-22
12.6 References	12-28
13 – Definition of efficiencies.....	13-1
13.1 Radiative efficiency (RE) of active region	13-2
13.2 Injection efficiency (IE) of carriers into active region	13-4
13.3 Internal quantum efficiency (IQE).....	13-5
13.4 Light-extraction efficiency (LEE)	13-6
13.5 External quantum efficiency (EQE)	13-6
13.6 Forward-voltage efficiency (VfE).....	13-7
13.7 Power-conversion efficiency (PCE)	13-8
13.8 Luminous efficacy of radiation (LER)	13-8
13.9 Luminous efficacy of a source (LES).....	13-10
13.10 Efficiency of wavelength-converter materials (phosphors)	13-11
13.11 References	13-13
14 – Designs for high internal quantum efficiency	14-1
14.1 Double heterostructures.....	14-1
14.2 Doping of active region.....	14-4
14.3 p-n junction displacement	14-6
14.4 Doping of the confinement regions.....	14-8
14.5 Non-radiative recombination	14-10
14.6 Lattice matching.....	14-11
14.7 References	14-14
15 – Designs for high light-extraction efficiency	15-1
15.1 Absorption of below-bandgap light in semiconductors	15-1
15.2 Double heterostructures.....	15-5
15.3 Shaping of LED dies	15-6
15.4 Textured semiconductor surfaces	15-10
15.5 GaN thin-film LEDs	15-12
15.6 Cross-shaped contacts and other contact geometries	15-14

15.7 Transparent substrate technology.....	15-15
15.8 Anti-reflection optical coatings.....	15-17
15.9 References	15-18
16 – Spontaneous emission from resonant cavities	16-1
16.1 Modification of spontaneous emission	16-1
16.2 Fabry–Perot resonators	16-3
16.3 Optical mode density in a one-dimensional resonator	16-6
16.4 Spectral emission enhancement.....	16-11
16.5 Integrated emission enhancement.....	16-13
16.6 Experimental emission enhancement and angular dependence	16-14
16.7 References	16-17
17 – Resonant-cavity light-emitting diodes	17-1
17.1 Introduction and history	17-1
17.2 RCLED design rules.....	17-3
17.3 GaInAs/GaAs RCLEDs emitting at 930 nm	17-6
17.4 AlGaInP/GaAs RCLEDs emitting at 650 nm	17-11
17.5 Newer developments in RCLED devices	17-14
17.6 References	17-14
18 – Photonic crystal light-emitting diodes	18-1
18.1 The photon recycling concept for LEDs	18-1
18.2 Thresholdless laser or single-mode LED	18-3
18.3 Photonic crystals	18-5
18.4 LEDs using photonic crystals.....	18-10
18.5 References	18-11
19 – Visible-spectrum LEDs	19-1
19.1 The GaAsP, GaP, GaAsP:N, and GaP:N material systems	19-1
19.2 The AlGaAs/GaAs material system	19-6
19.3 The AlGaN/P/GaAs material system	19-9
19.4 The GaInN material system.....	19-11
19.5 Optical characteristics of high-brightness LEDs.....	19-13
19.6 Electrical characteristics of high-brightness LEDs.....	19-15
19.7 Wavelength dependence of the efficiency in the visible spectrum	19-16
19.8 References	19-18
20 – The AlGaN material system	20-1
20.1 The AlGaN bandgap	20-1
20.2 Crystal structure of III–V nitrides and origin of polarization	20-2
20.3 Consequences of polarization effects in III–V nitrides	20-7
20.4 P-type doping in III–V nitrides.....	20-9
20.5 Epitaxial growth of GaN on sapphire	20-13
20.6 Dislocations in III–V nitrides.....	20-15
20.7 References	20-19

21 – Ultraviolet light emitters	21-1
21.1 The ultraviolet (UV) spectral range.....	21-1
21.2 UV and violet LEDs emitting at wavelengths longer than 360 nm	21-2
21.3 UV devices emitting at wavelengths shorter than 360 nm	21-4
21.4 Alternative approaches for UV emitters.....	21-8
21.5 References	21-9
22 – The efficiency droop.....	22-1
22.1 Introduction to the efficiency droop	22-1
22.2 The “ABC” model	22-3
22.3 The “ABC + f(n)” model.....	22-3
22.4 The function f(n) for drift-induced leakage	22-4
22.5 Role of “polarization charges and electric fields”	22-9
22.6 Role of “thermal roll-over”	22-10
22.7 Role of “carrier de-localization”	22-11
22.8 Role of Auger recombination.....	22-12
22.9 Strategies to overcome the efficiency droop	22-14
22.10 References	22-18
23 – Human eye sensitivity and photometric quantities	23-1
23.1 Light receptors of the human eye.....	23-1
23.2 Basic radiometric and photometric units	23-3
23.3 Eye sensitivity function	23-6
23.4 Luminous efficacy	23-9
23.5 Surface brightness.....	23-12
23.6 Circadian rhythm and circadian sensitivity.....	23-13
23.7 References	23-14
24 – Colorimetry.....	24-1
24.1 Colors of near-monochromatic emitters	24-2
24.2 Color-matching functions and chromaticity diagram.....	24-2
24.3 Color purity	24-10
24.4 LEDs in the chromaticity diagram	24-12
24.5 Relationship between chromaticity and color.....	24-13
24.6 The CIE Lab color space.....	24-14
24.7 The CIE UVW and CIE Luv color spaces.....	24-16
24.8 References	24-16
25 – Black-body radiation and color temperature	25-1
25.1 Incandescence (heat glow)	25-1
25.2 The planckian spectrum.....	25-2
25.3 The solar spectrum	25-3
25.4 Planckian radiation in the chromaticity diagram.....	25-5
25.5 Color temperature and correlated color temperature.....	25-6
25.6 Standardization of white light by means of color-temperature binning.....	25-9
25.7 References	25-11

26 – Color mixing.....	26-1
26.1 Subtractive color mixing	26-1
26.2 Additive color mixing	26-1
26.3 Calculation of color mixing	26-2
26.4 Implementation of color mixing	26-5
26.5 References	26-9
27 – Color rendering.....	27-1
27.1 Color rendering of physical objects	27-1
27.2 CIE uniform color space	27-6
27.3 Color-rendering index for planckian-locus illumination sources.....	27-9
27.4 Color-rendering index for non-planckian-locus illumination sources	27-10
27.5 Criticism of the CRI and alternative approaches	27-13
27.6 References	27-14
28 – White-light sources based on LED chips	28-1
28.1 Generation of white light with LED chips	28-1
28.2 Generation of white light by dichromatic sources	28-3
28.3 Generation of white light by trichromatic sources.....	28-7
28.4 Temperature dependence of trichromatic LED-based white-light source.....	28-11
28.5 Generation of white light by tetrachromatic and pentachromatic sources.....	28-14
28.6 References	28-14
29 – Phosphors.....	29-1
29.1 Introduction to phosphors.....	29-1
29.2 Types of phosphor materials	29-6
29.3 Phosphors used with LED chips	29-9
29.4 Phosphor body color	29-14
29.5 Phosphor structure and synthesis	29-16
29.6 Quenching by temperature and activator concentration	29-20
29.7 Phosphor particle size	29-23
29.8 Phosphor deposition	29-27
29.9 Phosphor sedimentation	29-29
29.10 Quantum-splitting phosphors.....	29-30
29.11 Location of phosphors in the CIE chromaticity diagram.....	29-31
29.12 References	29-32
30 – White LEDs based on phosphors.....	30-1
30.1 Efficiency of wavelength-converter materials.....	30-2
30.2 Wavelength-converter materials	30-4
30.3 White LEDs based on YAG:Ce phosphor	30-6
30.4 White LEDs based on phosphor blends	30-8
30.5 White LEDs using two LED chips and one phosphor	30-10
30.6 White LEDs using violet excitation sources	30-11
30.7 White LEDs using UV excitation sources.....	30-13
30.8 White LEDs based on semiconductor converters (PRS-LED)	30-15
30.9 White LEDs based on other methods	30-17

30.10 Spatial phosphor distributions.....	30-18
30.11 References	30-20
31 – Packaging: Types of packages for LEDs.....	31-1
31.1 Round 3 mm and 5 mm LED package.....	31-2
31.2 SMD LED injection molded package	31-3
31.3 SMD LED package fabricated by using a resin-molded body	31-5
31.4 Ceramic-base LED package	31-7
31.5 Side-view LED package.....	31-9
31.6 Other examples of LED packages.....	31-11
31.7 Chip-on-board LED packages (COBs)	31-12
31.8 Chip-scale LED packages (CSPs)	31-13
31.9 References	31-14
32 – Packaging: Materials used in packaging	32-1
32.1 Resin materials.....	32-1
32.2 Types of resins	32-2
32.3 Modification of resins by fillers	32-6
32.4 Epoxy resins	32-8
32.5 Silicone resins.....	32-11
32.6 Advanced resin structures	32-14
32.7 Degradation of resins under harsh operating conditions.....	32-16
32.8 References	32-17
33 – Packaging: Protection against electrostatic discharge (ESD)	33-1
33.1 The human body model and the IEC 61000-4-2 standard.....	33-1
33.2 The electrostatic discharge process.....	33-2
33.3 ESD-protection circuits	33-3
33.4 ESD-protection circuits integrated with the LED	33-5
33.5 References	33-6
34 – Packaging: Junction and carrier temperatures	34-1
34.1 Carrier temperature and high-energy slope of spectrum	34-1
34.2 Junction temperature and peak emission wavelength	34-3
34.3 Theory of temperature dependence of diode forward voltage	34-4
34.4 Measurement of junction temperature using forward voltage	34-8
34.5 Thermal considerations for pulsed measurements.....	34-11
34.6 Constant-current and constant-voltage DC drive circuits	34-13
34.7 References	34-14
35 – Communication: Optical communication	35-1
35.1 Types of optical fibers	35-1
35.2 Attenuation in silica and plastic optical fibers.....	35-3
35.3 Modal dispersion in fibers	35-4
35.4 Material dispersion in fibers	35-6
35.5 Numerical aperture of fibers	35-9
35.6 Coupling with lenses	35-11
35.7 Free-space optical communication.....	35-13

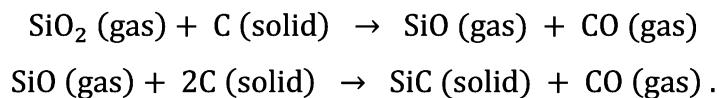
35.8 References	35-15
36 – Communication: Communication LEDs	36-1
36.1 LEDs for free-space communication	36-1
36.2 LEDs for fiber-optic communication	36-1
36.3 Surface-emitting Burrus-type communication LEDs emitting at 870 nm.....	36-2
36.4 Surface-emitting communication LEDs emitting at 1300 nm.....	36-3
36.5 Communication LEDs emitting at 650 nm	36-5
36.6 Edge-emitting superluminescent diodes (SLDs)	36-7
36.7 References	36-10
37 – Communication: LED modulation characteristics.....	37-1
37.1 Rise and fall times, 3-dB frequency, and bandwidth in linear circuit theory	37-1
37.2 Rise and fall time in the limit of large diode capacitance.....	37-4
37.3 Rise and fall time in the limit of small diode capacitance	37-5
37.4 Voltage dependence of the rise and fall times	37-6
37.5 Carrier sweep-out of the active region.....	37-7
37.6 Current shaping.....	37-8
37.7 Three-dB frequency	37-9
37.8 Eye diagram.....	37-10
37.9 Carrier lifetime and 3-dB frequency	37-11
37.10 References	37-12
Appendices	38-1
Appendix 1 – Frequently used symbols	38-1
Appendix 2 – Physical constants.....	38-5
Appendix 3 – Properties of III–V arsenides at 300 K.....	38-6
Appendix 4 – Properties of III–V nitrides at 300 K.....	38-7
Appendix 5 – Properties of III–V phosphides at 300 K	38-8
Appendix 6 – Properties of Si and Ge at 300 K	38-9
Appendix 7 – Periodic Table of Elements	38-10
Index.....	39-1
End of book.....	39-10

1

1 – History of SiC light-emitting diodes

1.1 Silicon carbide (SiC)

In 1891 Eugene G. Acheson established a commercial process (the so-called **Acheson process**) for an essentially new manmade material, silicon carbide (SiC). Acheson named this new material “carborundum”. The synthesis was accomplished in an electrically heated high-temperature furnace in which glass (silicon dioxide, SiO_2) and coal (carbon, C) reacted to form SiC according to the chemical reaction (Filsinger and Bourrie, 1990; Jacobson *et al.*, 1992)



SiC is an extremely stable material, mechanically as well as chemically. SiC, which has the same crystal symmetry as diamond, has a very high hardness. On the Mohs Hardness Scale, carborundum has a hardness of 9.0, pure SiC a hardness of 9.2~9.5, and diamond a hardness of 10.0. Because of its high hardness and because it could be synthesized in large quantities at low cost, carborundum was a material of choice for the abrasives industry.

SiC occurs only very rarely in nature. Its highest abundance occurs in some types of meteorites coming from outer space. Note that III-V semiconductors do not at all occur in nature. Photographs of pieces of SiC fabricated by the Acheson process are shown in **Figure 1.1**. SiC fabricated by the Acheson process was the first material to exhibit electroluminescence, as will be discussed in the next section.



Fig. 1.1: Silicon carbide (SiC) crystals fabricated by the Acheson process (left-hand-side: photograph made at Rensselaer Polytechnic Institute; right-hand-side: photograph after Wikipedia, 2007).

1.2 Henry Round's demonstration of the first LED

Starting early in the twentieth century, light emission from a solid-state material, caused by an electrical power source, was reported: a new phenomenon termed ***electroluminescence***. Because electroluminescence can occur at room temperature, it is fundamentally different from ***incandescence*** (or heat glow), which is the visible electromagnetic radiation emitted by a material heated to high temperatures, typically $> 650^{\circ}\text{C}$.

In 1907, Henry Joseph Round (1881–1966) tested SiC crystals for possible use as rectifying solid-state detectors, then called “crystal detectors”. Such crystal detectors were used in early radio receivers for the demodulation of amplitude-modulated radio-frequency signals. Crystal detectors (i.e. crystal–metal-point-contact structures) had first been demonstrated in 1906 and considered a low-cost alternative to expensive and power-hungry vacuum-tube diodes (the vacuum-tube diode or “Fleming valve” was first demonstrated in 1904).

Henry Round noticed that light was emitted from a SiC crystallite. The first light-emitting diode (LED) had been born! At that time, the material properties of SiC were poorly controlled, and the light-emission process was not understood. Nevertheless, Round reported his observations immediately to the editors of the journal *Electrical World*, as shown in **Figure 1.2** (Round, 1907). This remarkably short publication is titled “A Note on Carborundum” and has only two paragraphs. The publication reveals that he applied voltages varying from 10 V to 110 V and that he observed multiple emission colors including yellow, orange, and blue. Round (1907) writes:

“SIRS: – During an investigation of the unsymmetrical passage of current through a contact of carborundum and other substances a curious phenomenon was noted. On applying a potential of 10 volts between two points on a crystal of carborundum, the crystal gave out a yellowish light. Only one or two specimens could be found which gave a bright glow on such a low voltage, but with 110 volts a large number could be found to glow. In some crystals only edges gave the light and others gave instead of a yellow light green, orange or blue. In all cases tested the glow appears to come from the negative pole, a bright blue-green spark appearing at the positive pole. In a single crystal, if contact is made near the center with the negative pole, and the positive pole is put in contact at any other place, only one section of the crystal will glow and that the same section wherever the positive pole is placed.

There seems to be some connection between the above effect and the e.m.f. produced by a junction of carborundum and another conductor when heated by a direct or alternating current; but the connection may be only secondary as an obvious explanation of the e.m.f. effect is the thermoelectric one. The writer would be glad of references to any published account of an investigation of this or any allied phenomena.”

A Note on Carborundum.

To the Editors of Electrical World:

Sirs:—During an investigation of the unsymmetrical passage of current through a contact of carborundum and other substances a curious phenomenon was noted. On applying a potential of 10 volts between two points on a crystal of carborundum, the crystal gave out a yellowish light. Only one or two specimens could be found which gave a bright glow on such a low voltage, but with 110 volts a large number could be found to glow. In some crystals only edges gave the light and others gave instead of a yellow light green, orange or blue. In all cases tested the glow appears to come from the negative pole, a bright blue-green spark appearing at the positive pole. In a single crystal, if contact is made near the center with the negative pole, and the positive pole is put in contact at any other place, only one section of the crystal will glow and that the same section wherever the positive pole is placed.

There seems to be some connection between the above effect and the e.m.f. produced by a junction of carborundum and another conductor when heated by a direct or alternating current; but the connection may be only secondary as an obvious explanation of the e.m.f. effect is the thermoelectric one. The writer would be glad of references to any published account of an investigation of this or any allied phenomena.

NEW YORK, N. Y.

H. J. ROUND.

Fig. 1.2: Publication reporting on a “curious phenomenon”, namely the first observation of electroluminescence from a SiC (carborundum) light-emitting diode. The article indicates that the first LED was a Schottky diode rather than a p-n junction diode (after H. J. Round, *Electrical World* 49, 309, 1907).

Henry Round, born in 1881 in England, was a radio engineer and a prolific inventor who, by the end of his career, held 117 patents. He had joined the Marconi Company in 1902 not long after famed radio pioneer Guglielmo Marconi had made his first wireless transatlantic transmission. Henry Round became a personal assistant to Marconi and worked for some time in the United States before returning to his native England. A photograph of Henry Round is shown in **Figure 1.3.**



Henry Joseph Round

Fig. 1.3: Henry Joseph Round (1881–1966), a prolific inventor who was awarded over 100 patents, published the first-ever report of electroluminescence in 1907 in the journal *Electrical World* (photograph after Bodleian Library, University of Oxford, UK; see also “Marconi calling” Internet website, Image 2A, 2007).

Henry Round’s first light-emitting devices had rectifying current–voltage characteristics; that is, these first devices were indeed light-emitting *diodes* or LEDs. The light was produced by touching

the SiC crystal with metal electrodes so that a rectifying Schottky contact was formed. Schottky diodes are usually majority carrier devices. However, minority carriers can be created by either minority-carrier injection under strong forward-bias conditions, or avalanche multiplication under reverse-bias conditions.

The mechanism of light emission in forward-biased and reverse-biased Schottky diodes is shown in **Figure 1.4**, which displays the band diagram of a metal–semiconductor junction under (a) equilibrium, (b) moderate forward bias, (c) strong forward bias conditions, and (d) under strong reverse bias where avalanche multiplication occurs. The semiconductor shown in the figure is assumed to be lightly doped (n^- -type); the semiconductor can even be insulating (i-type). Such diodes are frequently called metal-insulator-semiconductor (MIS) Schottky diodes.

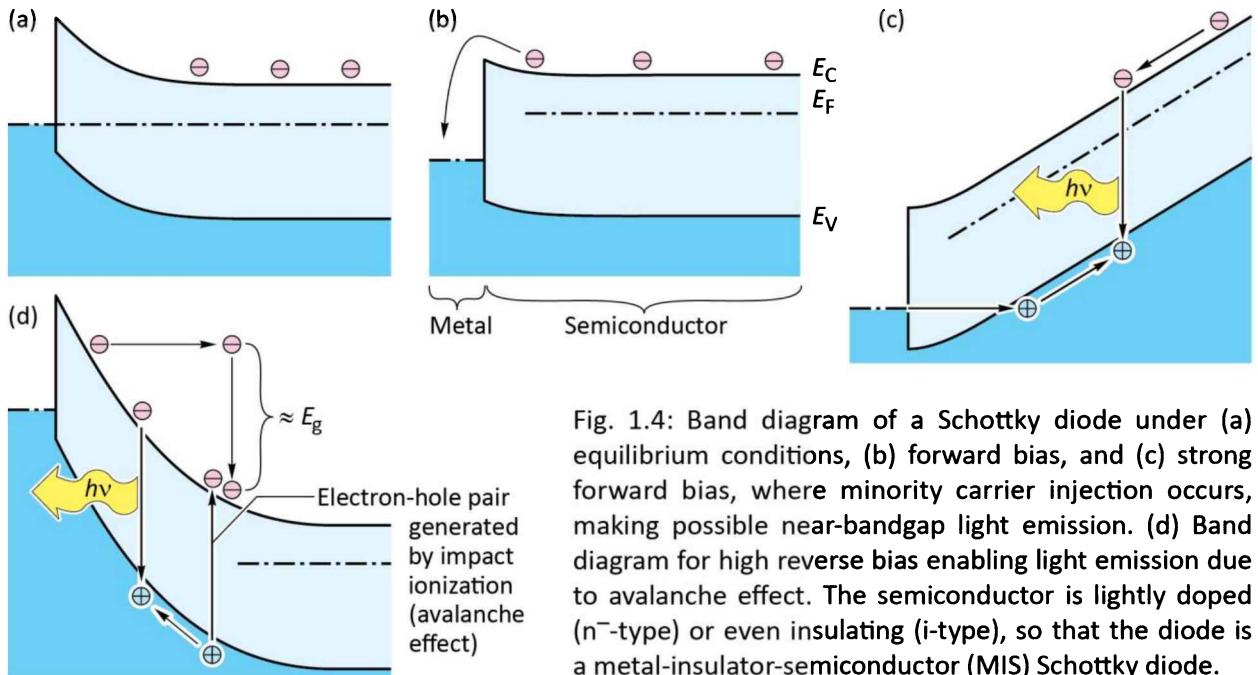


Fig. 1.4: Band diagram of a Schottky diode under (a) equilibrium conditions, (b) forward bias, and (c) strong forward bias, where minority carrier injection occurs, making possible near-bandgap light emission. (d) Band diagram for high reverse bias enabling light emission due to avalanche effect. The semiconductor is lightly doped (n^- -type) or even insulating (i-type), so that the diode is a metal-insulator-semiconductor (MIS) Schottky diode.

Under strong forward bias conditions, minority carriers are injected into the semiconductor by tunneling through the surface potential barrier. Light is emitted upon recombination of the injected minority carriers with the n-type majority carriers. The forward voltage required for minority carrier injection in Schottky diodes is larger than typical p-n junction forward voltages. Round (1907) reported operating voltages ranging between 10 and 110 V.

Light can also be generated in a Schottky diode under strong reverse-bias conditions as shown in **Figure 1.4 (d)**. At a sufficiently strong reverse bias, the avalanche effect occurs in which high-

energy carriers impact-ionize atoms of the semiconductor, that is, a valence electron is excited to the conduction band. In this process, holes are created in the valence band as well as electrons in the conduction band, which will eventually recombine thereby creating light. Additional light-generating processes in Schottky diodes under reverse-bias conditions have been reported by Eastman *et al.* (1964).

1.3 Oleg Lossev's research on SiC LEDs

Oleg Vladimirovich Lossev (1923, 1924a, 1924b, 1928), a brilliant Russian experimentalist who worked at the Nizhny Novgorod Radio Laboratory in the former Soviet Union, reported the first detailed investigations of the electroluminescence phenomenon observed with SiC metal–semiconductor rectifiers. Born in 1903, Lossev published his first paper on electroluminescence at the age of 20 years, in 1923, when he had not yet a formal degree. In the 1923 paper, Lossev reported seeing green light with the “naked eye” when reverse-biasing a SiC metal–semiconductor rectifier. In his 1924 paper published in the *Wireless World and Radio Review*, he showed the first photograph of SiC emitting electroluminescence. It is shown in *Figure 1.5*.

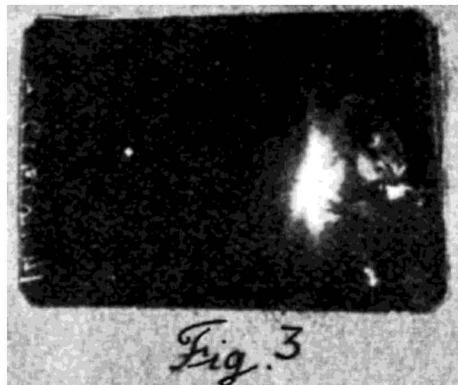


Fig. 1.5: First photograph of electroluminescence. The photograph was taken by Oleg Lossev in 1924 using a SiC metal–semiconductor LED (after O. V. Lossev, *Wireless World and Radio Review* 271, 93, 1924).

For forward bias, the current was higher but little or no light emission was observed. The main use of these rectifiers was in solid-state demodulation radio-circuits that did not employ vacuum tubes. Oleg Lossev found that luminescence occurred in some diodes when biased in the reverse direction; subsequently he found that in some diodes luminescence occurred when biased in forward *and* reverse directions. A detailed current-voltage characteristic revealing Lossev's the high degree of scientific rigor is shown in *Figure 1.6*.

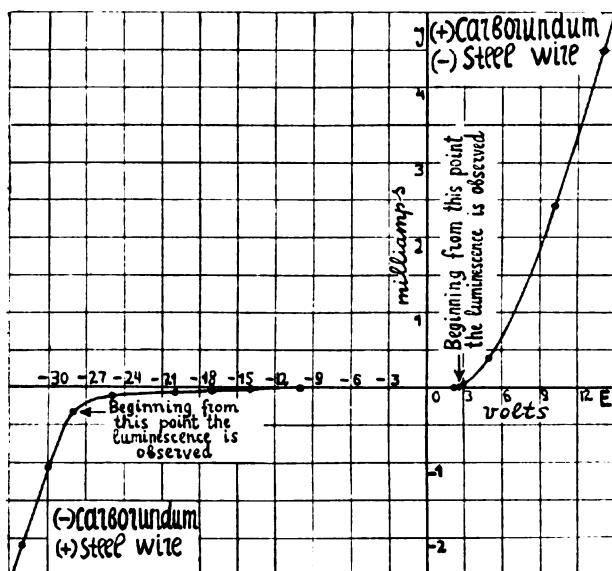


Fig. 1.6: Detailed current-voltage characteristic of SiC metal-semiconductor LED published in 1928 by Oleg Lossev (after O. V. Lossev, *Philosophical Magazine* 6, 1024, 1928).

Oleg Lossev was puzzled about the physical origin of the luminescence. He investigated whether light was generated by heat glow (incandescence) by testing the evaporation rate of a droplet of liquid benzene that he positioned on the luminous sample surface. He found, however, that the benzene evaporated very slowly and correctly concluded that the luminescence was not caused by incandescence. Lossev postulated that the process by which light was produced is “very similar to cold electronic discharge”. Lossev also found that the light could be switched on and off very rapidly, making the device suitable for what he called a “light relay”.

Oleg Lossev’s reports are the first detailed studies of electroluminescence originating from a semiconductor. In recognition of his brilliance, the Ioffe Institute awarded Lossev the degree of “Candidate”, the equivalent of a doctorate degree, in 1938, without a formal dissertation (Loebner, 1976). A photograph of Oleg Lossev is shown in **Figure 1.7**. He tragically died in 1942 during World War II in Leningrad (now St. Petersburg), at the age of only 39.



Fig. 1.7: Oleg Vladimirovich Lossev (1903–1942), brilliant Russian scientist and sole author of the first reports containing detailed observations, measurements, photographs, and analyses of SiC metal-semiconductor light-emitting diodes (courtesy of Alexander E. Yunovich, Moscow State University, 2008).

In the present context, the question as to the luminescence observed by Lossev originated from a unipolar device (Schottky contact) or a bi-polar device (p-n junction) is highly relevant. Although one would assume that the luminescence originated most likely from a unipolar device, there are indications to the contrary! Discussing Lossev's research, Lehovec *et al.* (1951) write:

“In the early literature, the light emission was related to the presence of an “active layer” on the surface of the crystal. Lossev removed the “active layer” by grinding and measured the resistance between two probes placed on the surface at various intermediate stages of the grinding process (O. V. Lossev, *Physikalische Zeitschrift*, volume 34, page 397, year 1933). The resistance was found to increase by several orders of magnitude with progressive removal of material and then to drop sharply after 12 μm had been removed. Probing the potential distribution while current was passing from the surface into the crystal showed the presence of a rectifying junction at the boundary of the active layer and the bulk silicon carbide. Yellow light emission occurs when the current passes through the boundary in the forward direction. In a later paper (O. V. Lossev, *Compt. Rend. Acad. Sci. U.R.S.S.* volume 29, page 363, year 1940) it is mentioned that the active layer consists of an n-type layer, whereas the bulk consists of p-type material.”

Thus, according to the report of Lehovec *et al.*, Lossev was the first person to associate light emission from a light-emitting diode with the boundary between a p-type layer and an n-type layer. This view was confirmed by Alexander E. Yunovich (Yunovich, 2008), who, due to his Russian nationality was able to understand the Russian writings of Lossev.

The study of a 1933 publication (Lossev, 1933) leaves little doubt that Lossev indeed performed experiments on p-n junctions and not just metal-semiconductor junctions. Tapered grinding of a SiC crystal surface allowed Lossev (1933) to attribute specific voltage drops to the (i) metal-semiconductor contacts, (ii) semiconductor bulk crystal, and (iii) pn-junction region (active region). Lossev (1933) concluded that the voltage drop (i) across the metal-semiconductor contacts is small, (ii) across the semiconductor bulk crystal is small, (iii) across the pn-junction region depends on the polarity of the applied voltage and is large for reverse-biased junctions and small for forward-biased junctions. This undoubtedly is the signature of a p-n junction. In addition, Lossev reported the emission of light from the pn junction region. Lossev calls this region the “active layer”.

In later years, Lossev (1940) measured the SiC LED emission spectrum and showed that the emission energy closely correlates with the absorption edge (bandgap energy) of SiC. Prior to the 1940 publication, Lossev (1933) related the photo-voltage (3.4 V), measured when illuminating the SiC p-n junction with a quartz-mercury direct-current light source, with the emission energy

(3.73 eV). Given this, Lossev may be recognized as the first researcher discovering the relationship between diode forward voltage, diode photo-voltage under illumination, semiconductor absorption edge (bandgap energy), and emission energy.

The Lossev 1933 and 1940 publications do not describe how the SiC p-n junctions were fabricated; questions in this regard are not readily answered by Lossev's 1933 and 1940 publications. Despite these uncertainties, it appears to be very clear from his published work that Oleg Vladimirovich Lossev should be considered the “father” of the modern light-emitting diode, the pn-junction light-emitting diode. This includes the first visible-spectrum pn-junction LED and the first UV pn-junction LED.

1.4 Kurt Lehovec's explanation of electroluminescence from p-n junctions

Kurt Lehovec *et al.* (1951, 1953) reported electroluminescence experiments on commercial bulk SiC crystals made by the Carborundum Company. The SiC samples emitted light in the wavelength range 450~650 nm with an external quantum efficiency of 10^{-6} . Assuming that their SiC crystals include a p-n junction, the authors proposed the first correct interpretation of electroluminescence originating from a current-injected p-n junction, namely the injection of minority-carriers into a majority-carrier region and the subsequent recombination of the two types of carriers, as shown in *Figure 1.8*. Lehovec's interpretation of electroluminescence was certainly inspired by and based upon the pioneering works of William B. Shockley, one of the three inventors of the transistor, Nobel Laureate, and certainly the spiritual father of microelectronics technology (Shockley, 1950).

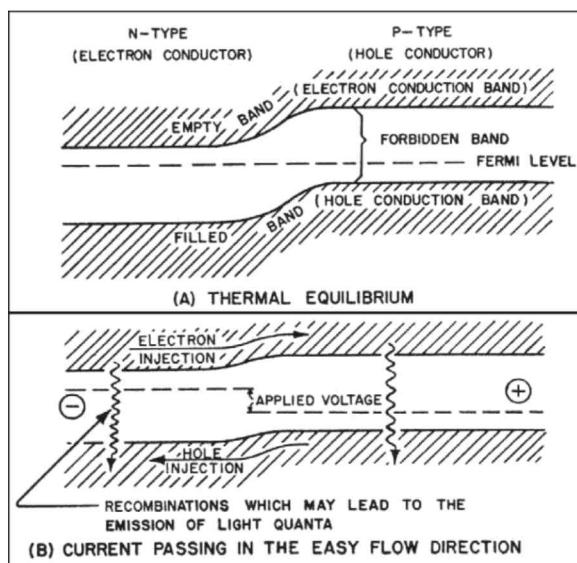


Fig. 1.8: First correct explanation of pn-junction electroluminescence in terms of minority-carrier injection illustrated in a band diagram (after Lehovec *et al.*, 1951).

However, it is not certain if the electroluminescence reported by Lehovec *et al.* (1951) indeed originated from a p-n junction or if it may have originated from a Schottky contact (metal-semiconductor contact). Given the employment of commercial SiC crystals and given the fact that the authors do not describe an intentional doping process, the presence of a p-n junction in the SiC crystal is somewhat unlikely yet not impossible.

A photograph of Kurt Lehovec, who is also known for his contributions to the fields of solar photo-voltaics and silicon integrated circuits, is shown in *Figure 1.9* (Lehovec, 2008). The pre-1960 history of LEDs was reviewed by Loebner (1976). The review contains many relevant pieces of information and is a valuable resource.



Fig. 1.9: Kurt Lehovec (1918 – 2012) provided the first correct explanation of light emission from pn junctions. He explained the electroluminescence of SiC pn-junction light-emitting diodes by minority carrier injection and subsequent recombination (after www.KurtLehovec.com, 2008).

1.5 Post 1950s developments in SiC LEDs

By the late 1960s, SiC films were grown by more sophisticated processes (Violin *et al.*, 1969). Blue light-emitting diodes were obtained from SiC p-n junction devices. Electrical-to-optical power-conversion efficiencies were 0.005% (Potter *et al.*, 1969). In the ensuing decades, blue SiC LEDs were substantially improved; yet efficiencies remained significantly below 1%. SiC, after all, has an indirect bandgap. Although many blue SiC LEDs were actually sold commercially in the early 1990s, they no longer are a viable product. In the end, the best SiC LEDs, emitting blue light at 470 nm, had an efficiency of only 0.03% (Edmond *et al.*, 1993). SiC, the material of the very first LED, could no longer compete with III–V semiconductors.

1.6 LED centennial in 2007

In 2007, the LED celebrated its centennial anniversary, recognizing Henry Joseph Round's demonstration of the first LED in 1907. Two articles, both published in the centennial year of

2007, recognized the importance of the date and the accomplishments of Henry Joseph Round and of those who subsequently advanced the art of LEDs (Schubert and Kim, 2007; Zheludev, 2007).

The centennial along with the strongly growing public interest in solid-state lighting made people curious about the experiments conducted by Round. In 2007, several groups re-enacted Round's seminal electroluminescence experiments using a SiC crystal, two wires, and a power supply (see, for example, Shah and Mont, 2007).

In 2010, a reenactment of Round's electroluminescence experiments was reported in the widely read journal *Popular Mechanics*. A 9 V battery served as power supply; two electrical probes connected the power supply to a SiC crystal. And indeed, a “faint colored glow” was observed (Gray, 2010). A photograph of an experiment conducted by Gray is shown in **Figure 1.10**. The author, Gray, indicated that the light emission was observed “sometimes” when the power supply was connected to the crystal. Furthermore, electrical rectification was found by the author thereby verifying that the device was indeed a diode, that is, an LED.

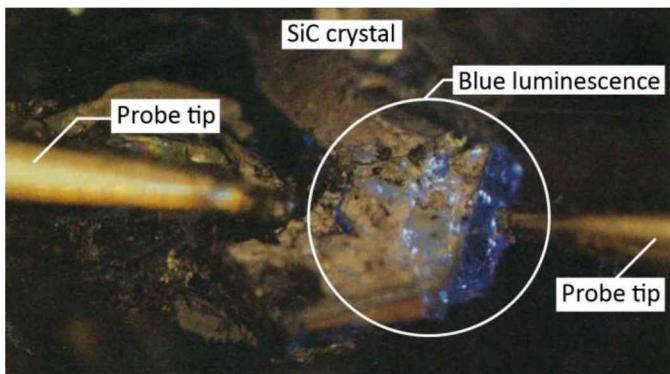


Figure 1.10: Reenactment of Henry Round's electroluminescence experiment by Theodore Gray. The author used a SiC crystal and a 9 V battery. He observed a “very faint colored glow” when connecting electrical probes to the crystal. Electrical rectification was found by the author thereby verifying that the device was a diode (after Gray, 2010).

1.7 References

- Eastman P. C., Haering R. R., and Barnes P. A. “Injection electroluminescence in metal–semiconductor tunnel diodes” *Solid-State Electronics* **7**, 879 (1964)
- Edmond J. A., Kong H. S., and Carter Jr. C. H. “Blue LEDs, UV photodiodes and high-temperature rectifiers in 6 H-SiC” *Physica B* **185**, 453 (1993)
- Filsinger D. H. and Bourrie D. B. “Silica to silicon: Key carbothermic reactions and kinetics” *Journal of the American Ceramic Society* **73**, 1726 (1990)
- Gray T. “A light mystery” *Popular Mechanics*, March issue, page 70 (2010)
- Jacobson N. S., Lee K. N., and Fox D. S. “Reactions of SiC and SiO₂ at elevated temperatures” *Journal of the American Ceramic Society* **75**, 1603 (1992)
- Lehovec K., Accardo C. A. and Jamgochian E. “Injected light emission in silicon carbide crystals” *Physical Review* **83**, 603 (1951)
- Lehovec K., Accardo C. A. and Jamgochian E. “Light emission produced by current injected into a green silicon carbide crystal” *Physical Review* **89**, 20 (1953)

- Lehovec K. < www.KurtLehovec.com > (accessed in June 2008)
- Loebner E. E. "Subhistories of the light-emitting diode" *IEEE Transactions on Electron Devices* **ED-23**, 675 (1976)
- Lossev O. V. "Wireless telegraphy and telephony" *Telegrafia i Telefonia*, No. **18**, page 61 (1923)
- Lossev O. V. "Wireless telegraphy and telephony" *Telegrafia i Telefonia*, No. **26**, page 403 (1924a)
- Lossev O. V. "Oscillating crystals" *Wireless World and Radio Review* **271**, 93 (1924b)
- Lossev O. V. "Luminous carborundum detector and detection effect and oscillations with crystals" *Philosophical Magazine* **6**, 1024 (1928)
- Lossev O. V. (note that there are two different spellings for Lossev's name: O. V. Lossev and O. W. Lossew; the present publication uses the latter spelling) "Über den lichtelektrischen Effekt in besonderer aktiver Schicht der Karborundkristalle" ("On the light-electric effect in particular active layer of carborundum crystals" translated from German) *Physikalische Zeitschrift* **34**, 397 (1933)
- Lossev O. V. (note that there are two different spellings for Lossev's name: O. V. Lossev and O. W. Lossew; the present publication uses the latter spelling) "Die spektrale Verteilung des Sperrsichtphotoeffektes an Karborundum-Einkristallen" ("The spectral distribution of the blocking-layer-photo-effect in carborundum single crystals" translated from German) *Compt. Rend. Acad. Sci. U.R.S.S.* **29**, 363 (1940)
- Potter R. M., Blank J. M., and Addamiano A. "Silicon carbide light-emitting diodes" *Journal of Applied Physics* **40**, 2253 (1969)
- Round H. J. "A note on carborundum" *Electrical World* **49**, 309 (1907)
- Schubert E. F. and Kim J. K. "Light-emitting diodes hit the centenary milestone" *Compound Semiconductors*, October issue, page 20 (2007)
- Shah J. M. and Mont F. W.; unpublished; experiments showing electroluminescence in SiC were conducted at Rensselaer Polytechnic Institute (located in Troy, New York) thereby reenacting Henry Joseph Round's 1907 experiments (2007)
- Shockley W. B. "Electrons and holes in semiconductors with applications to transistor electronics" (Van Nostrand, New York, 1950)
- Violin E. E., Kalnin A. A., Pasynkov V. V., Tairov Y. M., and Yaskov D. A. "Silicon Carbide – 1968" *2nd International Conference on Silicon Carbide*, published as a special issue of the *Materials Research Bulletin*, page 231 (1969)
- Wikipedia, the biggest multilingual free-content encyclopedia on the Internet < www.wikipedia.org > (accessed in September 2007)
- Yunovich A. E., Moscow State University, Russia, personal communication, (2008)
- Zheludev N. "The life and times of the LED – a 100-year history" *Nature Photonics* **1**, 189 (2007)

2**2 – History of III–V light-emitting diodes*****2.1 Compound semiconductors***

Compound semiconductors include III–V semiconductors (such as GaAs and GaN), II–VI semiconductors (such as ZnS and CdSe), as well as IV–IV semiconductors (such as SiGe and SiC). II–VI semiconductors are abundant in nature, for example in the form of pyrite (FeS_2), sphalerite (ZnS), chalcopyrite (CuFeS_2), and galena (PbS). It is thus not surprising that the first measurements on compound semiconductors were carried out on II–VI semiconductors.

In 1874, Karl Ferdinand Braun (1874), a German scientist, reported on the electrical conduction in metal–semiconductor structures where the semiconductor was a metal sulfide such as ZnS, PbS, FeS_2 , and CuFeS_2 . He found that the current–voltage characteristic was asymmetric and non-ohmic and thus exhibited some degree of rectification. His work is the first demonstration of the rectifying properties of metal-semiconductor contacts and the first scientific research in compound semiconductors.

It may be noted that in 1936 Georges Destriau (1936) reported on electroluminescence on thin phosphor layers located between two very thin dielectric layers and metal electrodes when applying a high alternating-current electric field. The phosphor layer was prepared by embedding ZnS powder in a dielectric matrix. Due to the high electric field in the ZnS layer, doped with the optically active element Mn, electrons are accelerated in the single ZnS crystallites, excite the Mn activator by raising an electron to a higher quantum level, from which it can return to its ground state by emitting a photon. Note, however, that the luminescence effects found by Destriau are not due to a rectifying metal–semiconductor or pn-junction luminescence and thus are not a light-emitting diode.

A significant leap forward in compound semiconductor technology was the first demonstration of III–V semiconductors. In contrast to II–VI semiconductors, III–V semiconductors do not occur in nature at all and thus were a novel substance that could only be made by humankind. The era of III–V compound semiconductors started in the early 1950s when this class of materials was postulated and demonstrated by Heinrich Welker (1951a, 1951b, 1952, and

1953). A patent, invented by Welker, discloses the conception of a new class of materials, i.e. III–V semiconductors (Welker, 1951a). Claim 1 of the patent reads:

“1. Electrical semiconductor device, such as a rectifier, controllable crystal amplifier, photo- or thermo-electrical device, and control devices, which are characterized by the use of a semiconductor that is a compound with an atomic ratio of 1:1 consisting of one of the elements boron (B), aluminum (Al), gallium (Ga), indium (In), which belong to group-III of the periodic system with one of the elements nitrogen (N), phosphorus (P), arsenic (As), antimony (Sb), which belong to group-V of the periodic system.” (Translated from German; Welker, 1951a)

The class of III–V compounds had been an unknown substance prior to the 1950s. The novel manmade III–V compounds proved to be optically very active and thus instrumental to modern LED and laser technology. From 1951 to 1961, Heinrich Welker headed the solid-state physics department of the Siemens-Schuckert Company, in Erlangen, Germany, where he developed III–V compounds. His research ultimately resulted in large-scale use of optoelectronic and microwave devices. A photograph of Heinrich Welker is shown in *Figure 2.1*¹



Fig. 2.1: Heinrich Welker, pioneer of III–V compounds. His research and development effort, carried out in the 1950s, led to a new class of materials, III–V compounds, which enabled the establishment of the fields of optoelectronics and microwave electronics (after Siemens Corporation, 2007).

Commercial bulk growth of the III–V compound GaAs commenced in 1954. In the mid-1950s, large single-crystal boules of GaAs were pulled from the melt. The sliced and polished wafers were used as substrates for the epitaxial growth of p-n junction diode structures, either by vapor-phase epitaxy (VPE) or liquid-phase epitaxy (LPE). Infrared (870~980 nm) LEDs and lasers based

¹ One may note that Heinrich Welker proposed and demonstrated *binary* III–V semiconductors. *Ternary* and *quaternary* III–V semiconductors were first proposed and demonstrated by Otto G. Folberth, to be discussed below in this chapter.

on GaAs were first reported in 1962 by groups working at RCA, GE, IBM, and MIT (Hall *et al.*, 1962; Nathan *et al.*, 1962; Pankove and Berkeyheiser, 1962; Pankove and Massoulie, 1962; Quist *et al.*, 1962). A comprehensive review of early light-emitting diodes made from III–V compounds was given by Bergh and Dean (1972).

Compound semiconductors include *binary* semiconductors such as GaN, *ternary* alloys such as $\text{Al}_x\text{Ga}_{1-x}\text{As}$, and *quaternary* alloys such as $\text{Ga}_{1-x}\text{In}_x\text{P}_y\text{As}_{1-y}$. A ternary semiconductor may be written as “ $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ” but also as “(AlGa)As” or simply as “AlGaAs”; the former style (i.e. “ $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ”) is most accurate; the latter styles (i.e. “(AlGa)As” and “AlGaAs”) are nevertheless frequently used.

2.2 The first III–V compound semiconductor LED

The advent of *infrared* (IR) LEDs made from III–V semiconductors dates back to 1955 when Braunstein (1955) reported the first electroluminescence from n-type GaAs and n-type GaSb. Braunstein’s LEDs were inefficient, not based on a pn-junction, and instead based on a rectifying metal-semiconductor contact (Schottky contact). The author found that the IR emission energy is approximately equal to the bandgap energy, i.e. $h\nu \approx E_g$.

The advent of *visible* LEDs made of a III–V semiconductors also dates back to 1955 when Wolff *et al.* (1955) demonstrated a single-crystal-GaP LED that emitted in the visible spectrum. The authors observed light emitted by the device to be of a “brilliant, deep red” as well as “orange” color. The abstract of the publication is shown in *Figure 2.2*.

Electroluminescence of GaP

G. A. WOLFF, R. A. HEBERT, AND J. D. BRODER
Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey
 (Received June 10, 1955)

Electroluminescence has been observed in GaP single crystals. Above a threshold current the intensity of the emitted light is proportional to the ac and dc passing through the crystals. The orange color of the light emission probably is due to the presence of Ga as activator, while the brilliant, deep red luminescence is observed in crystals where Zn is a possible activator. The temperature dependence of the luminescence is negligible in the range from liquid nitrogen temperature to 160°C. The impact excitation mechanism of Piper and Williams is proposed.

Fig. 2.2: Abstract of “Electroluminescence of GaP” reporting the first LED made of a III–V semiconductor, GaP (Wolff *et al.*, 1955).

Wolff *et al.* (1955) suggested two possibilities as origin of the electroluminescence: (i) Minority carrier generation at the metal–GaP-junction (Schottky junction) by means of impact ionization and subsequent radiative recombination, and (ii) pn-junction electroluminescence. The authors stated:

“It was found that in one out of ten cases luminescence did not occur in the vicinity of the cathode. From this, it may be concluded that recombination of injected minority carriers from a p-n junction captured by the activator system [...] is a second possibility for luminescence, although it is as yet impossible to prepare p-n junctions in GaP crystals to confirm this point.” (Wolff *et al.*, 1955)

Since GaP is an indirect-gap semiconductor, the luminescence was attributed to a transition via an energy level caused by an impurity (“activator” such as Zn) located within the bandgap of GaP. The Wolff *et al.* (1955) publication marks the beginning of III–V semiconductor LEDs, specifically the beginning of III–V visible LEDs.

By 1959, electroluminescence from a forward-biased pn junction had become a well-accepted and proven concept (Vavilov, 1959). Furthermore, it became apparent that the extraction of light from a semiconductor body with a relatively high refractive index was a problematic issue for LEDs. Vavilov (1959) stated:

“When electric excitation is used, i.e., injection across a p-n junction, an important problem is that of bringing about the emergence of the radiation from the specimen. In fact, for the simplest experimental geometry [i.e. for an LED with a rectangular cross section], a considerable part of the light flux is subject to total internal reflection, and does not reach the entrance slit of the spectrometer.”

During the following years, light emission from forward-biased GaP pn junctions became a common experiment. The associated quantum efficiencies were estimated to be about 1% at 80 K, and $\ll 1\%$ at 300 K (Gershenson and Mikulyak, 1962).

The first commercial availability of III–V pn-junction LEDs started in October 1962 when the Texas Instruments (TI) Company began offering for sale a GaAs infrared LED, the TI Model SNX-100 LED, for 130 US\$ per device (Rostky, 1997; Totzke, 2014). The technology of the device is described in a patent, co-invented by Biard and Pittman, filed in August 1962, and issued in 1966 (Biard and Pittman, 1966).

2.3 History of $\text{GaAs}_y\text{P}_{1-y}$ alloy semiconductors and LEDs

Whereas Heinrich Welker discovered binary III–V semiconductors, it was another important step to demonstrate *ternary* and *quaternary* III–V semiconductors, i.e. III–V **alloy semiconductors**. Otto Gert Folberth (Folberth 1955, 1959) first proposed ternary III–V alloy semiconductors and first demonstrated the complete miscibility of two binary compounds, such as GaAs and GaP,

which form the alloy $\text{GaAs}_y\text{P}_{1-y}$. Folberth (1955) stated:

“[In addition to InAs and InP,] GaAs and GaP also form mixed crystals without a miscibility gap. Our investigations were initially limited to crystals with more than 50% GaP. These crystals are transparent so that the change in E_g can be instantly seen. For $\text{Ga}(\text{As}_y\text{P}_{1-y})$ with $0 < y < 0.1$, the color of the samples is orange and barely distinguishable from GaP ($E_g = 2.24$ eV). For $0.1 < y < 0.3$ the color becomes more reddish, for $y \approx 0.4$ red, and for $y \approx 0.5$ dark red, and for $y > 0.6$ the samples are opaque. [...] In the quasi-binary systems InAs-InP and GaAs-GaP, miscibility-gap-less mixed-crystal formation occurs. Therefore, it is possible to demonstrate A^{III}B^V compounds with arbitrary E_g between 0.33 and 1.25 eV as well as 1.45 and 2.25 eV.”

(translated from German)

In addition, Folberth first proposed and demonstrated pn-junction devices made of III–V ternary alloy semiconductors (Folberth, 1958); Folberth (1958) stated:

“My invention relates generally to semiconductor devices. Such devices have a crystalline body of electrically semiconducting substance subjected to electric or magnetic fields, to corpuscular or wave radiation or to a plurality of such phenomena, for performing electrical, photo-electrical, optical or other physical effects. [...]

I found that the melting, zone-melting or other heat treatment does not result in decomposition or segregation of the mixed crystal, and also that the mixed crystals can be doped with substitutional impurities (lattice-defect atoms), for instance as needed for producing p-n junctions. Doping the crystal with an element of the second periodic group, preferably cadmium or zinc (acceptors), produces p-type conductance. Elements of the sixth group, preferably sulphur, selenium or tellurium, act as donors, i. e. produce n-type conductance.

I have found that among the compound mixed crystal semiconductors according to the invention those of the groups $\text{In}_1(\text{As}_y\text{P}_{1-y})$ and $\text{Ga}_1(\text{As}_y\text{P}_{1-y})$ are particularly advantageous.”

Ternary and quaternary III–V alloys have an advantage that is particularly relevant for LEDs: By changing the chemical composition of the alloy, its bandgap energy can be changed so as to attain a desired emission wavelength. Given his published accomplishments, Otto Gert Folberth can be credited with the invention and first demonstration of III–V alloy semiconductors, including n-type and p-type $\text{GaAs}_y\text{P}_{1-y}$ as well as pn-junction devices made thereof. A photograph of Otto Gert Folberth is shown in **Figure 2.3 (a)**.²

² One may note that Folberth has been credited by people familiar with his work but less so by the technical community at large. One of the people who credited Folberth was former IBM executive Karl Ganzhorn, who, on the occasion of Folberth's 80th birthday, called him as a “co-founder of semiconductor technology” (see Hans Barth in *Siebenbürgische Zeitung*, published on November 27, 2004). The article recognizes Folberth for his contributions to the field of red LEDs: “Red LEDs, used worldwide in equipment and machinery of all kinds are produced by means of

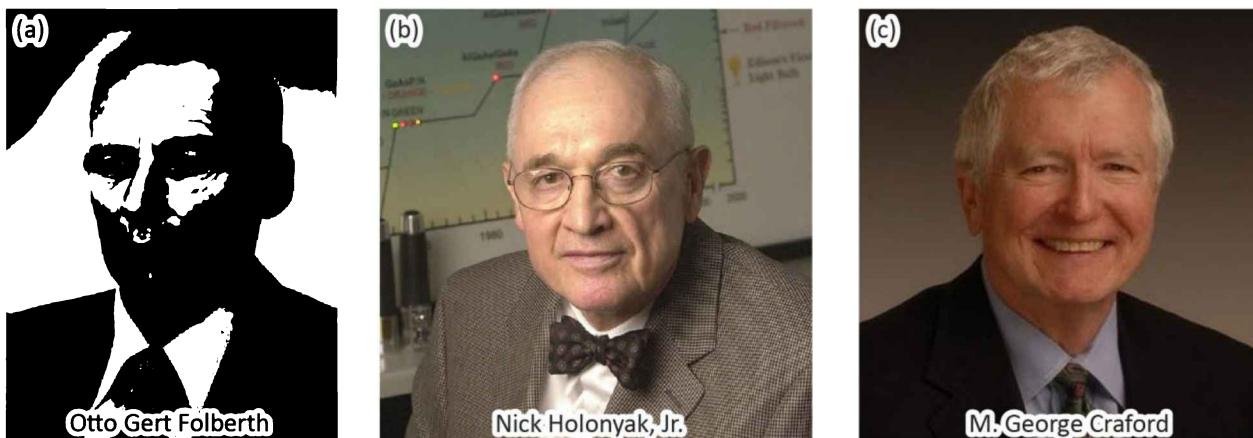


Fig. 2.3: (a) Otto Gert Folberth, inventor of GaAsP pn-junction devices (after Siebenbürgische Zeitung, 2004). (b) Nick Holonyak, Jr., LED pioneer who in 1962 demonstrated the first GaAsP pn-junction LED and laser (after <www.technology.gov>, 2007). (c) M. George Crawford, LED pioneer and leader at multiple companies (after University of Illinois, 2008).

Another important step in the development of $\text{GaAs}_y\text{P}_{1-y}$ (or simply GaAsP) LEDs occurred in 1962 when Holonyak and Bevacqua (1962) reported on the emission of coherent visible light from GaAsP p-n junctions at low temperatures (77 K) under pulsed current injection. The emission of coherent light was limited to low temperatures. At room temperature, the devices worked as LEDs and emitted incoherent visible red light. The 1962 publication is considered a milestone in the development of pn-junction LEDs made from III–V semiconductors emitting in the visible wavelength range (Holonyak, 1987).

Nick Holonyak Jr., who in 1962 worked at General Electric in Syracuse, New York, and who later joined the University of Illinois, had used a halogen-vapor-transport procedure for the synthesis of the GaAsP crystals. Holonyak (2000) recalled that when he first had made these LEDs, he had already envisioned many applications these new devices might have, including indicator lights, seven-segment numeric displays, and alphanumeric displays. A photograph of Nick Holonyak Jr. is shown in *Figure 2.3 (b)*. Nick Holonyak Jr. is also credited with the education of a new generation of engineers who became recognized leaders in the field of LEDs. These leaders include his former graduate student M. George Crawford whose contributions to the LED industry will be further discussed below.

However, despite the success of the Holonyak group, the goal of demonstrating a semiconductor laser, working at room temperature, remained elusive (Holonyak, 1963, 1964). It

a patent that names Otto G. Folberth as its inventor" (translated from German).

remained elusive for good reasons: The GaAsP material system grown on GaAs substrates has several problems.

Although excellent electrical junction characteristics were obtained (Holonyak *et al.*, 1963a), the optical properties degraded when phosphorus (P) was added to GaAs. When the phosphorus content in GaAsP reached about 45~50%, a strong decrease in the LED radiative efficiency was found. These difficulties were attributed to the direct–indirect transition of the bandgap of GaAsP (Holonyak *et al.*, 1963b, 1966; Pilkuhn and Rupprecht, 1964, 1965). It was determined that the 300 K efficiency of GaAsP alloy devices dropped to less than 0.005% when the phosphorus concentration exceeded 44% (Maruska and Pankove, 1967).

In 2014, within days after the Nobel Prize in physics had been awarded jointly to two Japanese researchers and one Japanese-American researcher (for the invention of the GaN blue LED), the News-Gazette (2014) reported: “Holonyak said he was disappointed and irritated” for not having been recognized as a Nobel laureate for his contributions to the field of LEDs. The IEEE Spectrum (2014) reported Holonyak as saying: “I find this one insulting.” The news articles purported that Holonyak made “the first visible LED” and “the first III–V semiconductor alloy”. Furthermore, Johnstone (2007) credits Holonyak’s work as follows: “The result was gallium arsenide phosphide, or GaAsP, the world’s first-ever semiconductor alloy.” However, such claims appear to negate contributions of Wolff *et al.* (1955) and Folberth (1955, 1958, 1959). The “world’s first-ever semiconductor alloy” was certainly demonstrated by Folberth (1955, 1958, 1959). One also may note that GaAs IR LEDs were already a commercial product prior to the December 1962 publication of Holonyak and Bevacqua (1962). Nevertheless, Holonyak’s pioneering “work stands on its own merits”, as Holonyak observed (News-Gazette, 2014), and there is not an iota of doubt about the value and impact of his contributions to technology and education.

The first commercial GaAsP LED was offered by the General Electric (GE) Company in the early 1960s. The LED emitted visible radiation in the red part of the spectrum. The manufactured quantities of the product were low, probably due to the high price, which was 260 US\$ for a single LED. The product was offered in the Allied Radio catalog, a widely distributed catalog for amateur radio electronics (Rostky, 1997).

A GaAsP LED manufactured in the 1960s is shown in *Figure 2.4* (Schanda, 2007). This early LED was packaged in a transparent glass tube. The devices did not have a transparent organic

encapsulation material surrounding the semiconductor chip, which became customary only at a later time. The cathode terminal of an LED is frequently marked, e.g. by a colored ring (“cathode mark”), as shown in the figure.

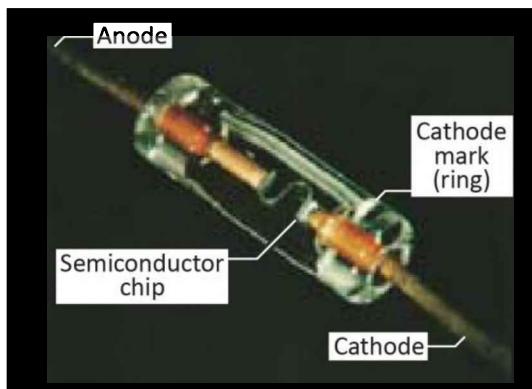
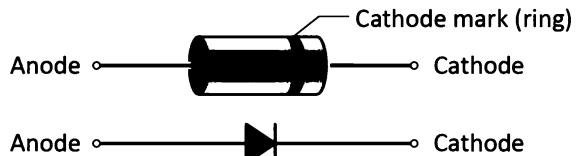


Fig. 2.4: Photograph of an early GaAsP red-emitting LED packaged in a glass tube. This device was manufactured in the 1960s (courtesy of János Schanda, 2007). The cathode of an LED is frequently marked (“cathode mark”), e.g. by a colored ring.



The Monsanto Company was the first commercial entity to mass produce LEDs. In 1968, the company had set up a factory, produced low-cost GaAsP LEDs, and sold them to customers. The era of commercial solid-state lamps had started. In the period 1968~1970, sales were skyrocketing, doubling every few months (Rostky, 1997). The Monsanto LEDs were based on GaAsP p-n junctions grown on GaAs substrates emitting in the visible red wavelength range (Herzog *et al.*, 1969; Craford *et al.*, 1972).

Monsanto developed a friendly collaboration with Hewlett-Packard Company, expecting Hewlett-Packard would make LEDs while Monsanto would provide the raw material – GaAsP. In the mid-1960s, Monsanto sent one of its scientists from Saint Louis, Missouri, to Palo Alto, California, to help Hewlett-Packard develop the LED business using Monsanto’s GaAsP material. However, Hewlett-Packard felt nervous about depending on a single source for the GaAsP material. The informal relationship ended and Hewlett-Packard started growing its own GaAsP (Rostky, 1997).

For several years, from the late 1960s to the mid-1970s, the emerging market was in numeric LED displays, driven at first by calculators, then by wristwatches, following Hamilton Watch Company’s introduction of the Pulsar digital watch in 1972. For a while, the early contenders, Monsanto and Hewlett-Packard, took turns leaping into first place with a more advanced multiple-numeric or alphanumeric LED display (Rostky, 1997).

A key technical innovator and manager at Monsanto was M. George Craford, who has made numerous contributions to LEDs including the first demonstration of a yellow LED (Craford *et al.*,

1972). It employed an N-doped GaAsP active region grown on a GaAs substrate. When Monsanto sold off its optoelectronics business in 1979, Craford joined Hewlett-Packard and became the key person in the company's LED business. A profile of Craford, who for many years served as Chief Technical Officer, was published by Perry (1995). A photograph of M. George Craford is shown in **Figure 2.3 (c)**. In 1999, Hewlett-Packard spun off parts of its business (including the LED business) into Agilent Company which in turn co-founded Lumileds Lighting Company in 1999, as a joint venture with Philips Company. In 2005, Agilent sold its share of Lumileds to Philips.

It soon became clear that the large lattice mismatch between the GaAs substrate and the GaAsP epilayer resulted in a high density of dislocations (Wolfe *et al.*, 1965; Nuese *et al.*, 1966). As a result, the external efficiency of these LEDs was quite low, about 0.2% or less (Isihamatsu and Okuno, 1989). The importance of the growth conditions and thickness of a *buffer layer* was realized by Nuese *et al.* (1969) who pointed out that a thick graded GaAsP buffer layer yields improved brightness red LEDs. It is understood today that the thick graded buffer layer reduces the high dislocation density in the GaAsP epitaxial layer originating near the GaAsP-epilayer-to-GaAs-substrate boundary.

The direct–indirect transitions as well as the high dislocation density limit the efficiency attainable with GaAsP / GaAs LEDs. Although this material system is of historical interest, it is rarely employed at the present time.

2.4 History of GaAs and AlGaAs infrared and red LEDs

A sustained research effort on GaAs and AlGaAs/GaAs devices started in the early 1960s at the IBM Thomas J. Watson Research Center in Yorktown Heights, located about an hour's drive north of New York City. The IBM team consisted of well-known researchers such as Jerry Woodall, Hans Rupprecht, Manfred Pilkuhn, Marshall Nathan, and others.

Woodall (2000) recalls that his work centered on the bulk crystal growth of GaAs used to fabricate semi-insulating substrates for Ge device epitaxy, and n-type substrates to fabricate injection lasers via Zn diffusion. At that time, the GaAs-based injection laser had already been demonstrated at IBM, GE, and MIT Lincoln Laboratories. Rupprecht's interests were in impurity-diffusion theory and experiment along with experimental investigations into the newly discovered injection laser. Rupprecht was associated with a laser device physics group headed by Marshall Nathan, a co-inventor of the first injection laser (Nathan *et al.*, 1962).

As Woodall developed a technique that lead to state-of-the-art horizontal Bridgman GaAs crystals, Rupprecht fabricated the materials into lasers and characterized them. This collaboration paid off immediately and continuous-wave (cw) operation of GaAs lasers at 77 K was attained (Rupprecht *et al.*, 1963). They then learned of the liquid-phase epitaxy (LPE) technique pioneered by Herb Nelson at the RCA Laboratories in Princeton. The employment of LPE to grow GaAs lasers resulted in the achievement of 300 K lasers with lower threshold current densities than for Zn-diffused lasers. Stimulated by papers found in a literature search, Woodall set out to grow GaAs p-n junction diodes by using Si as an amphoteric dopant, i.e. Si atoms on Ga sites acting as donors and Si atoms on As sites acting as acceptors. This was an interesting idea, as hitherto LPE had been used to grow epilayers with only a single conductivity type.

The LPE conditions to form Si-doped p-n junctions were found very quickly. Si-doped GaAs p-n junctions were formed by cooling a Ga-As-Si melt from 900 to 850°C to form Si donors and Si acceptors at the two temperatures, respectively. By examining the cross section of the chemically stained epitaxial layer, the lower layer, grown at 900°C, was identified as being an n-type layer and the upper layer, grown at 850°C, as a p-type layer. No loss in crystal quality was found in the regions of lower temperature growth. Furthermore, owing to band tailing effects caused by the highly doped, compensated region of the p-n junction, the LED emission occurred at 900~980 nm, far enough below the GaAs band edge (870 nm), so that the bulk GaAs substrate and the GaAs epilayer did not absorb much of the emitted light but acted as a transparent “window layer”. LED external quantum efficiencies as high as 6% were attained, a major breakthrough in LED technology (Rupprecht *et al.*, 1966). Rupprecht (2000) stated: “Our demonstration of the highly efficient GaAs LED is a typical example of a discovery made by serendipity.” The quantum efficiency of the amphotERICALLY doped GaAs LEDs was five times greater than that of GaAs p-n junctions fabricated by Zn diffusion. Si acceptor levels are deeper than Zn acceptor levels so that the emission from the compensated Si-doped active region occurs at longer wavelengths where GaAs is more transparent.

Being in the LED research business, the IBM group wondered if this doping effect could be extended to a crystal host with visible emission. There were two candidates, GaAsP and AlGaAs. Whereas Rupprecht tried to do GaAsP epitaxy via LPE, Woodall set up an apparatus for AlGaAs. It was difficult to form good quality GaAsP epilayers by LPE due to the 3.6% lattice mismatch between GaP and GaAs. AlGaAs had problems of its own: “AlGaAs is lousy material” was the

pervasive opinion at that time, because, as Woodall (2000) stated, “aluminum loves oxygen”. The “love” promotes the incorporation of the “luminescence killer” oxygen in AlGaAs; in particular, in the vapor-phase epitaxy (VPE) process, but less so in the LPE process.

Without the support of IBM management, Rupprecht and Woodall “went underground” with their research, conducting the LPE AlGaAs epitaxial growth experiments after regular working hours and on the weekends. Woodall designed and built a “vertical dipping”-type LPE apparatus, using graphite and alumina melt containers. As an undergraduate student Woodall had majored in metallurgy at MIT and he remembered something about phase diagrams. He made an “intelligent guess” to select the Al concentrations for the LPE melts. He added Si to the melt for the first experiment, saturated the melt and then “dipped” the GaAs substrate while cooling the melt from about 925 to 850°C. Finally, the substrate and epilayer were withdrawn from the melt, and the apparatus was returned to 300 K. Although no Si-doped p-n junction was observed, a 100 µm thick high-quality layer of AlGaAs had been grown with a bandgap in the red portion of the visible spectrum (Rupprecht *et al.*, 1967, 1968).

AlGaAs visible LEDs were also grown on GaP, a lattice-mismatched yet transparent substrate. **Figure 2.5** shows micrographs of the structure. When AlGaAs was grown on GaP substrates, the thermodynamics of LPE made the initially grown material Al-richer due to the Al distribution coefficient in the melt. As a result, the high-Al-content AlGaAs acts as a transparent window layer for light emitted from the low-Al-content AlGaAs active region (Woodall *et al.*, 1972).

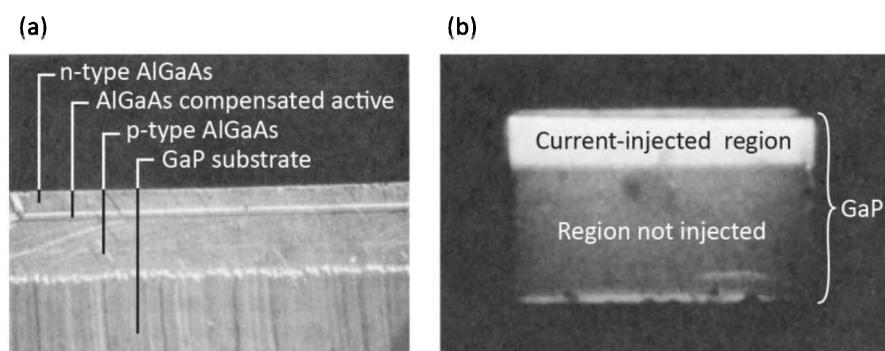


Fig. 2.5: (a) Cross section micrograph of AlGaAs LED grown on transparent GaP substrate. (b) Electroluminescence originating from current-injected region located under stripe-shaped contact viewed through transparent GaP substrate (after Woodall *et al.*, 1972).

Pilkuhn, also an “IBM’er” who had worked with Rupprecht on GaAsP LEDs and lasers (Pilkuhn and Rupprecht, 1965), had built a small battery-powered circuit with an LED emitting visible red light, which he showed to his colleagues and management at IBM (Pilkuhn, 2000). The reactions ranged from “nice but useless” to “great and useful”. However, it was soon realized that the latter was true, i.e. that LEDs were extremely useful devices. The first application of the GaAsP LEDs was as

indicator lights on circuit boards, where the LEDs indicated the status and proper function of the circuit board. LEDs were also used to show the status of the data processing unit of the classic IBM System/360 mainframe computer shown in *Figure 2.6*.



Fig. 2.6: Control panel of the classic 1964 mainframe computer “IBM System/360”. The computer initially used high-voltage gas-discharge lamps to indicate the status of the arithmetic unit. In later models, the status-indicator lamps were replaced by red LEDs. The cabinet-sized IBM System/360 was able to execute 34 500 instructions per second and had a main memory of 64 kbytes.

According to Rostky (1997), the first commercial GaAs LED was offered by the Texas Instruments Company in the early 1960s. The LED emitted infrared radiation near 870 nm. The manufacturing quantities of the product were low, probably caused by the high price for one LED, which reportedly was 130 US\$.

At the present time, infrared GaAs/AlGaAs LEDs are widely used in video and audio remote controls and as sources for local-area communication networks. In addition, red AlGaAs/AlGaAs LEDs are used as high-brightness visible LEDs with efficiencies higher than the GaAsP/GaAs red LEDs but lower than the AlGInP/GaAs red LEDs.

2.5 History of GaP and GaAsP LEDs doped with optically active impurities

Ralph Logan’s and his co-workers’ pioneering work on GaP LEDs was done while working at AT&T Bell Laboratories in Murray Hill, New Jersey, in the early 1960s, where they developed a manufacturing process for GaP-based red and green LEDs. At that time, semiconductors had been employed to demonstrate both bipolar and field-effect transistors for switching and amplifying electrical currents. Engineers and scientists back then also began to realize that semiconductors would be perfectly suitable for light-emitting devices.

Logan (2000) recalls that his interest was stimulated by the first reports of GaP p-n junction LEDs by Allen *et al.* (1963) and Grimmeiss and Scholz (1964) (for a review, see Grimmeiss and Allen, 2006). These devices emitted red light at a useful efficiency so that the light could be clearly

seen with the naked eye under ambient daylight conditions. The Grimmeiss–Scholz junctions had been reported to be made by alloying Sn, an n-type dopant, into p-type GaP. Prior to the Grimmeiss–Scholz junctions, Wolff *et al.* (1955) had reported electroluminescence emanating from a reverse-biased GaP Schottky contact, which was attributed to impact ionization and subsequent recombination of carriers.

GaP is an indirect-gap semiconductor that does not emit significant amounts of light due to the requirement of momentum conservation in optical transitions. **Figure 2.7** shows the band diagram of GaP illustrating that the band extremum points occur at different values in momentum space. If GaP is doped with an ***optically active isoelectronic impurity*** such as N, strong optical transitions are obtained, as shown by Thomas *et al.* (1965), because the impurity levels are smeared out in momentum space. GaP doped with optically active impurities is a wonderful example of a practical device based on the ***Heisenberg uncertainty principle***, which predicts that an impurity with a strongly localized wavefunction in position space (small Δx) will have a delocalized level in momentum space (large Δk), so that optical transitions can occur via the deep-level state.

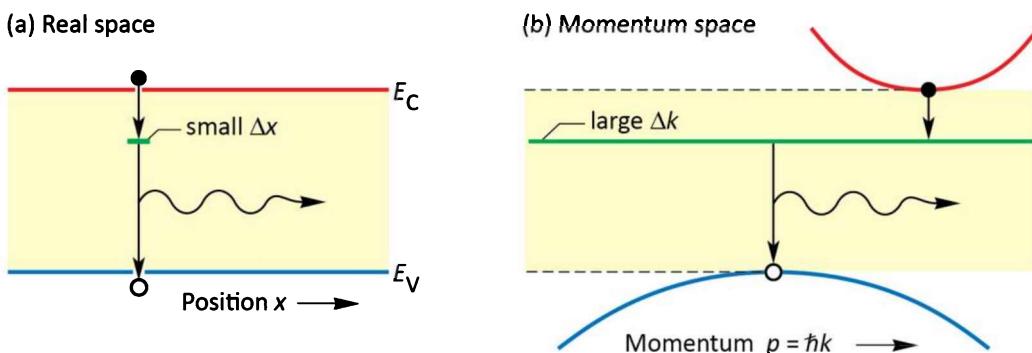


Fig. 2.7: (a) Real-space and (b) momentum-space optical transitions in GaP doped with an optically active impurity such as O or N, emitting in the red and green parts of the spectrum, respectively. GaP LEDs employ the ***uncertainty principle*** ($\Delta x \Delta p \geq h/2\pi$) which predicts that an electron wave function localized in real space is delocalized in momentum space, thereby making momentum-conserving (vertical) transitions possible.

The growth of GaP was accomplished by using platelets grown from a solution containing Ga and P. The platelets had lateral dimensions of $0.5\text{ cm} \times 1\text{ cm}$ and were grown to a thickness of about 1 mm. This was the standard initial method of growing GaP and overcame the problem of dealing with the P overpressure required when growing GaP at high temperatures. No one at Bell Laboratories could immediately reproduce the remarkable results reported by Grimmeiss and

Scholz. However, a big research effort in electroluminescence was launched at Bell Laboratories as a result.

In the solution growth of the GaP platelets, the dopants used were Zn and O (the latter from Ga_2O_3), but it was not generally realized that ordinary ambient air usually contains enough S to be a good n-dopant. The growth kinetics of the platelets had the intriguing result that the compensated melts generally produced an n-type layer on one platelet surface so that a p-n junction was formed under the GaP platelet surface. This was thought to explain Grimmeiss' results. Logan *et al.* (1967a) published these findings at once.

Logan's research group also reported the first demonstration of reproducible growth of efficient LEDs (Logan *et al.*, 1967b). These junctions were formed by growing an n-type GaP layer onto Zn-O-doped GaP wafers that were polished out of large solution-grown wafers with typical sizes of $2.5 \times 2.5 \text{ cm}^2$. Logan *et al.* found that post-growth annealing in the range $400\text{--}725^\circ\text{C}$ could increase the LED efficiency by as much as an order of magnitude, yielding efficiencies exceeding 2%. The annealing was thought to diffuse the Zn to the O atoms, thereby increasing the density of isoelectronic Zn-O centers that mediated the electroluminescence.

At the end of the 1960s, ingots of GaP grown from melts at high temperature and pressure were becoming available, suitable for being cut into real substrates as we know them today. Green LEDs were formed with efficiencies as high as 0.6% by doping the GaP with N isoelectronic impurities (Logan *et al.*, 1968, 1971). The N was added in the form of GaN to the growth melts used to form the p-n junctions. While the external quantum efficiency of green LEDs is less than for the red LEDs, the sensitivity of the human eye to green light is more than 10 times higher than in the red, so the apparent brightness of the LEDs is comparable.

Other research laboratories such as IBM, RCA, and GE also looked into the possibility of making visible LEDs that were more efficient than those made of GaAsP. Research on GaP LEDs was pursued at IBM's Thomas J. Watson Research Center in Yorktown Heights in New York State. Manfred Pilkuhn and co-workers demonstrated an LPE-grown red GaP LED doped with Zn and O. The picture of a GaP LED with top and bottom contacts is shown in *Figure 2.8*. The *IBM Research Journal* proudly reported that “brilliant red light” was emitted from the p-n junction. Note that in the 1960s, monochromatic colors were mostly generated by filtering incandescent light, so that the narrow-spectral-width LED light appeared to the observer as an impressively pure and “brilliant” color.

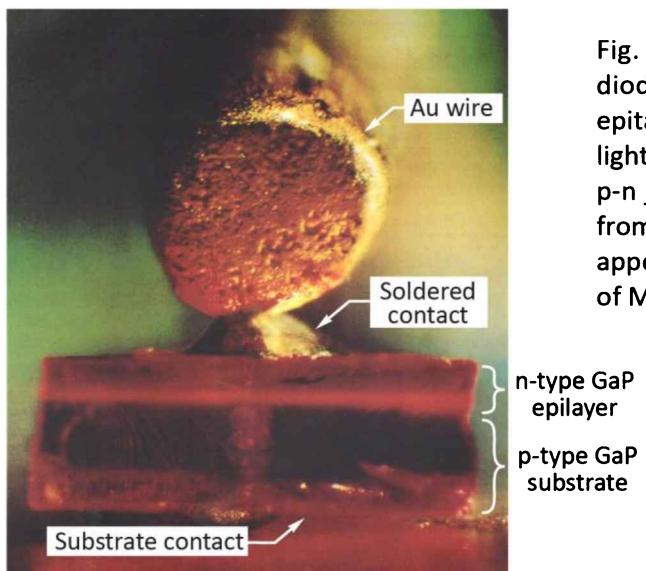


Fig. 2.8: GaP light-emitting diode, grown by liquid-phase epitaxy, emitting “brilliant red light” from the Zn- and O-doped p-n junction region (reproduced from *IBM Research Journal* that appeared in the 1960s; courtesy of Manfred H. Pilkuhn, 2000).

The active regions of Pilkuhn’s GaP LEDs were co-doped with acceptors, e.g. Zn acceptors, and donors, e.g. Te, S, or Se donors, so that light was generated predominantly by donor–acceptor pair recombination processes. The energy of the light was below the bandgap of GaP. It was also found that co-doping of GaP with Zn and O resulted in a particularly large wavelength shift so that the emission occurred in the red wavelength range (Foster and Pilkuhn, 1965). Oxygen in GaP is neither donor nor acceptor but was identified as a deep level (see, for example, Pilkuhn, 1981).

Logan and co-workers, and their management at AT&T Bell Laboratories immediately realized that there were many possible applications for LEDs. Indicator lamps were becoming useful in the telephone business. All such lights used at that time in the USA operated using 110 V. An example is the “Princess” telephone, which was intended to be used in bedrooms – the dial lit up when the telephone was picked up from its cradle. The “Princess” was a prestigious telephone and the latest fad in the 1960s but had to be installed near a 110 V outlet. A service call to the local telephone company was needed if the bulb burned out. If LEDs were to replace the 110 V light bulbs, the telephone line could power the LEDs and a 110 V outlet would no longer be needed. In addition, GaP LEDs had an expected lifetime exceeding 50 years when used in telephones (much longer than 110 V light bulbs) thereby promising substantial cost savings for the “Bell System” or simply “Ma Bell”, as the telephone company was called at that time.

More important was the multi-line “key telephone”. This is the multi-line telephone used in large offices mostly by operators and secretaries where indicator lamps tell which line is being called and busy. To switch the telephone lines and the 110 V indicator lamps, a remote switch

was used with dozens of wires to each telephone. Installing and servicing these telephones was very costly. In present-day telephones, the LED indicator lamps are powered over the telephone line. A compatible circuit inside the telephone handles the switching of the indicator lamps and telephone lines. The savings in telephone manufacturing, installation, and service were impressive.

With the demonstration of the reproducible growth of efficient green N-doped GaP LEDs and red Zn-O-codoped LEDs, both of which were about equally bright and useful, the Bell Laboratories Development Department decided to manufacture the LEDs at its Reading, Pennsylvania, facility.

Telephone lines typically operate with a line voltage of approximately 40 V DC with currents of several mA. The only effect of inserting an LED into this circuit is to reduce the drive voltage by approximately 2 V, a negligible effect, while the efficient LED makes a good indicator lamp. As a result, many telephone models were equipped with an illuminated dial pad. Both red and green LED illumination was available, and green was the final choice made by telephone designers.

Figure 2.9 shows a 1990 version of the AT&T Trimline telephone – still using GaP:N green LEDs for the dial pad illumination. Red and green LEDs were also incorporated in the multi-line “key telephones”.



Fig. 2.9: AT&T telephone set (“Trimline” model) with the dial pad illuminated by two green N-doped GaP LEDs. The illuminated dial pad was one of the first applications of green GaP:N LEDs.

Should the reader ever be near Murray Hill, New Jersey, visiting the Bell Laboratories Museum, located at 600 Mountain Avenue, should be considered. Many technical artifacts, including Logan *et al.*’s green GaP:N LED, are displayed in the museum.

The Monsanto team applied N-doping (nitrogen-doping) to GaAsP to attain emission in the red, orange, yellow, and green wavelength range (Groves *et al.*, 1971; Crawford *et al.*, 1972; for a

review see Duke and Holonyak, 1973). Many parameters, such as the emission and the absorption wavelength and the solubility of N in GaAsP and GaP were investigated. A useful growth method was vapor-phase epitaxy (VPE), because it allowed for N-doping in the vicinity of the p-n junction only. This resulted in less absorption of light in the layers adjoining the p-n junction and higher overall LED efficiencies (Groves *et al.*, 1977, 1978a, 1978b). Even today, GaP:N is in use for green emitters used in low-brightness applications such as indicator lights.

The first digital wristwatch with an LED display was released in 1972 by the Hamilton watch company. The watch became an instant furor and only its high price prevented it from becoming widely distributed. The digital Pulsar watch with an integrated calculator was released in 1975 and is shown in *Figure 2.10*.

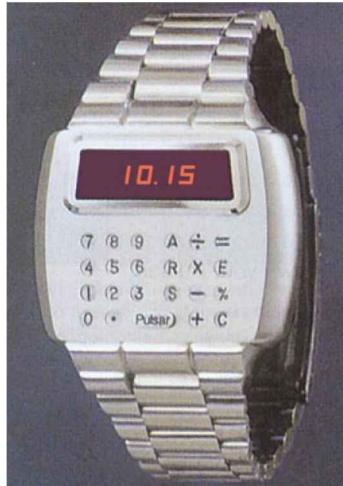


Fig. 2.10: Pulsar calculator watch with LED display released in 1975 by Hamilton Company. The first Pulsar LED watch (without calculator) was offered in 1972. It was based on GaAsP LEDs (after Seiko, 2004).

Another early application of LEDs was the numeric display in pocket calculators. *Figure 2.11* shows two programmable calculators of the mid-1970s, the Texas Instruments' SR-56 and the Hewlett-Packard's HP-67. Both used red GaAsP LEDs in the seven-segment numeric display. All calculators using LED displays shared a significant problem: the display could not be read under bright outdoor conditions because the light emitted by the LEDs was simply too dim; furthermore, the power consumption of LED displays was high so that the rechargeable battery running the calculator had to be recharged on a regular basis. Digital wrist watches using LEDs shared the same problem. Liquid crystal displays (LCDs), introduced at the end of the 1970s, had much lower power requirements. Consequently, LCDs totally replaced LED displays in calculators and watches by the beginning of the 1980s.

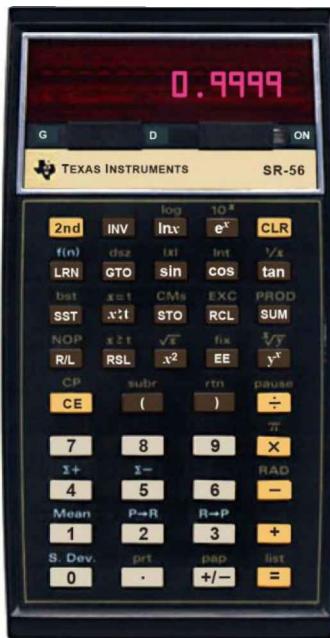
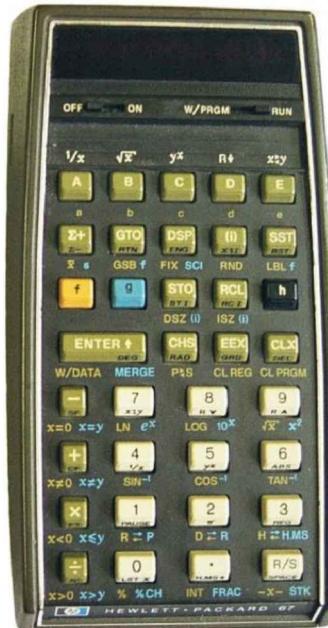


Fig. 2.11: Programmable pocket calculators Model SR-56 of the Texas Instruments Company and Model HP-67 of the Hewlett-Packard Company, both first manufactured in 1976. Seven-segment numeric characters composed of GaAsP LEDs were used in the display. The SR-56 came with a “huge” program memory of 100 steps. The HP-67 came with a magnetic card reader and had several freely programmable keys.



2.6 History of AlGaInP visible LEDs

The AlGaInP material system is suited for high-brightness emission in the red (625 nm), orange (610 nm) and yellow (590 nm) spectral range and today is the dominant material system for high-brightness emitters for these colors and associated wavelengths. **Figure 2.12** shows some of the common signage applications of red and yellow AlGaInP LEDs.



Fig. 2.12: Examples of red and amber AlGaInP/GaAs LEDs used in signage applications.

The AlGaInP material system was first developed in Japan for visible lasers (Kobayashi *et al.*, 1985; Ohba *et al.*, 1986; Ikeda *et al.*, 1986; Itaya *et al.*, 1990). Efforts started with AlGaInP/GaInP double-heterostructure lasers using $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$ as the active material, which is lattice matched to

GaAs substrates. The bandgap energy of lattice-matched GaInP is approximately 1.9 eV (650 nm), making the material suitable for visible lasers emitting in the red. These lasers are used, for example, in laser pointers and in digital video disc (DVD) players.

The addition of Al to the GaInP active region allows one to attain shorter emission wavelengths including the orange and yellow spectral region. However, $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$ becomes an indirect semiconductor at Al compositions of $x \approx 0.53$, so that the radiative efficiency strongly decreases at wavelengths near and, in particular, below 600 nm. Consequently, AlGaNp is not suited for high-efficiency emission at wavelengths below 570 nm.

Subsequent to the AlGaNp laser development that occurred in the early 1980s, AlGaNp LED development started at the end of the 1980s (Kuo *et al.*, 1990; Fletcher *et al.*, 1991; Sugawara *et al.*, 1991). In contrast to the AlGaNp laser structures, the LED structures typically employ current-spreading layers so that the entire p-n junction plane of the LED chip lights up and not just the region below the top ohmic contact. Further improvements were attained by using multiple quantum well (MQW) active regions (Huang and Chen, 1997), coherently strained MQW active regions (Chang and Chang, 1998a, 1998b), distributed Bragg reflectors (Huang and Chen, 1997; Chang *et al.*, 1997), transparent GaP substrate technology (Kish and Fletcher, 1997), and chip-shaping (Krames *et al.*, 1999). Comprehensive reviews of the AlGaNp material system and AlGaNp LEDs were published by Stringfellow and Crawford (1997), Mueller (2000), Krames *et al.*, (2002), Gessmann and Schubert (2004), and Dupuis and Krames (2008).

2.7 History of GaN metal–semiconductor emitters

In the late 1960s, the Radio Company of America (RCA) was one of the premier manufacturers of color televisions, using cathode ray tubes (CRT) with three electron guns to display images. At RCA's central research laboratory in Princeton, New Jersey, James Tietjen had become the director of the Materials Research Division, and he wanted to develop a flat-panel television display that could be hung on the wall like a painting. To create a full color image, the display must contain red, green, and blue pixels. Tietjen realized that red LEDs using GaAsP and green LEDs using GaP:N technology were already available. All that was needed for a flat TV based on LEDs was a bright blue LED.

In May 1968, Tietjen approached a young man in his group, Herbert Paul Maruska, and challenged him to find a method for growing single-crystal films of GaN, which Tietjen felt would

yield blue LEDs. Maruska had been growing GaAsP red LEDs using the metal-halide vapor-phase epitaxy (MHVPE) approach. He gained much experience with the promises and perils of III–V compounds including phosphorus, a pyrophoric substance. On a day in 1968, phosphorus caused a garbage truck to catch fire on New Jersey's Route 1 in Princeton shortly after picking up some phosphorus-containing laboratory waste at the RCA Laboratories. The driver of the truck decided to immediately return the burning and smoking load to RCA and dump it on the front lawn of the research laboratories (Maruska, 2000).

When Maruska started working on GaN, he first went to the library at Princeton University and thoroughly studied copies of all the old papers on GaN from the 1930s and 1940s (Juza and Hahn, 1938). GaN had been prepared as a powder by reacting ammonia with liquid gallium metal at elevated temperatures. He chose sapphire as the substrate because it is a robust material that is not reactive with ammonia. Unfortunately, he misinterpreted the results of Lorenz and Binkowski (1962), who had reported the decomposition of GaN in vacuum at temperatures as low as 600°C. All of his early GaN films were grown at temperatures below 600°C to prevent decomposition, and hence were polycrystalline. Finally, in March 1969, Maruska realized that in an ammonia environment, growth rather than decomposition would occur, and thus he raised the furnace temperature to 850°C, the temperature typically used for GaAs growth. The sapphire appeared to be uncoated, because the GaN film was clear and had a specular surface. He rushed down to the RCA analytical center, and a Laue pattern revealed that the deposit was indeed the first single-crystal film of GaN (Maruska and Tietjen, 1969).

Maruska found that all of the GaN films were n-type without intentional doping. He sought to find a p-type dopant so that he could make a p-n junction. Zinc seemed to be an appropriate acceptor because it worked for GaAs and GaP. With heavy Zn concentrations, GaN films proved to be insulating. But they were never conducting p-type (Maruska, 2000).

During 1969, Jacques Pankove spent a sabbatical year at Berkeley University writing his classic textbook, *Optical Processes in Semiconductors*. When he returned to RCA Laboratories in January 1970, he immediately became interested and strongly involved in the new GaN films. Pankove *et al.* undertook a study of optical absorption and photoluminescence of thin-film GaN (Pankove *et al.*, 1970a, 1970b). The first example of electroluminescence from GaN was announced at RCA in the summer of 1971 (Pankove *et al.* 1971a). The sample consisted of an insulating Zn-doped layer which was contacted with two surface probes, and blue light centered at 475 nm was emitted.

Pankove and co-workers then made a device consisting of an undoped n-type region, an insulating Zn-doped layer, and an indium surface contact (Pankove *et al.*, 1971b, 1972). This **metal–insulator–semiconductor (MIS) diode** was the first current-injected GaN light emitter, and it emitted green and blue light.

The RCA team speculated that magnesium might be a better choice of p-type dopant than zinc. They began growing Mg-doped GaN films using the MHVPE technique, and in July 1972, obtained blue and violet emission centered at 430 nm as shown in *Figure 2.13* (Maruska *et al.*, 1972). One of these Mg-doped blue light MIS emitters continued to emit light even in 2008, more than 35 years after its fabrication (Maruska, 2008). Maruska *et al.* (1973) described these efforts in a paper entitled “Violet luminescence of Mg-doped GaN”. Note that the GaN films, even though Mg doped, did not exhibit p-type conductivity, so that the luminescence in these films was probably mediated by minority carrier injection or impact ionization in the high-field insulating regions of the films. Pankove and the RCA team offered a model for the operation of these devices based on impact ionization and Fowler–Nordheim tunneling, because the characteristics were virtually independent of temperature (Pankove and Lampert, 1974; Maruska *et al.*, 1974a, 1974b). Of course, these devices were inefficient, and as a consequence, Tietjen, who had stimulated the work, now terminated it by ordering “stop this garbage” – words that Maruska (2000) still vividly remembers.

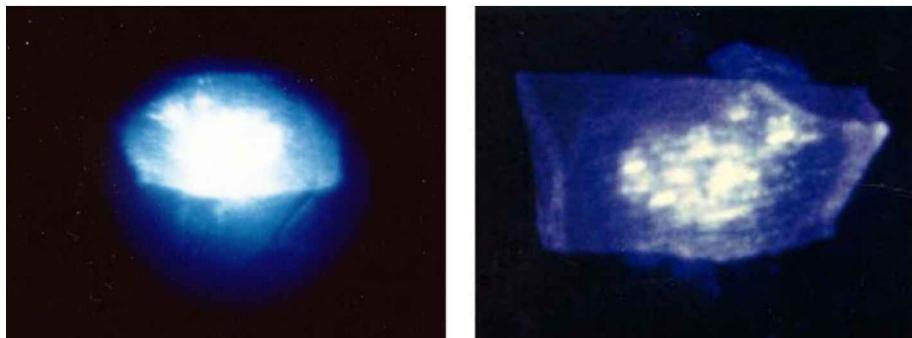


Fig. 2.13: Blue light emission found in 1972 caused by recombining electron–hole pairs created in a highly resistive GaN structure doped with Si and Mg (courtesy of H. Paul Maruska, 2000 and 2009).

A 1972 photograph of Maruska is shown in *Figure 2.14*. After a long career in science and technology, he moved to Florida and retired. The highlights of his career, such as the first demonstration of a violet LED, fabricated by him at Stanford University, still serve as a source of joy. Having grown the first epitaxial single-crystal GaN, Maruska called himself the “Father of GaN” (Maruska, 2009; Maruska and Rhines, 2015).

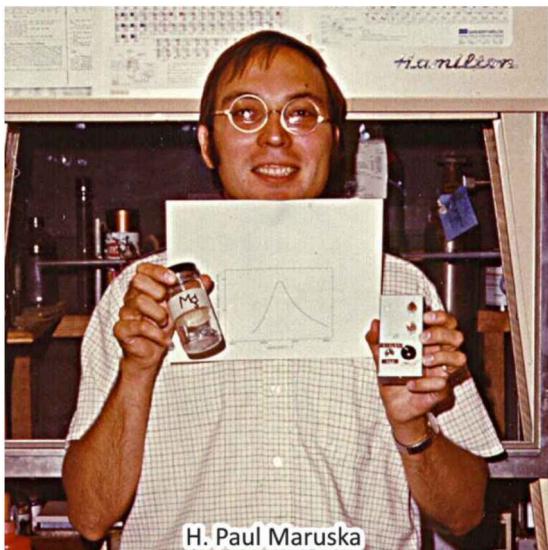


Fig. 2.14: H. Paul Maruska, in 1972 at Stanford University, where he was on leave from the Princeton RCA Laboratories, proudly displaying a GaN LED, its emission spectrum, and the new Mg acceptor used to make the i-type region of the MIS device. In April 1972, Maruska had decided that Mg might be a better choice as p-type dopant than Zn. He began growing Mg-doped GaN films using hydride vapor-phase epitaxy, and in July 1972, got a violet LED, emitting at 430 nm, that was bright enough to be seen in a well-lit room (courtesy of H. Paul Maruska, 2009).

2.8 History of GaN p-n junction LEDs

After the research efforts of Pankove and co-workers had ended, work on GaN virtually ceased. In 1982, only a single paper was published on GaN. However, a Japanese team of researchers that included Isamu Akasaki and Hiroshi Amano of Nagoya University located in Nagoya, Japan, refused to give up. This team of researchers made ***three key contributions*** to the field of GaN that allowed III–V nitride semiconductors to become the most widely used compound semiconductor.

The ***first key contribution*** is the demonstration of a thin AlN buffer layer that allowed for the growth of high-quality GaN epitaxial films on sapphire substrates (Amano *et al.*, 1986). The buffer layer, with a preferable thickness of a few tens of nm, accommodates the structural and lattice-constant mismatch between sapphire substrate and the epitaxial GaN so that crack-free GaN epitaxial layers can be attained. Subsequently, it was demonstrated that a GaN buffer layer, with a preferable thickness of 20 nm, is equally suited to accommodate the structural and lattice-constant mismatch between the sapphire substrate and the epitaxial GaN (Nakamura *et al.*, 1991).

The ***second key contribution*** is the demonstration of true p-type doping and p-type conductivity in GaN (Amano *et al.*, 1989). The stubborn Mg acceptors, which had been passivated by hydrogen, were activated by *low-energy electron-beam irradiation*, LEEBI. That is, the electron-beam irradiation heated the Mg-doped GaN sample so much that the passivating hydrogen was driven out of the sample thereby de-passivating the Mg acceptors and enabling p-type conductivity. The p-type conductivity of GaN was elucidated by strong blue

electroluminescence originating from only that region of the GaN sample that was previously treated by electron-beam irradiation. A photograph of electroluminescence from de-passivated GaN:Mg is shown in **Figure 2.15 (a)** (Amano *et al.*, 1989; Amano, 2012).

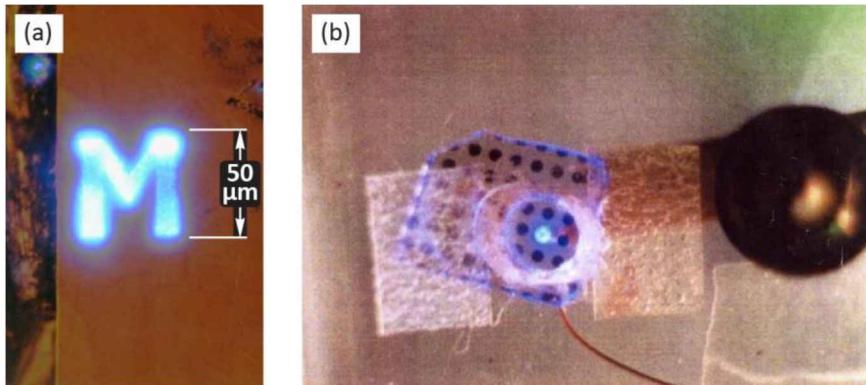


Fig. 2.15: (a) Strong blue electroluminescence originating from a GaN:Mg region treated by LEEBI, which caused Mg acceptors to be de-passivated. (b) First GaN pn-junction LED; the device emitted in the blue and UV spectrum (Amano *et al.*, 1989; courtesy of Hiroshi Amano, 2012).

Subsequent to the attainment of p-type doping, the first GaN p-n-homojunction LED was reported by Amano *et al.* (1989). The LED, which emitted light in the ultraviolet (UV) and blue spectral range, was grown on a sapphire substrate. The result was presented by Amano *et al.* (1989) in the *Japanese Journal of Applied Physics* and by Akasaki *et al.* at the “GaAs and Related Compounds” conference held in Karuizawa, Japan in 1992. The LED had an efficiency of approximately 1%. This was a surprisingly high value for the highly dislocated GaN material grown on the mismatched sapphire substrate. It was also the first demonstration that nitride LED efficiencies are not affected by dislocations in the same adverse manner as III–V arsenide and phosphide light emitters. A photograph of the first GaN pn-junction LED is shown in **Figure 2.15 (b)** (Amano, 2012)

Subsequently, it was shown that a high-temperature ***post-growth annealing*** of Mg-doped GaN also activates Mg dopants in GaN (Nakamura *et al.*, 1994a). High annealing temperatures activate the Mg acceptors by out-diffusion of the passivating hydrogen. Superlattice doping (Schubert *et al.*, 1996) further enhances the activation efficiency of deep acceptors. These p-type doping breakthroughs opened the door to efficient p-n junction LEDs and laser diodes. Today, Mg-doping of GaN is the basis for all nitride-based LEDs and laser diodes.

The ***third key contribution*** in the field of GaN is the epitaxial growth of $\text{Ga}_{1-x}\text{In}_x\text{N}$, and the demonstration of $\text{Ga}_{1-x}\text{In}_x\text{N}/\text{GaN}$ double heterostructures quantum wells (briefly called GaN/GaN DHs and QWs) for the light-emitting active region of an LED (Matsuoka *et al.*, 1992;

Nakamura *et al.*, 1993a; 1994b). The epitaxial growth of the mixed crystal GaInN is non-trivial since GaN inherently requires a high growth temperature whereas InN requires a much lower growth temperature at which pyrolysis of the precursors, particularly the NH₃ precursor, is inefficient. Prior to the advent of GaInN/GaN QWs, the active region consisted of Si-Zn co-doped GaN which has a far lower radiative efficiency than GaInN/GaN QWs.

Names closely associated with the commercialization of GaN LEDs and lasers are those of Shuji Nakamura, Takashi Mukai, and Nichia Company, Japan. At Nichia, a team of researchers that included Shuji Nakamura and Takashi Mukai has made numerous seminal contributions to the development of GaN growth, LEDs, and lasers (Johnstone, 2007). Contributions included the first viable blue and green GaInN double-heterostructure LED (Nakamura *et al.*, 1993a, 1993b, 1994b) that achieved efficiencies of 10% (Nakamura *et al.*, 1995), and the first pulsed and cw GaInN/GaN current-injection ***blue laser*** operating at room temperature (Nakamura *et al.*, 1996). Initially, a particular design, the two-flow organometallic vapor-phase epitaxy (OMVPE) growth-system design was used (Nakamura *et al.*, 1991). However, the use of two-flow OMVPE at Nichia has been discontinued (Mukai, 2005). Detailed accounts of the team's contributions were given by Nakamura and Fasol (1997) and Nakamura *et al.* (2000) in the book *The Blue Laser Diode* and by the Nichia Company in the booklet *Remarkable Technology* (Nichia, 2004).

Blue LEDs made by the Nichia Company are shown in **Figure 2.16**. GaN-based high-brightness blue and green LEDs were commercially available from the Nichia Company starting in about 1994 and 1995, respectively.

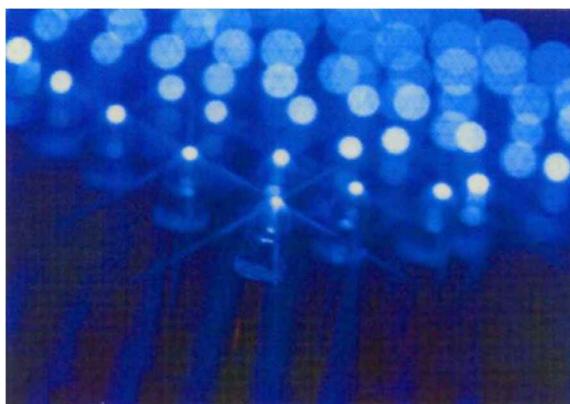


Fig. 2.16: Array of GaInN/GaN blue LEDs manufactured by the Nichia Company, Japan (after Nakamura and Fasol, 1997).

A common application of high-brightness GaInN green LEDs is traffic signals as shown in **Figure 2.17**. The earlier mentioned GaP:N green LEDs are not suited for this application due to their much lower brightness.

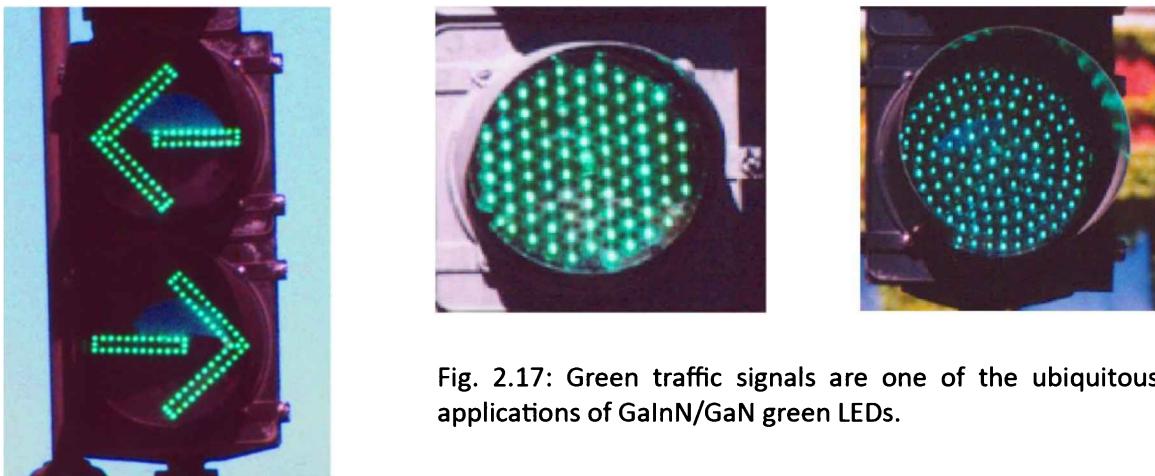


Fig. 2.17: Green traffic signals are one of the ubiquitous applications of GaN/GaN green LEDs.

In 1990, when Nakamura entered the field of GaN devices while working for the Nichia Company, he was a 36-year-old engineer without a Ph.D., not a single publication, and no conference contribution (Nakamura and Fasol, 1997). At the end of the 1990s, he had become a Professor at the University of California in Santa Barbara and a consultant for the Cree Lighting Company, a fierce competitor of Nichia. In 2001, he criticized the Nichia Company and Japanese society. In the book entitled *Breakthrough with Anger*, Nakamura (2001) stated, “There is something wrong with this country. Industry and universities are terribly sick.”

Photographs of the most widely recognized pioneers of GaN materials, GaN p-type doping, GaN LEDs, and GaN lasers, Isamu Akasaki, Hiroshi Amano, and Shuji Nakamura, are shown in **Figure 2.18**. The researchers were received numerous awards for their seminal research in III–V nitride technology and LEDs including the 2014 Nobel Prize in physics.



Fig. 2.18: (a) Isamu Akasaki, (b) Hiroshi Amano, and (c) Shuji Nakamura, pioneers of III–V nitride devices. Their research with colleagues at Meijo University and Nichia Company led to blue, green, and white LEDs ((a) courtesy of Isamu Akasaki, 2009; (b) courtesy of Hiroshi Amano, 2012; (c) courtesy of Tampere University of Technology, Finland, 2006).

The Royal Swedish Academy of Sciences (2014) published a 10-page Scientific Background of the 2014 Nobel Prize in Physics. The introductory paragraphs of the Scientific Background read as follows:

“Light-emitting diodes (LEDs) are narrow-band light sources based on semiconductor components, with wavelengths ranging from the infrared to the ultraviolet. The first LEDs were studied and constructed during the 1950s and 1960s in several laboratories. They emitted light at different wavelengths, from the infrared to the green. However, emitting blue light proved to be a difficult task, which took three more decades to achieve. It required the development of techniques for the growth of high-quality crystals as well as the ability to control p-doping of semiconductors with high bandgap, which was achieved with gallium-nitride (GaN) only at the end of the 1980s. The development of efficient blue LEDs also required the production of GaN-based alloys with different compositions and their integration into multilayer structures such as heterojunctions and quantum wells.

The invention of efficient blue LEDs has led to white light sources for illumination. When exciting a phosphor material with a blue LED, light is emitted in the green and red spectral ranges, which, combined with the blue light, appears as white. Alternatively, multiple LEDs of complementary colors (red, green and blue) can be used together. Both of these technologies are used in today's high-efficiency white electroluminescent light sources. These light sources, with very long lifetimes, have begun to replace incandescent and fluorescent lamps for general lighting purposes. Since lighting represents 20-30% of our electrical energy consumption, and since these new white light sources require ten times less energy than ordinary light bulbs, the use of efficient blue LEDs leads to significant energy savings, of great benefit to mankind.

This year's Nobel Prize in Physics honors the inventors of efficient blue LEDs: I. Akasaki, H. Amano and S. Nakamura.”

2.9 References

- Akasaki I., Amano H., Itoh K., Koide N., and Manabe K. “GaN based UV/blue light-emitting devices” *GaAs and Related Compounds conference, Institute Physics Conference Series* **129**, 851 (1992)
- Allen J. W., Moncaster M. E., and Starkiewicz J. “Electroluminescent devices using carrier injection in gallium phosphide” *Solid State Electronics* **6**, 95 (1963)
- Amano H., Sawaki N., Akasaki I., and Toyoda Y. “Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer” *Applied Physics Letters* **48**, 353 (1986)
- Amano H., Kito M., Hiramatsu K., Akasaki I. “P-type conduction in Mg-doped GaN treated with low-energy electron beam irradiation (LEEBI)” *Japanese Journal of Applied Physics* **28**, L2112 (1989)
- Amano Hiroshi, personal communication. The provision of photographs showing electroluminescence from LEEBI-treated GaN:Mg and of the first GaN pn-junction LED is gratefully acknowledged (2012)
- Bergh A. A. and Dean P. J. “Light-Emitting Diodes” *Proceedings of the IEEE* **60**, 156 (1972)
- Biard J. R. and Pittman G. E. “Semiconductor radiant diode” US Patent 3,293,513; filed August 8, 1962; issued December 20 (1966)
- Braun K. F. “Über die Stromleitung durch Schwefelmetalle” (“On the conduction of current through metal sulfides”, translated from German) Editor J. C. Poggendorf, *Annalen der Physik und Chemie*, **153**, 556 (1874)

- Braunstein R. "Radiative transitions in semiconductors" *Physical Review* **99**, 1892 (1955)
- Chang S. J., Chang C. S., Su Y. K., Chang P. T., Wu Y. R., Huang K. H. and Chen T. P. "AlGaN_P multiquantum well light-emitting diodes" *IEE Proceedings Optoelectronics* **144**, 1 (1997)
- Chang S. J. and Chang C. S. "AlGaN_P–GaN_P compressively strained multiquantum well light-emitting diodes for polymer fiber applications" *IEEE Photonics Technology Letters* **10**, 772 (1998a)
- Chang S. J. and Chang C. S. "650 nm AlGaN_P/GaN_P compressively strained multi-quantum well light-emitting diodes" *Japanese Journal of Applied Physics* **37**, L653 (1998b)
- Craford M. G., Shaw R. W., Herzog A. H., and Groves W. O. "Radiative recombination mechanisms in GaAsP diodes with and without nitrogen doping" *Journal of Applied Physics* **43**, 4075 (1972)
- Destriau G. "Scintillations of zinc sulfides with alpha-rays" *Journal de Chimie Physique* **33**, 587 (1936)
- Duke C. B. and Holonyak Jr. N. "Advances in light-emitting diodes" *Physics Today*, December issue, p. 23 (1973)
- Dupuis R. D. and Krames M. R. "History, development, and applications of high-brightness visible light-emitting diodes" *Journal of Lightwave Technology* **26**, 1154 (2008)
- Fletcher R. M., Kuo C., Osentowski T. D., and Robbins V. M. "Light-emitting diode with an electrically conductive window" US Patent 5,008,718 (1991)
- Folberth O. G. "Mischkristallbildung bei A_{III}B_V-Verbindungen" ("Formation of mixed crystals in A_{III}B_V-compounds", translated from German) *Zeitschrift für Naturforschung* **10a**, 502, 1955
- Folberth O. G. "Mixed crystal semiconductor devices" US Patent 2,858,275 issued October 28 (1958)
- Folberth O. G. "Zur Frage der Existenz tetraedrischer Phasen" ("On the question of the existence of tetrahedral phases" translated from German) *Zeitschrift für Naturforschung* **14a**, 94 (1959)
- Foster L. M. and Pilkuhn M. "Electroluminescence near bandgap in GaP containing shallow donor and acceptor levels" *Applied Physics Letters* **7**, 65 (1965)
- Gershenson M. and Mikulyak R. M. "Light emission from forward biased p-n junctions in gallium phosphide" *Solid-State Electronics* **5**, 313 (1962)
- Gessmann T. and Schubert E. F. "High-efficiency AlGaN_P light-emitting diodes for solid-state lighting applications" *Journal of Applied Physics* **95**, 2203 (2004)
- Grimmeiss H. G. and Scholz H. J. "Efficiency of recombination radiation in GaP" *Physics Letters* **8**, 233 (1964)
- Grimmeiss H. G. and Allen J. W. "Light emitting diodes – How it started" *Journal of Non-Crystalline Solids* **352**, 871 (2006)
- Groves W. O., Herzog A. H., and Craford M. G. "The effect of nitrogen doping on GaAs_{1-x}P_x electroluminescent diodes" *Applied Physics Letters* **19**, 184 (1971)
- Groves W. O. and Epstein A. S. "Epitaxial deposition of III–V compounds containing isoelectronic impurities" US Patent 4,001,056 issued Jan. 4 (1977)
- Groves W. O., Herzog A. H., and Craford M. G. "Process for the preparation of electroluminescent III–V materials containing isoelectronic impurities" US Patent Re. 29,648 issued May 30 (1978a)
- Groves W. O., Herzog A. H., and Craford M. G. "GaAsP electroluminescent device doped with isoelectronic impurities" US Patent Re. 29,845 issued Nov. 21 (1978b)
- Hall R. N., Fenner G. E., Kingsley J. D., Soltys T. J., and Carlson R. O. "Coherent light emission from GaAs junctions" *Physical Review Letters* **9**, 366 (1962)
- Herzog A. H., Groves W. O., and Craford M. G. "Electroluminescence of diffused GaAs_{1-x}P_x diodes with low donor concentrations" *Journal of Applied Physics* **40**, 1830 (1969)
- Holonyak Jr. N. "Active region in visible-light diode laser" *Electronics* **36**, 35 (1963)
- Holonyak Jr. N. "Laser action in Ga(AsP) and GaAs" *Proceedings of the IEEE* **52**, 104 (1964)
- Holonyak Jr. N. "Semiconductor alloy lasers – 1962" *IEEE Journal of Quantum Electronics* **QE-23**, 684 (1987)
- Holonyak Jr. Nick, personal communication (2000)
- Holonyak Jr. N. and Bevacqua S. F. "Coherent (visible) light emission from Ga(As_{1-x}P_x) junctions" *Applied Physics Letters* **1**, 82 (1962)
- Holonyak Jr. N., Bevacqua S. F., Bielan C. V., Carranti F. A., Hess B. G., and Lubowski S. J. "Electrical

- properties of Ga(AsP) p-n junctions" *Proceedings of the IEEE* **51**, 364 (1963a)
- Holonyak Jr. N., Bevacqua S. F., Bielan C. V., and Lubowski S. J. "The "direct–indirect" transition in Ga(As_{1-x}P_x) p-n junctions" *Applied Physics Letters* **3**, 47 (1963b)
- Holonyak Jr. N., Nuese C. J., Sirkis M. D., and Stillman G. E. "Effect of donor impurities on the direct–indirect transition in Ga(AsP)" *Applied Physics Letters* **8**, 83 (1966)
- Huang K.-H. and Chen T.-P. "Light-emitting diode structure" US Patent 5,661,742 (1997)
- IEEE Spectrum article authored by Rachel Courtland titled "No Nobel for the father of the LED" *IEEE Spectrum*, posted on the Internet < <http://spectrum.ieee.org/tech-talk/semiconductors/devices/no-nobel-for-the-father-of-the-led> > on October 8 (2014)
- Ikeda M., Nakano K., Mori Y., Kaneko K. and Watanabe N. "MOCVD growth of AlGaN_xP at atmospheric pressure using triethylmetals and phosphine" *Journal of Crystal Growth* **77**, 380 (1986)
- Isihamatsu S. and Okuno Y. "High efficiency GaAlAs LED" *Optoelectronics – Device Technology* **4**, 21 (1989)
- Itaya T., Ishikawa M., and Uematsu Y. "636 nm room temperature cw operation by heterobarrier blocking structure InGaAlP laser diodes" *Electronics Letters* **26**, 839 (1990)
- Johnstone B. "Brilliant! Shuji Nakamura and the Revolution in Lighting Technology" (Prometheus Books, Amherst NY, 2007)
- Juza R. and Hahn H. "On the crystal structure of Cu₃N, GaN and InN (translated from German)" *Zeitschrift für anorganische und allgemeine Chemie* **239**, 282 (1938)
- Kish F. A. and Fletcher R. M. "AlGaN_xP light-emitting diodes" in *High Brightness Light-Emitting Diodes* edited by G. B. Stringfellow and M. G. Crawford, Semiconductors and Semimetals **48**, p. 149 (Academic Press, San Diego, 1997)
- Kobayashi K., Kawata S., Gomyo A., Hino I. and Suzuki T. "Room-temperature cw operation of AlGaN_xP double-heterostructure visible lasers" *Electronics Letters* **21**, 931 (1985)
- Krames M. R., Ochiai-Holcomb M., Höfler G. E., Carter-Coman C., Chen E. I., Tan I.-H., Grillot P., Gardner N. F., Chui H. C., Huang J.-W., Stockman S. A., Kish F. A., Crawford M. G., Tan T. S., Kocot C. P., Hueschen M., Posselt J., Loh B., Sasser G., and Collins D. "High-power truncated-inverted-pyramid (Al_xGa_{1-x})_{0.5}In_{0.5}P/GaP light-emitting diodes exhibiting > 50% external quantum efficiency" *Applied Physics Letters* **75**, 2365 (1999)
- Krames M. R., Amano H., Brown J. J., and Heremans P. L. "High-efficiency light-emitting diodes" Special Issue of *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 185 (2002)
- Kuo C. P., Fletcher R. M., Osentowski T. D., Lardizabel M. C., Crawford M. G., and Robbins V. M. "High performance AlGaN_xP visible light-emitting diodes" *Applied Physics Letters* **57**, 2937 (1990)
- Logan Ralph A., personal communication (2000)
- Logan R. A., White H. G., and Trumbore F. A. "P-n junctions in compensated solution grown GaP" *Journal of Applied Physics* **38**, 2500 (1967a)
- Logan R. A., White H. G., and Trumbore F. A. "P-n junctions in GaP with external electroluminescence efficiencies ~2% at 25°C" *Applied Physics Letters* **10**, 206 (1967b)
- Logan R. A., White H. G., Wiegmann W. "Efficient green electroluminescence in nitrogen-doped GaP p-n junctions" *Applied Physics Letters* **13**, 139 (1968)
- Logan R. A., White H. G., and Wiegmann W. "Efficient green electroluminescent junctions in GaP" *Solid State Electronics* **14**, 55 (1971)
- Lorenz M. R. and Binkowski B. B. "Preparation, stability, and luminescence of gallium nitride" *Journal of the Electrochemical Society* **109**, 24 (1962)
- Maruska H. P., personal communication. The provision of photographs including those of blue-emitting GaN MIS LEDs is gratefully acknowledged (2000, 2008, and 2009)
- Maruska H. P. and Pankove J. I. "Efficiency of GaAs_{1-x}P_x electroluminescent diodes" *Solid State Electronics* **10**, 917 (1967)
- Maruska H. P. and Tietjen J. J. "The preparation and properties of vapour-deposited single-crystalline GaN" *Applied Physics Letters* **15**, 327 (1969)
- Maruska H. P., Rhines W. C., Stevenson D. A. "Preparation of Mg-doped GaN diodes exhibiting violet electroluminescence" *Mat. Res. Bull.* **7**, 777 (1972)

- Maruska H. P., Stevenson D. A., Pankove J. I. "Violet luminescence of Mg-doped GaN (light-emitting diode properties)" *Applied Physics Letters* **22**, 303 (1973)
- Maruska H. P., Anderson L. J., Stevenson D. A. "Microstructural observations on gallium nitride light-emitting diodes" *Journal of the Electrochemical Society* **121**, 1202 (1974a)
- Maruska H. P. and Stevenson D. A. "Mechanism of light production in metal–insulator–semiconductor diodes; GaN:Mg violet light-emitting diodes" *Solid State Electronics* **17**, 1171 (1974b)
- Maruska H. P. and Rhines W. C. "A modern perspective on the history of semiconductor nitride blue light sources" *Solid-State Electronics* **111**, 32 (2015)
- Matsuoka T., Yoshimoto N., Sasaki T., Katsui A. "Wide-gap semiconductor InGaN and InGaAlN grown by MOVPE" *Journal of Electronic Materials* **21**, 157 (1992)
- Mueller G. (Editor) *Electroluminescence I Semiconductors and Semimetals* **64** (Academic Press, San Diego, 2000)
- Mukai, Takashi, personal communication (2005)
- Nakamura S. "GaN growth using GaN buffer layer" *Japanese Journal of Applied Physics* **30**, L1705 (1991)
- Nakamura S., Senoh M., and Mukai T. "Highly p-type Mg doped GaN films grown with GaN buffer layers" *Japanese Journal of Applied Physics* **30**, L1708 (1991)
- Nakamura S., Senoh M., and Mukai T. "P-GaN/n-InGaN/n-GaN double-heterostructure blue-light-emitting diodes" *Japanese Journal of Applied Physics* **32**, L8 (1993a)
- Nakamura S., Senoh M., and Mukai T. "High-power InGaN/GaN double-heterostructure violet light-emitting diodes" *Applied Physics Letters* **62**, 2390 (1993b)
- Nakamura S., Iwasa N., and Senoh M. "Method of manufacturing p-type compound semiconductor" US Patent 5,306,662 (1994a)
- Nakamura S., Mukai T., and Senoh M. "Candela-class high-brightness InGaN/AlGaN double-heterostructure blue-light-emitting diodes" *Applied Physics Letters* **64**, 1687 (1994b)
- Nakamura S., Senoh M., Iwasa N., Nagahama S. "High-brightness InGaN blue, green, and yellow light-emitting diodes with quantum well structures" *Japanese Journal of Applied Physics* **34**, L797 (1995)
- Nakamura S., Senoh M., Nagahama S., Iwasa N., Yamada T., Matsushita T., Sugimoto Y., and Kiyoku H. "Room-temperature continuous-wave operation of InGaN multi-quantum-well structure laser diodes" *Applied Physics Letters* **69**, 4056 (1996)
- Nakamura S. and Fasol G. *The Blue Laser Diode* (Springer, Berlin, 1997)
- Nakamura S., Fasol G., and Pearton S. J. *The Blue Laser Diode* second edition (Springer, Berlin, 2000)
- Nakamura S. *Breakthrough with Anger* (Shueisha, Tokyo, 2001); See also *Compound Semiconductors* **7**, No. 7, 25 (Aug. 2001) and **7**, No. 9, 15 (Oct. 2001)
- Nathan M. I., Dumke W. P., Burns G., Dill Jr. F. H., and Lasher G. J. "Stimulated emission of radiation from GaAs p-n junctions" *Applied Physics Letters* **1**, 62 (1962)
- News-Gazette article authored by Julie Wurth titled "UPDATED: Nobel snub for former UI prof" *The News-Gazette*, posted on the Internet <<http://www.news-gazette.com/news/local/2014-10-07/updated-nobel-snub-former-ui-prof.html>> on October 7 (2014)
- Nichia Company *Remarkable Technology* edited by I. Matsushita and E. Shibata (Nichia Company, Tokushima, Japan, 2004)
- Nuese C. J., Stillman G. E., Sirkis M. D., and Holonyak Jr. N. "Gallium arsenide-phosphide: crystal, diffusion, and laser properties" *Solid State Electronics* **9**, 735 (1966)
- Nuese C. J., Tietjen J. J., Gannon J. J., and Gossenberger H. F. "Optimization of electroluminescent efficiencies for vapor-grown GaAsP diodes" *Journal of the Electrochemical Society: Solid State Science* **116**, 248 (1969)
- Ohba Y., Ishikawa M., Sugawara H., Yamamoto T., and Nakanisi T. "Growth of high-quality InGaAlP epilayers by MOCVD using methyl metalorganics and their application to visible semiconductor lasers" *Journal of Crystal Growth* **77**, 374 (1986)
- Pankove J. I. and Berkeyheiser J. E. "A light source modulated at microwave frequencies" *Proceedings of the IRE*, **50**, 1976 (1962)
- Pankove J. I. and Massoulie M. J. "Injection luminescence from GaAs" *Bull. Am. Physical Soc.* **7**, 88 (1962)

- Pankove J. I., Berkeyheiser J. E., Maruska H. P., and Wittke J. "Luminescent properties of GaN" *Solid State Communications* **8** 1051 (1970a)
- Pankove J. I., Maruska H. P., and Berkeyheiser J. E. "Optical absorption of GaN" *Applied Physics Letters* **17**, 197 (1970b)
- Pankove J. I., Miller E. A., Richman D., and Berkeyheiser J. E. "Electroluminescence in GaN" *Journal of Luminescence* **4**, 63 (1971a)
- Pankove J. I., Miller E. A., and Berkeyheiser J. E. "GaN electroluminescent diodes" *RCA Review* **32**, 383 (1971b)
- Pankove J. I., Miller E. A., and Berkeyheiser J. E. "GaN blue light-emitting diodes" *Journal of Luminescence* **5**, 84 (1972)
- Pankove J. I. and Lampert M. A. "Model for electroluminescence in GaN" *Physical Review Letters* **33**, 361 (1974)
- Perry T. S. "M. George Craford" *IEEE Spectrum*, February issue, p. 52 (1995)
- Pilkuhn M. H. and Rupprecht H. "Light emission from $\text{GaAs}_x\text{P}_{1-x}$ diodes" *Transactions of the Metallurgical Society AIME* **230**, 282 (1964)
- Pilkuhn M. H. and Rupprecht H. "Electroluminescence and lasing action in $\text{GaAs}_x\text{P}_{1-x}$ " *Journal of Applied Physics* **36**, 684 (1965)
- Pilkuhn M. H. "Light-emitting diodes" in *Handbook of Semiconductors* edited by T. S. Moss, **4**, edited by C. Hilsum, p. 539 (1981)
- Pilkuhn Manfred H., personal communication. The provision of a photograph showing a GaP:Zn-O LED is gratefully acknowledged (2000)
- Quist T. M., Rediker R. H., Keyes R. J., Krag W. E., Lax B., McWhorter A. L. and Zeigler H. J. "Semiconductor maser of GaAs" *Applied Physics Letters* **1**, 91 (1962)
- Rostky G. "LEDs cast Monsanto in unfamiliar role" *Electronic Engineering Times (EETimes)* on the internet, see <http://eetimes.com/anniversary/designclassics/monsanto.html>, Issue 944, March 10 (1997)
- Royal Swedish Academy of Sciences, Stockholm, Sweden "Scientific Background on the Nobel Prize in Physics 2014" posted on the Internet < http://www.nobelprize.org/nobel_prizes/physics/laureates/2014/advanced-physicsprize2014.pdf > on October 7 (2014)
- Rupprecht H., Pilkuhn M., and Woodall J. M. "Continuous stimulated emission from GaAs diodes at 77 K" (First report of 77 K cw laser) *Proceedings of the IEEE* **51**, 1243 (1963)
- Rupprecht H., Woodall J. M., Konnerth K., and Pettit D. G. "Efficient electro-luminescence from GaAs diodes at 300 K" *Applied Physics Letters* **9**, 221 (1966)
- Rupprecht H., Woodall J. M., and Pettit G. D. "Efficient visible electroluminescence at 300 K from AlGaAs p-n junctions grown by liquid phase epitaxy" *Applied Physics Letters* **11**, 81 (1967).
- Rupprecht H., Woodall J. M., Pettit G. D., Crowe J. W., and Quinn H. F. "Stimulated emission from AlGaAs diodes at 77 K" *Quantum Electron.* **4**, 35 (1968)
- Rupprecht Hans, personal communication (2000)
- Schanda János, personal communication. The provision of photographs showing an early GaAsP LED and a large-scale color display is gratefully acknowledged (2007)
- Schubert E. F., Grieshaber W., and Goepfert I. D. "Enhancement of deep acceptor activation in semiconductors by superlattice doping" *Applied Physics Letters* **69**, 3737 (1996)
- Seiko Company "Pulsar - it's all in the details" www.pulsarwatches-europe.com (2004)
- Stringfellow G. B. and Craford M. G. (Editors) *High Brightness Light-Emitting Diodes* Semiconductors and Semimetals **48** (Academic Press, San Diego, 1997)
- Sugawara H., Ishikawa M., Kokubun Y., Nishikawa Y., and Naritsuka S. "Semiconductor light-emitting device" US Patent 5,048,035 (1991)
- Thomas D. G., Hopfield J. J., and Frosch C. J. "Isoelectronic traps due to nitrogen in gallium phosphide" *Physical Review Letters* **15**, 857 (1965)
- Totzke D. "ECE professor [James R. Biard] leads way to Nobel Prize" *Currents – Texas A&M University ECE Magazine* < <http://engineering.tamu.edu/electrical/news/currents> > page 14, November (2014)
- Vavilov V. S. "Radiative recombination in semiconductors" *Soviet Physics Uspekhi* **2**, 455 (1959)

- Welker H. "Electrical semiconductor device" (translated from German) German patent DE 970 420 C; filed March 11, 1951; published December 18, 1952; issued September 4, 1958 (1951a)
- Welker H. "Semiconductor devices and methods of their manufacture" US patent 2,798,989; filed March 10, 1952 claiming priority of 1951; issued July 9, 1957 (1951b)
- Welker H. "On new semiconducting compounds" (translated from German) *Zeitschrift für Naturforschung* **7a**, 744 (1952)
- Welker H. "On new semiconducting compounds II" (translated from German) *Zeitschrift für Naturforschung* **8a**, 248 (1953)
- Wolfe C. M., Nuese C. J., and Holonyak Jr. N. "Growth and dislocation structure of single-crystal Ga(AsP)" *Journal of Applied Physics* **36**, 3790 (1965)
- Wolff G. A., Herbert R. A., and Broder J. D. "Electroluminescence of GaP" *Physical Review* **100**, 1144 (1955)
- Woodall Jerry M., personal communication (2000)
- Woodall J. M., Potemski R. M., Blum S. E., and Lynch R. "AlGaAs LED structures grown on GaP substrates" *Applied Physics Letters* **20**, 375 (1972)

3

3 – History of white light-emitting diodes

3.1 The advent of the white LEDs

Traditional LEDs emit monochromatic light. However, white LEDs, by definition, emit polychromatic light. Therefore, white LEDs are a significant departure from traditional LEDs. In the spirit of this significant departure, the phrase “solid-state lighting” is frequently employed for the field of white LEDs. While traditional LEDs, i.e. monochromatic LEDs, created mostly their own new markets, the implication of the phrase “solid-state lighting” is that LEDs are used to replace conventional lighting sources: Incandescent lamps (Thomas Edison’s light bulb) and fluorescent lamps. Therefore, the phrase “solid-state lighting” is employed for white LEDs that are used in applications traditionally served by conventional white-light sources (incandescent and fluorescent lamps). Solid-state lighting includes general lighting in homes and offices, street lighting, automotive lighting, and backlighting in liquid-crystal displays (LCDs).

High-efficiency red, orange, and yellow devices (AlGaN_P LEDs) had been developed in the 1980s. High-efficiency violet, blue, and green devices (GaInN LEDs) had been developed in the 1990s. Therefore, in the mid-1990s, monochromatic high-efficiency devices, covering the entire visible spectrum, were available. As a result, the generation of white light by LEDs had been enabled. There are several viable approaches for white LEDs:

A ***first approach*** is a multi-LED-chip approach in which the light emitted from three LED chips emitting the three primary colors (red, green, and blue) is mixed to generate white light (Stinson, 1991). This approach is illustrated in ***Figure 3.1***. Inspection of the figure reveals that the lamp needs to have four lead electrodes so that each of the three LED chips can be injected with an appropriate current. The approach became feasible with the demonstration of highly efficient blue and green LEDs (Nakamura *et al.*, 1994; 1995). Nakamura *et al.* (1995) stated: “By combining high-power and high-brightness blue InGaN SQW LED, green InGaN SQW LED and red GaAlAs LED, many kinds of applications, such as LED full-color displays and LED white lamps for use in place of light bulbs or fluorescent lamps, are now possible with characteristics of high reliability, high durability and low energy consumption.”

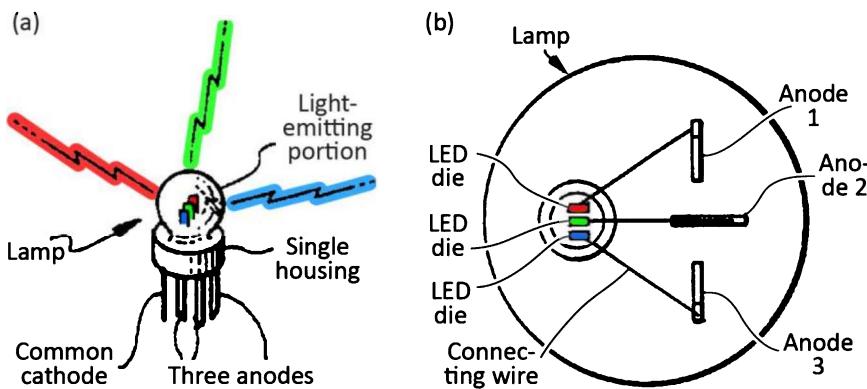


Fig. 3.1: (a) Perspective view and (b) top view of multi-LED-chip white LED consisting of a red (R), green (G), and blue (B) LED chip (or die). Optical mixing of the RGB emission components results in white light (adapted from Stinson, 1991).

A **second approach** is based on using a UV or violet LED chip and a phosphor that absorbs the UV or violet light and converts it to a broad-band white light. Due to its similarity to conventional fluorescent lamps³, this approach has been proposed multiple times prior to the advent of GaN-based blue LEDs⁴. Tabuchi (1973) disclosed an LED structure, shown in **Figure 3.2** that includes an LED chip and a phosphor that is coated on the inside of a transparent cover. Tabuchi (1973) stated: “For example, it goes without saying that a near UV light emitting devices with GaN can be employed and that an ordinary UV-visible light conversion phosphor can be utilized.” Another research group, Stevenson *et al.* (1974) stated: “violet light [from an LED chip] may be converted to lower frequencies [...] using organic and inorganic phosphors. Such a conversion is appropriate [...] to produce light in a spectral range of greater sensitivity for the human eye. By use of different phosphors, all the primary colors may be developed [and] may be used for color display systems”. Yet another research group, Tokailin and Hosokawa (1992) disclosed using an organic UV LED and an organic phosphor converting the UV light to white light: “An electroluminescent element [...] which emits a near ultraviolet ray of light and a fluorescent material part which absorbs the ultraviolet light [...] and emits a fluorescence in a visible light range from blue to red [...] to be used as an element for emitting white light.” That is, in each of these proposals, a current-injected pn-junction emitting near UV or violet light is used to excite a phosphor that emits visible light in the range from blue to red, i.e. white light. However, the proposals of Tabuchi, Stevenson *et al.*, and Tokailin and Hosokawa (1992) were not followed up by practical demonstrations. One may note that this approach can suffer from a low efficiency, because the

³ The now-obsolete Hg-based fluorescent lamp converts deep the UV light emitted by a Hg-vapor source ($\lambda = 254 \text{ nm}$) into visible white light by means of a phosphor.

⁴ The general use of phosphors in conjunction with LED chips dates back to 1970 when a phosphor was used to up-convert infrared light to visible light using two-photon absorption (Potter *et al.*, 1970).

down-conversion of UV or violet light involves a relatively large wavelength shift (large Stokes shift).

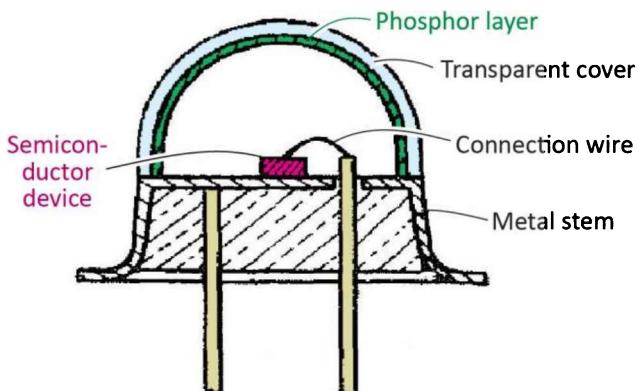


Fig. 3.2: LED structure according to Tabuchi (1973) having an LED chip exciting a phosphor that is coated on the inside of a transparent cover. The arrangement of the phosphor is reminiscent of a fluorescent lamp that has the inside of a glass tube coated with a phosphor (adapted from Tabuchi, 1973).

A **third and most successful approach** is a blue-LED-chip-plus-phosphor combination with the phosphor emitting green and red light. The blue light from the LED chip is *partially* absorbed by the phosphor while the other part of the blue light is transmitted through the phosphor. As a result, the blue light (from LED chip), green light (from phosphor), and red light (from phosphor) together form white light (Shimizu, 1995 and 1996). This concept is now known as the “partial conversion” concept. Shimizu (1995) stated: “A part of the [blue] light is absorbed by the phosphor, its wavelength is converted at the same time, and then the light is radiated.” After an extensive search, numerous tests, and after the realization that organic phosphors could not operate reliably under the harsh operating conditions of a blue LED, a very suitable inorganic phosphor was identified by Shimizu *et al.* (1996, 1999): Cerium-doped yttrium-aluminum garnet phosphor (YAG:Ce). This phosphor is able to absorb blue light and emits red and green light⁵ (Blasse and Bril, 1967). The approach, first demonstrated by Shimizu *et al.* (1996, 1999), was further described by Bando *et al.* (1996, 1998) and subsequently reviewed by Nakamura and Fasol (1997). The approach evolved into a pervasive commercial success, in part due to its efficiency, simplicity, and the requirement of only one LED chip driven by only one power supply.

One may note that the concept of combining an LED chip with a phosphor dates back to 1970 when Potter *et al.* (1970) demonstrated the up-conversion of infrared (IR) radiation emitted by a GaAs LED chip to visible light by using a phosphor. **Figure 3.3** shows (a) a GaAs LED chip mounted on a header and (b) the same chip covered by a blob of LaF₂:Tm or (La_{0.8}Yb_{0.2})F₂:Tm phosphor

⁵ The mixture of red and green light appears yellow to the human eye. YAG:Ce is therefore frequently referred to as a “yellow phosphor”.

dispersed in a binder, polystyrene (Potter *et al.*, 1970). In the up-conversion process, the exciting photon energy is smaller than the phosphor-emission photon energy, that is $h\nu_1 < h\nu_2$, so that two exciting photons are required per phosphor-emitted photon. The up-conversion process is inefficient, even when used in conjunction with a high photon density. The down-conversion process ($h\nu_1 > h\nu_2$), used in white LEDs, is generally much more efficient.

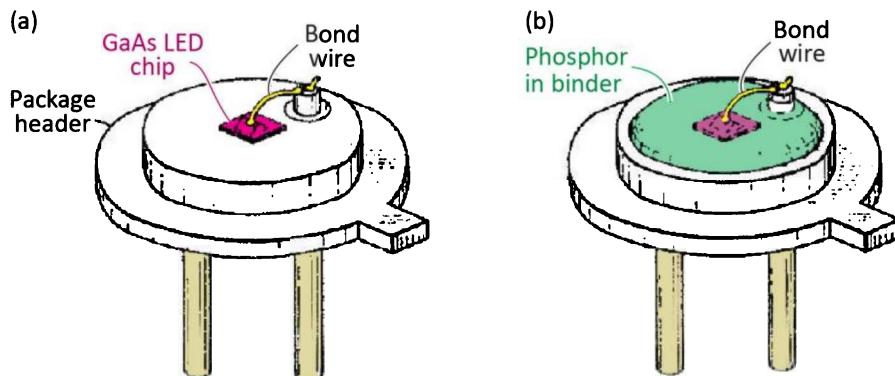


Fig. 3.3: (a) GaAs-based infrared (IR) LED chip mounted on header. (b) LED chip coated by a phosphor dispersed in a binder for up-converting IR radiation to visible light (adapted from Potter *et al.*, 1970).

During the year 1996, several publications and patent filings emerged that illustrate the pursuit for the white LED. Notable contributions to the field of white LEDs are, in chronological order:

January 12, 1996: Yoshinori Shimizu (Shimizu, 1996) of the Nichia Company publishes a Japanese Patent Application Publication that discloses a planar white light source for white backlighting of a full-color display. The planar light source uses a blue LED whose light is partially absorbed by organic phosphors (a red-emitting and a green-emitting phosphor). Together, the blue light from the LED chip and the green and red light from the phosphors create white light. However, the organic phosphors proposed by Shimizu (1996) are now known to not be sufficiently stable under the high-radiation conditions of blue LEDs⁶.

March 26, 1996: Baretz and Tischler (2003) of the ATMI Company file a patent application that discloses a semiconductor LED chip and a phosphor with the light emitted by the chip being fully converted by a phosphor so as to generate white light. The inventors suggest that organic phosphors would be preferable over inorganic phosphors: “[Further, it is most desirable to utilize organic fluorescent materials with extremely short radiative lifetimes...](#)”

⁶ The radiation intensity of present-day blue LED chips can be estimated to $0.25 \text{ W/mm}^2 = 250 \text{ kW/m}^2$ which is about 250 times (!) the radiation intensity of solar radiation (about 1 kW/m^2). This illustrates that phosphors used in conjunction with blue LEDs are exposed to operating conditions that are extremely harsh so that some phosphors, particularly organic phosphors, will degrade rapidly.

June 26, 1996: Reeh *et al.* (2006) of the Osram Company and the Fraunhofer Society, file a patent application that discloses the partial conversion of blue LED-chip light by a yellow organic phosphor. However, organic phosphors lack the chemical stability needed to withstand the high radiation intensities occurring in the vicinity of a GaInN blue LED chip.

July 29, 1996: Yoshinori Shimizu and co-workers at the Nichia Company filed a Japanese patent application that uses a garnet phosphor, specifically yttrium aluminum garnet, $\text{Y}_3\text{Al}_5\text{O}_{12}$, doped with the optically active rare-earth atom cerium (YAG:Ce). When excited in the blue wavelength range (e.g. at 450 nm), YAG:Ce has a broad emission band that spans from green to red. The YAG:Ce emission appears yellow to the observer, since the mixing of green and red light appears yellow to the human eye. Accordingly, YAG:Ce is frequently referred to as a yellow phosphor. Shimizu *et al.* (1996, 1999) demonstrated what would become a pervasively successful strategy in white LEDs: A blue LED-chip, the use of partial conversion, and the YAG:Ce phosphor. Some of the chip's blue light is converted to green and red light while the remaining part of the blue light is transmitted through the phosphor so that the blue, green, and red lights mix to form white light. The YAG:Ce phosphor exhibits superior chemical stability and can withstand the high radiation intensity of the blue LED chip. As the inventors pointed out, the YAG:Ce phosphor has to endure a blue light intensity that exceeds the intensity of sunlight by more than a factor of ten. The Sun's radiation intensity is about 1 kW/m²; however, the radiation intensity of a blue LED can be 100 mW/mm² = 100 kW/m², i.e. 100 times higher than the radiation intensity of the Sun! YAG:Ce proved to be able to operate under such unprecedently harsh conditions. Furthermore, Shimizu *et al.* (1996, 1999) found that white LEDs using pure YAG:Ce has an unpleasant green-yellowish tint. Accordingly, the inventors worked to chemically modify the $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$ by adding gadolinium, $(\text{Y}_{1-x}\text{Gd}_x)_3\text{Al}_5\text{O}_{12}:\text{Ce}$, as well as gallium, $\text{Y}_3(\text{Al}_{1-x}\text{Ga}_x)_5\text{O}_{12}:\text{Ce}$. Addition of Gd shifts the peak emission wavelength to longer wavelengths, thereby giving the phosphor a stronger red component. Similarly, addition of Ga shifts the peak emission wavelength to shorter wavelengths, thereby giving the phosphor a stronger green component. This allowed the new white LED to have a variety of color temperatures including daylight white (5 000~6 500 K) as well as incandescent-like white (2 500~3 300 K). Shimizu *et al.* (1996, 1999) can be considered the first disclosure of a viable white LED. The first white LEDs manufactured by the Nichia Company had a luminous efficacy of 5 lm/W; white LEDs with efficiencies as high as 12 lm/W were demonstrated by the team in 1996 (Shimizu *et al.*, 1996, 1999). This was a remarkable result: It was a world record for the efficacy of white LEDs and exceeded the efficiency of low-wattage incandescent lamps! The “blue-LED chip plus YAG:Ce phosphor” approach is still being practiced at the present time, decades after the invention.

September 13, 1996: A new era in lighting began. An article in the Japanese newspaper Nikkei Sangyo Shimbun (1996), shown in *Figure 3.4*, announced a new type of white light source, based on a blue LED and containing a yellow cerium-doped yttrium aluminum garnet (YAG:Ce) phosphor. The white LED was reported to be efficient (5 lm / W), low cost, and have a predicted lifespan of 50 000 hours. What was announced in a relatively short newspaper article was a novel light source that was set to revolutionize the world.

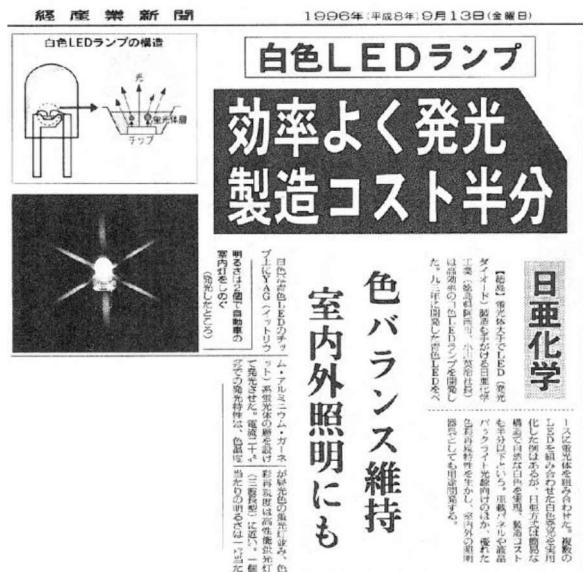


Fig. 3.4: Article of the Japanese newspaper Nikkei Sangyo Shimbun on the impending commencement of manufacturing of white LEDs by Nichia Company: “White LED lamp: Light emission with high luminous efficiency halving production costs”. The white LED is based on a blue LED chip and a YAG (yttrium aluminum garnet) phosphor. A start of full-scale production of the white LED in early October 1996 is announced. A lifespan of 50 000 hours or more is predicted, 10 and 50 times longer than the lifespan of fluorescent lamps and incandescent lamps, respectively. The white LED has a reported efficiency of 5 lm/W (Nikkei Sangyo Shimbun, 1996).

September 20, 1996: Reeh *et al.* (2006) of the Osram Company filed a patent application that includes partial conversion of blue-LED light and the use of YAG:Ce phosphor.

October 1996: Full-scale production of the phosphor-converted white LED started at the Nichia Company in Anan, Japan.

November 29, 1996: Technical details of Nichia’s white LED were presented at a technical meeting of the Institute of Phosphor Society (Japan) and the associated 264th Proceedings of the Institute of Phosphor Society (Japan). Nichia’s phosphor-converted white LED, including its emission spectrum, is shown in *Figure 3.5* (Bando *et al.*, 1996). The emission spectrum is continuous covering nearly the entire visible wavelength range. Compared with the “spiky” emission spectrum of the then-common compact fluorescent lamps (CFLs), the quality of the white light emitted by the LED was superior. Likewise, compared with the power consumption of the then-common incandescent light bulbs, the prospect of white LEDs, in terms of power consumption, was far superior. Thus, the foundation of a new LED white-light source, that was poised to replace conventional white-light sources, had been established.

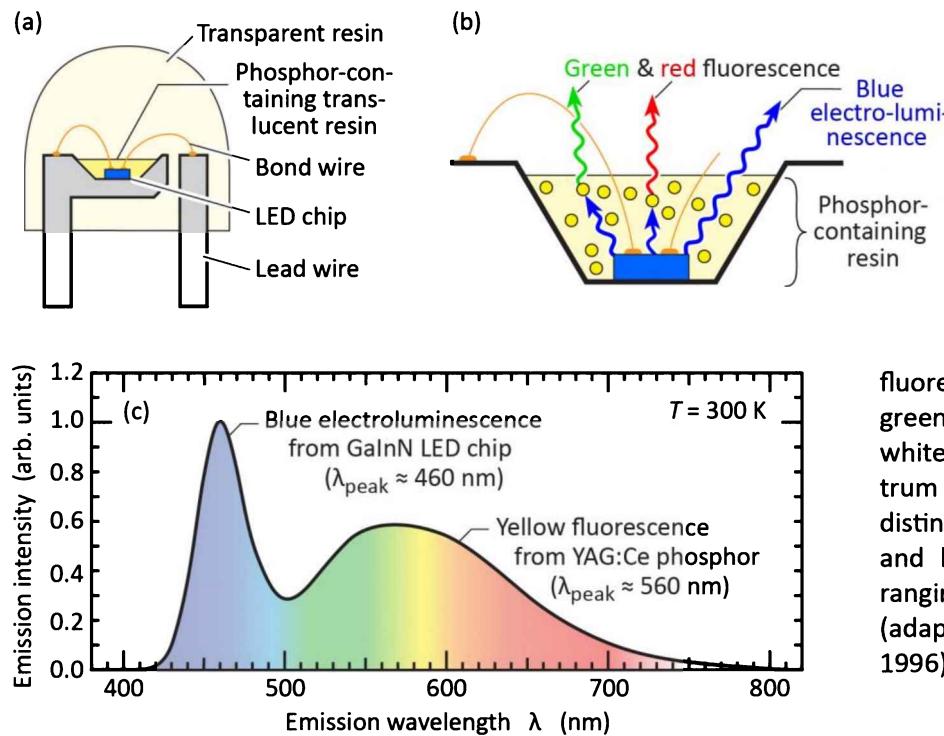


Figure 3.5: (a) Structure of first white LED. (b) InGaN LED chip emitting blue electroluminescence (EL) and YAG:Ce phosphor, excited by blue EL, emitting broadband yellow

fluorescence ranging from green to red thereby generating white light. (c) Emission spectrum of first white LED showing distinct blue EL line at 460 nm and broad fluorescence band ranging from green to red (adapted from Bando *et al.*, 1996).

April 1997: In the April issue of *Applied Physics A*, an Osram-Fraunhofer team reported on blue-emitting GaN chips in combination with luminescence down-converting inorganic YAG:Ce phosphor as well as organic fluorescent dye (Schlotter *et al.*, 1997). Using the LED-chip-phosphor combination, the team demonstrated LEDs emitting white light as well as mixed colors, such as cyan and magenta.

For phosphor-based white LEDs, it is useful to distinguish between *full conversion* and *partial conversion*. As shown in **Figure 3.6 (a)**, for full conversion, all primary radiation (exciting radiation from the LED chip) is converted to secondary light by means of the phosphor. If the primary radiation is UV light, the full conversion concept is reminiscent of the conventional fluorescent lamp in which all primary radiation (Hg-vapor radiation) is converted to secondary radiation. Tabuchi (1973), Stevenson (1974), and Baretz and Tischler (2003) proposed to use the full-conversion concept to produce white light. In this case, the phosphor itself must emit white light. **Figure 3.6 (b)** shows partial conversion in which only a part of the primary light (usually blue light) is converted to red and green light. In this case, the phosphor emission can be limited to red and green light, while the blue light is generated by the LED chip. Partial conversion has a fundamentally higher efficiency, since it inherently involves lower wavelength-conversion losses (lower Stokes-shift losses). Shimizu (1996), Schlotter *et al.* (1997), Shimizu (1999), and Reeh *et al.*, (2006) demonstrated the partial conversion concept for white LEDs.

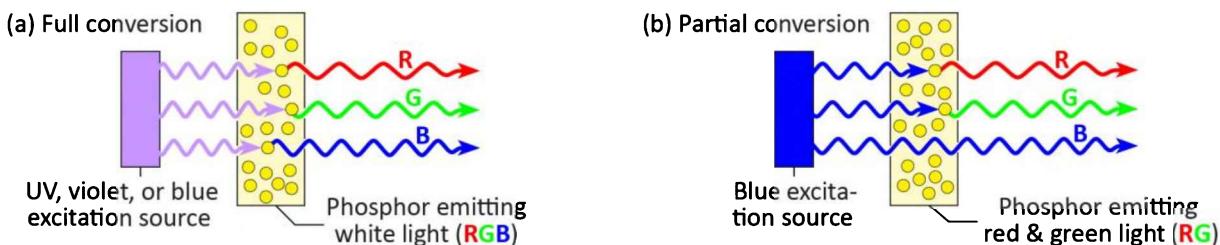


Figure 3.6: (a) Full conversion and (b) partial conversion in a white LED. For full conversion, all primary light (excitation light) is converted to phosphor-based white fluorescence. For partial conversion, the primary blue light becomes part of the white light.

Undoubtedly, a new era in the generation of white light started in the summer of 1996, based on semiconductor pn-junction diodes and phosphors. This innovation in the generation of white light can certainly be compared to the development of the incandescent lightbulb and fluorescent tube that occurred decades earlier. Photographs of key innovators in the field of white LEDs are shown in **Figure 3.7**, including Yoshinori Shimizu, Yasunobu Noguchi and Kensho Sakano of the Nichia Company (Anan-Shi, Tokushima, Japan) and Peter Schlotter of the Fraunhofer Institute for Applied Solid-State Physics (Freiburg, Germany).

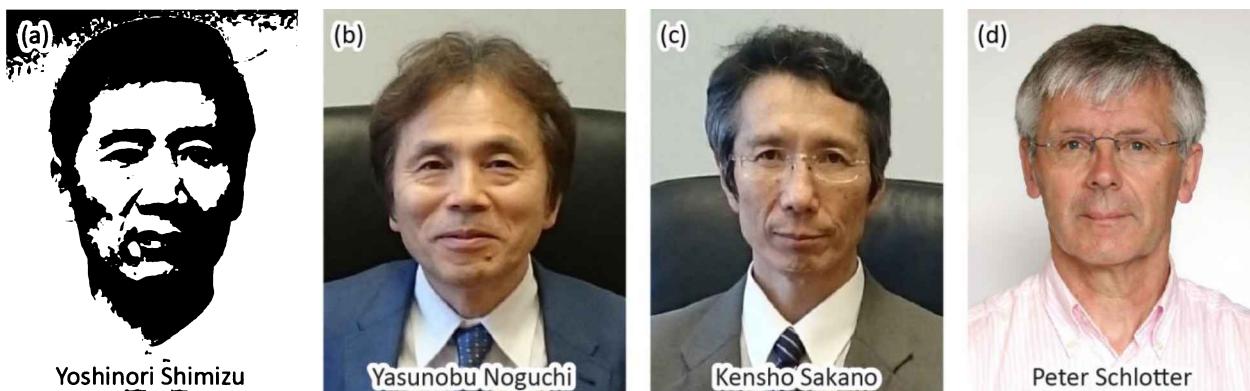


Fig. 3.7: (a) Yoshinori Shimizu, pioneer of white LEDs and senior manager in the Engineering Department at Nichia. (b) Yasunobu Noguchi, pioneer who developed YAG:Ce phosphors for white LEDs. In 2015, he was head of a Nichia Phosphor Department. (c) Kensho Sakano fabricated and tested hundreds of white LED prototypes. (d) Peter Schlotter (Fraunhofer Institute), first author of a highly-cited article on white LEDs (Schlotter et al., 1997) and co-inventor of a family of patents related to white LEDs.

In retrospect, it is easy to forget that there was substantial skepticism regarding white LEDs. There were numerous concerns about white LEDs, particularly from the established lighting community. Opinions such as “Poor color rendering of objects, narrow output cones, and lower

total lumen outputs make solid-state lamps impractical for general illumination” were common in the early years of white LEDs (Peralta and Ruda, 1998). The view that LEDs were to be relegated to niche applications was not uncommon either: “The limitations of commercial LEDs do not preclude their use in niche applications” (Peralta and Ruda, 1998). Even in 2008, when it was abundantly clear that white LEDs are revolutionizing general lighting (Narukawa *et al.*, 2006), there were skeptical voices: “I do not see a major step toward change in general illumination. To say LEDs will change everything, I don’t [believe] it. I think a lot of it is hype” (Rea, 2008).

3.2 Transition from LEDs to LED-based lighting

Traditional LEDs were used for indicator lights, signage, and displays. Prior to about 2000, illumination, i.e. the general lighting of homes, offices, and streets, was not the common purpose of LEDs. This was about to change in 2000 when the vision that LEDs can be used, on a large scale, for lighting applications, was articulated for the first time.

A report issued by Sandia National Laboratories and authored by Haitz *et al.* (2000), addressed the potential energy savings enabled by LED lighting (called “semiconductor lighting” in the report), in a quantitative way. The authors concluded that by a large-scale introduction of solid-state lighting, the electrical power consumed by lighting would decrease by more than 50% and the global annual electrical energy savings would exceed 1 000 TWhr. The authors also discussed the emission of the global warming gas CO₂ which is generated through the production of electricity from coal-, oil-, and natural gas-burning power plants. The authors concluded that the large-scale introduction of solid-state lighting and the associated electricity savings would reduce carbon dioxide emissions by annually approximately 200 Mt (Megatons). Additional analyses, that are consistent with the major conclusions of Haitz *et al.*, include discussions by Bergh *et al.* (2001) and Schubert *et al.* (2006).

For general lighting applications, the *total power* emitted by an LED is of importance. Whereas conventional light sources such as incandescent light bulbs can be easily scaled up to provide a high luminous flux, LEDs, historically (i.e. during the pre-2000 time), have been mostly low-power devices. The historical development of the luminous flux per LED package, measured in lumens, is shown in *Figure 3.8* (Krames *et al.*, 2000).

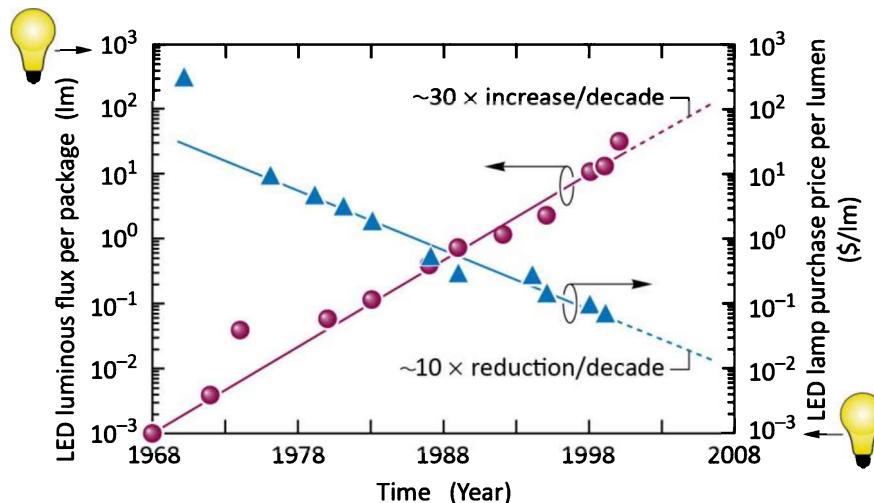
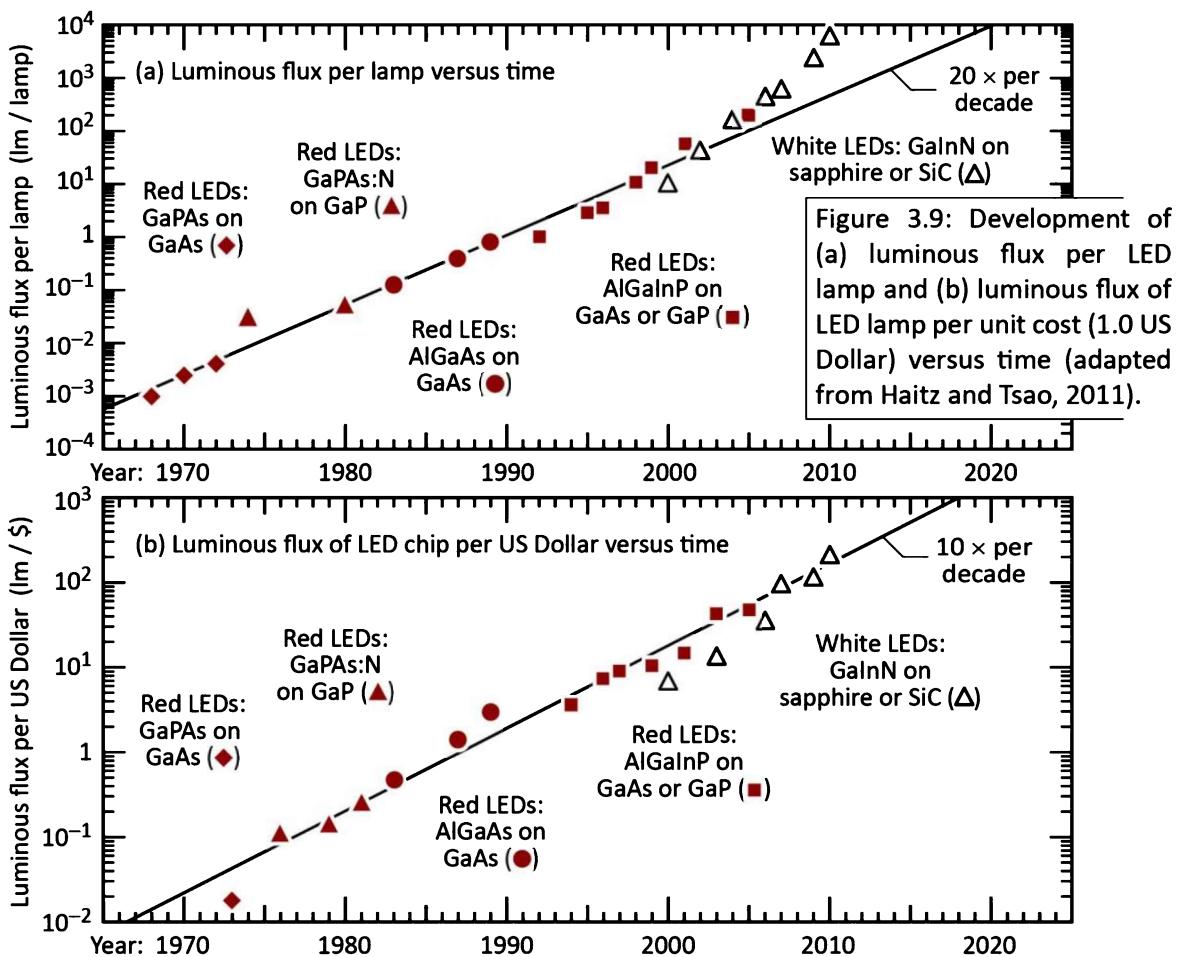


Fig. 3.8: LED luminous flux per package and LED lamp purchase price per lumen versus year. Also shown are the values for a 60 W incandescent tungsten-filament light bulb with a luminous efficiency of $\sim 13 \text{ lm/W}$ and a luminous flux of 800 lm with an approximate price of 1.00 US\$ (after Krames *et al.*, 2000).

The figure shows that the luminous flux per LED package has increased by about four orders of magnitude over a period of 30 years. For comparison, the figure also shows a 60-W-incandescent-bulb's luminous flux and approximate purchase price. The figure illustrates continued progress in the performance and manufacturing cost of LEDs. Note that the cost shown in the figure is just the purchase price of the lamp and does not include the cost for the electrical energy consumed by the lamp. The cost of the electrical energy of incandescent lamps typically is much higher than its purchase price. That is, efficient LED light sources have a significant energy cost advantage over incandescent lamps.

The advancement of LED efficiency has been compared to the advancement made in Si integrated circuits where the performance increase versus time has been characterized by “Moore’s law”. This “law” states that the performance of Si integrated circuits doubles approximately every 18 months. In this context, the term “**Haitz’s observation**”, also called “Haitz’s law”, may be mentioned (Haitz *et al.*, 2000; Steigerwald *et al.*, 2002; Editorial, 2006; Haitz and Tsao, 2011). Haitz’s observation is that the amount of light (luminous flux) generated by one LED lamp increases by a factor of 30 per decade (Krames *et al.*, 2000; Haitz *et al.*, 2000) or by a factor of 20 per decade (Haitz and Tsao, 2011), while the amount of light (luminous flux) generated by one LED lamp available at a certain price increases by a factor of 10 per decade (Haitz and Tsao, 2011). A graphical representation of Haitz’s observation is shown in **Figure 3.9** (adapted from Haitz and Tsao, 2011).



It should be noted that the data presented in Haitz's observation were gathered selectively and restricted to certain LED material systems, emission colors, color temperatures, and time frames. Haitz and Tsao (2011) indeed admit that their selection of data points "is not very scientific". Indeed, the selectivity by which the data was gathered led to the visually convincing dependence shown in **Figure 3.9**. Despite its shortcomings, Haitz's observation is a useful articulation of the long-term trend towards LEDs with higher luminous flux as well as the trend towards lower purchase cost per luminous-flux unit.

Jeffrey Y. Tsao, a pioneer in solid-state lighting at US Sandia National Laboratories (Albuquerque, New Mexico), co-authored the Haitz *et al.* (2000) report. Tsao pointed out a correlation between the cost of light, consumption of light, and productivity of human society. He stated: "throughout its history, lighting technology has made tremendous progress: the efficiency with which power is converted into usable light has increased 2.8 orders of magnitude [i.e. by a factor of 631] over three centuries. This progress has, in turn, fueled large increases in the consumption of light and productivity of human society" (Tsao, 2010; Tsao *et al.*, 2010).

Yet another dimension of solid-state light sources was pointed out by Schubert and Kim (2005), in an article entitled: “Solid-state light sources getting smart”. The authors showed that solid-state sources can be controlled (by either design or by means of real-time control), much more so than conventional light sources (i.e. incandescent and fluorescent sources). The specific parameters that can be controlled in LEDs are the (i) emission spectrum, (ii) polarization direction, (iii) color temperature, (iv) spatial emission pattern, and (v) temporal modulability. This opens up the possibility for solid-state light sources to be tailored for specific applications, and for the possibility to establish new functionalities of solid-state sources, so that they can serve for new, useful purposes. For example, the communication of information by means of lighting sources had not been an option for conventional lighting sources whereas it is an option for solid-state lighting sources. Thus, the controllability of LEDs along with their lower energy consumption and their positive environmental effects are the key advantages of solid-state sources.

The temporal evolution of the luminous efficiency of light sources is shown in *Figure 3.10*. The figure shows the following light sources: Incandescent lamp with C filament demonstrated in 1879 by Thomas A. Edison; Incandescent lamp with ductile W filament demonstrated in 1911 by William D. Coolidge; Linear fluorescent lamp introduced in 1937 at the New York World’s Fair; Compact fluorescent lamp introduced in 1985 by Osram Company (after Kane and Sell, 2002). These conventional light sources and their typical efficiencies are as follows:

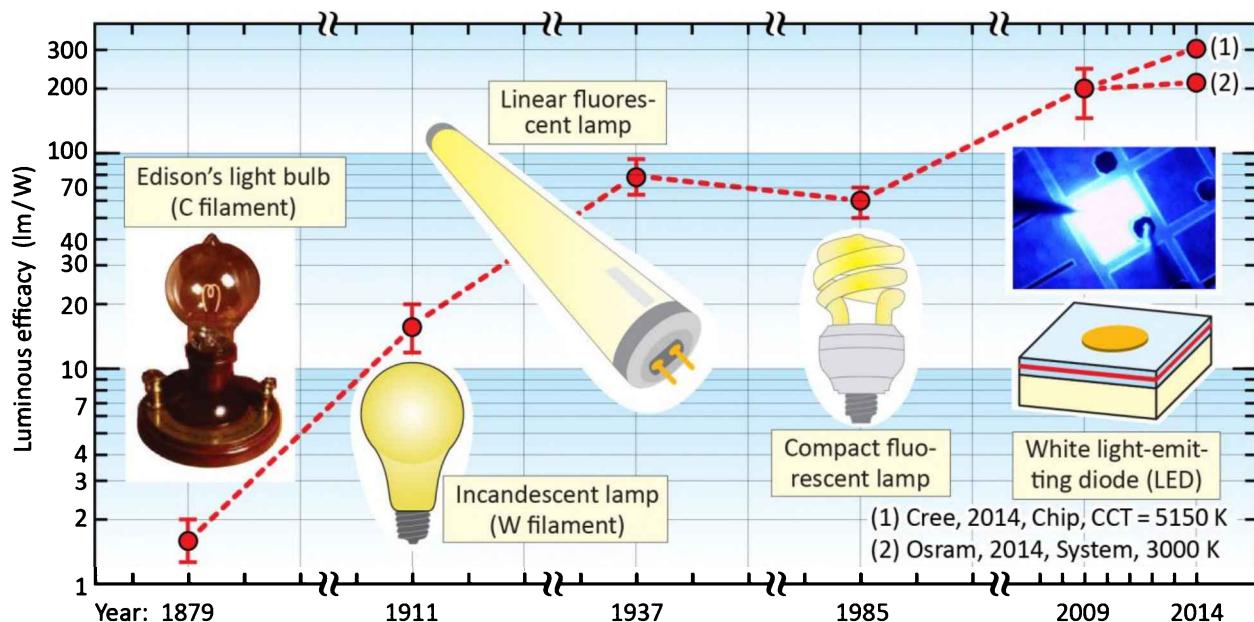


Fig. 3.10: Temporal development of the luminous efficiency of different types of lamps. The 2014 points represent: (1) White LED device performance (Cree Company, 2014; 303 lm/W; CCT = 5150 K), and (2) White LED device and system performance (Osram Company, 2014; 215 lm/W (device); 205 lm/W (system); CCT = 3000 K).

Incandescent: C (carbon) filament light bulb	1.2~2.0 lm/W
Incandescent: W (tungsten) filament light bulb	10~18 lm/W
Incandescent: W (tungsten) filament halogen bulb	16~24 lm/W
Fluorescent: Linear fluorescent lamp (LFL)	65~95 lm/W
Fluorescent: Compact fluorescent lamp (CFL)	50~70 lm/W

In 2010, reported efficiencies of white LEDs were in the 150~250 lm/W range. Very high efficiencies were reported by several companies including the following: The Nichia Company announced a laboratory-result efficiency of 249 lm/W for a white LED injected with a very low current (Mukai, 2009; Narukawa *et al.*, 2010). The Cree Company announced a laboratory-result efficiency of 208 lm/W for a white LED with a correlated color temperature of about 4 600 K at an injection current of 350 mA (Cree, 2010).

On March 26, 2014, the Cree Company announced a laboratory-result efficiency of 303 lm/W for a white LED lamp (excluding power supply) with a correlated color temperature of 5 150 K at an injection current of 350 mA (Cree, 2014).

On March 28, 2014, the Osram Company announced an lamp efficacy of 215 lm/W and a system efficiency (including power supply) of 205 lm/W for a white LED lamp system with a color temperature of 3 000 K (Osram, 2014).

The improvement in luminous efficiency of visible-spectrum LEDs has been truly breathtaking. The historical development of the luminous efficiency of visible-spectrum LEDs is shown in *Figure 3.11* (Craford, 1997, 1999, updated 2000 and 2012). The chart illustrates the modest beginnings of visible-spectrum LED technology which started in the 1960s. If the progress from 1960 to 2000 is assumed to be continuous, then the LED luminous efficiency has doubled every 4 years. The following types of LEDs are shown in *Figure 3.11*:

- GaAsP red LEDs grown on GaAs substrates. The GaAsP/GaAs material system is lattice mismatched so that an abundance of misfit dislocations occurs in GaAsP epitaxial films. As a result, these LEDs have a low luminous efficiency (on the order of only 0.1 lm/W).
- GaP LEDs doped with radiative recombination centers. Pure GaP is an indirect semiconductor and therefore has a low radiative efficiency. However, when doped with isoelectronic impurities such as N or co-doped with Zn and O, radiative transitions in the red and green spectral range occur via these centers.
- GaAsP/GaAs red LEDs doped with N. Again, a mismatched materials system with low efficiency due to the abundance of misfit dislocations.
- AlGaAs/GaAs red LEDs. These LEDs employ GaAs quantum well active regions.
- AlGaAs/AlGaAs double heterostructure red LEDs using AlGaAs active regions and AlGaAs

barriers.

- AlGaNp/GaAs LEDs with absorbing GaAs substrates
- AlGaNp/GaP LEDs and with wafer-bonded transparent GaP substrates.
- AlGaNp/GaP orange LEDs with truncated inverted-pyramid-shaped dies with efficiencies exceeding 100 lm/W as reported by Krames *et al.* (1999).
- GaInN LEDs emitting in the blue and green wavelength range.
- GaInN/GaN white LEDs and LED light bulbs.

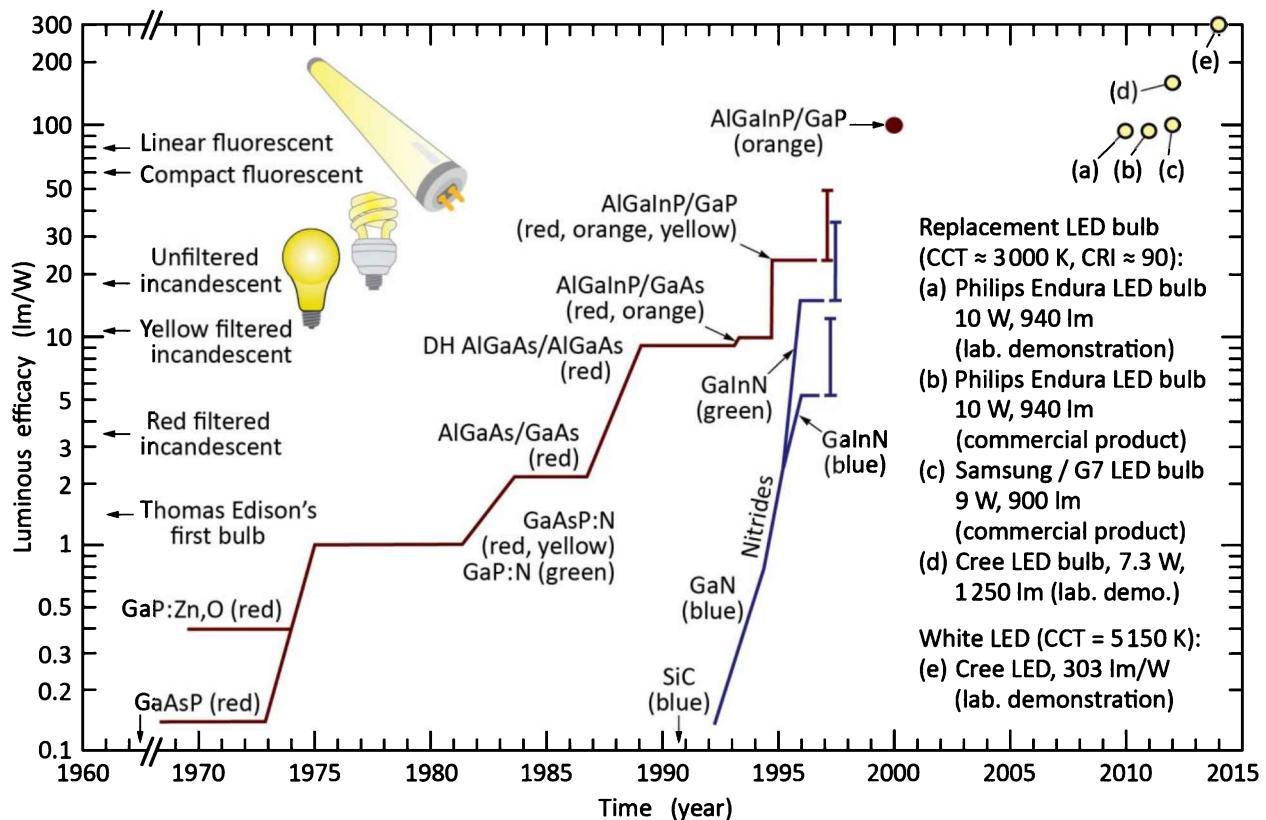


Fig. 3.11: Development of luminous efficacy of LEDs and conventional sources. Luminous efficacy of LED light bulbs available as laboratory demonstrations and commercial products (2010–2012). (Adapted from Crawford, 1997, 1999; updated 2000, 2012, 2014).

Figure 3.11 also shows the luminous efficiency of conventional (pre-LED) sources including Edison's first light bulb (1.4 lm/W) and red and yellow filtered incandescent lamps. The figure reveals that LEDs outperform filtered red and yellow incandescent lights by a large margin.

3.3 LEDs entering new fields of applications

As devices with higher power capabilities have become available, new application areas have emerged constantly. **Figure 3.12** shows the use of LEDs integrated into medical goggles worn by a surgeon during an operation (Shimada *et al.*, 2001, 2003). The LED-based light source promises

substantial weight savings and fulfills the stringent requirements of high-quality color rendition required during medical operations.

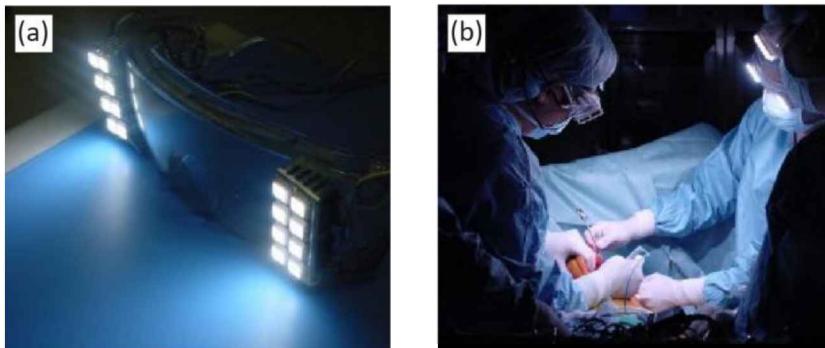


Fig. 3.12: (a) First goggle with integrated white LEDs used for (b) illumination during medical surgery (after Shimada *et al.*, 2001; Shimada *et al.*, 2003).

LED-based automotive headlights were first introduced by the Audi car company in 2004 using Lumileds Lighting's Luxeon devices. The car is shown in *Figure 3.13*.



Fig. 3.13: First automotive daytime running lights based on LEDs (Audi car company, 2004).

The use of LEDs in large-scale display and signage applications continues unabated. A seven-story high display and an animated pedestrian traffic signal are shown in *Figure 3.14* and *Figure 3.15*, respectively.



Fig. 3.14: LED display consisting of 18 million LEDs covering front of building, located in New York City (2005).



Fig. 3.15: Pedestrian sign indicating number of seconds left to cross street, located in Taipei, Taiwan (2005).

LEDs have also entered the domain of the arts. **Figure 3.16** shows the accent-lighted Stone Bridge across the Danube River located in Regensburg, Germany. A line of 21 900 bright red LEDs stretches across the bridge, thereby emphasising the link between the two banks of the river. Each LED stands for one month since the city was founded in the year AD 179 (Osram Opto Semiconductors, 2004). **Figure 3.17** shows the accent-lighted entrance of the Science Based Industrial Park located in Hsin Chu, Taiwan, an industrial park housing numerous companies active in semiconductor technology.

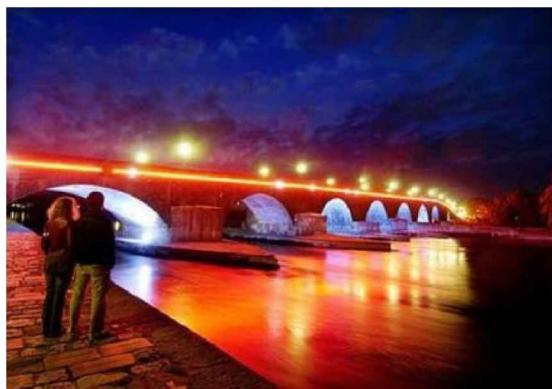


Fig. 3.16: Stone Bridge located in Regensburg, Germany, illuminated by LEDs (after *Focus Magazine*, Munich, 2004).



Fig. 3.17: Accent lighting at Science-Based Industrial Park in Hsin Chu, Taiwan (courtesy of K. R. Wang and L.-W. Tu, 2005).

Starting in 2006, LEDs were used to illuminate entire buildings enhancing their appearance during night. An example is shown in **Figure 3.18** where the six columns of the neo-classical Brandenburg gate in Berlin, Germany, are illuminated with blue, green, yellow, orange, red, and violet LEDs.



Fig. 3.18: The Brandenburg Gate in Berlin, Germany, illuminated with LEDs (after *Focus Magazine*, <www.focus.de> Munich, 2006).

The 2008 Olympic summer games held in Beijing, China, provided numerous examples of large-area LED displays that allow spectators in the sports arena to get an additional view of the events. An example of a large LED display used during the 2008 Olympics is shown in *Figure 3.19*.



Fig. 3.19: Large full-color LED display with 2400×792 pixel resolution used in the 2008 Olympic Games in Beijing, China. The display was assembled of 5.7 million red, green, and blue LEDs (courtesy of János Schanda, 2007).

3.4 LED replacement light bulbs

LED replacement light bulbs were widely introduced in 2010 and during that year became available in stores worldwide. Typical color temperatures for these lamps currently are 2 700 K (incandescent) and 5 000 K (daylight). Photographs of retrofit LED lamps are shown in *Figure 3.20* (Sylvania, 2010; Philips, 2010).

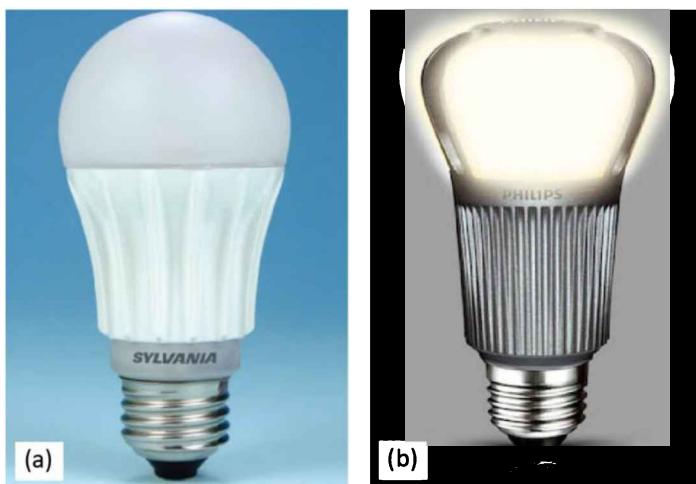


Fig. 3.20: (a) Sylvania A19 LED lamp consuming 8 W for replacement of a 40 W incandescent lamp. The LED lamp offers a color-rendering index of 82 with enhanced red tone rendering and lasts 16 times longer than an incandescent lamp. The LED lamp contains no Hg and is free of UV and IR radiation minimizing discoloration and fading of materials. (b) Philips Endura LED lamp consuming 12 W and emitting a luminous flux of 806 lm for replacement of a 60 W incandescent lamp (after Sylvania, 2010 and Philips, 2010).

There are clear advantages of LED lamps over their conventional counterparts, including:

- High luminous efficiency and, consequentially, low power consumption (67 lm/W, Philips, 2010; 94 lm/W, Philips, 2011; commercial LED light bulbs with source efficacies > 100

lm/W, by multiple manufacturers, 2015)

- High quality color rendition (high color rendering index, *CRI*)
- Available in a wide range of correlated color temperatures (2 700~6 000 K)
- No Hg (mercury) is contained in the lamp (in contrast to fluorescent lamps)
- No UV and IR radiation leading to the bleaching of fading of colored objects
- Long lifetime, e.g. 25 000 hours of operation (at 70% lumen maintenance)

These clear advantages make LED lamps a convincing proposition so that LED lamps are expected to displace virtually all conventional light sources.

3.5 References

- Audi car company “Lumileds’ Luxeon LEDs blaze new trails with Audi A8 6.0, the first production vehicle with LED daytime running lights” press release by Lumileds Lighting company, January 8 (2004)
- Bando K., Noguchi Y., Sakano K., and Shimizu Y. “Development and application of high-brightness white LEDs” (in Japanese) *Technical Digest, Phosphor Research Society*, 264th Meeting, November 29 (1996)
- Bando K., Sakano K., Noguchi Y., and Shimizu Y. “Development of high-bright and pure-white LED lamps” *Journal of Light and Visual Environments* **22**, 2 (1998)
- Baretz B. and Tischler M. A. “Solid-state white light emitter and display using the same” US Patent 6,600,175 B1 having a priority date of March 26, 1996 (2003)
- Bergh A., Crawford M. G., Duggal A., and Haitz R. “The promise and challenge of solid-state lighting” *Physics Today*, December 2001, page 42 (2001)
- Blasse G. and Bril A. “A new phosphor for flying-spot cathode-ray tubes for color television: Yellow-emitting $\text{Y}_3\text{Al}_5\text{O}_{12}\text{-Ce}^{3+}$ ” *Applied Physics Letters* **11**, 53 (1967)
- Crawford M. H. “LEDs for solid-state lighting: Performance challenges and recent advances” *IEEE Journal of Selected Topics in Quantum Electronics* **15**, 1028, (2009)
- Cree Company (press release) “Cree breaks 200 lumen per watt efficacy barrier” (February 3, 2010)
- Cree Company (press release) “Cree first to break 300 lumens-per-watt barrier” Internet page: <<http://www.cree.com/News-and-Events/Cree-News/Press-Releases/2014/March/300LPW-LED-barrier>> (March 26, 2014)
- Editorial in *Nature Photonics* **1**, 23 (2006)
- Haitz R., Kish F., Tsao J., and Nelson J. “The case for a national research program on semiconductor lighting” Sandia Report SAND-2000-1612 (2000)
- Haitz R. and Tsao J. Y. “Solid-state lighting: ‘The case’ 10 years after and future prospects” *Physica Status Solidi A* **208**, 17 (2011)
- Kane R. and Sell H. (editors) “Revolution in lamps: A chronicle of 50 years of progress (2nd edition)” (Fairmont Press, Lilburn, Georgia, 2002)
- Krames M. R., Christenson G., Coffins D., Cook L. W., Crawford M. G., Edwards A., Fletcher R. M., Gardner N., Goetz W., Imler W., Johnson E., Kern R. S., Khare R., Kish F. A., Lowery C., Ludowise M. J., Mann R., Maranowski M., Maranowski S., Martin P. S., O’Shea J., Rudaz S., Steigerwald D., Thompson J., Wierer J. J., Yu J., Basile D., Chang Y.-L., Hasnain G., Hueschen M., Killeen K., Kocot C., Lester S., Miller J., Mueller G., Mueller-Mach R., Rosner J., Schneider R., Takeuchi T., and Tan T. S. “High-brightness AlGaN light emitting diodes” *Proceedings of the SPIE* **3938**, 2 (2000)
- Mukai T. presentation at *SPIE Photonics West* (January 29, 2009); see also *Compound Semiconductor* “Nichia’s 249 lm/W LEDs mark tech shift” page 5 (March 2009)
- Nakamura S., Mukai T., and Senoh M. “Candela-class high-brightness InGaN/AlGaN double-heterostructure blue-light-emitting diodes” *Applied Physics Letters* **64**, 1687 (1994)
- Nakamura S., Senor M., Iwasa N., Nagarama S.-I., Yamada T., and Mukai T. “Super-bright green InGaN

- single-quantum-well-structure light-emitting diodes" *Japanese Journal of Applied Physics* **34**, 1332 (1995)
- Nakamura S., Fasol G. "The blue laser diode: GaN based light emitters and lasers" (Springer-Verlag, Berlin, Germany, 1997)
- Narukawa Y. "White-light LEDs" *OSA Optics and Photonics News*, April 2004 issue, page 25 (2004)
- Narukawa Y., Narita J., Sakamoto T., Deguchi K., Yamada T. and Mukai T. "Ultra-high efficiency white light emitting diodes" *Japanese Journal of Applied Physics* **45**, L1084 (2006)
- Narukawa Y., Ichikawa M., Sanga D., Sano M., and Mukai T. "White light emitting diodes with super-high luminous efficacy" *Journal of Physics D: Applied Physics* **43**, 354002 (2010)
- Nikkei Sangyo Shimbun "White LED lamp: Light emission with high luminous efficiency halving production costs" (in Japanese); *Nikkei Sangyo Shimbun* (Nikkei Industrial Journal) published in Japan by Nikkei Inc. (September 13, 1996)
- Osram Opto Semiconductors "LEDs bridge time and space" press release, June 17 (2004)
- Osram Company (press release) "Osram constructs the world's most efficient LED lamp [with 205 lm/W system efficiency]" Internet page: http://www.osram.com/osram_com/press/press-releases/_trade_press/2014/osram-constructs-the-worlds-most-efficient-led-lamp/index.jsp (March 28, 2014)
- Peralta S. and Ruda H. "Applications for advanced solid-state lamps" *IEEE Industry Applications Magazine*, July & August issue, page 31 (1998)
- Philips Lumileds "Philips launches 60 W-replacement LED lamp that consumes just 12 W" in *Semiconductor Today* <<http://www.semiconductor-today.com>> April 16 (2010)
- Potter R. M., Pike P., and Galginaitis S. V. "Light-emitting phosphor-diode combination" US Patent 3,529,200 (1970)
- Rea M. S. as quoted in the article titled "Fans of LEDs say this bulb's time has come" *New York Times*, Technology Section, published on July 28 (2008)
- Reeh U., Höhn K., Stath N., Waitl G., Schlotter P., Schneider J., Schmidt R. "Light-radiating semiconductor component with luminescence conversion element" US patent 7,078,732 B1 having priority dates of June 26, 1996 and September 20, 1996 (2006)
- Schanda János, personal communication. The photographs of an early GaAsP LED and a large-scale color display are gratefully acknowledged (2007)
- Schlotter P., Schmidt R., and Schneider J. "Luminescence conversion of blue light emitting diodes" *Applied Physics* **A64**, 417 (1997)
- Schubert E. F. and Kim J. K. "Solid-state light sources getting smart" *Science* **308**, 1274 (2005)
- Schubert E. F., Kim J. K., Luo H., and Xi J.-Q. "Solid-state lighting – a benevolent technology" *Reports on Progress in Physics* **69**, 3069 (2006)
- Shimada J., Kawakami Y., and Fujita S. "Medical lighting composed of LEDs arrays for surgical operation" *SPIE Photonics West: Light Emitting Diodes: Research, Manufacturing, and Applications*, 4278, 165, San Jose, January 24~25 (2001)
- Shimada J., Kawakami Y., and Fujita S. "Development of lighting goggle with power white LED modules" *SPIE Photonics West: Light Emitting Diodes: Research, Manufacturing, and Applications*, 4996, 174, San Jose, January 28~29 (2003)
- Shimizu Y. "Planar light source" Japanese Publication of Unexamined Patent Application H07-176794, July 14, 1995 (1995)
- Shimizu Y. "Planar light source" Japanese Patent Application Publication H08-7614 January 12, 1996 (1996)
- Shimizu Y., Sakano K., Noguchi Y., and Moriguchi T. "Light-emitting diode and display device using the same" Japanese priority patent applications to US patent 5,998,925 filed in 1996: JPA H8-198585 (filed July 29, 1996), JPA H8-244339 (filed September 17, 1996), JPA H8-245381 (filed September 18, 1996), and JPA H8-359004 (filed December 27, 1996)
- Shimizu Y., Sakano K., Noguchi Y., and Moriguchi T. "Light-emitting device having a nitride compound semiconductor and a phosphor containing a garnet fluorescent material" US Patent 5,998,925 (1999)
- Steigerwald D. A., Bhat J. C., Collins D., Fletcher R. M., Holcomb M. O., Ludowise M. J., Martin P. S., and

- Rudaz S. L. "Illumination with solid state lighting technology" *IEEE Journal on Selected Topics in Quantum Electronics* **8**, 310 (2002)
- Stevenson D. A., Rhines W. C., and Maruska H. P. "Gallium nitride metal-semiconductor junction light emitting diode" US Patent 3,819,974 (1974)
- Stinson J. W. "Variable color light emitting diode" US Patent 4,992,704 (1991)
- Sylvania "LED A-line lamps" Internet address: <www.Sylvania.com> (2010)
- Tabuchi S. "Light emitting semiconductor apparatus" Japanese Utility Model Patent Application Publication No. S50-79379, November 24, 1973 (1973)
- Tokailin H. and Hosokawa C. "Electroluminescent element" US Patent 5,126,214 (1992)
- Tsao J. Y. "Solid-state lighting: Science, technology and economic perspectives" Plenary Presentation *SPIE Photonics West*, January 23~28 (2010)
- Tsao J. Y., Coltrin M. E., Crawford M. H., and Simmons J. A. "Solid-state lighting: An integrated human factors, technology, and economic perspective" *Proceedings of the IEEE* **98**, 1162 (2010)

4

4 – Radiative and non-radiative recombination

Electrons and holes in semiconductors recombine either radiatively, i.e. accompanied by the emission of a photon, or non-radiatively. In light-emitting devices, the former is clearly the preferred process. However, non-radiative recombination can, under practical conditions, never be reduced to zero. Thus, there is competition between radiative and non-radiative recombination. Maximization of the radiative process and minimization of the non-radiative process can be attained in a number of ways which will be discussed below.

4.1 Radiative electron–hole recombination

Any undoped or doped semiconductor has two types of free carriers, electrons and holes. Under equilibrium conditions, i.e. without external stimuli such as light or current, the law of mass action teaches that the product of the electron and hole concentrations is, at a given temperature, a constant, i.e.

$$n_0 p_0 = n_i^2 \quad (4.1)$$

where n_0 and p_0 are the equilibrium electron and hole concentrations and n_i is the intrinsic carrier concentration. The validity of the law of mass action is limited to non-degenerately doped semiconductors (see, for example, Schubert, 1993).

Excess carriers in semiconductors can be generated either by absorption of light or by an injection current. The total carrier concentration is then given by the sum of equilibrium and excess carrier concentrations, i.e.

$$n = n_0 + \Delta n \quad \text{and} \quad p = p_0 + \Delta p \quad (4.2)$$

where Δn and Δp are the excess electron and hole concentrations, respectively.

Next, we consider recombination of carriers. The band diagram of a semiconductor with electrons and holes is shown in *Figure 4.1*. We are interested in the rate at which the carrier concentration decreases and denote the recombination rate as R . Consider a free electron in the conduction band. The probability that the electron recombines with a hole is proportional to the

hole concentration, that is, $R \propto p$. The number of recombination events will also be proportional to the concentration of electrons, as indicated in **Figure 4.1**. Thus, the recombination rate is proportional to the product of electron and hole concentrations, that is, $R \propto n p$. Using a proportionality constant, the recombination rate per unit time per unit volume can be written as

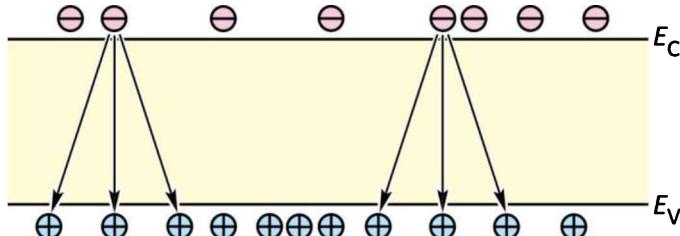


Fig. 4.1: Electron–hole recombination. The number of recombination events per unit time per unit volume is proportional to the product of electron and hole concentrations, i.e. $R \propto np$.

$$R = -\frac{dn}{dt} = -\frac{dp}{dt} = B n p . \quad (4.3)$$

This equation is the **bimolecular rate equation** and the proportionality constant B is called the **bimolecular recombination coefficient**. It has typical values of $10^{-11} \sim 10^{-9} \text{ cm}^3/\text{s}$ for direct-gap III–V semiconductors. The bimolecular recombination coefficient will be calculated in a subsequent section using the van Roosbroeck–Shockley model.

4.2 Radiative recombination for low-level excitation

Next, we discuss the recombination dynamics as a function of time. Consider a semiconductor subjected to photoexcitation. The equilibrium and excess electron and hole concentrations are n_0 , p_0 , Δn , and Δp , respectively. Since electrons and holes are generated and annihilated (by recombination) in pairs, the steady-state electron and hole excess concentrations are equal,

$$\Delta n(t) = \Delta p(t) . \quad (4.4)$$

Using the bimolecular rate equation, the recombination rate is given by

$$R = B [n_0 + \Delta n(t)] [p_0 + \Delta p(t)] . \quad (4.5)$$

For the case of **low-level excitation**, the photogenerated carrier concentration is much smaller than the majority carrier concentration, i.e. $\Delta n \ll (n_0 + p_0)$. Using $\Delta n(t) = \Delta p(t)$, one obtains from Eqn. (4.5)

$$R = B n_i^2 + B (n_0 + p_0) \Delta n(t) = R_0 + R_{excess} \quad (4.6)$$

The first summand on the right-hand side of the equation can be identified as the *equilibrium recombination rate* (R_0) and the second term as the *excess recombination rate* (R_{excess}).

The time-dependent carrier concentration can be calculated from the rate equation

$$\frac{dn(t)}{dt} = G - R = (G_0 + G_{\text{excess}}) - (R_0 + R_{\text{excess}}) \quad (4.7)$$

where G_0 and R_0 are the equilibrium generation and recombination rates, respectively.

Next, we assume that the semiconductor has been illuminated with light and excess carriers are generated. At the time $t = 0$, the illumination is switched off (i.e. $G_{\text{excess}} = 0$) as indicated in **Figure 4.2**. The recombination rate can then be calculated by insertion of Eqn. (4.6) into Eqn. (4.7) and using $G_0 = R_0$. This yields the differential equation

$$\frac{d}{dt} \Delta n(t) = -B(n_0 + p_0) \Delta n(t). \quad (4.8)$$

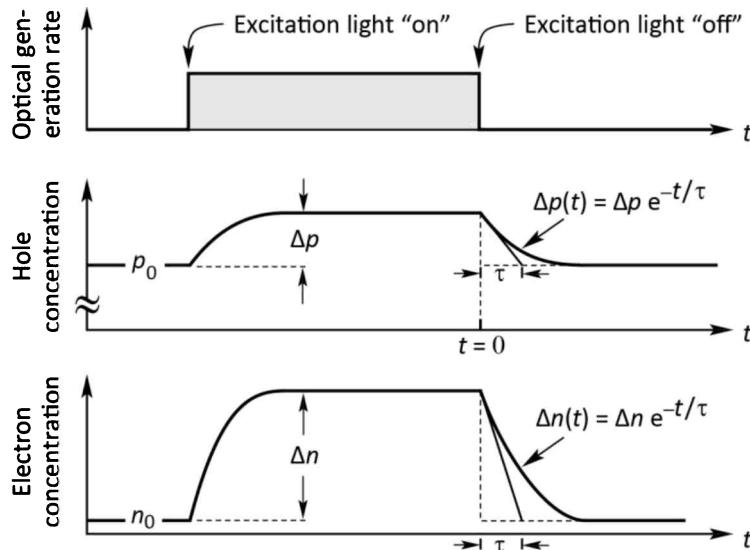


Fig. 4.2: Carrier concentration as a function of time before, during, and after an optical excitation pulse. The semiconductor is assumed to be p-type and thus $p_0 \gg n_0$. Electrons and holes are generated in pairs, thus $\Delta p = \Delta n$. Under low-level excitation as shown here, it is $\Delta n \ll p_0$. In most practical cases the equilibrium minority carrier concentration is extremely small so that $n_0 \ll \Delta n$.

The solution of the differential equation can be obtained by separation of variables. One obtains

$$\Delta n(t) = \Delta n_0 e^{-B(n_0+p_0)t} \quad (4.9)$$

where $\Delta n_0 = \Delta n(t = 0)$. Rewriting the result as

$$\Delta n(t) = \Delta n_0 e^{-t/\tau} \quad (4.10)$$

allows one to identify the **carrier lifetime** τ as

$$\tau = \frac{1}{B(n_0 + p_0)} . \quad (4.11)$$

For semiconductors with a specific doping type, Eqn. (4.11) reduces to

$$\tau_n = \frac{1}{B p_0} = \frac{1}{B N_A} \quad (\text{for p-type semiconductors}) \quad (4.12)$$

and

$$\tau_p = \frac{1}{B n_0} = \frac{1}{B N_D} \quad (\text{for n-type semiconductors}) \quad (4.13)$$

where τ_n and τ_p are the electron and hole lifetimes, respectively. Using this result, the rate equation, Eqn. (4.8), can be simplified for semiconductors of a specific conductivity type. One obtains the **monomolecular rate equations**:

$$\frac{d}{dt} \Delta n(t) = -\frac{\Delta n(t)}{\tau_n} \quad (\text{for p-type semiconductors}) \quad (4.14)$$

and

$$\frac{d}{dt} \Delta p(t) = -\frac{\Delta p(t)}{\tau_p} \quad (\text{for n-type semiconductors}). \quad (4.15)$$

Figure 4.2 shows the majority and minority carrier concentrations in a p-type semiconductor as a function of time (similar considerations apply if an n-type semiconductor is chosen). Note that the figure shows the case of *low-level excitation* in which the photogenerated carrier concentration is much smaller than the majority carrier concentration. However, the photogenerated carrier concentration is much larger than the minority carrier concentration.

Once photoexcitation is terminated, the minority carrier concentration decays exponentially with a characteristic time constant denoted as the **minority carrier lifetime** τ . It is the mean time between generation and recombination of a minority carrier.

Note that the majority carrier concentration also decays with the same time constant τ . However, only a very *small fraction* of the majority carriers disappears by recombination, as illustrated in **Figure 4.2**. Thus, for low-level excitation, the average time it takes for a majority carrier to recombine is much longer than the minority carrier lifetime. For many practical purposes, the majority carrier lifetime can be assumed to be infinitely long.

Theoretical and experimental values for the minority carrier lifetime in GaAs as a function of the doping concentration are shown in **Figure 4.3** (Hwang, 1971; Nelson and Sobers, 1978a,

1978b; Ehrhardt *et al.*, 1991; Ahrenkiel, 1993). The theoretical line in the figure is calculated from Eqn. (4.10) using $B = 10^{-10} \text{ cm}^3/\text{s}$. In nominally undoped material, minority carrier lifetimes as long as 15 μs have been measured in GaAs at room temperature (Nelson and Sobers, 1978a, 1978b).

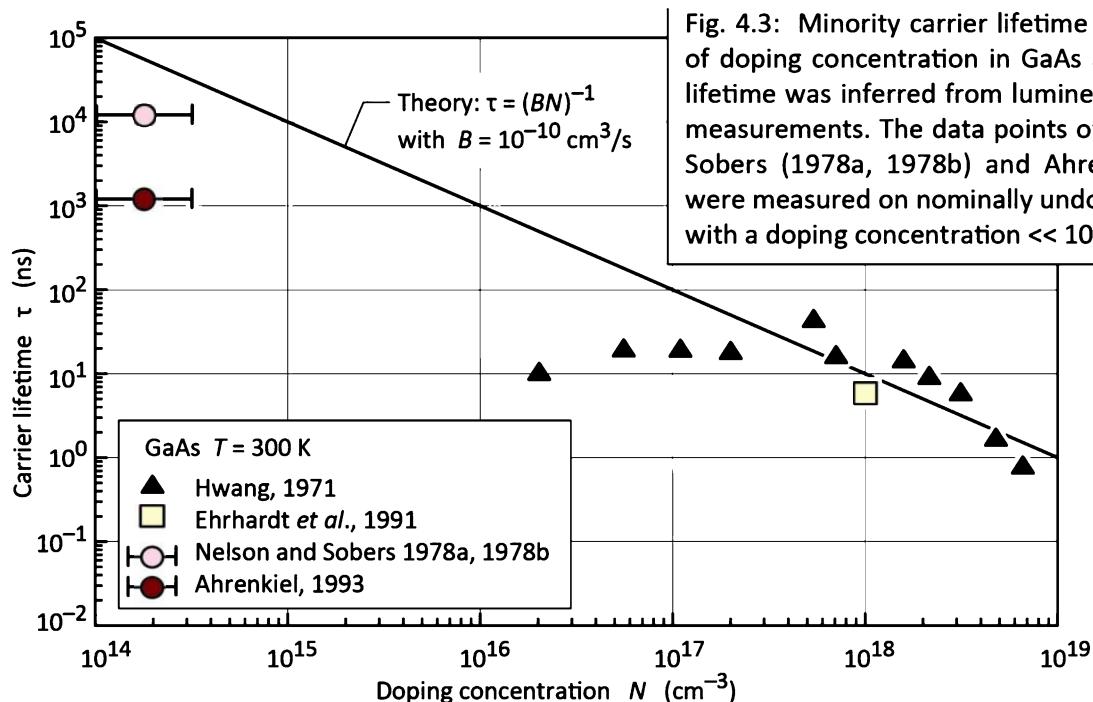


Fig. 4.3: Minority carrier lifetime as a function of doping concentration in GaAs at 300 K. The lifetime was inferred from luminescence decay measurements. The data points of Nelson and Sobers (1978a, 1978b) and Ahrenkiel (1993) were measured on nominally undoped material with a doping concentration $<< 10^{15} \text{ cm}^{-3}$.

Exercise: Minority carrier lifetimes. Calculate the minority carrier lifetime in p-type GaAs at doping concentrations of 10^{15} and 10^{18} cm^{-3} using a bimolecular recombination coefficient of $B = 10^{-10} \text{ cm}^3/\text{s}$. Assume that one could fabricate GaAs without any impurities. What is the carrier lifetime in intrinsic GaAs with an intrinsic carrier concentration of $2 \times 10^6 \text{ cm}^{-3}$?

Solution:

$\tau_n = 10 \mu\text{s}$	for $N_A = 10^{15} \text{ cm}^{-3}$
$\tau_n = 10 \text{ ns}$	for $N_A = 10^{18} \text{ cm}^{-3}$
$\tau = 2500 \text{ s}$	for undoped GaAs.

Discuss how the modulation speed of communication LEDs is affected by the radiative lifetime and the doping concentration.

Solution: When the injection current of an LED is switched off instantaneously at $t = 0$ and the radiative lifetime is τ , then light emission from the LED essentially ceases after the time $t = \tau$. Thus, an LED cannot be switched “on” or “off” faster than the time τ . Assuming that both the switch-on and switch-off time is τ , then the maximum modulation frequency is given by $f \approx 1/(2\tau)$.

We have shown that the radiative lifetime τ depends strongly on the doping concentration. Thus, by highly doping the active region of a device, the radiative lifetime τ is shortened and the maximum modulation speed is increased.

4.3 Radiative recombination for high-level excitation

For the case of **high-level excitation**, the photogenerated carrier concentration is larger than the equilibrium carrier concentration, i.e. $\Delta n \gg (n_0 + p_0)$. The bimolecular rate equation (Eqn. 4.3) is then given by

$$\frac{d \Delta n(t)}{dt} = -B \Delta n^2 . \quad (4.16)$$

Solving this differential equation by the separation-of-variables method and using the initial condition $\Delta n(0) = \Delta n_0$ yields the solution

$$\Delta n(t) = \frac{1}{B t + \Delta n_0^{-1}} . \quad (4.17)$$

This solution is, in contrast to the low-density approximation, a *non-exponential* carrier decay.

In an exponential decay, it takes the time constant τ for the carrier concentration to decrease from Δn_0 to $\Delta n_0 e^{-1}$. Using the same definition for the non-exponential decay given by Eqn. (4.17), the “time constant” can be calculated from the slope of the decay by using the equation

$$\tau(t) = -\frac{\Delta n(t)}{\frac{d\Delta n(t)}{dt}} . \quad (4.18)$$

Using this definition for the non-exponential decay of Eqn. (4.17), one obtains the “time constant”

$$\tau(t) = t + \frac{1}{B \Delta n_0} . \quad (4.19)$$

Thus, for non-exponential decays, the “time constant” depends on time. Equation (4.19) shows that the minority carrier lifetime increases with time. For sufficiently long times, low-level excitation conditions will be reached and τ will approach the low-level value.

4.4 Bimolecular rate equations for quantum well structures

Quantum wells provide a means of confining the free carriers to a narrow quantum well region by using the two barrier regions cladding the quantum well. Assume that the well region has a thickness of L_{QW} . Assume further that the conduction band and valence band wells have carrier densities of n^{2D} and p^{2D} , respectively. The effective three-dimensional (3D) carrier concentration

for electrons and holes can be approximated by n^{2D}/L_{QW} and p^{2D}/L_{QW} , respectively. Using these values as the 3D carrier concentration, the recombination rate can be inferred from Eqn. (4.5), and it is given by

$$R = B (n^{2D}/L_{QW}) (p^{2D}/L_{QW}). \quad (4.20)$$

This equation illustrates one of the essential advantages of quantum-well and double heterostructures. A decrease of the quantum well thickness allows one to attain high 3D carrier concentrations (carriers per cm³). As a result, the carrier lifetime for radiative recombination is reduced, as inferred from Eqn. (4.11), and the radiative efficiency is increased.

For sufficiently small quantum well thicknesses, the wave function no longer scales with the physical well width. L_{QW} must be replaced by the carrier distribution width, which for sufficiently small well thicknesses is larger than L_{QW} , since the wave function will extend into the barriers. This effect should be considered for well thicknesses < 10 nm in the AlGaAs/GaAs material system.

4.5 Luminescence decay

The carrier decay in semiconductors can be measured by the decay of the luminescence after a short optical excitation pulse. The luminescence intensity is proportional to the recombination rate. Calculating the recombination rate for the low and high excitation cases (Eqns. 4.9 and 4.17), one obtains

$$R = -\frac{dn(t)}{dt} = \frac{\Delta n_0}{\tau} e^{-t/\tau} \quad (\text{for low excitation}) \quad (4.21)$$

and

$$R = -\frac{dn(t)}{dt} = \frac{-B}{(B t + \Delta n_0^{-1})^2} \quad (\text{for high excitation}). \quad (4.22)$$

Figure 4.4 illustrates the decay of the luminescence after optical excitation by a short pulse. For the case of low excitation density, the luminescence decay is exponential with a time constant τ ; the luminescence decay's time constant is equal to the carrier lifetime. For the case of high excitation, the decay is non-exponential. Non-exponential decay functions can be expressed by an exponential function *with a time-dependent time constant*, i.e. $\exp [-t/\tau(t)]$. In most cases, the time constant τ increases with time. This type of decay function is known as a **stretched exponential decay function**. It describes a *slower-than-exponential* decay.

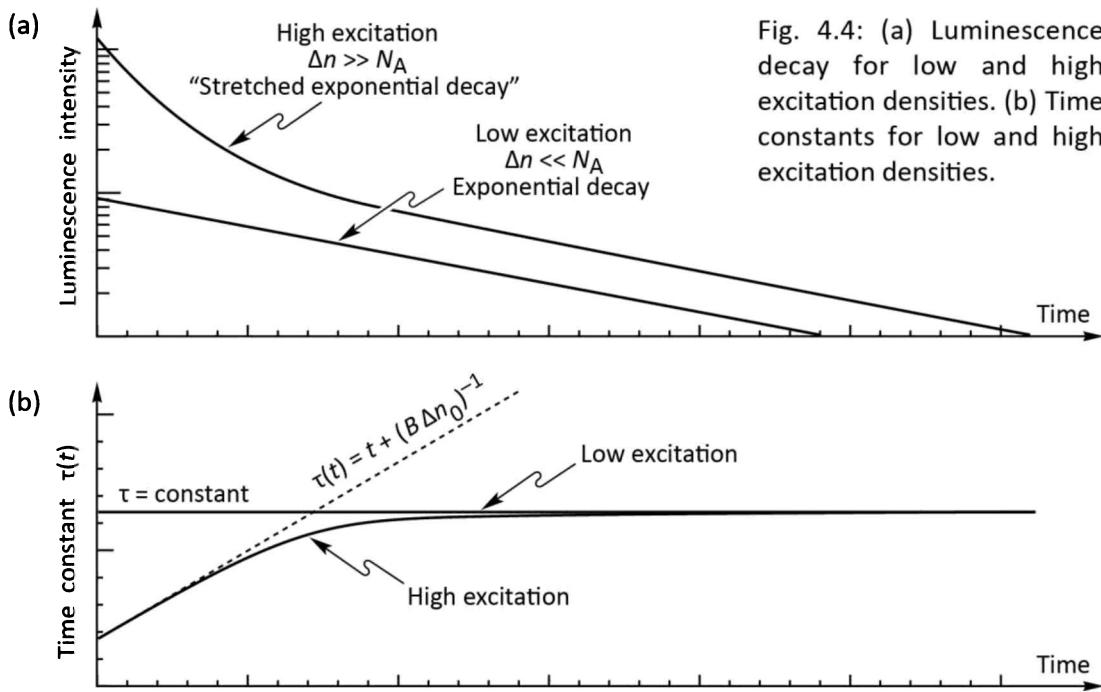


Fig. 4.4: (a) Luminescence decay for low and high excitation densities. (b) Time constants for low and high excitation densities.

A particularly well-known stretched exponential function is given by $\exp [-(t/\tau)^\beta]$, where β , the ***disorder parameter***, represents the disorder of the radiative material. For $\beta = 1$, no disorder exists in the material. For $\beta \approx \frac{1}{2}$, a strong disorder exists in the material and such values of β have been found in glasses (Phillips, 1996) as well as semiconductors. Non-exponential decays were first discovered and discussed by Friedrich Kohlrausch in the late 1800s and such decays are therefore also referred to as ***Kohlrausch decays***.

The recombination dynamics of carriers in LEDs is one of the factors that limits the time it takes to switch an LED on and off. The modulation speed of LEDs used for communication applications can be limited by the minority carrier lifetime. The carrier lifetime can be reduced by either a high doping of the active region or a high concentration of injected carriers in the active region. Heterostructures that confine free carriers to the small well region are frequently employed to obtain high carrier concentrations and thus short carrier lifetimes.

4.6 Shockley–Read–Hall non-radiative recombination

There are two basic recombination mechanisms in semiconductors, namely *radiative* recombination and *non-radiative* recombination. In a radiative recombination event, one photon with energy equal to the bandgap energy of the semiconductor is emitted, as illustrated in **Figure 4.5**. During non-radiative recombination, the electron energy is converted to vibrational energy of lattice atoms, i.e. phonons. Thus, the electron energy is converted to heat. For obvious reasons, non-radiative recombination events are unwanted in light-emitting devices.

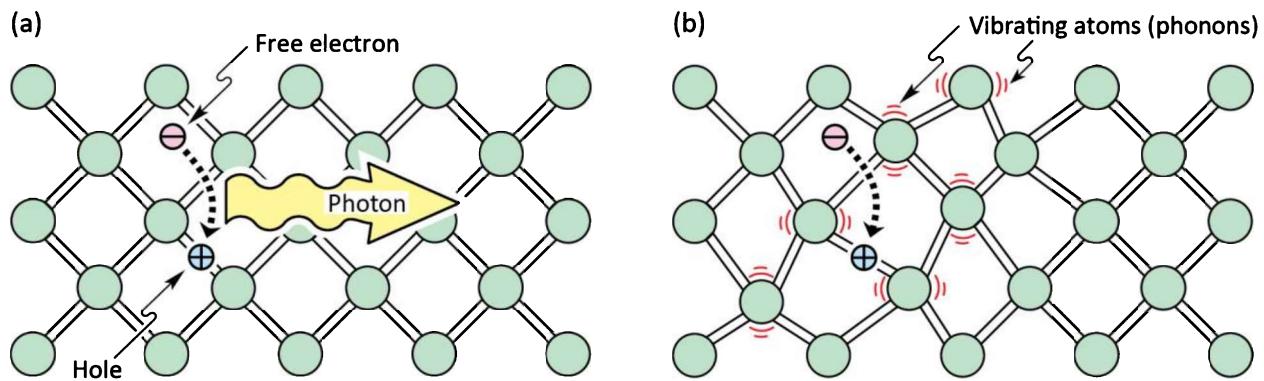


Fig. 4.5: (a) Radiative recombination accompanied by the emission of a photon with energy $h\nu \approx E_g$. (b) In non-radiative recombination events, the energy released during the electron-hole recombination is converted to phonons (adapted from Shockley, 1950).

There are several physical mechanisms by which non-radiative recombination can occur. Defects in the crystal structure are the most common cause for non-radiative recombination. These defects include unwanted foreign atoms, native defects, dislocations, and any complexes of defects, foreign atoms, or dislocations. In compound semiconductors, native defects include interstitials, vacancies, and antisite defects (Longini and Greene, 1956; Baraff and Schluter, 1985). All such defects have energy level structures that are different from substitutional semiconductor atoms. It is quite common for such defects to form one or several energy levels within the forbidden gap of the semiconductor.

Energy levels within the gap of the semiconductor are efficient recombination centers; in particular, if the energy level is close to the middle of the gap. The recombination of carriers via a trap level is shown schematically in *Figure 4.6*. Owing to the promotion of non-radiative processes, such ***deep levels*** or traps are also called ***luminescence killers***.

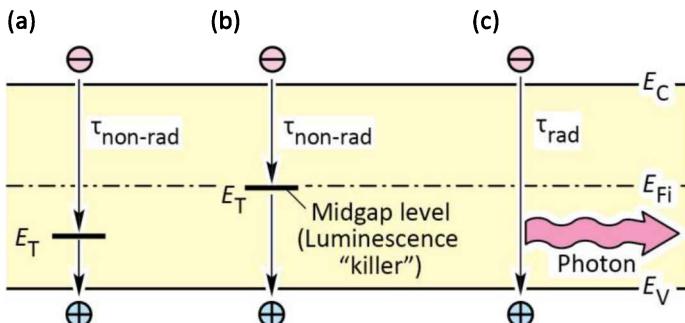


Fig. 4.6: Band diagram showing:
 (a) non-radiative electron-hole recombination via a deep level;
 (b) non-radiative recombination via a midgap-level which is particularly efficient and thus harmful; midgap-levels are referred to as "luminescence killers"; (c) radiative recombination.

The recombination of free carriers via deep levels was first analyzed by Shockley, Read, and Hall (Hall, 1952; Shockley and Read, 1952). The non-radiative recombination rate through a deep level

with trap energy E_T and concentration N_T is given by (Shockley and Read, 1952)

$$R_{SRH} = \frac{p_0 \Delta n + n_0 \Delta p + \Delta n \Delta p}{(N_T v_p \sigma_p)^{-1}(n_0 + n_1 + \Delta n) + (N_T v_n \sigma_n)^{-1}(p_0 + p_1 + \Delta p)} \quad (4.23)$$

where $\Delta n = \Delta p$; v_n and v_p are the electron and hole thermal velocities, and σ_n and σ_p are the capture cross sections of the traps. The quantities n_1 and p_1 are the electron and hole concentrations if the Fermi energy is located at the trap level. These quantities are given by

$$n_1 = n_i \exp\left(\frac{E_T - E_{Fi}}{kT}\right) \quad \text{and} \quad p_1 = n_i \exp\left(\frac{E_{Fi} - E_T}{kT}\right) \quad (4.24)$$

where E_{Fi} is the Fermi level in the intrinsic semiconductor.

The non-radiative lifetime of excess electrons can be deduced from the equation $R_{SRH} = \Delta n / \tau$. Consequently, the lifetime is given by

$$\frac{1}{\tau} = \frac{p_0 + n_0 + \Delta n}{(N_T v_p \sigma_p)^{-1}(n_0 + n_1 + \Delta n) + (N_T v_n \sigma_n)^{-1}(p_0 + p_1 + \Delta p)} \quad (4.25)$$

We now differentiate between majority and minority carriers and assume that the semiconductor is p-type. Then, $p_0 \gg n_0 p_0$ and $p_0 \gg p_1$. If we further assume a small deviation from equilibrium, i.e. $\Delta n \ll p_0$, then the minority carrier lifetime is given by

$$\frac{1}{\tau} = \frac{1}{\tau_{n_0}} = N_T v_n \sigma_n \quad (4.26)$$

If electrons were the majority carriers, the lifetime would be obtained in an analogous way, i.e.

$$\frac{1}{\tau} = \frac{1}{\tau_{p_0}} = N_T v_p \sigma_p \quad (4.27)$$

The results show that the Shockley–Read–Hall recombination rate is limited by the rate of capture of minority carriers. This result suggests itself since the capture of *majority* carriers is a much more *likely* event than the capture of *minority* carriers. Equation (4.25) can then be written as

$$\frac{1}{\tau} = \frac{p_0 + n_0 + \Delta n}{\tau_{p_0}(n_0 + n_1 + \Delta n) + \tau_{n_0}(p_0 + p_1 + \Delta p)} \quad (4.28)$$

For small deviations from equilibrium, i.e. $\Delta n \ll p_0$, the equation simplifies to

$$\tau = \tau_{n_0} \frac{p_0 + p_1}{p_0 + n_0} + \tau_{p_0} \frac{n_0 + n_1 + \Delta n}{p_0 + n_0} \approx \tau_{n_0} \frac{p_0 + p_1}{p_0 + n_0} . \quad (4.29)$$

Inspection of the equation reveals that the lifetime does not change for small deviations from equilibrium in an extrinsic semiconductor.

For further insight, we assume that the trap captures electrons and holes at the same rate, i.e. $v_n \sigma_n = v_p \sigma_p$ and $\tau_{n_0} = \tau_{p_0}$. One obtains from Eqn. (4.29)

$$\tau = \tau_{n_0} \left(1 + \frac{p_0 + p_1}{p_0 + n_0} \right) . \quad (4.30)$$

For the special case of intrinsic material, i.e. $n_0 = p_0 = n_i$, the equation simplifies to

$$\tau_i = \tau_{n_0} \left(1 + \frac{p_1 + n_1}{2 n_i} \right) = \tau_{n_0} \left[1 + \cosh \left(\frac{E_T - E_{Fi}}{kT} \right) \right] \quad (4.31)$$

where E_{Fi} is the intrinsic Fermi level, which is typically close to the middle of the gap. The cosh function has a minimum when the argument of the function is zero. Thus, the non-radiative lifetime is minimized if $E_T - E_{Fi}$ is zero; i.e. when the trap level is at or close to the midgap energy. For such midgap levels, the lifetime is given by $\tau = 2\tau_{n_0}$. This result demonstrates that deep levels are effective recombination centers if they are near the middle of the gap.

Inspection of Eqn. (4.31) also reveals the temperature dependence of Shockley–Read–Hall recombination: As T increases, the non-radiative recombination lifetime *decreases*. As a result, the radiative band-to-band recombination efficiency *decreases* at high temperatures. Consequently, the highest band-to-band radiative efficiencies of direct-gap semiconductors can be obtained at cryogenic temperatures. Since the cosh function in Eqn. (4.31) contains exponential terms, an exponential temperature dependence of the Shockley–Read–Hall lifetime (τ_{SRH}) should be expected. However, a more detailed study shows that a weaker-than-exponential dependence on temperature may be found: Schenk (1992) and Bleichner *et al.* (1996) found temperature dependences that range between $\tau_{SRH} \propto T^{-1}$ and $\tau_{SRH} \propto T^{-5/2}$. Nevertheless, it can be assumed that the SRH recombination process has a stronger temperature dependence than radiative recombination.

However, some devices are based on radiative recombination through a deep state. A well-known example of radiative recombination mediated by a deep level is N-doped GaP. It follows

from the Shockley–Read–Hall model that the deep-level recombination rate increases with increasing temperature.

In *indirect-gap* semiconductors such as GaP, radiative transitions are mediated by phonons. That is, radiative recombination must be accompanied by absorption or emission of a phonon. Since phonons are more abundant at high temperatures, radiative recombination (mediated by the absorption of a phonon) can increase with temperature.

In the vicinity of a deep level, the luminescence intensity decreases. A single ***point defect*** will be difficult to observe due to its relatively small effect. Frequently, however, defects group into clusters of defects or ***extended defects***. Such extended defects are, for example, threading dislocations and misfit dislocations occurring when epitaxial semiconductors are grown on mismatched substrates, i.e. substrates with a different lattice constant from that of the epitaxial layer. There are also many other types of extended defects. The luminescence-killing nature of extended defects is illustrated in **Figure 4.7**, which shows a cathodoluminescence micrograph of a GaAs layer measured at room temperature. The figure reveals several dark spots. Luminescence in the vicinity of the defects is reduced due to the non-radiative recombination channels so that the defects manifest themselves as dark spots. The size of the dark spots depends on the size of the defect and the minority carrier diffusion length.

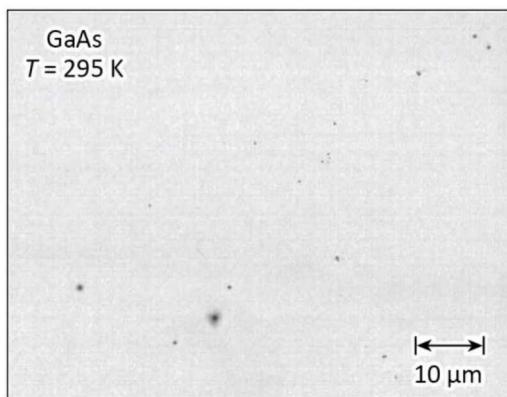


Fig. 4.7: Cathodoluminescence micrograph of a GaAs epitaxial layer. The dark spots are due to large clusters of non-radiative recombination centers (after Schubert, 1995).

While most deep-level transitions are non-radiative, some deep-level transitions are radiative. An example of a radiative deep-level transition in GaN is shown in **Figure 4.8** (Grieshaber *et al.*, 1996). The luminescence spectrum shows the band-to-band transition at 365 nm and a broad deep-level transition around 550 nm. The deep-level transition occurs near the yellow range of the visible spectrum. The yellow luminescence line has been shown to be due to Ga vacancies

(Neugebauer and Van de Walle, 1996; Schubert *et al.* 1997; Saarinen *et al.* 1997), a common point defect in n-type GaN.

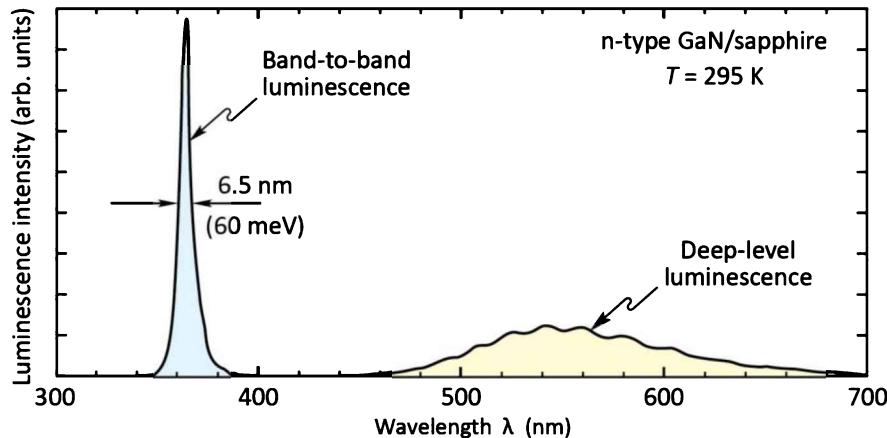


Fig. 4.8: Room-temperature photoluminescence spectrum of GaN with a band-to-band optical transition at 365 nm and a second transition at 550 nm identified as an optically active deep-level transition (after Grieshaber *et al.*, 1996).

Deep levels can be caused by ***native defects*** (group-III vacancies, group-V vacancies, group-III interstitials, and group-V interstitials), unwanted foreign impurities, dislocations, impurity–defect complexes, and combinations of different types of defects. Frequently it takes many years to unambiguously identify the atomic nature of a defect. For a review of defects in semiconductors see, for example, Pantelides (1992).

4.7 Auger non-radiative recombination

Auger recombination is a non-radiative process that occurs predominantly at high carrier concentrations in narrow-gap semiconductors. There are two types of Auger recombination, one involving two electrons and one hole (***eeh process***) and the other one requiring one electron and two holes (***ehh process***). These two processes are illustrated in **Figure 4.9**.

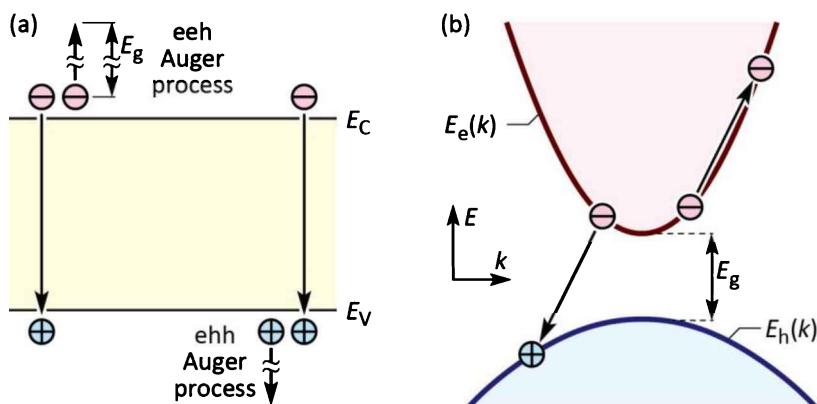


Fig. 4.9: Auger recombination process in (a) position and (b) momentum space. In the eeh process, subsequent to electron excitation high into the conduction band, the electron relaxes down to near the conduction-band edge by emission of longitudinal optical (LO) phonons. The relaxation process is known as thermalization.

During the eeh process, the energy that becomes available due electron–hole recombination (approximately E_g) is transferred to a second electron that is excited high into the conduction band. This highly excited electron subsequently loses its energy by means of ***thermalization***, i.e. emission of multiple phonons, mostly longitudinal optical (LO) phonons, until the electron is close to the band edge. A corresponding statement can be made for the ehh process.

The recombination rates due to the two Auger processes are given by

$$R_{Auger} = C_{eeh} n^2 p \quad (\text{eeh process}) \quad (4.32)$$

and

$$R_{Auger} = C_{ehh} n p^2 \quad (\text{ehh process}). \quad (4.33)$$

Auger recombination is proportional to the square of the carrier concentration (either p^2 or n^2) since two carriers of the same type (either two holes or two electrons) are required for the recombination process. The first process (see Eqn. 4.32) is more likely to happen in p-type semiconductors due to the abundance of holes. The second process (see Eqn. 4.33) is more likely in n-type semiconductors due to the abundance of electrons.

During Auger recombination, energy and momentum must be conserved. Owing to the differences in conduction and valence band structure in semiconductors, the two Auger coefficients C_{eeh} and C_{ehh} are generally different.

In the high-excitation limit, when non-equilibrium carriers have a higher concentration than equilibrium carriers, the Auger rate equations reduce to

$$R_{Auger} = (C_{eeh} + C_{ehh}) n^3 = C_{Auger} n^3 \quad (4.34)$$

where C_{Auger} is the ***Auger coefficient***.

Figure 4.10 illustrates an eeh-type Auger process in a two-band semiconductor in which an electron at $k = 0$ recombines with a hole. The figure uses a geometric construction to satisfy the energy and momentum conservation requirement. The geometric construction allows one to identify allowed Auger transitions and takes into account that the conduction band is populated by fewer electrons as the energy increases (as expressed by the gray shading). Transitions are allowed whenever one of the gray-shaded dispersion curves intersects the $E_e(k)$ curve, as illustrated in the figure.

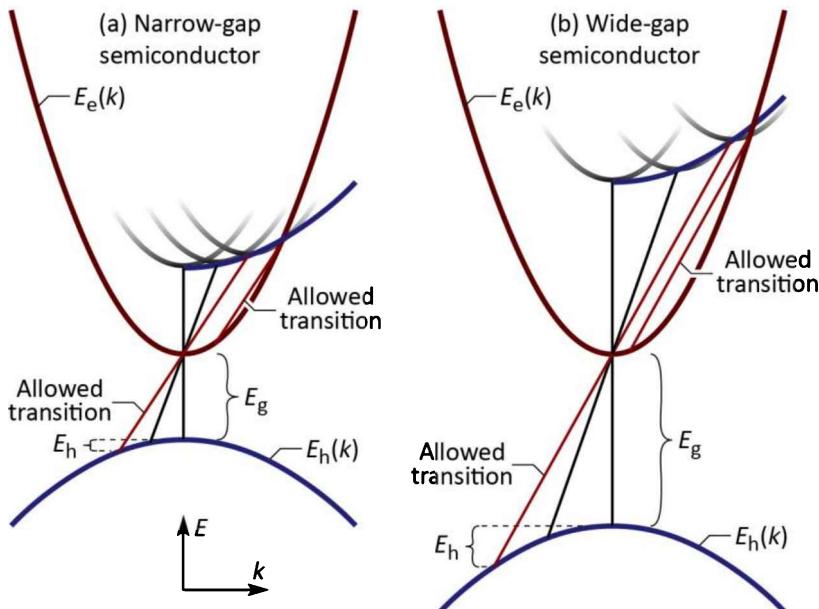


Fig. 4.10: Auger eeh-type recombination process in the E -versus- k diagram for an electron at the bottom of the conduction band recombining with a hole for a (a) narrow-gap and (b) wide-gap semiconductor. The geometric construction shown allows one to determine which transitions are allowed. Generally, eeh transitions in a wide-gap semiconductor require a hole with higher energy, E_h . However, high-energy holes are less abundant thereby making Auger recombination less likely in wide-gap semiconductors.

Figure 4.10 (a) and (b) show an Auger transition for a narrow-gap and a wide-gap semiconductor, respectively, with both semiconductors having the same band dispersions. Inspection of the figure reveals that the hole energies (E_h) increase as the bandgap energy increases. However, since the availability of holes decreases exponentially with energy, the probability of Auger transitions decreases exponentially as the energy gap increases.

A quantitative analysis confirms that Auger rates decrease as the bandgap energy increases. The calculated dependence of Auger recombination on bandgap energy, temperature, and carrier masses was given by (Beattie and Landsberg, 1959; Landsberg, 2003)

$$\tau_{Auger} \propto \left(\frac{E_g}{kT} \right)^{3/2} \exp \left(\frac{1 + 2 \frac{m_e}{m_h}}{1 + \frac{m_e}{m_h}} \right) \exp \left(\frac{E_g}{kT} \right) \quad (4.35a)$$

which yields the proportionality

$$C_{Auger} \propto \exp \left(-E_g / kT \right). \quad (4.35b)$$

Numerical values of the Auger coefficient can be determined by quantum mechanical calculations that take into account the band structure of the semiconductor (see, for example, Agrawal and Dutta, 1986; Landsberg, 2003; Hader 2008). Published values of the Auger coefficient are as follows:

InAsSb:	$C = 1.5 \times 10^{-26} \text{ cm}^6/\text{s}$	(Krier <i>et al.</i> 1999; $\lambda = 4.6 \mu\text{m}$)
InAs:	$C = 2.2 \times 10^{-27} \text{ cm}^6/\text{s}$	(Ioffe, 2010)
GaInAs:	$C = 3.2 \times 10^{-28} \text{ cm}^6/\text{s}$	(Hauser <i>et al.</i> , 1990)
GaInPAs:	$C = 10^{-28} \sim 10^{-29} \text{ cm}^6/\text{s}$	(Olshansky <i>et al.</i> , 1984; Agrawal & Dutta, 1986)
AlGaAs:	$C = 10^{-28} \sim 10^{-29} \text{ cm}^6/\text{s}$	(Olshansky <i>et al.</i> , 1984; Agrawal & Dutta, 1986)
GaInAs QW:	$C = 9.0 \times 10^{-29} \text{ cm}^6/\text{s}$	(Hauser <i>et al.</i> , 1990)
GaSb:	$C = 2.6 \times 10^{-29} \text{ cm}^6/\text{s}$	(Haug, 1987)
GaP:	$C = 1.0 \times 10^{-30} \text{ cm}^6/\text{s}$	(Levinshtein <i>et al.</i> 1996)
InP:	$C = 9.0 \times 10^{-31} \text{ cm}^6/\text{s}$	(Levinshtein <i>et al.</i> 1996)
GaInN QW:	$C = 3.5 \times 10^{-34} \text{ cm}^6/\text{s}$	(Hader <i>et al.</i> , 2008)
GaInN QW:	$C \leq 1.0 \times 10^{-31} \text{ cm}^6/\text{s}$	(Hangleiter, 2009)
GaInN:	$C \leq 1.0 \times 10^{-31} \text{ cm}^6/\text{s}$	(Delaney <i>et al.</i> , 2009; $\lambda \leq 450 \text{ nm}$)
GaInN:	$C < 1.0 \times 10^{-32} \text{ cm}^6/\text{s}$	(Bertazzi <i>et al.</i> , 2010)

In summary, Auger recombination can reduce the luminescence efficiency, particularly for (i) high-excitation conditions, (ii) high temperatures, and (iii) in narrow-gap semiconductors. However, Auger recombination can be assumed to be negligibly small for wide-gap III–V semiconductors under all typical LED operating conditions.

Semiconductor laser diodes employ inherently higher injection current densities than LEDs. Whereas LEDs typically use current densities $< 1 \text{ kA/cm}^2$, laser diodes use $\geq 1 \text{ kA/cm}^2$. Consequently, Auger recombination may need to be considered in laser diodes rather than in LEDs.

4.8 Non-radiative recombination at surfaces

Substantial non-radiative recombination can occur at semiconductor surfaces. Surfaces are a strong perturbation of the periodicity of a crystal lattice. Recall that the band diagram model is based on the strict periodicity of a lattice. Since this periodicity ends at a surface, the band diagram will need to be modified at a semiconductor surface. This modification includes the addition of electronic states within the forbidden gap of the semiconductor.

Next, we consider a semiconductor surface from a chemical point of view. Atoms at the surface cannot have the same bonding structure as bulk atoms due to the lack of neighboring atoms. Thus, some of the valence orbitals do not form a chemical bond. These partially filled electron orbitals, or **dangling bonds**, are electronic states that can be located in the forbidden gap of the semiconductor where they act as recombination centers. Depending on the charge

state of these valence orbitals, the states can be acceptor-like or donor-like states.

The dangling bonds may also rearrange themselves and form bonds between neighboring atoms in the same surface plane. This **surface reconstruction** can lead to a *locally* new atomic structure with state energies different from bulk atomic states. The surface bonding structure depends on the specific nature of the semiconductor surface. The energetic location of surface states is very difficult to predict, even with powerful theoretical models. Thus, phenomenological models of surface recombination are commonly used.

It has been shown that electronic states within the forbidden gap appear at semiconductor surfaces. Bardeen and Shockley (Shockley, 1950) pioneered the understanding of surface states and their role as recombination centers.

Next, we calculate the effect of surface recombination on the carrier distribution in a p-type semiconductor subjected to illumination. Assume that the illumination causes a uniform steady-state generation rate G . The one-dimensional continuity equation must be fulfilled at any point in the semiconductor. The continuity equation for electrons is given by

$$\frac{\partial \Delta n(x, t)}{\partial t} = G - R + \frac{1}{e} \frac{\partial}{\partial x} J_n \quad (4.36)$$

where J_n is the current density caused by electrons flowing to the surface. In the bulk of the uniform semiconductor, there is no dependence on space and thus the continuity equation reduces to $G = R$ under steady-state conditions. Using the recombination rate in the bulk as given by Eqn. (4.14), the excess carrier concentration in the bulk is given by $\Delta n_\infty = G\tau_n$ as indicated in **Figure 4.11**. Assuming that the electron current is a diffusion current of the form

$$J_n = e D_n \frac{\partial \Delta n(x, t)}{\partial x} \quad (4.37)$$

and inserting the diffusion current into Eqn. (4.36) yields the continuity equation for diffusive currents, i.e.

$$\frac{\partial \Delta n(x, t)}{\partial t} = G - \frac{\Delta n(x, t)}{\tau_n} + D_n \frac{\partial^2 \Delta n(x, t)}{\partial x^2}. \quad (4.38)$$

At the semiconductor surface, carriers will recombine rapidly due to surface states. The **boundary condition** at the surface is given by

$$eD_n \frac{\partial \Delta n(x, t)}{\partial x} \Big|_{x=0} = e S \Delta n(x, t) \Big|_{x=0} \quad (4.39)$$

where S is the surface recombination velocity. The boundary condition states that minority carriers diffusing to the surface will recombine at the surface. We assume that the generation rate is constant with time, and thus the minority carrier concentration has no time dependence. The steady-state solution to the differential equation with the above boundary condition is given by

$$n(x) = n_0 + \Delta n(x) = n_0 + \Delta n_\infty \left[1 - \frac{\tau_n S \exp(-x/L_n)}{L_n + \tau_n S} \right]. \quad (4.40)$$

The carrier concentration near a semiconductor surface is shown in *Figure 4.11* for different surface recombination velocities. For $S \rightarrow 0$, the minority carrier concentration at the surface is identical to the bulk value, i.e. $n(0) \rightarrow n_0 + \Delta n_\infty$. For $S \rightarrow \infty$, the minority carrier concentration at the surface approaches the equilibrium value, i.e. $n(0) \rightarrow n_0$.

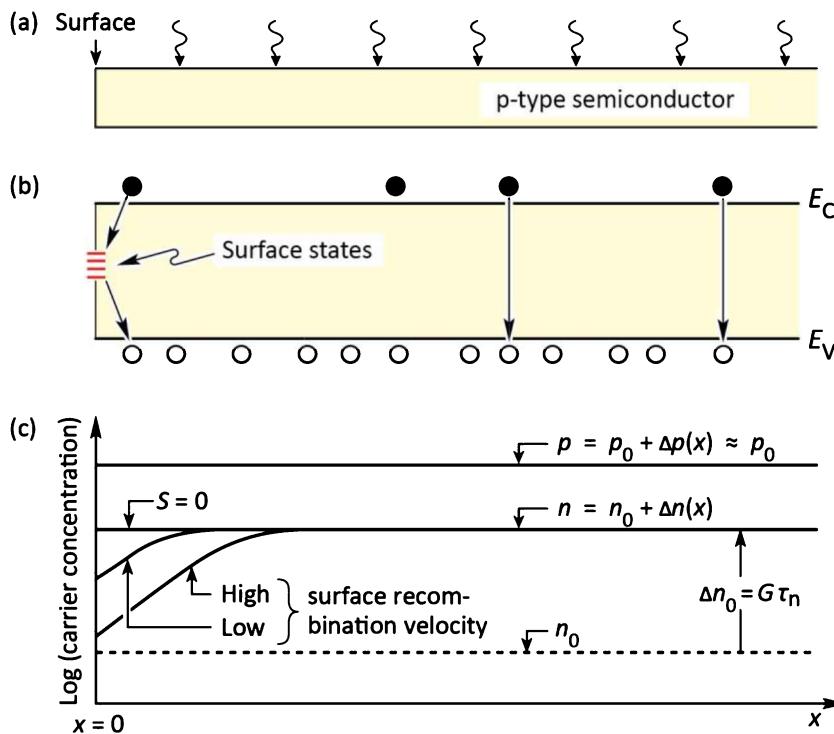


Fig. 4.11: (a) Illuminated p-type semiconductor, (b) band diagram, and (c) minority and majority carrier concentrations near the surface assuming uniform carrier generation due to illumination. The excess carrier concentrations are Δn and Δp .

Next, consider an undoped active region that is injected with carriers so that $\Delta n(0) = \Delta p(0) = \Delta n_{surface} = \Delta p_{surface} = n_{surface} = p_{surface}$. The surface recombination rate (per unit area) is given by

$$R_{surface} = S n_{surface} \quad (4.41)$$

where S is the surface recombination velocity. For a pn-junction device with the perimeter of the pn-junction plane ending at the semiconductor surface (as in a mesa-etched LED), the surface-leakage current is given by

$$I_{surface-leakage} = e R_{surface} A = e S n_{surface} A \quad (4.42)$$

where A is the area of the surface leakage, i.e. the length of the pn-junction-plane's perimeter multiplied by the thickness of the active region.

Surface recombination leads to (i) reduced luminescence efficiency, (ii) heating of the surface due to non-radiative recombination at the surface, and (iii) structural degradation of the surface. These effects are unwanted in electroluminescent devices. The surface recombination velocities for several semiconductors are summarized in Table 4.1. The data shown in the table show that GaAs has a particularly high surface recombination velocity.

Table 4.1: Surface recombination velocities of several semiconductors (GaN data after Tu *et al.*, 2000; Aleksiejunas *et al.*, 2003).

Semiconductor	Surface recombination velocity
GaAs	$S = 10^6$ cm/s
GaN	$S = 5 \times 10^4$ cm/s
InP	$S = 10^3$ cm/s
Si	$S = 10^1$ cm/s

Experimental evidence of surface recombination is illustrated in **Figure 4.12**, which shows the luminescence emanating from a stripe-like current-injection contact on a GaAs laser chip. The luminescence is viewed from the substrate side so that the stripe-like metal contact is “behind” the light-emitting region. **Figure 4.12** clearly reveals that the luminescence decreases in the near-surface region.

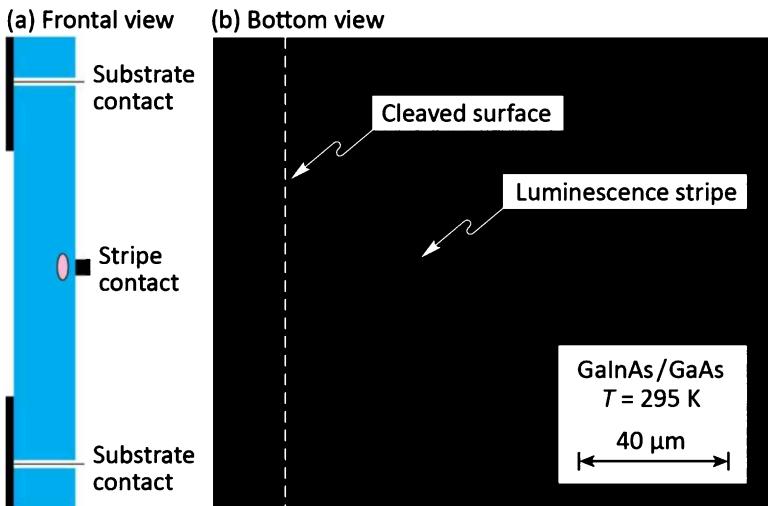


Fig. 4.12: (a) Schematic frontal view and (b) micrograph of a GaInAs/GaAs structure with a stripe-shaped top contact and contact window on substrate side under current injection conditions. The luminescence emanating from active region located below the stripe clearly decreases in the vicinity of the surface due to surface recombination.

Surface recombination can occur only when both types of carriers are present. It is important in the design of LEDs that the carrier-injected active region, in which naturally both types of carriers are present, *be far removed from any surface*. This can be achieved, for example, by carrier injection under a contact that is much smaller than the semiconductor die. Furthermore, the contact must be sufficiently far away from the side surfaces of the die. If the current flow is confined to the region below the contact, carriers will not “see” any semiconductor surfaces. Note that unipolar regions of a semiconductor device, e.g. the confinement regions, are not affected by surface recombination due to the lack of minority carriers.

Several passivation techniques have been developed to reduce the surface recombination in semiconductors, including treatments with sulfur, and other chemicals (Lipsanen *et al.*, 1999). In the field of GaN LEDs, SiO_2 and Si_3N_4 films have been used to passivate the exposed pn-junction surface (Yang *et al.*, 2005).

4.9 Competition between radiative and non-radiative recombination

So far, we have seen that several mechanisms for non-radiative recombination exist, including Shockley–Read–Hall, Auger, and surface recombination. Even though non-radiative recombination can be reduced, it can never be totally eliminated. For example, surface recombination can be drastically reduced by device designs that spatially separate the active region from any surfaces. However, even if the separation is large, a few carriers will still diffuse to the surface and recombine there.

Just as for surface recombination, non-radiative bulk recombination and Auger recombination can never be totally avoided. Any semiconductor crystal will have some native

defects. Even though the concentration of these native defects can be low, it is never zero. Thermodynamic considerations predict that if an energy E_a is needed to create a specific point defect in a crystal lattice, the probability that such a defect does indeed form at a specific lattice site, is given by the Boltzmann factor, i.e. $\exp(-E_a/kT)$. The product of the concentration of lattice sites and the Boltzmann factor gives the concentration of defects. A native point defect or extended defect may form a deep state in the gap and thus be a non-radiative recombination center.

Exercise: Concentration of point defects. Assume that the energy required to move a substitutional lattice atom into an interstitial position is $E_a = 1.1$ eV. What is the equilibrium concentration of interstitial defects of a simple cubic lattice with lattice constant $a_0 = 2.5$ Å?

Solution: The concentration of lattice atoms of a simple cubic lattice is given by $N = a_0^{-3} = 6.4 \times 10^{22}$ cm⁻³. The concentration of interstitial defects under equilibrium conditions at room temperature is then given by

$$N_{\text{defect}} = N \exp(-E_a/kT) = 2.7 \times 10^4 \text{ cm}^{-3}.$$

Note that the calculated concentration of defects is small when compared to the typical concentrations of electrons and holes. If the defect discussed here forms a level in the gap, non-radiative recombination through the defect level can occur.

Another issue is the chemical purity of semiconductors. It is difficult to fabricate materials with impurity levels lower than the *parts per billion* (ppb) range. Thus, even the purest semiconductors contain impurities in the 10¹² cm⁻³ range. Some elements may form deep levels and thus reduce the luminescence efficiency.

In the 1960s, when the first III–V semiconductors had been demonstrated, the radiative luminescence efficiencies at room temperature were very low, typically a fraction of 1%. At the present time, high-quality bulk semiconductors and quantum well structures can have radiative efficiencies exceeding 90%, and in some cases even 99%. This remarkable progress is due to improved crystal quality, and reduced defect and impurity concentrations.

Next, we calculate the radiative efficiency in a semiconductor with non-radiative recombination centers. If the radiative lifetime is denoted as τ_r , and the non-radiative lifetime is denoted as τ_{nr} , then the total probability of recombination is given by the sum of the radiative and non-radiative probabilities:

$$\tau^{-1} = \tau_r^{-1} + \tau_{nr}^{-1}. \quad (4.43)$$

The relative probability of radiative recombination is given by the radiative probability over the total probability of recombination. Thus, the probability of radiative recombination or **radiative efficiency** is given by

$$\eta_{radiative} = \frac{\tau_r^{-1}}{\tau_r^{-1} + \tau_{nr}^{-1}}. \quad (4.44)$$

The radiative efficiency gives the ratio of the number of light quanta emitted from a semiconductor active region to the number of charge quanta undergoing recombination in the active region. Note that not all photons emitted from the active region may escape from the semiconductor due to the light-escape problem, reabsorption in the substrate, or other reabsorption mechanisms.

4.10 Recombination in direct-gap and indirect-gap semiconductors

LEDs are most commonly made of direct-gap semiconductors. In such semiconductors, the band extrema (CB minimum and VB maximum) occur at the same momentum value (k -value), as shown in **Figure 4.13 (a)**. Accordingly, an electron can recombine with a hole without momentum change thereby satisfying the momentum conservation requirement (the momentum of the simultaneously emitted photon, $p = \hbar k$, is negligibly small). The emission of the photon satisfies the energy-conservation requirement. Note that the energy- and momentum-conservation requirement applies to all elementary processes in solid-state materials.

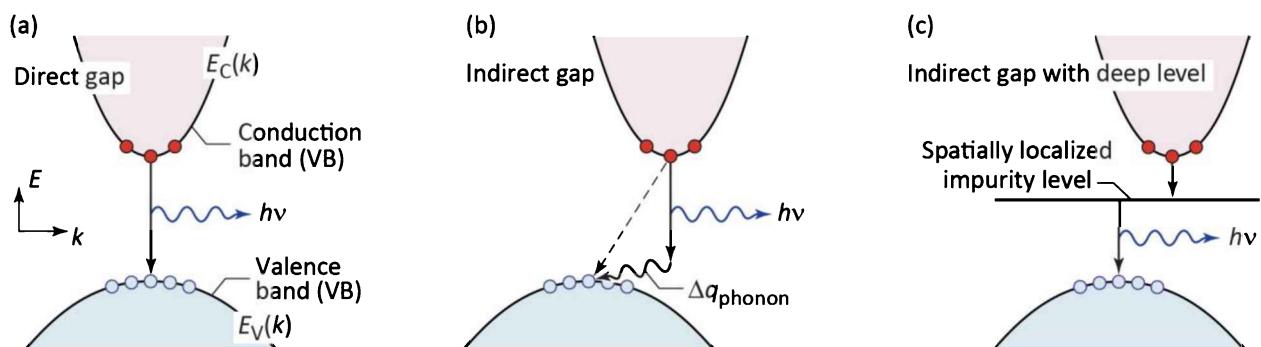


Fig. 4.13: Band structure for a (a) direct-gap, (b) indirect-gap, and (c) indirect-gap semiconductor with a deep isoelectronic impurity level. Indirect radiative transitions have a much smaller probability because a phonon is required in the recombination process.

There are also indirect-gap semiconductors, the most notable being silicon (Si). In such semiconductors, the band extrema occur at different momenta as shown in *Figure 4.13 (b)*. As a result, momentum conservation can no longer be satisfied by only the recombining electron and hole; that is, a phonon is required to satisfy the momentum-conservation requirement. The probability of a recombination process is strongly reduced if a third particle, a phonon, must participate in the process. Hence, the radiative lifetime in indirect-gap semiconductors is very long. Consequently, the always-competing non-radiative Shockley-Read-Hall (SRH) recombination becomes dominant in indirect-gap semiconductors and radiative recombination is negligibly small. For this reason, indirect-gap semiconductors, such as Si, are unsuitable for LEDs.

In order to overcome the difficulty encountered with indirect-gap semiconductors, doping with impurities, frequently isoelectronic impurities, is employed to form a deep level inside the forbidden gap, as illustrated in *Figure 4.13 (c)*. Spatially localized impurity states are de-localized in k -space⁷, as illustrated in the figure. As a consequence, a sequential transition from the CB, via the impurity level, to the VB, is enabled. The necessary momentum change is provided by the impurity. For some impurities, one of the transitions is radiative, as illustrated in the figure. The emission of a photon enables energy conservation for the radiative transition. N-doped GaP and N-doped GaAsP are primary examples of an indirect-gap semiconductor whose radiative efficiency can be strongly enhanced by doping (Thomas *et al.*, 1965a; Thomas *et al.*, 1965b; Thomas and Hopfield, 1966; Hopfield *et al.*, 1966; Dean and Faulkner, 1969; Dean, 1970).

4.11 References

- Agrawal G. P. and Dutta N. K. *Long Wavelength Semiconductor Lasers* (John Wiley and Sons, New York, 1986)
- Ahrenkiel R. K. "Minority-carrier lifetime in III–V semiconductors" in *Minority Carriers in III–V Semiconductors: Physics and Applications* edited by R. K. Ahrenkiel and M. S. Lundstrom, Semiconductors and Semimetals **39**, 40 (Academic Press, San Diego, 1993)
- Aleksiejunas R., Sudzius M., Malinauskas T., Vaitkus J., Jarasiunas K., and Sakai S. "Determination of free carrier bipolar diffusion coefficient and surface recombination velocity of undoped GaN epilayers" *Applied Physics Letters* **83**, 1157 (2003)
- Baraff G. A. and Schluter M. "Electronic structure, total energies, and abundances of the elementary point

⁷ The delocalization in k -space is taught by the Heisenberg uncertainty principle: $\Delta x \Delta p \geq \hbar$. If an impurity atom has a spatially localized state, i.e. the electron is tightly bound to the impurity atom, then the state has a small Δx and, based on the uncertainty principle, a large Δp (or, since $p = \hbar k$, a large Δk). A state localized in position space therefore is delocalized in k -space. Spatial localization of impurity states is particularly common for deep states and less common for shallow states.

- defects in GaAs" *Physical Review Letters* **55**, 1327 (1985)
- Beattie A. R. and Landsberg P. T. "Auger effect in semiconductors" *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* **249**, 16 (1959)
- Bertazzi F., Goano M., and Bellotti E. "A numerical study of Auger recombination in bulk InGaN" *Applied Physics Letters* **97**, 231118 (2010)
- Bleichner H., Jonsson P., Keskitalo N., and Nordlander E. "Temperature and injection dependence of the Shockley–Read–Hall lifetime in electron irradiated n-type silicon" *Journal of Applied Physics* **79**, 9142 (1996)
- Dean P. J. and Faulkner R. A. "The nitrogen isoelectronic trap in phosphorus-rich gallium arsenide phosphide" *Applied Physics Letters* **14**, 210 (1969)
- Dean P. J. "Recombination processes associated with 'deep states' in gallium phosphide" *Journal of Luminescence* **1**, 398 (1970)
- Delaney K. T., Rinke P., and Van de Walle C. G. "Auger recombination rates in nitrides from first principles" *Applied Physics Letters* **94**, 191109 (2009)
- Ehrhardt A., Wetling W., and Bett A. "Transient photoluminescence decay study of minority carrier lifetime in GaAs heteroface solar cell structures" *Applied Physics A (Solids and Surfaces)* **A53**, 123 (1991)
- Grieshaber W., Schubert E. F., Goepfert I. D., Karlicek R. F. Jr., Schurman M. J., and Tran C. "Competition between band gap and yellow luminescence in GaN and its relevance for optoelectronic devices" *Journal of Applied Physics* **80**, 4615 (1996)
- Hader J., Moloney J. V., Pasenow B., Koch S. W., Sabathil M., Linder N., and Lutgen S. "On the importance of radiative and Auger losses in GaN-based quantum wells" *Applied Physics Letters* **92**, 261103 (2008)
- Hall R. N. "Electron–hole recombination in germanium" *Physical Review* **87**, 387 (1952)
- Hangleiter A. "What can we learn from lasers about Auger recombination" The author concluded that an upper limit of Auger recombination in GaInN QWs is $C \leq 1.0 \times 10^{-31} \text{ cm}^6 \text{ s}^{-1}$; *International Conference on Nitride Semiconductors (ICNS)*, Jeju, Korea, October 18~23 (2009)
- Haug A. "Auger recombination in quantum well gallium antimonide" *Journal of Physics C: Solid State Physics* **20**, 1293 (1987)
- Hausser S., Fuchs G., Hangleiter A., Streubel K., and Tsang W. T. "Auger recombination in bulk and quantum well InGaAs" *Applied Physics Letters* **56**, 913 (1990)
- Hopfield J. J., Thomas D. G., and Lynch R. T. "Isoelectronic donors and acceptors" *Physical Review Letters* **17**, 312 (1966)
- Hwang C. J. "Doping dependence of hole lifetime in n-type GaAs" *Journal of Applied Physics* **42**, 4408 (1971)
- Ioffe Internet web site "InAs – Indium Arsenide" <<http://www.ioffe.ru/SVA/NSM/Semicond/InAs/>> (2010)
- Krier A., Gao H. H., Sherstnev V. V. and Yakovlev Y. "High power 4.6 μm light emitting diodes for CO detection" *Journal of Physics D: Applied Physics* **32**, 3117 (1999)
- Landsberg P. T. "Recombination in semiconductors" (Cambridge University Press, Cambridge UK, 2003)
- Levinshtein M., Rumyantsev S., and Shur M. (editors) "Handbook series on semiconductor parameters, Volume 2" (World Scientific Publishing, Singapore, 1996)
- Lipsanen H., Sopanen M., Ahopelto J., Sandman J., and Feldmann J. "Effect of InP passivation on carrier recombination in $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ surface quantum wells" *Japanese Journal of Applied Physics* **38**, 1133 (1999)
- Longini R. L. and Greene R. F. "Ionization interaction between impurities in semiconductors and insulators" *Physical Review* **102**, 992 (1956)
- Nelson R. J. and Sobers R. G. "Interfacial recombination velocity in AlGaAs/GaAs heterostructures" *Applied Physics Letters* **32**, 761 (1978a)
- Nelson R. J. and Sobers R. G. "Minority-carrier lifetime and internal quantum efficiency of surface-free GaAs" *Applied Physics Letters* **49**, 6103 (1978b)
- Neugebauer J. and Van de Walle C. "Gallium vacancies and the yellow luminescence in GaN" *Applied*

- Physics Letters* **69**, 503 (1996)
- Olshansky R., Su C. B., Manning J., and Powazinik W. "Measurement of radiative and non-radiative recombination rates in InGaAsP and AlGaAs light sources" *IEEE Journal of Quantum Electronics* **QE-20**, 838 (1984)
- Pantelides S. T. (Editor) *Deep Centers in Semiconductors* (Gordon and Breach, Yverdon, Switzerland, 1992)
- Phillips J. C. "Stretched exponential relaxation in molecular and electronic glasses" *Reports on Progress in Physics* **59**, 1133 (1996)
- Saarinen K. et al. "Observation of native Ga vacancies in GaN by positron annihilation" *Physical Review Letters* **79**, 3030 (1997)
- Schenk A. "A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon" *Solid-State Electronics* **35**, 158 (1992)
- Schubert E. F. *Doping in III-V Semiconductors* (Cambridge University Press, Cambridge UK, 1993)
- Schubert E. F., unpublished (1995)
- Schubert E. F., Goepfert I., and Redwing J. M. "Evidence of compensating centers as origin of yellow luminescence in GaN" *Applied Physics Letters* **71**, 3224 (1997)
- Shockley W. *Electrons and Holes in Semiconductors* (D. Van Nostrand Company, New York, 1950)
- Shockley W. and Read W. T. "Statistics of the recombinations of holes and electrons" *Physical Review* **87**, 835 (1952)
- Thomas D. G., Hopfield J. J., and Augustyniak W. M. "Kinetics of radiative recombination at randomly distributed donors and acceptors" *Physical Review* **140**, A202 (1965a)
- Thomas D. G., Hopfield J. J., and Frosch C. J. "Isoelectronic traps due to nitrogen in gallium phosphide" *Physical Review Letters* **15**, 857 (1965b)
- Thomas D. G. and Hopfield J. J. "Isoelectronic traps due to nitrogen in gallium phosphide" *Physical Review* **150**, 680 (1966)
- Tu L. W., Kuo W. C., Lee K. H., Tsao P. H., and Lai C. M., Chu A. K., and Sheu J. K. "High-dielectric-constant Ta_2O_5 / n-GaN metal-oxide-semiconductor structure" *Applied Physics Letters* **77**, 3788 (2000)
- Yang H. S., Han S. Y., Hlad M., Gila B. P., Baik K. H., Pearton S. J., Jang S., Kang B. S., and Ren F. "Comparison of Surface Passivation Layers on InGaN/GaN MQW LEDs" *Journal of Semiconductor Technology and Science* **5**, 131 (2005)

5

5 – Theory of radiative recombination

In this chapter, the theory of radiative recombination is first discussed in terms of a rigorous quantum mechanical model. Subsequently recombination is discussed in terms of a semi-classical model based on equilibrium generation and recombination. This model was developed by van Roosbroeck and Shockley (1954). Finally, the Einstein model of spontaneous and stimulated transitions in a two-level atom is discussed.

5.1 Quantum mechanical model of recombination

Spontaneous recombination based on quantum mechanics was discussed by Bebb and Williams (1972), Agrawal and Dutta (1986), Dutta (1993), Thompson (1980), and others. The quantum mechanical calculation for the spontaneous emission rate is based on the induced emission rate given by **Fermi's Golden Rule** which gives the transition probability per unit time (called transition rate) from a quantum mechanical state j to a state m

$$W_{j \rightarrow m} = \frac{d}{dt} |a'_m(t)|^2 = \frac{2\pi}{\hbar} |H'_{mj}|^2 \rho(E = E_j + \hbar\omega_0) \quad (5.1)$$

where H'_{mj} is the **transition matrix element**. For the one-dimensional case with a dependence on only the spatial variable x , the matrix element that connects the (initial) j th state with the (final) m th state via the perturbation hamiltonian H' , is given by

$$H'_{mj} = \langle \psi_m^0 | H' | \psi_j^0 \rangle = \int_{-\infty}^{\infty} \psi_m^{0*}(x) A(x) \psi_j^0(x) dx . \quad (5.2)$$

For the derivation of Fermi's Golden Rule, the perturbation hamiltonian H' is assumed to have a harmonic time dependence, i.e. $H' = A(x) [\exp(i\omega_0 t) + \exp(-i\omega_0 t)]$, as expected for the excitation by a harmonic wave of a photon. Equation (5.2) indicates that a necessary condition for recombination is **spatial overlap** between electron and hole wave functions. This is intuitively clear, because spatially separated electrons and holes will not be able to recombine.

For optical transitions between the conduction and valence band, the electron momentum

must be conserved because the photon momentum ($p = \hbar k$) is negligibly small. The conservation of momentum condition is known as the ***k-selection rule***. For the remainder of this section, we closely follow the quantum mechanical analysis given by Agrawal and Dutta (1986) and the reader is referred to that reference for a comprehensive discussion.

The average matrix element for the Bloch states, $|M_b|$ can be derived using the four-band Kane model (Kane, 1957), which takes into account the conduction, heavy-hole, light-hole, and split-off band. In bulk semiconductors, $|M_b|^2$ is given by (Kane, 1957; Casey and Panish, 1978)

$$|M_b|^2 = \frac{m_e^2 E_g (E_g + \Delta)}{12 m_e^* (E_g + 2\Delta/3)} \quad (5.3)$$

where m_e is the free-electron mass, E_g is the band gap energy, and Δ is the spin-orbit splitting. For GaAs, using $E_g = 1.424$ eV, $\Delta = 0.33$ eV, $m_e^* = 0.067 m_e$, we get $|M_b|^2 = 1.3 m_e E_g$.

Taking into account the *k*-selection rule, the total spontaneous emission rate per unit volume is given by

$$\begin{aligned} r_{sp}(E) &= \frac{4\pi \bar{n} e^2 E}{m_e^2 \epsilon_0 h^2 c^3} |M_b|^2 \frac{(2\pi)^3}{V} 2 \left(\frac{V}{(2\pi)^3} \right)^2 \frac{1}{V} \times \\ &\times \Sigma \int \dots \int f_c(E_C) f_v(E_V) d^3\vec{k}_C d^3\vec{k}_V \delta(\vec{k}_C - \vec{k}_V) \delta(E_i - E_f - E) \end{aligned} \quad (5.4)$$

where f_C and f_V are the Fermi factors for electrons and holes, and the term $\delta(\vec{k}_C - \vec{k}_V)$ ensures satisfaction of the *k*-selection rule. The factor 2 arises from the two spin states. In Eqn. (5.4), Σ stands for the sum over the three valence bands (heavy-hole, light-hole, and split-off bands). For definiteness, we first consider transitions involving electrons and heavy holes. The integrals in Eqn. (5.4) can be evaluated with the following result:

$$r_{sp}(E) = \frac{2 \bar{n} e^2 E |M_b|^2}{\pi m_e^2 \epsilon_0 h^2 c^3} \left(\frac{2 m_r}{\hbar^2} \right)^{3/2} \sqrt{E - E_g} f_c(E_C) f_v(E_V) \quad (5.5)$$

where

$$E_C = (m_r / m_e^*) (E - E_g) \quad (5.6)$$

$$E_V = (m_r / m_{hh}^*) (E - E_g) \quad (5.7)$$

$$m_r = \frac{m_e^* m_{hh}^*}{m_e^* + m_{hh}^*} \quad (5.8)$$

and m_{hh}^* is the effective mass of the heavy hole. Equation (5.5) gives the spontaneous emission rate at the photon energy E . To obtain the total spontaneous emission rate, a final integral should be carried out over all possible energies. Thus, the total spontaneous emission rate per unit volume due to electron–heavy-hole transitions is given by

$$R = \int_{E_g}^{\infty} r_{sp}(E) dE = A |M_b|^2 I \quad (5.9)$$

where

$$I = \int_{E_g}^{\infty} \sqrt{E - E_g} f_c(E_C) f_v(E_V) dE \quad (5.10)$$

and A represents the remaining constants in Eqn. (5.5). A similar equation holds for the electron–light-hole transitions if we replace m_{hh}^* by the effective light-hole mass m_{lh}^* .

The quantum mechanical absorption coefficient $\alpha(E)$ derived by Agrawal and Dutta (1986) using a similar analysis is given by

$$\alpha(E) = \frac{e^2 h |M_b|^2}{4\pi^2 \epsilon_0 m_e^2 c \bar{n} E} \left(\frac{2 m_r}{\hbar^2} \right)^{3/2} \sqrt{E - E_g} [1 - f_c(E_C) - f_v(E_V)]. \quad (5.11)$$

Although the quantum mechanical model of recombination is most appropriate and accurate, it can be time-consuming and awkward to deal with. The following sections analyze recombination in semi-classical terms that usually are more convenient to work with.

5.2 The van Roosbroeck–Shockley model

The van Roosbroeck–Shockley model allows one to calculate the spontaneous radiative recombination rate under equilibrium and non-equilibrium conditions. To calculate the recombination rate, the model requires the knowledge of only a few basic parameters, namely the bandgap energy, the absorption coefficient, and the refractive index. All of these parameters can be determined by simple, well-known experimental methods.

Consider a semiconductor with an absorption coefficient $\alpha(v)$ given in units of cm^{-1} . A photon is generated in the semiconductor by electron–hole recombination and is subsequently

absorbed, as illustrated in **Figure** 5.1. The mean distance a photon of frequency ν travels before being absorbed is simply $\alpha(\nu)^{-1}$. The time it takes for a photon to be absorbed is given by

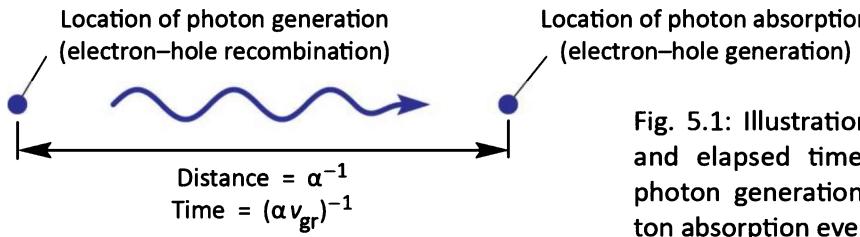


Fig. 5.1: Illustration of distance and elapsed time between a photon generation and a photon absorption event.

$$\tau(\nu) = \frac{1}{\alpha(\nu) v_{gr}} \quad (5.12)$$

where v_{gr} is the group velocity of photons propagating in the semiconductor. The group velocity of photons is given by

$$v_{gr} = \frac{d\omega}{dk} = \frac{dv}{d(1/\lambda)} = c \frac{dv}{d(\bar{n} \nu)} \quad (5.13)$$

where \bar{n} is the refractive index. Inserting the group velocity into Eqn. (5.12) yields

$$\frac{1}{\tau(\nu)} = \alpha(\nu) v_{gr} = \alpha(\nu) c \frac{dv}{d(\bar{n} \nu)} . \quad (5.14)$$

This equation gives the *inverse photon lifetime* or *photon absorption probability* per unit time. The product of the absorption probability and the photon density yields the photon absorption rate per unit time per unit volume.

Under equilibrium conditions, the density of photons per unit volume in a medium with refractive index \bar{n} is given by Planck's **black-body radiation formula**

$$N(\lambda) d\lambda = \frac{8\pi}{\lambda^4} \frac{1}{e^{h\nu/kT} - 1} d\lambda \quad (5.15)$$

from which we can readily obtain $N(\nu) dv$, the number of photons having frequencies in the interval ν and $\nu + dv$. We have $\lambda = c/(\bar{n} \nu)$, so that

$$d\lambda = -\frac{c}{(\bar{n} \nu)^2} \frac{d(\bar{n} \nu)}{d\nu} dv . \quad (5.16)$$

Inserting this value in Eqn. (5.15) yields Planck's black-body photon distribution as a function of

frequency

$$N(\nu) \, d\nu = \frac{8\pi \nu^2 \bar{n}^2}{c^3} \frac{d(\bar{n} \nu)}{d\nu} \frac{1}{e^{h\nu/kT} - 1} \, d\nu. \quad (5.17)$$

The absorption rate per unit volume in the frequency interval ν and $\nu + d\nu$ is given by the photon density divided by the mean lifetime of photons

$$R_0(\nu) = \frac{N(\nu)}{\tau(\nu)} = \frac{8\pi \nu^2 \bar{n}^2}{c^3} \frac{d(\bar{n} \nu)}{d\nu} \frac{1}{e^{h\nu/kT} - 1} \alpha(\nu) c \frac{d\nu}{d(\bar{n} \nu)}. \quad (5.18)$$

Integration over all frequencies yields the absorption rate per unit volume

$$R_0 = \int_0^\infty R_0(\nu) \, d\nu = \int_0^\infty \frac{8\pi \nu^2 \bar{n}^2}{c^2} \frac{\alpha(\nu)}{e^{h\nu/kT} - 1} \, d\nu \quad (5.19)$$

which is the celebrated **van Roosbroeck–Shockley equation**. The van Roosbroeck–Shockley equation can be simplified by writing the absorption coefficient as

$$\alpha = \alpha_0 \sqrt{(E - E_g) / E_g}. \quad (5.20)$$

The square-root dependence of the absorption coefficient is motivated by the proportionality of the absorption coefficient and the density of states, which in turn follows a square-root dependence on energy. Note that α_0 is the absorption coefficient at $h\nu = 2E_g$. Approximate values of α_0 for several semiconductors are given in **Table 5.1**.

Table 5.1: Bimolecular recombination coefficient at 300 K for different semiconductors as calculated from the energy gap, absorption coefficient, and refractive index at the bandgap energy. The spontaneous lifetime is given by $B^{-1} N_{D,A}^{-1}$ and it is calculated for a majority carrier concentration of 10^{18} cm^{-3} .

Material	E_g (eV)	α_0 (cm^{-1})	\bar{n} (–)	R_0 ($\text{cm}^{-3} \text{s}^{-1}$)	n_i (cm^{-3})	B ($\text{cm}^3 \text{s}^{-1}$)	τ_{spont} (s)
GaAs	1.42	2×10^4	3.3	7.9×10^2	2×10^6	2.0×10^{-10}	5.1×10^{-9}
InP	1.35	2×10^4	3.4	1.2×10^4	1×10^7	1.2×10^{-10}	8.5×10^{-9}
GaN	3.4	2×10^5	2.5	8.9×10^{-30}	2×10^{-10}	2.2×10^{-10}	4.5×10^{-9}
GaP	2.26	2×10^3	3.0	1.0×10^{-12}	1.6×10^0	3.9×10^{-13}	2.6×10^{-6}
Si	1.12	1×10^3	3.4	3.3×10^6	1×10^{10}	3.2×10^{-14}	3.0×10^{-5}
Ge	0.66	1×10^3	4.0	1.1×10^{14}	2×10^{13}	2.8×10^{-13}	3.5×10^{-6}

The van Roosbroeck–Shockley equation can be simplified further by neglecting the frequency dependence of the refractive index and using the refractive index value at the band edge. One obtains

$$R_0 = 8\pi c \bar{n}^2 \alpha_0 \sqrt{\frac{kT}{E_g}} \left(\frac{kT}{c h}\right)^3 \int_{x_g}^{\infty} \frac{x^2 \sqrt{x - x_g}}{e^x - 1} dx \quad (5.21)$$

where $x = hv/(kT) = E/(kT)$ and $x_g = E_g/(kT)$. Owing to the strong increase of the exponential function with x , only a small range of energies close to the bandgap contributes to the integral. The integral has no simple analytical solution and it needs to be evaluated by a numerical method.

Under equilibrium conditions, the carrier generation rate (photon absorption rate) is equal to the carrier recombination rate (photon emission rate). Thus, the van Roosbroeck–Shockley model provides the *equilibrium* recombination rate. As discussed earlier, the bimolecular rate equation, which applies to both equilibrium and non-equilibrium conditions, gives the number of recombination events occurring per unit volume per unit time:

$$R = B n p . \quad (5.22)$$

Next, we use the van Roosbroeck–Shockley model to calculate the **bimolecular recombination coefficient** B . Under equilibrium conditions, it is $R = R_0 = B n_i^2$. Thus, the bimolecular recombination coefficient is related to the equilibrium recombination rate by

$$B = R_0 / n_i^2 . \quad (5.23)$$

Table 5.1 shows the bimolecular recombination coefficient for different semiconductors as calculated from Eqns. (5.21) and (5.23). All material parameters used in the calculation are given in the table. The calculated results reveal that $B = 10^{-9} \sim 10^{-11} \text{ cm}^3/\text{s}$ for direct-gap III–V semiconductors. This calculated result agrees well with experimental results. GaP, Si, and Ge, all indirect-gap semiconductors, have a much smaller bimolecular recombination coefficient compared with direct-gap III–V semiconductors.

There are several other ways to calculate the bimolecular recombination coefficient. An early calculation by Hall (1960) using a two-band model produced a bimolecular recombination coefficient of

$$B = 5.8 \times 10^{-13} \frac{\text{cm}^3}{\text{s}} \left(\frac{m_h^*}{m_e} + \frac{m_e^*}{m_e} \right)^{-\frac{3}{2}} \left(1 + \frac{m_e}{m_h^*} + \frac{m_h^*}{m_e} \right) \left(\frac{300 \text{ K}}{T} \right)^{\frac{3}{2}} \left(\frac{E_g}{1 \text{ eV}} \right)^2 \bar{n} \quad (5.24)$$

where m_e^* , m_h^* , and m_e are the electron effective mass, hole effective mass, and free electron mass, respectively. Garbuzov (1982) described a simple quantum mechanical calculation for direct-gap semiconductors and obtained the following expression for the bimolecular recombination coefficient:

$$B = 3.0 \times 10^{-10} \frac{\text{cm}^3}{\text{s}} \left(\frac{300 \text{ K}}{T} \right)^{3/2} \left(\frac{E_g}{1.5 \text{ eV}} \right)^2. \quad (5.25)$$

All of the methods described here to calculate B give reasonably similar results. The B coefficient is in the $10^{-10} \text{ cm}^3/\text{s}$ range when calculated for GaAs at 300 K using Eqns. (5.21) and (5.23) ~ (5.25).

5.3 Temperature and doping dependence of recombination

The temperature dependence of the recombination probability is elucidated in **Figure 5.2**, which shows a parabolic $E(k)$ relationship at low and high temperatures. Inspection of the figure reveals that the number of carriers per dk interval decreases with increasing temperature. As radiative recombination requires momentum conservation and the recombination probability of an electron is proportional to the number of holes available at equal momentum, the recombination probability decreases with increasing temperature. This trend is confirmed by Eqns. (5.24) and (5.25), both of which display a $T^{-3/2}$ dependence of the bimolecular recombination coefficient.

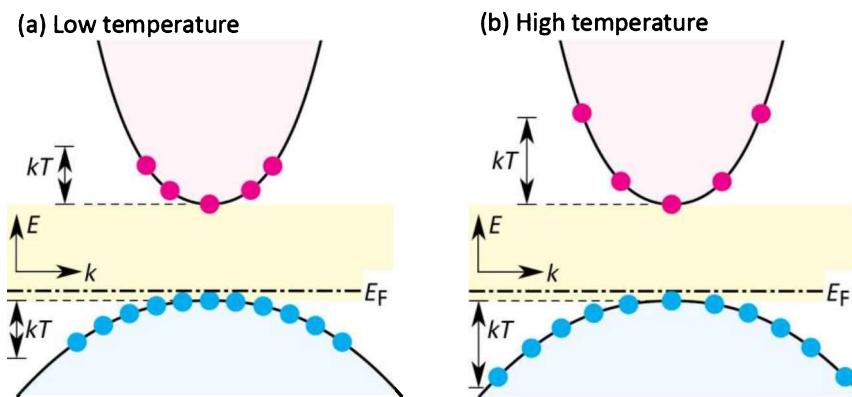


Fig. 5.2: Carrier distribution at (a) low and (b) high temperatures. Recombination probability decreases at high temperatures due to reduced number of carriers per dk interval.

The doping-concentration dependence of the recombination probability is elucidated in **Figure 5.3**, which shows $E(k)$ for non-degenerate and degenerate doping concentrations. Inspection of the figure reveals that the number of holes per dk interval remains constant in the degenerate doping regime. Thus, the recombination probability does not increase in the degenerate doping regime.

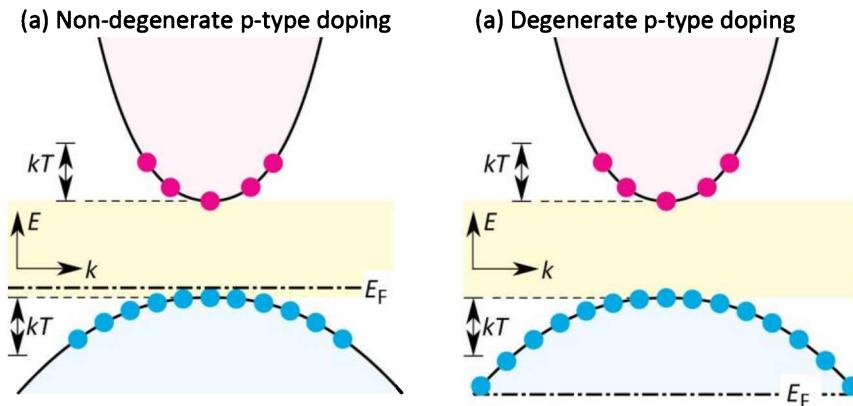


Fig. 5.3: Carrier distribution in (a) a non-degenerately and (b) a degenerately doped p-type semiconductor. Degenerate doping does not increase the overlap between electrons and holes with equal momentum.

This contention is confirmed by the quantum mechanical calculation of the bimolecular recombination coefficient that is shown versus doping concentration in **Figure 5.4** (Waldron, 2002). Inspection of the figure reveals that the bimolecular recombination coefficient saturates in the degenerate doping regime. The van Roosbroeck–Shockley model does not exhibit this characteristic as its validity is limited to the non-degenerate case.

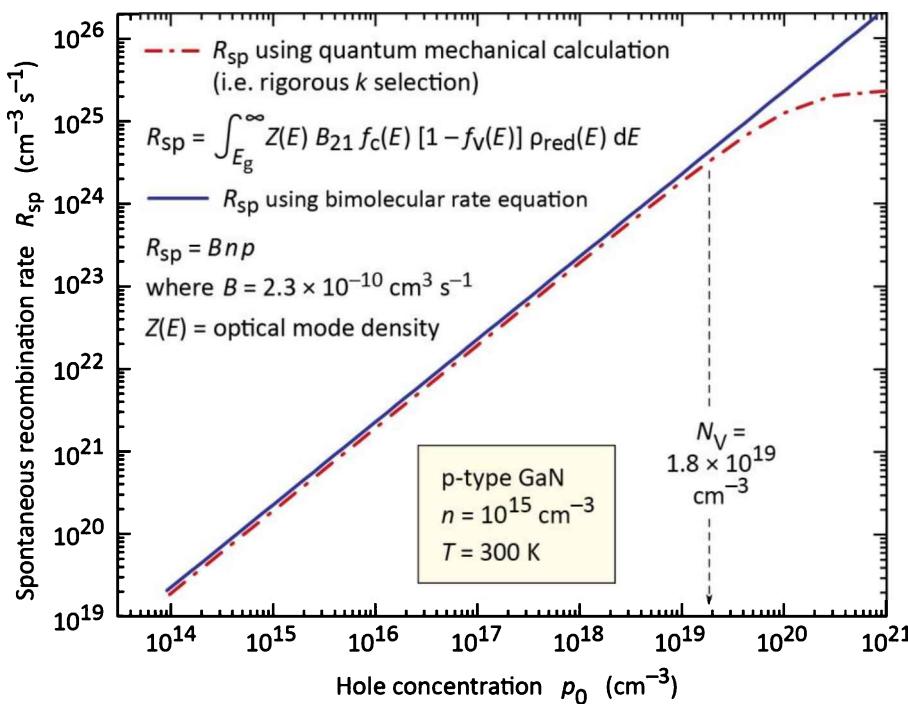


Fig. 5.4: Calculated spontaneous recombination rate in GaN at 300 K as a function of p-type doping concentration using a classical and quantum mechanical approach. The quantum mechanical approach (employing rigorous k selection) exhibits saturation in the degenerate doping regime (after Waldron, 2002).

The bimolecular rate equation, $R = Bnp$, applies to *dilute* carrier concentrations, i.e. to non-degenerately doped semiconductors. Thus, the bimolecular recombination coefficient, B , applies to semiconductors with non-degenerate carrier concentrations. In this case, the bimolecular recombination coefficient is independent of the carrier concentration. For very high carrier concentrations, however, the bimolecular recombination coefficient decreases due to an increasing momentum mismatch between electrons and holes. The bimolecular recombination coefficient in the degenerate regime can be expressed as

$$B|_{high\ concentrations} = \frac{B}{1 + (n/N_c)} \quad (5.26)$$

where N_c is a constant that is approximately equal to the effective density-of-states in the conduction band. The equation reveals that the bimolecular recombination coefficient is reduced at high carrier concentrations. Further discussions of the bimolecular recombination coefficient at high concentrations can be found in the literature (Agrawal and Dutta, 1986; Olshansky *et al.*, 1984; Hader *et al.*, 2008; David and Grundmann, 2010).

Exercise: *Radiative efficiency.* Analyze the temperature dependence of the radiative lifetime based on the van Roosbroeck–Shockley model and the non-radiative lifetime based on the Shockley–Read–Hall model and predict the temperature dependence of the radiative efficiency in semiconductors.

Solution: The *radiative* recombination rate (van Roosbroeck–Shockley rate) has a weak temperature dependence, and it depends on temperature according to $R = Bnp$, where $B \propto T^{-3/2}$, as concluded from Eqns. (5.24) and (5.25). The concentrations n and p mostly depend on the excitation strength (injection current) and can be assumed to be temperature independent. Thus, it is $R \propto \tau_{radiative}^{-1} \propto T^{-3/2}$.

The *non-radiative* recombination rate (Shockley–Read–Hall) has a strong temperature dependence and it increases very rapidly with increasing temperature. The Shockley–Read–Hall recombination rate includes the term $\cosh[(E_T - E_{Fi})/(kT)]$. Because $\cosh x = (1/2)(e^x + e^{-x})$ and $(E_T - E_{Fi})$ is either < 0 or > 0 (depending on the location of the trap with respect to the intrinsic Fermi level, E_{Fi}), one of the exponential functions (e^x or e^{-x}) dominates, so that an approximately exponential temperature dependence is obtained. Thus, it is $R_{SRH} \propto \tau_{non-radiative}^{-1} \propto \exp[-1/(kT)]$.

The exponential dependence (Shockley–Read–Hall: $R_{SRH} \propto \exp[-1/(kT)]$) is much stronger than the power-law dependence (van Roosbroeck–Shockley: $R \propto T^{-3/2}$), so that the radiative efficiency of semiconductors strongly decreases with increasing temperature.

5.4 The Einstein model

The first theory of optical transitions was developed by Albert Einstein. The Einstein model includes *spontaneous* and *stimulated* or *induced transitions*. *Spontaneous transitions* occur

spontaneously, that is, without an external stimulus. In contrast, *stimulated transitions* are induced by an external stimulus, namely a photon. Thus, the induced transition rates are proportional to the photon density or radiation density.

The coefficients A and B describe spontaneous and stimulated transitions in an atom with two quantized levels. These transitions are schematically illustrated in **Figure 5.5**. Denoting the two levels as “1” and “2”, Einstein postulated the probability per unit time for the downward transition ($2 \rightarrow 1$) and upward transition ($1 \rightarrow 2$) as

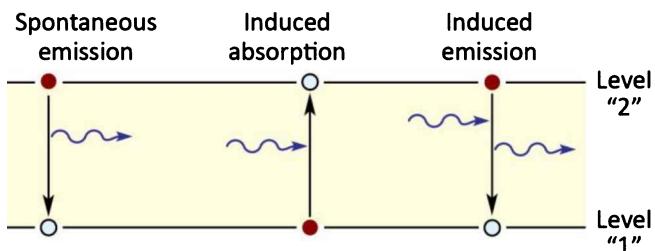


Fig. 5.5: Spontaneous emission, induced absorption, and induced emission events in the two-level atom model. Initially occupied and unoccupied levels are indicated by solid and open circles, respectively.

$$W_{2 \rightarrow 1} = B_{2 \rightarrow 1} \rho(v) + A \quad (5.27)$$

and

$$W_{1 \rightarrow 2} = B_{1 \rightarrow 2} \rho(v) . \quad (5.28)$$

The downward transition probability (per atom) has two terms, namely the induced term and the spontaneous term. The induced term, $B_{2 \rightarrow 1} \rho(v)$, is proportional to the radiation density $\rho(v)$. The spontaneous downward transition probability is a constant A . Note that the spontaneous lifetime is given by $\tau_{spont} = A^{-1}$. The upward probability is just $B_{1 \rightarrow 2} \rho(v)$.

The Einstein A coefficient in an atom corresponds to the bimolecular recombination coefficient in a semiconductor. In an atom, the concentration terms (i.e. n and p in the bimolecular rate equation $R = B n p$) do not come into play, since for a downward transition to occur, the upper level must be occupied (one electron, “ $n = 1$ ”) and the lower level must be unoccupied (one hole, “ $p = 1$ ”).

Einstein showed that $B = B_{2 \rightarrow 1} = B_{1 \rightarrow 2}$. Thus, stimulated absorption and stimulated emission are complementary processes. He also showed that the ratio of the coefficients at a frequency v in an isotropic and homogeneous medium with refractive index \bar{n} is a constant given by $A/B = 8\pi \bar{n}^3 h v^3/c^3$.

The equivalence of $B_{2 \rightarrow 1}$ and $B_{1 \rightarrow 2}$ can also be shown by quantum mechanical considerations, namely by using Fermi's Golden Rule. However, an in-depth analysis of the Einstein model goes beyond the scope of the present discussion.

5.5 References

- Agrawal G. P. and Dutta N. K. *Long-Wavelength Semiconductor Lasers*, Chapter 3 (Van Nostrand Reinhold, New York, 1986)
- Casey Jr. H. C. and Panish M. B. *Heterostructure Lasers*, Part A, Cap. 3 (Academic Press, New York, 1978)
- Bebb H. B. and Williams E. W. "Photoluminescence I: Theory" in *Transport and Optical Phenomena* edited by R. K. Willardson and A. C. Beer, Semiconductors and Semimetals **8**, p. 181 (Academic Press, New York, 1972)
- David A. and Grundmann M. J. "Droop in InGaN light-emitting diodes: A differential carrier lifetime analysis" *Applied Physics Letters* **96**, 103504 (2010)
- Dutta N. K. "Radiative transitions in GaAs and other III–V compounds" in *Minority Carriers in III–V Semiconductors: Physics and Applications* edited by R. K. Ahrenkiel and M. S. Lundstrom, Semiconductors and Semimetals **39**, p. 1 (Academic Press, San Diego, 1993)
- Garbuзов D. Z. "Radiation effects, lifetimes and probabilities of band-to-band transitions in direct A_3B_5 compounds of GaAs type" *Journal of Luminescence* **27**, 109 (1982)
- Hader J., Moloney J. V., Pasenow B., Koch S. W., Sabathil M., Linder N., and Lutgen S. "On the importance of radiative and Auger losses in GaN-based quantum wells" *Applied Physics Letters* **92**, 261103 (2008)
- Hall R. N. "Recombination processes in semiconductors" *Proc. Inst. Electr. Eng.* **106 B**, Suppl. 17, 983 (1960)
- Kane E. O. "Band structure of indium antimonide" *Journal of Physics and Chemistry of Solids* **1**, 249 (1957)
- Olshansky R., Su C. B., Manning J., and Powazinik W. "Measurement of radiative and non-radiative recombination rates in InGaAsP and AlGaAs light sources" *IEEE Journal of Quantum Electronics* **QE-20**, 838 (1984)
- Thompson G. H. B. *Physics of Semiconductor Laser Devices* (John Wiley and Sons, New York, 1980)
- Van Roosbroeck W. and Shockley W. "Photon-radiative recombination of electrons and holes in germanium" *Physical Review* **94**, 1558 (1954)
- Waldron E. L. "Optoelectronic properties of AlGaN/GaN superlattices" Ph.D. Dissertation, Boston University (2002)

6

6 – Recombination modeled by the ABCCD-model

Recombination in a semiconductor active region of an LED can be modeled by four mathematical terms. They are the (i) non-radiative Shockley-Read-Hall recombination in the active region, (ii) radiative recombination in the active region, (iii) Auger recombination in the active region, and (iv) leakage of carriers out of the active region; these four terms are represented by an A , B , and C coefficient, and an $f(n)$ term, respectively. As will be shown, the $f(n)$ term can be expressed by a C and D coefficient so that the coefficients are A , B , C (Auger), C (leakage) and D which will be referred to as the ABCCD model.

6.1 Relation between optical and electrical carrier generation

In a semiconductor active region, subjected to a steady-state optical excitation, the optical carrier-generation rate, $G_{optical}$, is equal to the carrier-recombination rate, R , that is

$$G_{optical} = R \quad (6.1)$$

where the generation and recombination rate have units of $\text{cm}^{-3} \text{ s}^{-1}$. In a semiconductor diode with an active region subjected to a steady-state electrical injection current I , the carrier injection rate, $G_{electrical}$, is equal the carrier-recombination rate, that is

$$G_{electrical} = \frac{I}{e} \frac{1}{V_{active}} = \frac{I}{e} \frac{1}{A d_{active}} = R \quad (6.2)$$

where V_{active} , A , and d_{active} are the volume of the active region, the current-injected area, and the thickness of the active region, respectively. Comparison of the above equations shows that the optical generation rate and the electrical generation rate can be viewed as equivalent, that is $G_{optical} = G_{electrical}$.

6.2 Free carrier concentrations in the active region

The active region of an LED, under equilibrium conditions (no injection current) will have an

equilibrium electron and hole concentration, n_0 and p_0 , respectively. Under optical or electrical injection conditions and neglecting carrier leakage out of the active region, an excess electron and hole concentration of $\Delta n = \Delta p$ is present in the active region. The total carrier concentrations of the active region thus are: $n = n_0 + \Delta n$ and $p = p_0 + \Delta p$. For an active region that is *either* undoped (or lightly doped) *or* for active regions with a very small volume, the following inequalities are valid: $n_0 \ll \Delta n$ and $p_0 \ll \Delta p$. Therefore,

$$n \approx \Delta n = \Delta p \approx p. \quad (6.3)$$

This result, that is, $n \approx p$, is valid under virtually all optical and electrical carrier-injection conditions; the result will be used in subsequent sections.

6.3 Recombination expressed in terms of the A, B, and C coefficients and f(n) term

Next, we analyze three recombination channels, namely:

- Non-radiative recombination via deep levels (Shockley-Read-Hall recombination),
- Radiative recombination (bimolecular recombination), and
- Non-radiative recombination via the Auger process

Shockley-Read-Hall (SRH) recombination

First, we discuss Shockley-Read-Hall (SRH) recombination; it is schematically illustrated in **Figure 6.1** and involves a deep level (“recombination center” or “trap”) through which recombination occurs. Assuming that the deep level has acceptor-like characteristics, its charge state can be neutral and negative. Recombination can then be described by the reaction equations

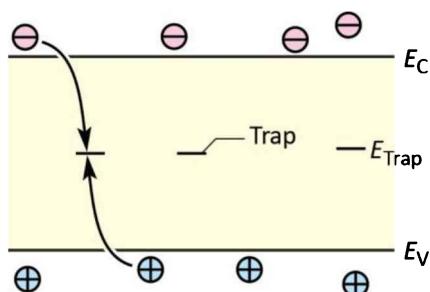
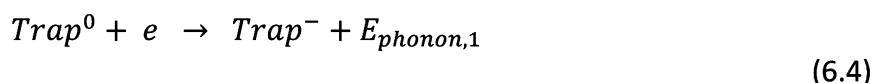
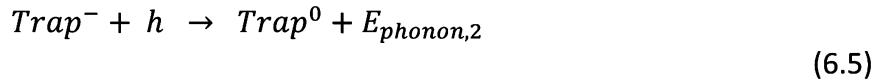


Fig. 6.1: Illustration of the Shockley-Read-Hall (SRH) recombination process. This non-radiative recombination process is taken into account by the A coefficient of the ABC model.



and



where $E_{\text{phonon},1}$ and $E_{\text{phonon},2}$ symbolize the sum of the energies of multiple phonons emitted during the two parts of the SRH process; it is $E_{\text{phonon},1} = E_C - E_T$ and $E_{\text{phonon},2} = E_T - E_V$.

Either one of the two reactions, Eqn. (6.4) or Eqn. (6.5), can be the slower one to occur; the slower one will limit the SRH-recombination rate. To simplify our subsequent analysis (and without loss of general validity), we assume that Eqn. (6.4) is the limiting (slower) process. As a result, the trap will be mostly in its neutral state, that is, $N_T^0 \approx N_T$, where N_T denotes the total trap concentration.

According to the **law of mass action**, the probability for a reaction to occur is given by the mathematical product of the concentrations of a reaction's initial reagents. Thus, the probability for the process given in Eqn. (6.4) is given by

$$\text{Probability} \propto [\text{Trap}^0] \times [e] \quad (6.6)$$

where the brackets “[...]" indicate the concentration of the species contained in the brackets. Writing the above concentrations in a notation that is more common in the field of semiconductors, we write

$$\text{Probability} \propto N_T n . \quad (6.7)$$

Given this probability, we can write the SRH non-radiative recombination rate as

$$R_{SRH} = \underbrace{\sigma_e v_{th} N_T}_A n \quad (6.8)$$

where σ_e is the trap's capture-cross-section for electron capture and v_{th} is the thermal velocity of electrons. The equation shows that R_{SRH} is directly proportional to the trap concentration.

Defining $A = \sigma_e v_{th} N_T$, we can write

$$R_{SRH} = A n \quad (6.9)$$

where A is proportional to the trap concentration.

The A coefficient has a unit of s^{-1} . The inverse of A thus has a unit of time and is referred to as the **SRH lifetime**

$$\tau_{SRH} = A^{-1} \quad (6.10)$$

The SRH lifetime can vary between 1 ns and 1 μ s for poor-quality (high trap content) and high-quality (low trap content) semiconductors, respectively. Accordingly, the A coefficient can vary between 10^6 s $^{-1}$ and 10^9 s $^{-1}$ for high-quality and poor-quality semiconductors, respectively.

Radiative recombination

Radiative recombination is associated with the emission of a photon; it is schematically illustrated in **Figure 6.2** and involves an electron and a hole. Radiative recombination can be described by the reaction equation



where $h\nu$ symbolizes the energy of the photon emitted.

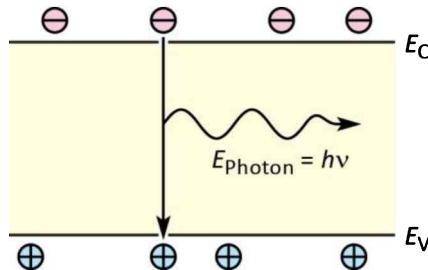


Fig. 6.2: Illustration of the radiative recombination process. This process is associated with the B coefficient of the ABC model.

According to the **law of mass action**, the probability for a reaction to occur is given by the mathematical product of the concentrations of a reaction's initial reagents. Thus, the probability for the process is given by

$$\text{Probability} \propto [e] \times [h]. \quad (6.12)$$

Writing the above concentrations in the more common notation, yields

$$\text{Probability} \propto n p. \quad (6.13)$$

Given this probability, we write the radiative recombination rate as

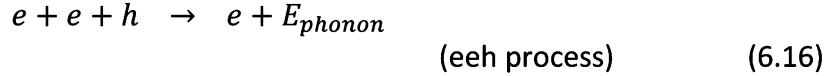
$$R_{\text{radiative}} = B n p \quad (6.14)$$

where B is a proportionality constant, namely the **bimolecular recombination coefficient**. Since, for an undoped active region (or at high excitation densities), it is $n = p$, we can write

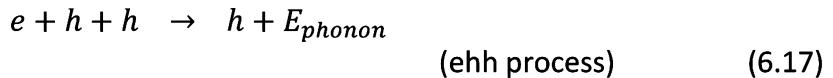
$$R_{\text{radiative}} = B n^2. \quad (6.15)$$

Auger recombination

Next, we discuss non-radiative **Auger recombination**; it is schematically illustrated in **Figure 6.3** and involves three carriers, either (i) two electrons and one hole or (ii) one electron and two holes (Landsberg, 2003). Auger recombination can be described by the reaction equations



and



where E_{phonon} symbolizes the sum of the energies of multiple phonons emitted when the electron excited highly into the conduction band (or the hole excited deep into the valence band) thermalizes to near the band edge by multiple phonon emission. Energy conservation requires $E_{\text{phonon}} \approx E_g$.

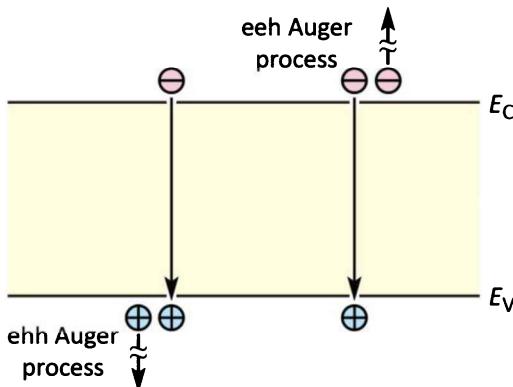


Fig. 6.3: Illustration of the Auger recombination process. This process is associated with the C coefficient in the ABC model. In the eeh process, subsequent to electron excitation high into the conduction band, the electron relaxes down to near the conduction-band edge by multiple longitudinal optical (LO) phonon emission. The relaxation process is known as thermalization. A similar thermalization occurs for holes in the ehh process.

According to the **law of mass action**, the probability for a reaction to occur is given by the mathematical product of the concentrations of a reaction's initial reagents. Thus, the probability for the Auger process is given by

$$\text{Probability} \propto [e] \times [e] \times [h] \quad (\text{eeh process}) \quad (6.18)$$

$$\text{Probability} \propto [e] \times [h] \times [h] \quad (\text{ehh process}) \quad (6.19)$$

Writing the above concentrations in the more common notation, and using the proportionality constants C_{eeh} and C_{ehh} yields

$$R_{\text{eeh}} = C_{\text{eeh}} n n p \quad (\text{eeh process}) \quad (6.20)$$

$$R_{ehh} = C_{ehh} n p p \quad (\text{ehh process}) \quad (6.21)$$

For the conditions considered here, it is $n \approx p$. Using $C = C_{eeh} + C_{ehh}$, we can write the Auger recombination rate as

$$R_{Auger} = C n^3 \quad (6.22)$$

where C is a proportionality constant called the **Auger coefficient**.

Carrier leakage expressed by $f(n)$ term

For carriers to recombine in the active region, they (i) must be injected into the active region and, once injected, (ii) must not escape from the active region. However, in practical devices, carriers may not be captured by the active region and carriers may escape from the active region. Carriers may be transported over a quantum well (QW) without being captured into the QW, especially in asymmetrically shaped QWs whose exit barrier has a lower energy than the entrance barrier. Escape of carriers out of a QW may be caused by a weak carrier confinement by the QW's confining barriers. Combined, these two effects (i.e. lack of capture and escape) are referred to as **carrier leakage** out of the active region; the effect is illustrated in **Figure 6.4**. Carriers that leak out of the active region will recombine, most likely non-radiatively, in the confinement regions or at the semiconductor–metal contacts.

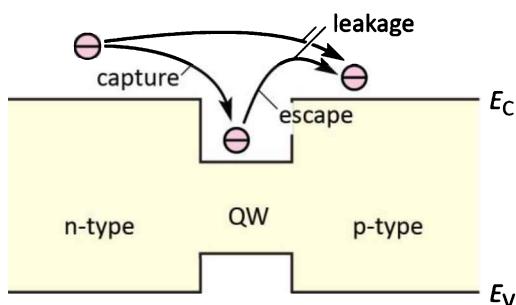


Fig. 6.4: Illustration of electron capture by a quantum well (QW) and electron escape from a QW. Electron leakage out of the active region is the result of either (i) lack of capture or (ii) escape.

The functional dependence of the carrier leakage on the carrier concentration depends on the specifics of the active region. Therefore, carrier leakage is expressed by the general function $f(n)$. This function can be assumed to strongly increase with carrier concentration.

Total recombination

The total recombination rate is given by the sum of SRH, radiative, Auger recombination, and carrier leakage, that is

$$R = R_{SRH} + R_{radiative} + R_{Auger} + R_{leakage} = An + Bn^2 + Cn^3 + f(n) \quad (6.23)$$

These four terms can be used to fully describe the total recombination rate in the active region.

6.4 Determination of the A, B, and C coefficients in absence of carrier leakage

In this section, we neglect carrier leakage and thus assume $f(n) = 0$. Then, the electrical injection rate must be equal to the recombination rate, that is

$$G_{electrical} = \frac{I}{e} \frac{1}{V_{active}} = R = An + Bn^2 + Cn^3 \quad (6.24)$$

where I is the injection current, e is the elementary charge, and V_{active} is the volume of the active region. The light output power (LOP) of a device is given by

$$LOP = Bn^2 h\nu V_{active} \quad (6.25)$$

where $h\nu$ is the photon energy. In the absence of carrier leakage, the radiative efficiency (RE) is equal to the internal quantum efficiency, which is given by

$$IQE = RE = \frac{Bn^2 V_{active}}{I/e} \quad (6.26)$$

Figure 6.5 shows an exemplary calculation of the three recombination mechanisms for a hypothetical LED. Inspection of the figure reveals that SRH non-radiative recombination dominates at low currents, followed by a regime where radiative recombination dominates. At high currents, the Auger term (Cn^3) term dominates. The regime in which Auger recombination dominates depends on the magnitude of the Auger coefficient and may not be reached for common LED devices.

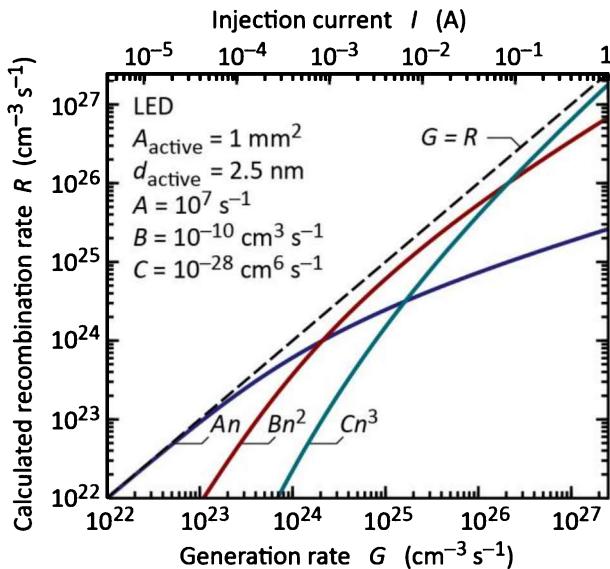


Fig. 6.5: Calculated SRH non-radiative, radiative, and Auger recombination rate as a function of the carrier generation rate in the active region (or carrier injection current; top abscissa). The figure shows that the SRH, radiative, and Auger process dominates at low, medium, and high carrier injection rates.

Note that in the calculation, we have considered a GaN-based LED and chosen realistic A and B coefficients, that is $A = 10^7 \text{ s}^{-1}$ and $B = 10^{-10} \text{ cm}^3 \text{s}^{-1}$. However, for illustrative purposes (i.e. to show the three mechanisms), we have chosen an unrealistically high Auger coefficient, that is, $C = 10^{-28} \text{ cm}^6 \text{s}^{-1}$. We note that it is generally accepted that the Auger coefficient in GaN is orders of magnitude lower than the chosen value.

It is well known that the Auger coefficient decreases exponentially with increasing bandgap energy (Beattie and Landsberg, 1959). As a result, the Auger coefficient can be very small, particularly for wide bandgap semiconductors. Harder *et al.* (2008) and Hader (2010) found an Auger coefficient of $3.5 \times 10^{-34} \text{ cm}^6/\text{s}$ for GaInN quantum wells. Delaney *et al.* (2009) found an Auger coefficient of about $10^{-31} \text{ cm}^6/\text{s}$ for GaInN with an energy gap of 2.7 eV corresponding to blue wavelengths ($\lambda = 459 \text{ nm}$). Based on the smallness of the Auger coefficient and the typically low current densities employed with LEDs (as compared to lasers), it is widely believed that Auger recombination plays a negligible role in wide-gap III–V nitride LEDs.

In the *absence* of carrier leakage (injection efficiency = $I/E = 100\%$), the A , B , and C coefficients can be determined by the following steps:

- Measure the light-output power (LOP) of the LED as a function of the injection current (I). That is, measure the LOP -versus- I curve.
- From the LOP values, determine the corresponding n values:
 - We use the equation $LOP = IE \times LEE \times Bn^2 h\nu V_{\text{active}}$.
 - We assume that $I/E = 100\%$.
 - We assume that the value of the light-extraction efficiency (LEE) can be estimated with reasonable accuracy.

- We assume that the design of the LED active region is known, so that the volume of the active region, V_{active} , is known.
- We use the known numerical value of the B coefficient ($B \approx 10^{-10} \text{ cm}^3 \text{ s}^{-1}$).
- Then, in the equation $LOP = IE \times LEE \times Bn^2 h\nu V_{\text{active}}$, all terms are known except n .
- Therefore, we can determine the value of n for each value of LOP .
- From the injection current (I) values, determine the corresponding R values:
 - The injection current and the recombination rate are related by: $I = e V_{\text{active}} R$.
 - Therefore, we can determine the value of R for each value of I .
- Using the results given above, plot the experimental R -versus- n curve.
- Fit the theoretical R -versus- n curve (i.e. $R = An + Bn^2 + Cn^3$) to the experimental R -versus- n curve.
- As a result of this procedure, the A and C coefficients are obtained.
- **Note:** Even if the values of B , V_{active} , and LEE are not known with high accuracy, the procedure nevertheless allows one to attain good estimates of the A and C coefficient.

It is instructive to consider the RE -versus- n curve. Based on the ABC model, the curve is described by the equation

$$RE = \frac{Bn^2}{An + Bn^2 + Cn^3} \quad (6.27)$$

Based on the equation, **Figure 6.6** displays (a) a family of RE -versus- n curves, where the A and C coefficient is changed (while B is kept constant) and (b) a family of RE -versus- n curves where the B coefficient is changed (while A and C is kept constant).

There are several points that are noteworthy regarding the ABC model and the associated RE -versus- n curves shown in **Figure 6.6**. *Firstly*, the curves display even symmetry when plotted on a logarithmic abscissa scale. This can be easily inferred from a visual inspection of the curves; the even symmetry was also proven in a mathematically rigorous manner (Dai *et al.*, 2011). *Secondly*, visual inspection also reveals that the width of the RE -versus- n curves becomes wider as the peak efficiency increases. This property allows one to determine the RE from the measured width of the RE -versus- n curve (Lin *et al.*, 2012). *Thirdly*, based on the ABC model, the maximum or peak radiative efficiency is related to the A , B , and C coefficients by (Lin *et al.*, 2012a and 2012b)

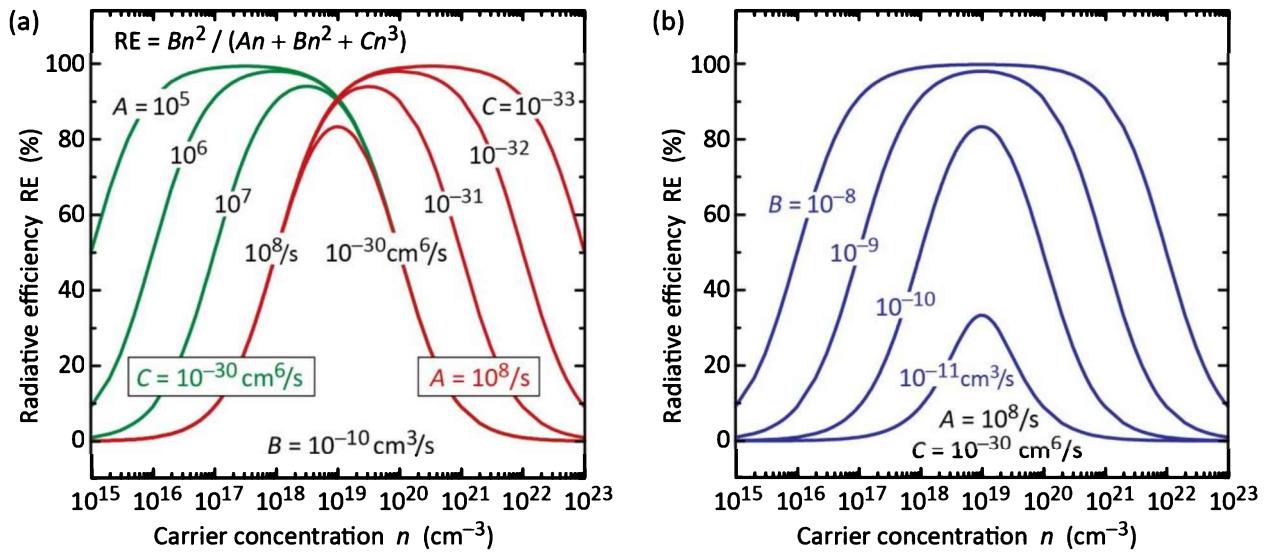


Fig. 6.6: Calculated radiative efficiency of a semiconductor material as a function of the carrier concentration for different A , B , and C coefficients. In figure (a), B is kept constant. In figure (b), A and C are kept constant.

$$RE_{peak} = \frac{1}{1 + 2\sqrt{AC} / B} . \quad (6.28)$$

Fourthly, the carrier concentration at which the radiative efficiency has its maximum is given by (Lin *et al.*, 2012a and 2012b)

$$n_{peak-RE} = \sqrt{A / C} . \quad (6.29)$$

The above discussion shows that the ABC is a simple yet powerful model. It has been widely employed for GaN-based LEDs to explain recombination dynamics in these devices (Piprek 2010; David and Grundmann, 2010a and 2010b; David and Gardner 2010c; Dai *et al.*, 2011; Scheibenzuber *et al.*, 2011). A noteworthy refinement to the ABC model was implemented by David and Grundmann (2010a and 2010b) by taking into account phase-space filling. It was shown, however, that the quantitative effect of phase-space filling is relatively small (Dai *et al.*, 2011).

6.5 Carrier leakage and the ABCCD model

In this section, we do not neglect carrier leakage and thus assume $f(n) \neq 0$. Motivated by the fact that carrier leakage is a widely found phenomena in III–V nitride LEDs, the recombination rate can be written as (Schubert *et al.*, 2007)

$$R = A_{SRH} n + Bn^2 + C_{Auger} n^3 + f(n) \quad (6.30)$$

where the function $f(n)$ represents the carrier leakage term that becomes increasingly relevant at high carrier concentrations. The function $f(n)$ is a mathematical function that depends on the physical mechanism causing the carrier leakage.

An analytic model for $f(n)$, based on drift-induced electron leakage out of the active region was developed by Lin *et al.* (2012a). The authors showed that drift-induced electron leakage can occur in pn junctions that are strongly asymmetric, that is, in pn junctions in which the electron concentration (as well as the electron mobility) is much higher than the hole concentration (as well as the hole mobility). When such asymmetric junctions are operated under high-injection conditions, an electric field evolves in the p-type neutral cladding layer. This field extracts electrons from the active region. For drift-induced electron leakage, the function $f(n)$ has the form (Lin *et al.*, 2012a)

$$f(n) = C_{DL} n^3 + D_{DL} n^4 \quad (6.31)$$

where the subscript “DL” refers to “drift leakage” and C_{DL} and D_{DL} are given by

$$C_{DL} = \frac{\delta \mu_n}{\mu_p p_{p0}} B \quad \text{and} \quad D_{DL} = \left(\frac{\delta \mu_n}{\mu_p p_{p0}} \right)^2 B . \quad (6.32)$$

The quantity δ is the ratio of the electron concentration in the barrier adjoining the active region to the electron concentration in the active region; δ has a value of $\delta \ll 1$; Lin *et al.* (2012a) deduced a value of $\delta \approx 10^{-3}$. The numerical value of C_{DL} was found to be on the order of 10^{-29} to $10^{-28} \text{ cm}^6 \text{ s}^{-1}$ for GaInN LEDs. Note that this value (C_{DL}) is orders of magnitude greater than the Auger coefficient (C_{Auger}).

Taking into account electron leakage out of the active region, the recombination rate is given by

$$R = A_{SRH} n + Bn^2 + C_{Auger} n^3 + C_{DL} n^3 + D_{DL} n^4 . \quad (6.33a)$$

This equation can be referred to as the **ABCCD-model**. If $C_{DL} \gg C_{Auger}$, the recombination rate can be written as

$$R \approx A_{SRH} n + Bn^2 + C_{DL} n^3 + D_{DL} n^4 . \quad (6.33b)$$

Accordingly, the internal quantum efficiency is given by

$$IQE = \frac{B n^2}{A_{SRH} n + B n^2 + C_{DL} n^3 + D_{DL} n^4} \quad (6.34)$$

A numerical example of a calculated IQE-versus-current curve is shown in *Figure 6.7*. Inspection of the figure reveals that the efficiency for very low currents approaches zero. The efficiency then reaches a peak and finally decreases gradually as the current increases. The regime of decreasing efficiency is referred to as the “**efficiency droop**” region.

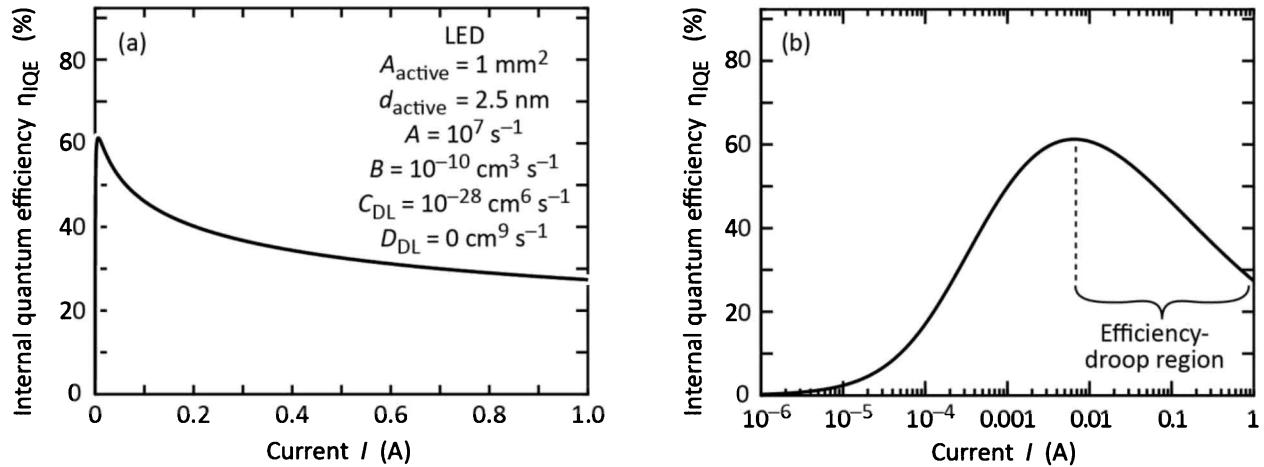


Fig. 6.7: Calculated IQE as a function of the injection current on (a) linear and (b) logarithmic abscissa using SRH recombination (An), radiative recombination (Bn^2), and a third-order non-radiative term representing drift-induced electron leakage ($C_{DL}n^3$).

6.6 References

- Beattie A. R. and Landsberg P. T. “Auger effect in semiconductors” *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* **249**, 16 (1959)
- Dai Q., Shan Q., Wang J., Chhajed S., Cho J., Schubert E. F., Crawford M. H., Koleske D. D., Kim M.-H., and Park Y. “Carrier recombination mechanisms and efficiency droop in GaInN/GaN light-emitting diodes” *Applied Physics Letters* **97**, 133507 (2010)
- Dai Q., Shan Q., Cho J., Schubert E. F., Crawford M. H., Koleske D. D., Kim M.-H., and Park Y. “On the symmetry of efficiency-versus-carrier-concentration curves in GaInN/GaN light-emitting diodes and relation to droop-causing mechanisms” *Applied Physics Letters* **98**, 033506 (2011)
- David A. and Grundmann M. J. “Droop in InGaN light-emitting diodes: A differential carrier lifetime analysis” *Applied Physics Letters* **96**, 103504 (2010a)
- David A. and Grundmann M. J. “Influence of polarization fields on carrier lifetime and recombination rates in InGaN-based light-emitting diodes” *Applied Physics Letters* **97**, 033501 (2010b)
- David A. and Gardner N. F. “Droop in III-nitrides: Comparison of bulk and injection contributions” *Applied Physics Letters* **97**, 193508 (2010c)
- Delaney K. T., Rinke P., and Van de Walle C. G. “Auger recombination rates in nitrides from first principles” *Applied Physics Letters* **94**, 191109 (2009)
- Hader J., Moloney J. V., Pasenow B., Koch S. W., Sabathil M., Linder N., and Lutgen S. “On the importance of radiative and Auger losses in GaN-based quantum wells” *Applied Physics Letters* **92**, 261103 (2008)
- Hader J. “On the importance of radiative and Auger losses in GaN-based quantum wells” *March Meeting*

- of the American Physical Society, Portland, Oregon, March 15~19 (2010)
- Landsberg P. T. "Recombination in semiconductors" (Cambridge University Press, Cambridge UK, 2003)
- Lin G.-B., Meyaard D., Cho J., Schubert E. F. Shim H., and Sone C. "Analytic model for the efficiency droop in semiconductors with asymmetric carrier-transport properties based on drift-induced reduction of injection efficiency" *Applied Physics Letters* **100**, 161106 (2012a)
- Lin G.-B., Shan Q., Birkel A. J., Cho J., Schubert E. F., Crawford M. H., Westlake K. R. and Koleske D. D. "Method for determining the radiative efficiency of GaInN quantum wells based on the width of efficiency-versus-carrier-concentration curve" *Applied Physics Letters* **101**, 241104 (2012b)
- Piprek J. "Efficiency droop in nitride-based light-emitting diodes" *Physica Status Solidi A* **207**, 2217 (2010)
- Scheibenzuber W. G., Schwarz U. T., Sulmoni L., Dorsaz J., Carlin J.-F., and Grandjean N. "Recombination coefficients of GaN-based laser diodes" *Journal of Applied Physics* **109**, 093106 (2011)
- Schubert M. F., Chhajed S., Kim J. K., Schubert E. F., Koleske D. D., Crawford M. H., Lee S. R., Fischer A. J., Thaler G., and Banas M. A. "Effect of dislocation density on efficiency droop in GaInN/GaN light-emitting diodes" *Applied Physics Letters* **91**, 231114 (2007)

7

7 – Current–voltage characteristics

In this chapter, we summarize the electrical characteristics of p-n junctions. We will strongly rely on the Shockley equation. We note that a detailed derivation of the results, e.g. the Shockley equation, will not be provided in this chapter.

7.1 Diode current–voltage characteristic based on Shockley equation

We consider an *abrupt p-n junction* with a donor concentration of N_D and an acceptor concentration of N_A . All dopants are assumed to be fully ionized so that the free electron concentration is given by $n = N_D$ and the free hole concentration is given by $p = N_A$. It is further assumed that no compensation of the dopants occurs by unintentional impurities and defects.

In the vicinity of an unbiased p-n junction, electrons originating from donors on the n-type side diffuse over to the p-type side where they encounter many holes with which they recombine. A corresponding process occurs with holes that diffuse to the n-type side. As a result, a region near the p-n junction is depleted of free carriers. This region is known as the **depletion region**.

In the absence of free carriers in the depletion region, the only charge in the depletion region is from ionized donors and acceptors. These dopants form a space charge region, with donors being located on the n-type side and acceptors on the p-type side. The space charge region produces a potential that is called the **built-in voltage** (also called **diffusion voltage**). The built-in voltage is given by

$$V_{\text{built-in}} = V_{bi} = \frac{kT}{e} \ln \frac{N_A N_D}{n_i^2} \quad (7.1)$$

where N_D and N_A are the donor and acceptor concentrations, respectively, and n_i is the intrinsic carrier concentration of the semiconductor. The built-in voltage is shown in the band diagram of **Figure 7.1**. The built-in voltage represents the barrier that free carriers must overcome in order to reach the neutral region of opposite conductivity type.

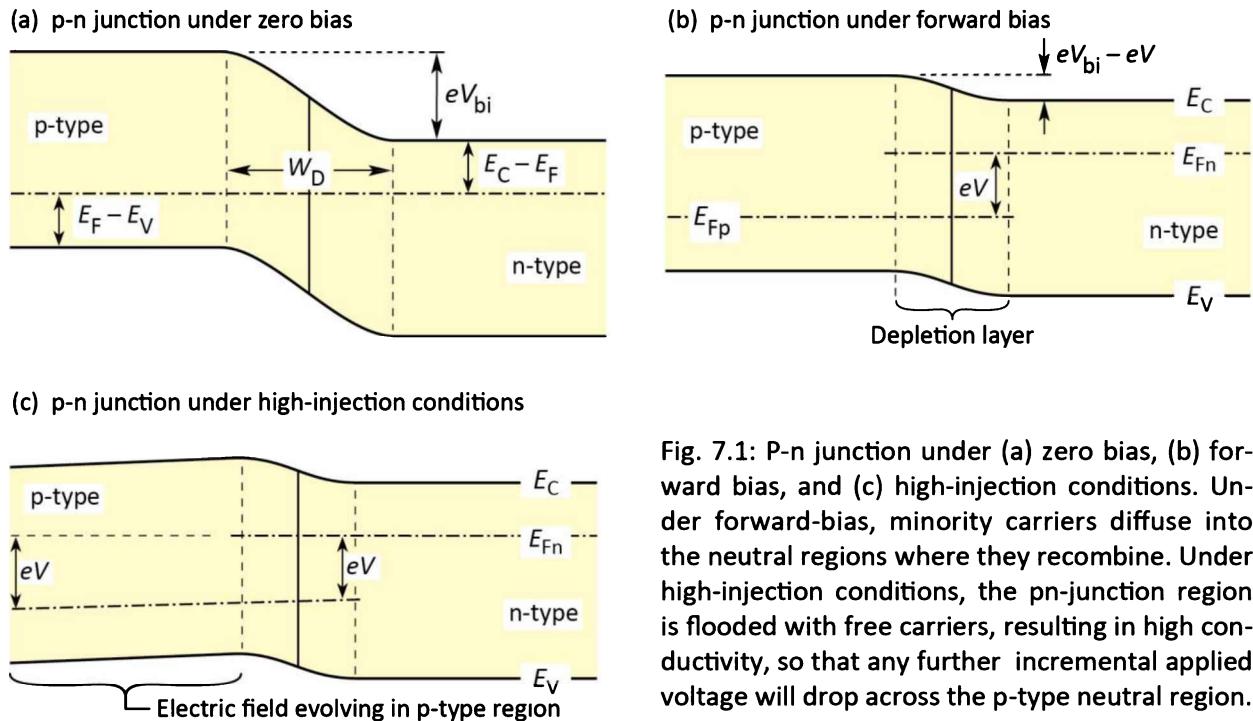


Fig. 7.1: P-n junction under (a) zero bias, (b) forward bias, and (c) high-injection conditions. Under forward-bias, minority carriers diffuse into the neutral regions where they recombine. Under high-injection conditions, the pn-junction region is flooded with free carriers, resulting in high conductivity, so that any further incremental applied voltage will drop across the p-type neutral region.

The width of the depletion region, the charge in the depletion region, and the built-in voltage are related by the Poisson equation. It is therefore possible to determine the depletion layer width from the built-in voltage. The depletion layer width is given by

$$W_{Depletion} = W_D = \sqrt{\frac{2\epsilon}{e} (V_{bi} - V) \left(\frac{1}{N_A} + \frac{1}{N_D} \right)} \quad (7.2)$$

where $\epsilon = \epsilon_r \epsilon_0$ is the semiconductor's permittivity and V is the applied voltage.

Upon application of a bias voltage to the p-n junction, the voltage drops across the highly resistive depletion region. This region is highly resistive due to the fact that it is depleted of free carriers. An external bias therefore decreases or increases the p-n junction barrier for forward or reverse bias, respectively. Under **forward-bias conditions**, shown in *Figure 7.1(b)*, electrons and holes are injected into the neutral regions and current flow *increases*. The carriers diffuse into the neutral regions of opposite conductivity type where they will eventually recombine and, for radiative recombination, emit a photon.

Under **high-injection conditions**, shown in *Figure 7.1(c)*, the depletion region becomes flooded with free carriers and its resistivity decreases so much that the resistivity of the neutral regions becomes relevant; the p-type neutral region is frequently more resistive as the n-type region. As a result, any further incremental voltage applied to the diode will drop across the p-

type neutral region, as shown in the figure.

William B. Shockley was the first to derive the current–voltage (*I*–*V*) characteristic of a p–n junction diode. The equation describing the *I*–*V* characteristic is therefore called the **Shockley equation**. It applies to low-injection conditions and is given by

$$I = eA \left(\sqrt{\frac{D_p}{\tau_p}} \frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}} \frac{n_i^2}{N_A} \right) (e^{eV/kT} - 1) \quad (7.3)$$

where A is the cross-sectional area of the diode, and $D_{n,p}$ and $\tau_{n,p}$ are the electron and hole diffusion constants and the electron and hole minority-carrier lifetimes, respectively.

Under reverse-bias conditions, the diode current saturates and the saturation current is given by the factor preceding the exponential function in the Shockley equation. Thus, the diode *I*–*V* characteristic can be written as

$$I = I_S (e^{eV/kT} - 1) \quad \text{with} \quad I_S = eA \left(\sqrt{\frac{D_p}{\tau_p}} \frac{n_i^2}{N_D} + \sqrt{\frac{D_n}{\tau_n}} \frac{n_i^2}{N_A} \right). \quad (7.4)$$

Under typical forward-bias conditions, the diode voltage is $V \gg kT/e$, and thus $[exp(eV/kT) - 1] \approx exp(eV/kT)$. Using Eqn. (7.1), the Shockley equation can be rewritten, for forward-bias conditions, as

$$I = eA \left(\sqrt{\frac{D_p}{\tau_p}} N_A + \sqrt{\frac{D_n}{\tau_n}} N_D \right) (e^{e(V-V_{bi})/kT}) \quad (7.5)$$

The exponent of the exponential function in the equation illustrates that the current strongly increases as the diode voltage approaches the built-in voltage, i.e. $V = V_{bi}$. The voltage at which the current strongly increases is called the **threshold voltage** or **turn-on voltage** and this voltage is given by $V_{th} = V_{turn-on} \approx V_{bi}$.

The band diagram of a p–n junction, shown in *Figure 7.1*, also illustrates the separation of the Fermi level from the conduction and valence band edge. The difference in energy between the Fermi level and the band edges can be inferred from Boltzmann statistics and is given by

$$E_C - E_F = -kT \ln \frac{n}{N_c} \quad (\text{for the n-type side}) \quad (7.6)$$

and

$$E_F - E_V = -kT \ln \frac{p}{N_v} \quad (\text{for the p-type side}). \quad (7.7)$$

The band diagram shown in **Figure 7.1** illustrates that the following sum of energies is zero:

$$eV_{bi} - E_g + (E_F - E_V) + (E_C - E_F) = 0. \quad (7.8)$$

In highly doped semiconductors, the separation between the band edges and the Fermi level is small compared with the bandgap energy, i.e. $(E_C - E_F) \ll E_g$ on the n-type side and $(E_F - E_V) \ll E_g$ on the p-type side. Furthermore, these quantities depend only weakly (i.e. logarithmically) on the doping concentration, as inferred from Eqns. (7.6) and (7.7). Thus, the third and fourth summand of Eqn. (7.8) can be neglected and the built-in voltage can be approximated by the bandgap energy divided by the elementary charge. Accordingly, we obtain

$$V_{bi} \leq V_{th} \leq E_g/e. \quad (7.9)$$

This approximation works well for all light-emitting diodes.

A detailed forward current–voltage curve (*I*–*V* curve) of a pn-junction diode is shown in **Figure 7.2**. Inspection of the curve reveals different regimes of the *I*–*V* curve and band diagrams of these regimes, all of which will be discussed below.

(1) **Shockley equation:** The Shockley equation predicts a diode current given by Eqn. (7.3).

This current is based on drift and diffusion and therefore may be referred to as “drift-diffusion current”. For voltages $V > 4kT/e$ (i.e. > 100 mV), the exponential term of the Shockley equation, i.e. $[\exp(eV/kT) - 1]$, can be approximated by $\exp(eV/kT)$. That is, in this regime, when plotting $\ln I$ versus V , the Shockley equation is represented by a straight line, as shown in **Figure 7.2**. The Shockley equation is frequently modified to include a “diode ideality factor” so that the exponential dependence, $I \propto \exp(eV/kT)$, becomes $I \propto \exp(eV/n_{ideal}kT)$ (Sah *et al.*, 1957). That is, the original Shockley equation has a diode ideality factor of $n_{ideal} = 1.0$. Deviations from the Shockley equation will be discussed below.

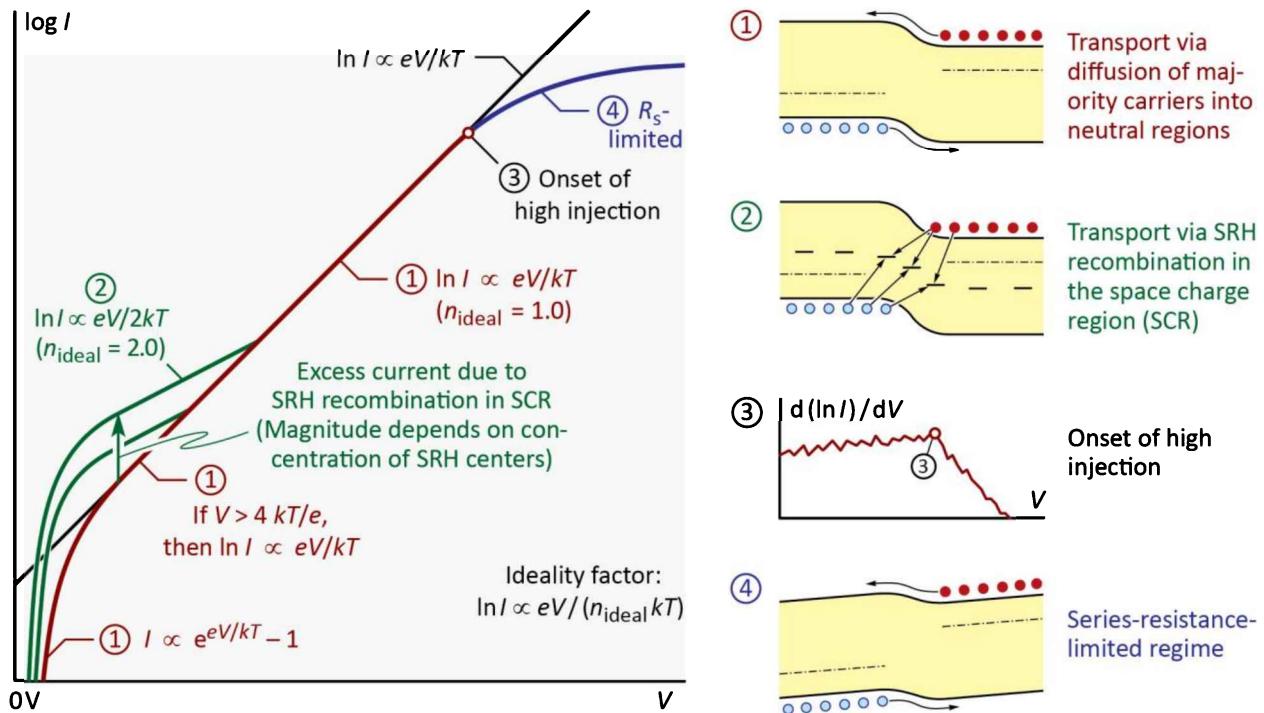


Fig. 7.2: Regions of interest in the forward I – V characteristic of a pn-junction diode based on Shockley's drift-diffusion theory. An "ideality factor", n_{ideal} , is frequently used to characterize the diode I – V characteristic.

(2) **Deviation at low current levels – SRH recombination in the space-charge region:** When $V \ll V_{bi}$, then the Shockley equation predicts a very small current. However, a second transport channel occurs that is not taken into account in the Shockley equation: Recombination of electrons and holes through deep levels (i.e. Shockley-Read-Hall or SRH recombination) in the space-charge region (SCR). This additional transport mechanism results in an excess current so that the total current is greater than the drift-diffusion current predicted by the Shockley equation, as shown in **Figure 7.2**. It has been shown that the SRH-recombination current in the SCR of a diode follows the proportionality $I \propto \exp(eV/2kT)$ so that the diode ideality factor is $n_{\text{ideal}} = 2.0$ (Sah *et al.*, 1957; Lorenz and Pilkuhn, 1966). The magnitude of the SRH current depends on the deep-level concentration and thus can vary widely among different diodes. For sufficiently large voltages, the excess current becomes, relative to the drift-diffusion current, increasingly smaller, and, as a result, becomes negligible at some point.

(3) **Onset of high-level injection:** The validity of the Shockley equation is limited to the low-injection regime. Assuming a n^+p junction ($N_D > N_A$), the low-level injection condition can be written as $\Delta n \ll 0.1 p_{p0}$, where Δn is the injected electron concentration, and p_{p0} is

the equilibrium hole concentration in the p-type region. Once this condition is violated, we enter the high-injection regime and the I – V curve deviates from the Shockley-equation. The onset of high injection is a specific point in the I – V characteristic. This specific point can be identified by the distinct kink (i.e. change in slope) that is found when plotting $d(\ln I)/dV$ as a function of V , as illustrated in **Figure 7.2**. The distinct kink is the onset-of-high-injection point.

- (4) ***Deviation at high current levels – Series-resistance-limited regime:*** Starting at the high-injection point, *part* of an incremental applied voltage drops across the space-charge region but also across the neutral regions (to be specific, predominantly across the more resistive neutral region). For sufficiently large voltages, *all* of an incremental applied voltage drops across the neutral regions. This is the series-resistance-limited regime. In this regime, the $(\ln I)$ -versus- V curve becomes a logarithmic curve, as illustrated in **Figure 7.2**.

William B. Shockley was a pioneer of semiconductor physics, co-inventor of the transistor, and the first to derive the current–voltage (I – V) characteristic of a p-n junction. Due to his many unique contributions to semiconductor physics and technology, including the co-invention of the bipolar transistor, he is considered a genius and the key mind in the early development of this field. A photograph of William B. Shockley, while teaching at Stanford University (Palo Alto, California), is shown in **Figure 7.3**.

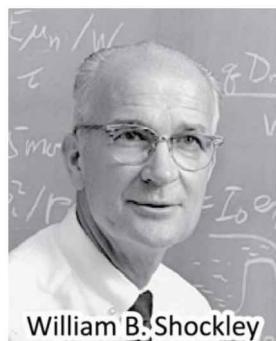


Fig. 7.3: William B. Shockley, pioneer of semiconductor physics, co-inventor of the bipolar transistor, and the first to derive the I – V characteristic of the p-n junction diode. Due to his many contributions to semiconductors, he is considered a genius and the key mind in the early development of the field. The photograph shows him while teaching at Stanford University (Palo Alto, California).

7.2 Diode forward voltage

The energy of photons emitted from a semiconductor with energy gap E_g is given by the bandgap energy, i.e.

$$h\nu \approx E_g . \quad (7.10)$$

In an ideal diode, every electron injected into the active region will generate a photon. Conservation of energy thus requires that the energy with which an electron is injected is equal to the photon energy. Thus, energy conservation requires

$$V_f = h\nu/e \approx E_g/e . \quad (7.11)$$

where V_f is the forward voltage of the diode. That is, the voltage applied to the LED is equal to the photon energy divided by the elementary charge.

A diode series resistance will cause the forward voltage to be slightly different from the ideal value, i.e. $h\nu/e$. That is, if the diode has a significant series resistance, an additional voltage drop occurs. The additional resistance can be caused by (i) *contact resistance*, (ii) *resistances caused by abrupt heterostructures*, and (iii) *bulk resistance* occurring particularly in materials with low carrier concentrations or low carrier mobilities. A voltage drop of magnitude IR_s occurs at the series resistance thereby increasing the drive voltage.

Several diode I – V characteristics of semiconductors made from different materials are shown in **Figure 7.4** along with the bandgap energy of these materials. The experimental threshold voltages shown in the figure, and the comparison with the bandgap energy of these materials, indicates that the energy gap and the threshold voltage indeed agree reasonably well.

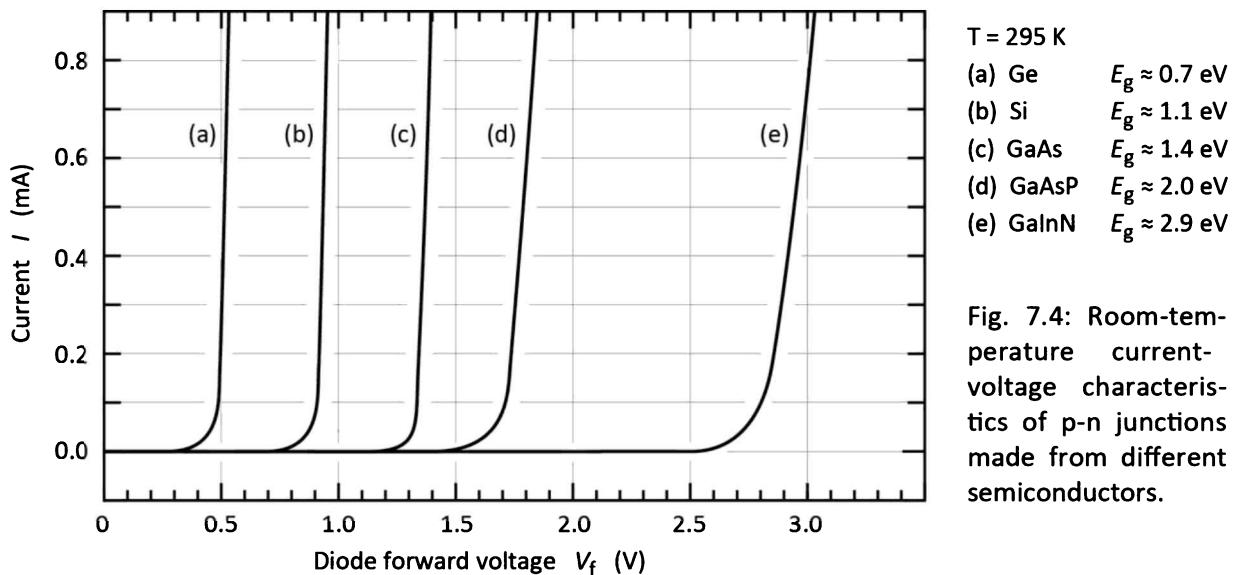
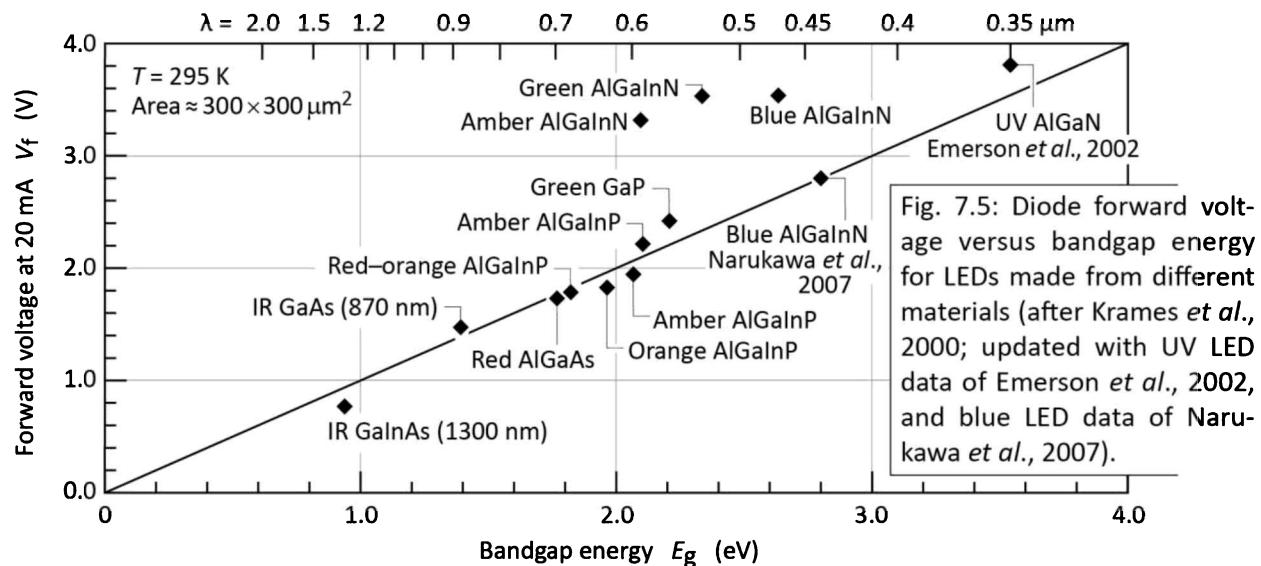


Fig. 7.4: Room-temperature current-voltage characteristics of p-n junctions made from different semiconductors.

The forward diode voltage at a diode current of 20 mA versus bandgap energy for LEDs emitting in the ultraviolet, visible, and infrared wavelength range is shown in *Figure 7.5* (Krames *et al.*, 2000; Emerson *et al.*, 2002; Narukawa *et al.*, 2007). The solid line illustrates the expected forward diode voltage. The line equals the bandgap energy divided by the elementary charge. Inspection of the figure reveals that most semiconductor LEDs follow the solid line, except for early LEDs based on III–V nitrides. This peculiarity is due to several reasons. Firstly, large bandgap discontinuities occur in the nitride material system, which cause an additional voltage drop. Secondly, the contact technology is less mature in the nitride material system, which causes an additional voltage drop at the ohmic contacts. Thirdly, the p-type conductivity in bulk GaN is generally low. Lastly, a parasitic voltage drop can occur in the n-type buffer layer. More recently, Narukawa *et al.* (2007) showed that blue AlGaN LEDs can nevertheless have very low forward voltages that are consistent with the theoretical expectation.



Assuming a chip area of $250 \mu\text{m} \times 250 \mu\text{m}$ and a current of 20 mA, the current density used in *Figure 7.5* to characterize the forward voltage is 32 A/cm^2 . Typical current densities in LEDs range from 30 A/cm^2 in low-power devices to 300 A/cm^2 in high-power devices.

Exercise: Drive voltages of LEDs. Calculate the approximate forward diode voltage of LEDs emitting in the blue, green, and red parts of the visible spectrum. Also calculate the forward diode voltage of LEDs emitting at 870 nm and 1 550 nm.

Solution:	Emission color	Wavelength	Photon energy	Drive voltage
	Blue	470 nm	2.6 eV	2.6 V
	Green	550 nm	2.2 eV	2.2 V
	Red	650 nm	1.9 eV	1.9 V
	IR	870 nm	1.4 eV	1.4 V
	IR	1550 nm	0.8 eV	0.8 V

7.3 Diode ideality factor

The Shockley equation gives the expected theoretical I – V characteristic of a p–n junction. Under typical forward-bias conditions, the diode voltage is $V \gg kT/e$, and thus $[\exp(eV/kT) - 1] \approx \exp(eV/kT)$, so that the diode I – V characteristic is given by

$$I = I_S e^{eV/(kT)} . \quad (7.12)$$

However, to describe experimentally measured characteristics, the following equation is used:

$$I = I_S e^{eV/(n_{ideal} kT)} \quad (7.13)$$

where n_{ideal} is the **diode-ideality factor**. This factor should have a value of 1.0 and is used to describe the deviation of a diode from ideal behavior.

Taking the logarithm of Eqn. (7.13), forming a derivative with respect to V , and solving the resulting equation for n_{ideal} yields

$$n_{ideal} = \frac{e}{kT} \left(\frac{d}{dV} \ln \frac{I}{I_S} \right)^{-1} = \frac{e}{kT} \left(\frac{d}{dV} \ln I \right)^{-1} . \quad (7.14)$$

Based on the equation, the ideality factor can be determined by plotting the value of n_{ideal} as a function of current. Such a plot is shown in **Figure 7.6**. When measuring the ideality factor, the following points should be considered:

- The range of voltage for which the ideality factor is measured need to fulfill the condition $V \gg kT/e$, i.e. $V > 100$ mV. Only for such voltages the approximation $[\exp(eV/kT) - 1] \approx \exp(eV/kT)$, used in the above derivation, is valid.
- At low forward voltages, the measured ideality factor can have unusually high values that are limited by the diode parallel resistance (shunt resistance).
- At high forward voltages, the measured ideality factor can have unusually high values that are limited by the diode series resistance. This limitation usually restricts the voltage

range within which the ideality factor should be measured to $V < V_{turn-on}$.

Due to these reasons, the ideality factor will have a minimum as a function of current, as can be seen in **Figure 7.6**. This minimum value is the appropriate ideality factor since it is least limited by either the diode parallel or series resistance.

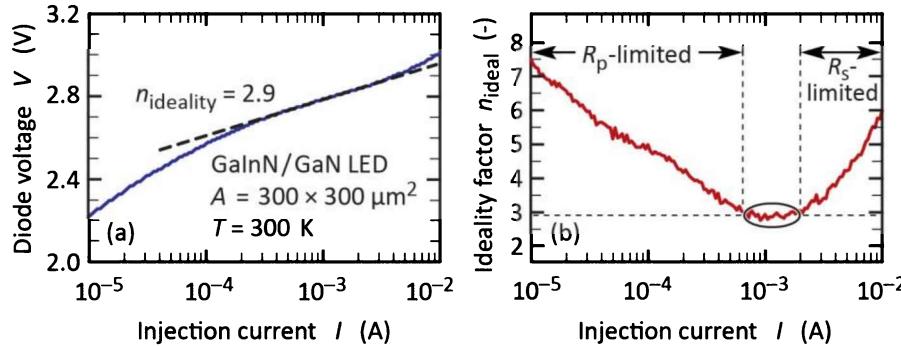


Fig. 7.6: Measurement of (a) diode voltage and (b) diode-ideality factor for a 5-QW GaInN/GaN LED. In the indicated ranges, the measurement is limited by the diode parallel resistance, R_p , and series resistance, R_s .

For a diode that follows the Shockley theory, i.e. when the current is dominated by diffusion of injected minority carriers into the neutral regions of the p-n junction, the ideality factor has a value of $n_{\text{ideal}} = 1.0$. However, it has been shown that if the forward current is dominated by recombination of carriers in the space-charge region, then the ideality factor has a value of $n_{\text{ideal}} = 2.0$ (Lorenz and Pilkuhn, 1966). The space-charge-recombination current dominates in the presence of high concentrations of recombination centers. A rigorous derivation of the $n_{\text{ideal}} = 2.0$ case is not performed here; the reader is referred to the literature (Sah *et al.*, 1957; Lorenz and Pilkuhn, 1966; Rhoderick and Williams, 1988).

For a very-high-quality diode, the ideality factor has a value of unity ($n_{\text{ideal}} = 1.0$). For diodes made of III–V arsenides and phosphides, the ideality factor indeed has typical values of $n_{\text{ideal}} = 1.1 \sim 1.5$, although, values as high as $n_{\text{ideal}} = 2.0$ have been found for such diodes.

For GaInN/GaN multi-quantum-well (MQW) diodes, the ideality factors can be much greater than 2.0. Values as high as $n_{\text{ideal}} = 7.0$ have been found for GaInN/GaN diodes. Various models have been proposed to explain the unusually high ideality factors in GaN-based diodes (Casey *et al.*, 1996; Perlin *et al.*, 1996; Li *et al.*, 2001; Cao *et al.*, 2002; Shah *et al.*, 2003; Mayes *et al.*, 2004; Zhu *et al.*, 2009).

Casey *et al.* (1996) investigated InGaN/AlGaN LEDs and found a diode-ideality factor of 6.8 which the authors considered unrealistically high. The high ideality factor was attributed to a tunneling current rather than a diffusion current or space-charge recombination current where

the ideality factor varies between 1.0 and 2.0, respectively.

Shah *et al.* (2003) identified multiple individual hetero-junctions and metal-semiconductor junctions in GaN-based diodes, all of which can have a rectifying characteristic. The measured ideality factor then is the sum of multiple ideality factors, each being associated with the multiple individual hetero-junctions of the device.

Subsequently, Zhu *et al.* (2009) attributed the high ideality factors in GaInN/GaN multiple-quantum well LEDs to the presence of multiple GaN quantum-barrier layers that are triangular in shape. Each of the triangular-shaped quantum barriers was proposed to act as a Schottky-like barrier with a rectifying effect. As a consequence, the measured ideality factor is the sum of the ideality factors associated with the multiple individual quantum barriers of the MQW LED. Given the presence of large polarization fields in GaN-based devices, the occurrence of such rectifying triangular barriers indeed appears realistic.

7.4 Diode parasitic resistances

Frequently, a diode has ***unwanted*** or ***parasitic resistances***. The effect of a series resistance and a parallel resistance is shown in **Figure 7.7 (a)**. A series resistance can be caused by excessive contact resistance or by the resistance of the neutral regions. A parallel resistance can be caused by any channel that bypasses the p-n junction. This bypass can be caused by damaged regions of the p-n junction or by surface imperfections.

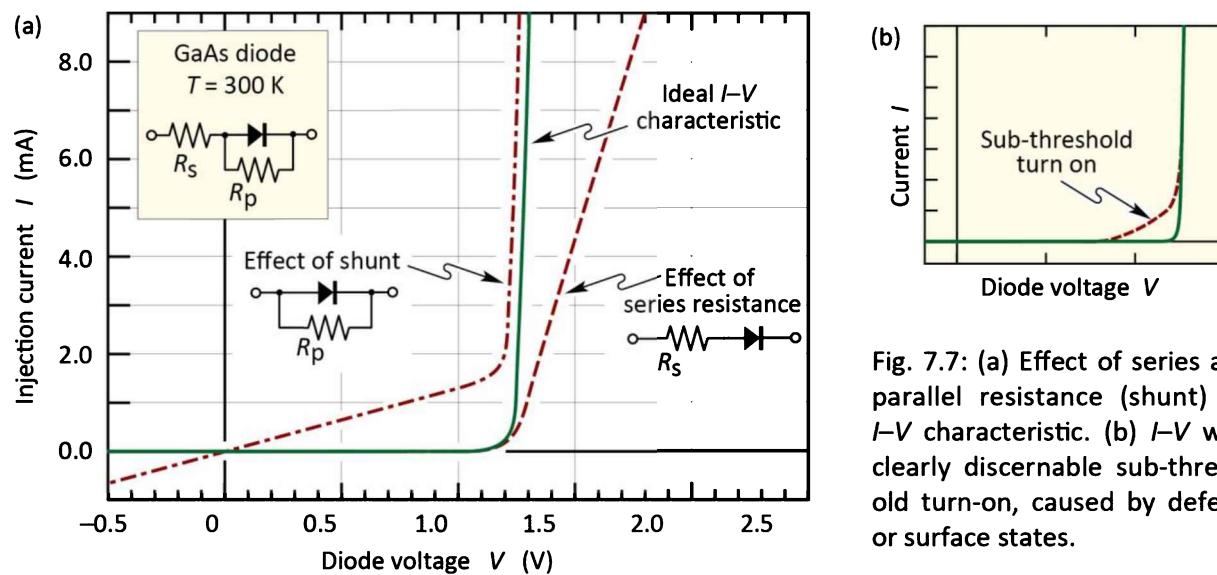


Fig. 7.7: (a) Effect of series and parallel resistance (shunt) on I - V characteristic. (b) I - V with clearly discernable sub-threshold turn-on, caused by defects or surface states.

The diode I – V characteristic, as given by the Shockley equation, needs to be modified in order to take into account parasitic resistances. Assuming a shunt with resistance R_p (parallel to the ideal diode) and a series resistance R_s (in series with the ideal diode and the shunt), the I – V characteristic of a forward-biased p–n junction diode is given by

$$I - \frac{V - IR_s}{R_p} = I_S e^{e(V - IR_s) / (n_{ideal} kT)}. \quad (7.15)$$

For $R_p \rightarrow \infty$ and $R_s \rightarrow 0$, this equation reduces to the Shockley equation.

Occasionally, the diode turn-on is distributed over a range of voltages rather than occurring abruptly at the threshold voltage. Both types of turn-on are shown in *Figure 7.7 (b)*. The non-abrupt turn-on is referred to as **sub-threshold turn-on** or **premature turn-on**. The sub-threshold current can be caused by carrier transport through surface states or deep levels in the bulk of the semiconductor.

7.5 Measurement of diode parasitic resistances

The diode parallel resistance can be evaluated near the origin of the I – V diagram where $V \ll E_g/e$. For this voltage range, the p–n junction current can be neglected and the parallel resistance is given by

$$R_p \approx \left. \frac{dV}{dI} \right|_{near\ origin}. \quad (7.16)$$

Note that in any reasonable diode, the parallel resistance is much larger than the series resistance so that the series resistance need not be taken into account when evaluating the parallel resistance.

The series resistance can be evaluated at a high voltage where $V > E_g/e$. For sufficiently large voltages, the diode I – V characteristic becomes linear and the series resistance is given by the tangent to the I – V curve, as shown in *Figure 7.8 (a)*

$$R_s \approx \left. \frac{dV}{dI} \right|_{at\ voltages\ exceeding\ turn-on}. \quad (7.17)$$

However, it may not be practical to evaluate the diode resistance at high voltages due to device heating effects. For this case, the following procedure will be suitable.

For devices with a high parallel resistance ($R_p \rightarrow \infty$), the diode I - V characteristic given in Eqn. (7.15), can be written as

$$I = I_S e^{e(V - IR_S) / (n_{ideal} kT)} . \quad (7.18)$$

Solving the equation for V and then differentiating V with respect to I yields

$$\frac{dV}{dI} = R_S + \frac{n_{ideal} kT}{e} \frac{1}{I} \quad (7.19)$$

where the second summand on the right-hand side of the equation represents the differential p-n junction resistance. Multiplication of the equation by I allows one to identify the series resistance of the diode as the slope of a ($I dV/dI$)-versus- I plot, as shown in **Figure 7.8 (b)**.

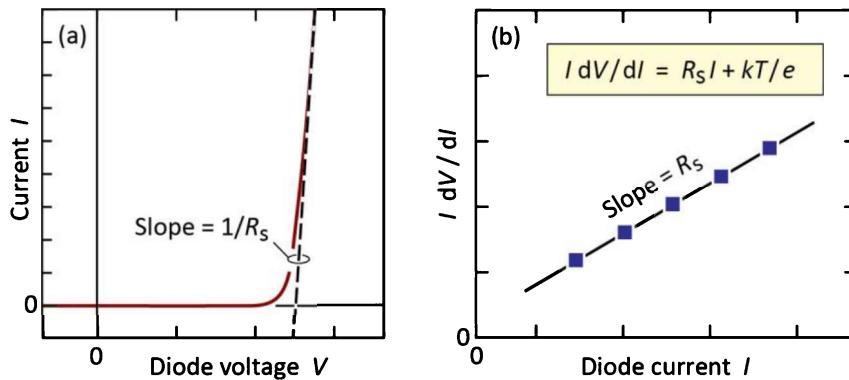
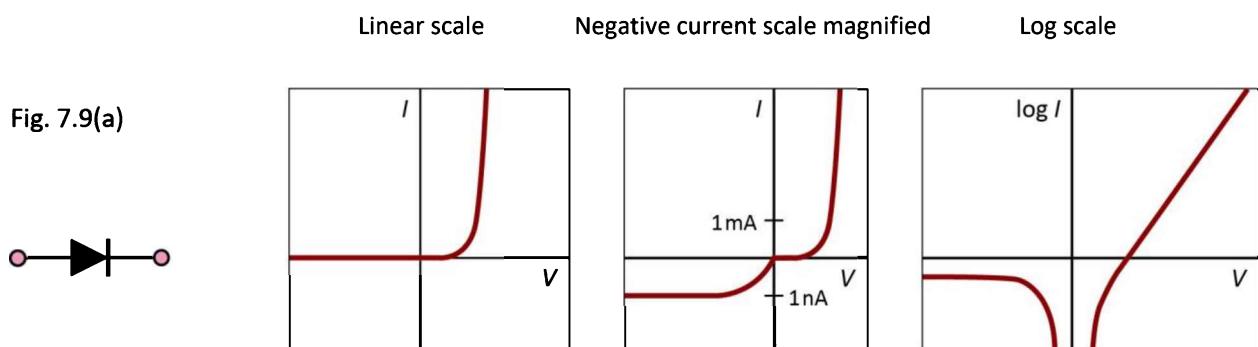


Fig. 7.8: Methods for evaluating diode series resistance. (a) Tangent for $V > V_{th}$ provides R_s . (b) Equation shown as inset is valid for forward bias ($V \gg kT/e$).

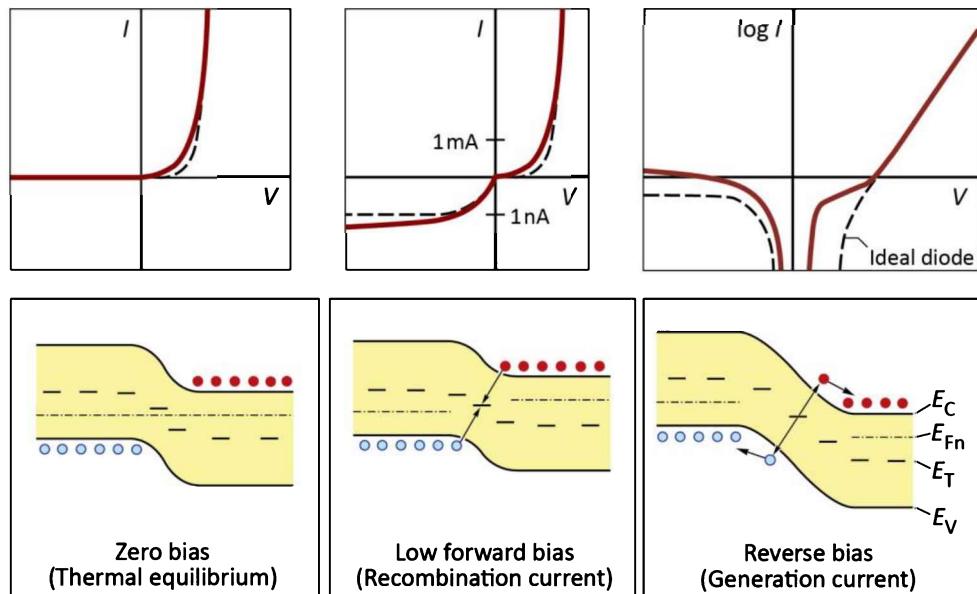
7.6 Diagnosis of potential problems in the current-voltage characteristic

Detailed inspection of the diode I-V characteristic on a linear as well as logarithmic scale allows for the diagnosis of potential problems such as shunts, series resistances, premature turn-on, and parasitic diodes. **Figure 7.9** shows a number of parasitic effects that can occur in diodes. The diagrams may allow the reader to diagnose and identify specific problems in diodes.



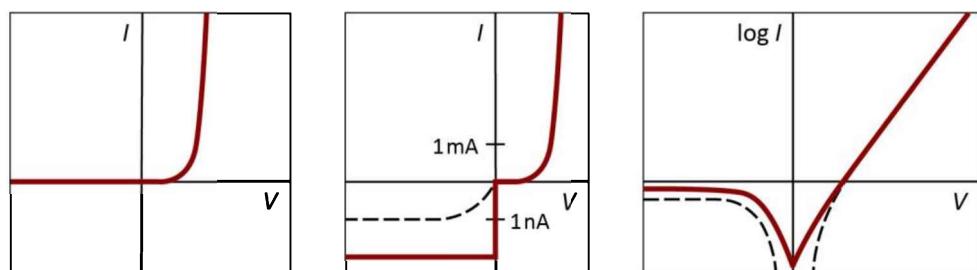
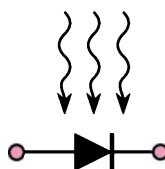
Ideal diode: The ideal diode I - V characteristic is given by the Shockley equation.

Fig. 7.9(b)



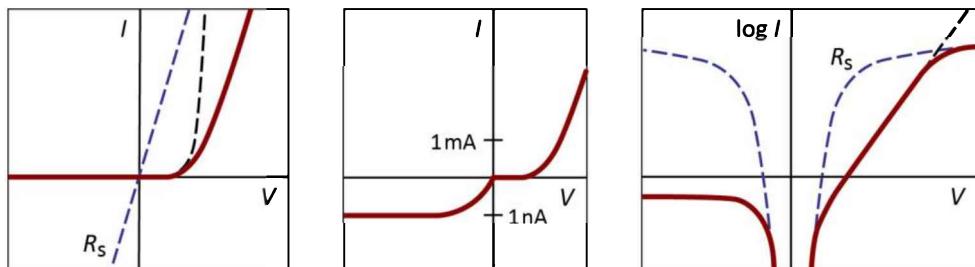
Depletion region generation and recombination: The Shockley diode equation does not account for carrier generation and recombination events in the depletion region. However, in practical diodes, there are trap levels in the depletion region, which make such events likely. Carrier generation and recombination causes an excess current for both forward and reverse bias. In the forward-bias regime, the excess current is due to the recombination of minority carriers in the depletion region. This recombination current dominates only at low voltages and gives an ideality factor of 2.0. At higher voltages, the diffusion current dominates resulting in an ideality factor of 1.0. In the reverse-bias regime, the excess current is due to the generation of carriers in the depletion region. Under the influence of the electric field in the depletion region, generated carriers drift to the neutral regions. This generation current keeps increasing with reverse voltage due to the increasing depletion-layer width.

Fig. 7.9(c)



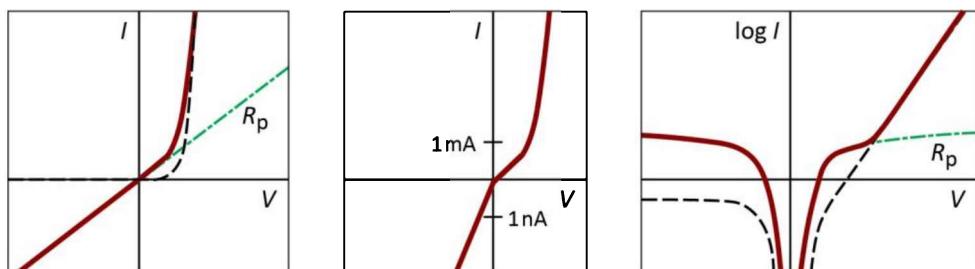
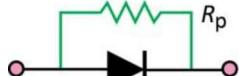
Photocurrent: In a practical measurement within a lighted room, a diode located in a transparent package generates a photocurrent. Therefore, measurements need to be carried out in the dark. Switching off room lights or covering the setup with a dark cloth helps in reducing the photocurrent. In the dark, at zero voltage, there should be zero current. However, a very small non-zero current (e.g. 10^{-12} A) is frequently measured. The non-zero current is usually due to the limited accuracy of the measurement instrument. The best instruments will measure a current of about 10^{-15} A at zero bias, even if the measurement is carried out in total darkness ($10^{-15} \text{ A} = 1 \text{ atto ampere} = 10^{-3} \text{ pA}$).

Fig. 7.9(d)



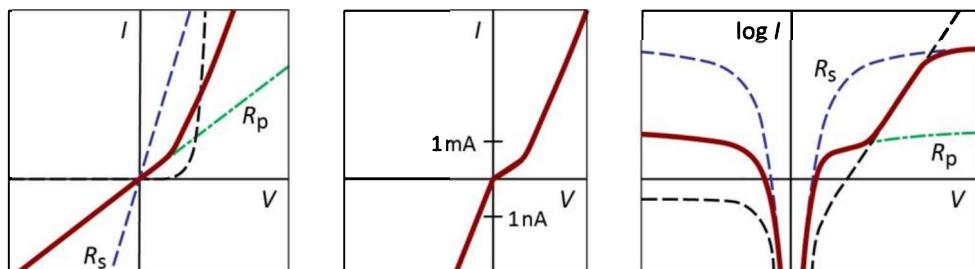
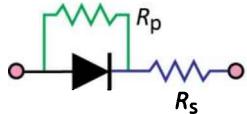
Diode with series resistance: A diode with a series resistance shows a deviation from the exponential behavior at high forward currents. According to Kirchhoff's voltage law, the voltages across the diode and the resistor add up. Note that a simple resistor has a linear and a logarithmic shape of the I - V characteristic when plotted on a linear and semi-logarithmic scale, respectively.

Fig. 7.9(e)



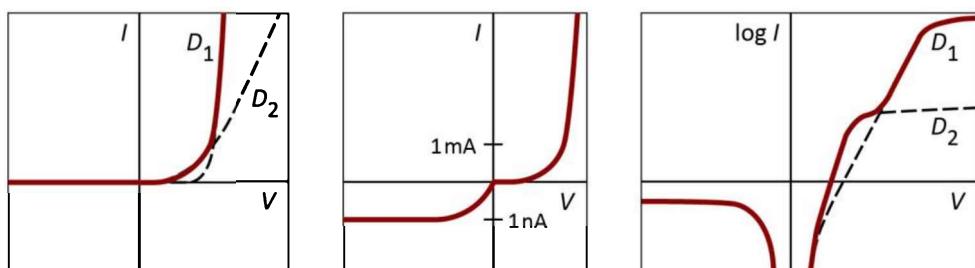
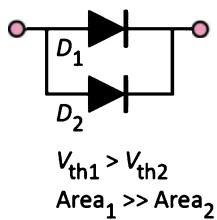
Diode with parallel resistance (shunt): According to Kirchhoff's current law, the currents through diode and resistor add up. Note that the forward “hump” seen on the semi-logarithmic plot has about the same level as the reverse saturation current. This is a characteristic by which a shunt can be identified.

Fig. 7.9(f)



Diode with series and parallel resistance (shunt): Effects of shunt and series resistance found at low and high currents, respectively.

Fig. 7.9(g)



Parasitic diode with lower barrier height and smaller area than main diode: Such diodes display premature turn-on caused by leakage through either surface states at the perimeter

of the diode chip or defective regions within the p-n junction plane that have a lower barrier height than the main p-n junction. Note that the forward “hump” on the semi-logarithmic plot has much higher level than the reverse saturation current, which is not the case for diodes with a shunt.

Exercise: Critical points of diode current–voltage characteristics. The I – V characteristics of diodes are frequently characterized in terms of **four critical points**, namely *forward voltage one*, V_{f1} , *forward voltage two*, V_{f2} , *forward voltage three*, V_{f3} , and *reverse saturation current*, I_s , specified at the operating current (e.g. 100 mA), a small forward current (e.g. 100 μ A), a very small forward current (e.g. 10 μ A), and at negative bias (e.g. -5 V), respectively. The critical points are shown in **Figure 7.10**.

- (a) Explain the relevance of the critical points.
- (b) Two GaInN diodes have the following data: (1) $V_{f1} = 3.2$ V, $V_{f2} = 2.5$ V, $V_{f3} = 2.3$ V, $I_s = 0.8$ μ A; (2) $V_{f1} = 3.4$ V, $V_{f2} = 2.0$ V, $V_{f3} = 1.8$ V, $I_s = 0.8$ μ A. Which device has the more favorable characteristics?
- (c) What is the threshold voltage, V_{th} , of a diode?

Solution: (a) V_{f1} should be as *low* as possible; high values of V_{f1} indicate a high series resistance. The forward voltage two, V_{f2} , should be as *high* as possible (as close to V_{f1} as possible), since low values of V_{f2} indicate an excessive sub-threshold leakage current. The same argument applies to V_{f3} . The reverse saturation current should be as low as possible as high values of I_s indicate excessive leakage paths (e.g. surface leakage or bulk leakage mediated by surface states, bulk point defects and dislocations). Low values of V_{f1} , high values of V_{f2} and V_{f3} , and low values of I_s are consistently correlated with high device reliability. (b) Device (1) has more favorable characteristics due to lower series resistance and lower sub-threshold leakage. (c) The threshold voltage, V_{th} , is the approximate voltage at which the I – V characteristic has a kink, i.e. where the diode becomes conductive. The threshold voltage is frequently identified by plotting the tangent to the I – V curve, as indicated in **Figure 7.10**.

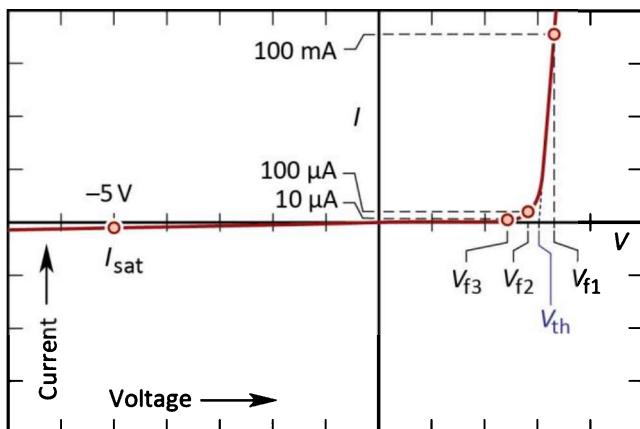


Fig. 7.10: Critical points of diode I – V characteristic, namely “forward voltage one”, V_{f1} (measured at operating current, e.g. 100 mA), “forward voltage two”, V_{f2} (measured at low current, e.g. 100 μ A), “forward voltage three”, V_{f3} (measured at very low current, e.g. 10 μ A), and reverse saturation current (measured at reverse bias, e.g. -5.0 V). The threshold voltage, V_{th} , is the approximate voltage where the I – V characteristic has a kink, i.e. where the diode becomes conductive.

7.7 Ohmic contacts

Ohmic metal-to-semiconductor contacts have a linear I - V characteristic and a low contact resistance so that the current flows unimpededly between the metal contact and the semiconductor. It is useful to distinguish between ohmic contacts with vertical current flow geometry and lateral current-flow geometry. The two geometries are illustrated in *Figure 7.11*.

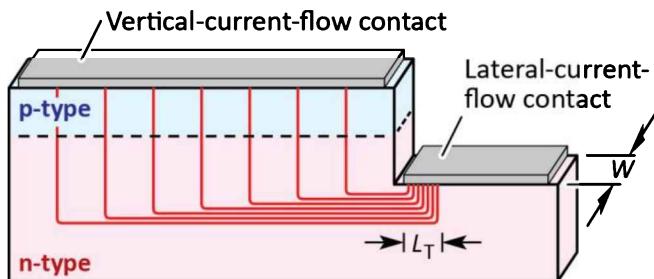


Fig. 7.11: Vertical-current-flow contact and lateral-current-flow contact. For the vertical-current-flow contact, the entire contact area is desirably used. For the lateral-contact-flow contact, only a stripe with dimension L_T (transfer length) is used.

In a **vertical-current-flow contact**, the current flows perpendicular to the contact plane, as illustrated for the p-type contact in the figure. That is, the current flows uniformly across the entire area of the contact. Accordingly, the total resistance of the contact, R_c , is given by

$$R_c = \rho_c / A \quad (7.20)$$

where ρ_c is the specific contact resistance (having the units $\Omega \text{ cm}^2$), and A is the contact area.

In contrast, in a **lateral-current-flow contact** the current flows parallel to the contact plane. Naturally, the current flow bunches at the edge of the contact, as illustrated in *Figure 7.11* for the n-type contact. The length over which the current flow transfers from the metal to the semiconductor (i.e. the length over which the current crowds or bunches) is called the transfer length, L_T (Berger, 1972; Schroder, 2006). As a consequence, the effectively used contact area, i.e. $L_T \times W$, can be smaller than the actual contact area (as illustrated in the figure). Accordingly, the total resistance of the contact, R_c , is given by

$$R_c = \rho_c / (L_T W) . \quad (7.21)$$

The most common method to measure the specific contact resistance is the **Transmission Line Model** or **Transfer Length Model** (TLM) (Berger, 1972; Schroder, 2006). The TLM method uses a contact configuration shown in *Figure 7.12 (a)*, i.e. a series of rectangular contacts separated by gaps with different sizes. In order for the current to flow uniformly between the contacts, a mesa is frequently employed, as shown in the figure. The TLM method is based on the approximation

$W_{\text{mesa}} \approx W$. The literature gives guidance on corrections that can be made, if $W_{\text{mesa}} > W$ (Berger, 1972; Schroder, 2006).

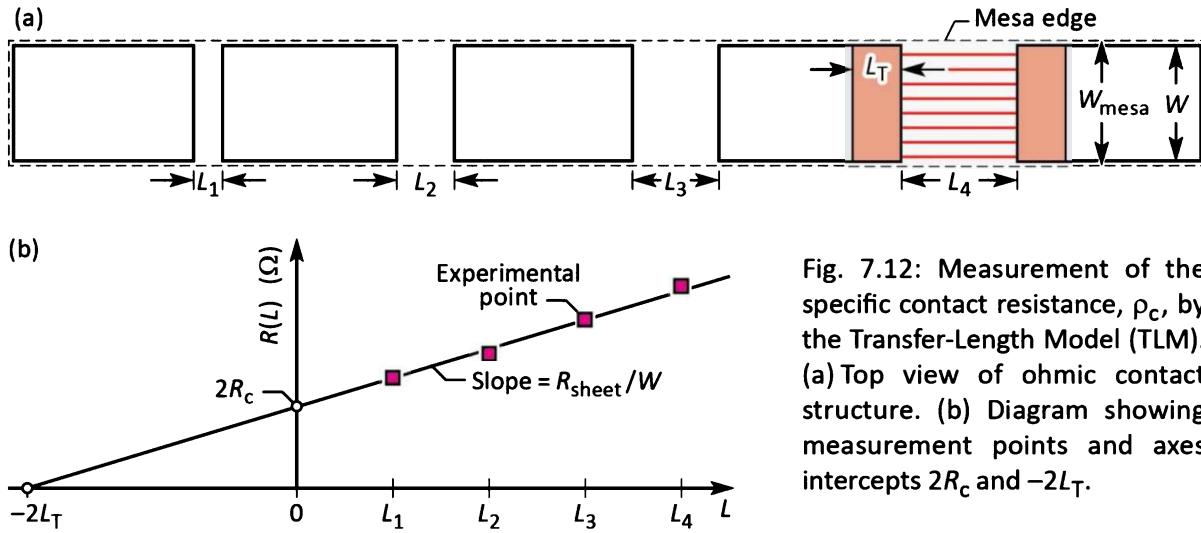


Fig. 7.12: Measurement of the specific contact resistance, p_c , by the Transfer-Length Model (TLM). (a) Top view of ohmic contact structure. (b) Diagram showing measurement points and axes intercepts $2R_c$ and $-2L_T$.

The resistance measured between adjacent contacts depends on their separation L and is given by (Berger, 1972; Schroder, 2006)

$$R(L) = \frac{R_{\text{sheet}}}{W} (L + 2L_T) \quad (7.22)$$

where R_{sheet} is the sheet resistance of the semiconductor and L is the distance between the contacts (size of the gap). For $L \rightarrow 0$, the resistance reduces to

$$R(0) = \frac{R_{\text{sheet}}}{W} 2L_T = 2R_c \quad (7.23)$$

where R_c is the contact resistance. Solving the equation for the sheet resistance yields

$$R_{\text{sheet}} = \frac{R_c W}{L_T} \quad (7.24)$$

In this equation, R_c and L_T are obtained from experiments, i.e. from the axes intercepts of the straight line fitted to the experimental data, as shown in **Figure 7.12 (b)**. Therefore, the TLM method allows one to determine the sheet resistance. Because the sheet resistance is obtained from the TLM measurement and may be known independently (e.g. from an independent sheet resistance measurement), the comparison of the sheet resistances can be used to confirm the reasonableness of the results.

The specific contact resistance can be obtained by taking into account that the effective area of a planar contact is $W \times L_T$. Accordingly, the specific contact resistance is given by (Berger, 1972; Schroder, 2006)

$$\rho_c = R_c W L_T = R_{sheet} L_T^2 \quad (7.25)$$

The above-described procedure is a standard procedure for assessing the contact resistance in LEDs and other devices. Several modifications to the TLM method are found in the literature (see, for example, Chen, 2004; Schroder, 2006). A very common modification is the **circular TLM method**. It does not require a mesa etch. The reader is referred to the literature for a detailed discussion of the circular TLM method (Schroder, 2006).

Typical good values of the specific contact resistances are less than $10^{-5} \Omega\text{cm}^2$ for III–V nitrides, phosphides, arsenides, and antimonides. However, wide-bandgap semiconductors are generally more challenging in forming low-resistance ohmic contacts. For III–V nitrides, typical specific contact resistances are in the $10^{-4} \sim 10^{-6} \Omega\text{cm}^2$ range. Furthermore, p-type ohmic contacts to GaN generally have a larger contact resistance than n-type contacts; this is due to the lower mobility of holes and their higher effective mass.

7.8 Physics of ohmic contacts

Much of the understanding of metal-semiconductor contacts is based on the pioneering work of Walter Schottky who analyzed rectifying metal-semiconductor contacts, also referred to as Schottky contacts. The starting point of the **Schottky model** is the work function in the metal and the electron affinity in the semiconductor (Schottky, 1938, 1939, 1942). Removing an electron from an atom, i.e. ionizing the atom, requires energy, specifically the ionization energy, to overcome the electrostatic attraction between the electron and the ion. Likewise, it takes a minimum energy to remove an electron from a metal. This energy is the **metal work function**, Φ_{metal} . It is defined as the energy difference between a free electron (i.e. an electron outside the crystal, e.g. in vacuum) and the Fermi energy, i.e. $\Phi_{metal} = \Phi_{vacuum} - E_F$. Similarly, for a semiconductor, it takes energy to remove an electron from the bottom of the conduction band and lift it to the vacuum level. This energy is the **semiconductor electron affinity**, χ_{semi} . It is defined as the energy difference between the vacuum level and the conduction-band energy, i.e. $\chi_{semi} = \Phi_{vacuum} - E_C$.

Bringing the two materials in contact, energy conservation requires that a barrier is formed at the metal semiconductor junction, as illustrated in *Figure 7.13 (a)*. This barrier, called **Schottky barrier**, is given by

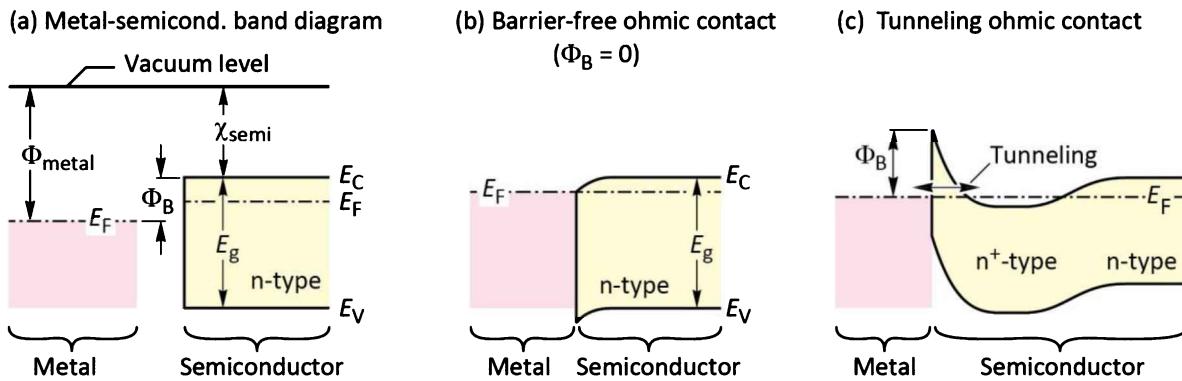


Fig. 7.13: (a) Metal-semiconductor band diagram (before physical contact). (b) Barrier-free contact enabled by the match of metal work function (Φ_{metal}) and semiconductor electron affinity (χ_{semi}). (c) Tunneling contact enabled by a high doping concentration.

$$\Phi_B = \Phi_{metal} - \chi_{semi} \quad (7.26)$$

As a result of the Schottky barrier, electrons in the conduction band of the semiconductor can easily propagate out of the semiconductor. However, in the reverse direction, a barrier is encountered. Accordingly, the contact is rectifying and has a very high resistance for $V \rightarrow 0$ (i.e. very low voltages).

Making a low-resistance ohmic contact requires that charge carriers can overcome the Schottky barrier. There are two basic strategies that allow carriers to overcome the Schottky barrier, as discussed below.

(i) Ohmic contact having reduced barrier height enabled by metal with proper work function

Following the Schottky model, the Schottky barrier vanishes, if, for an n-type semiconductor $\Phi_{metal} \leq \chi_{semi}$. The particular situation $\Phi_{metal} = \chi_{semi}$ is shown in *Figure 7.13 (b)*. Correspondingly, for a p-type semiconductor, if $\Phi_{metal} \geq \chi_{semi} + E_g$, the Schottky barrier vanishes. Therefore, the metal used for an ohmic contact to an n-type semiconductor should have a low work function. Correspondingly, the metal used for an ohmic contact to a p-type semiconductor should have a high work function.

These considerations work exceptionally well for III–V nitrides such as GaN. We first consider ohmic contacts to n-type GaN: Given that the electron affinity of GaN is $\chi_{\text{GaN}} = 4.1 \text{ eV}$, and given that the work function of Al is $\Phi_{\text{Al}} \approx 4.1 \text{ eV}$, an Al-based contact is an excellent choice for an ohmic contact to n-type GaN (Luther *et al.*, 1997). Next, we consider ohmic contacts to p-type GaN: A metal should have a work function that is equal to $\chi_{\text{GaN}} + E_g = 4.1 \text{ eV} + 3.4 \text{ eV} = 7.5 \text{ eV}$. However, such high work function metals are not available. Pt and Pd have work functions in the range of 5.1 to 5.9 eV. Indeed, good Pt-based ohmic contacts to p-type GaN have been reported (Jang and Seong, 2000). It is generally believed that the Schottky model works well for GaN because GaN has a low surface state density (and a low electrical activity of dislocations) and a resulting unpinned surface Fermi level (Moustakas, 2013); this has been attributed to the more ionic character of the GaN chemical bond when compared to the more covalent GaAs chemical bond.

(ii) Ohmic contact with thin Schottky barrier enabled by high doping concentration

Due to the exceptionally high surface state density in many III–V semiconductors, particularly phosphides and arsenides, the surface Fermi level is pinned near the middle of the bandgap. As a result, the Schottky barrier height is $\Phi_B \approx \frac{1}{2} E_g$, i.e. independent of the metal work function. The useful strategy to overcome the barrier is the high doping of the near-surface-region of the semiconductor. The concept of a highly doped contact is shown in *Figure 7.13 (c)*. As the doping concentration increases, the barrier becomes thinner and quantum-mechanical tunneling becomes the dominant transport mechanism (Rideout, 1975).

This type of ohmic contact enables low-resistance contacts and is employed in III–V phosphides and arsenides. Common n-type dopants (contained in the metal layer of n-type contacts) include Si and Ge (e.g. AuGe-Ni-Au). Common p-type dopants (contained in the metal layer of p-type contacts) include Zn and Be (e.g. AuBe and AuZn).

7.9 Multilayer ohmic contacts

Multilayer metal contacts are motivated by the different requirements for a contact. The bottom metal layer, i.e. the layer that is in physical contact with the semiconductor, establishes the ohmic contact between the metal and the semiconductor. The top metal layer generally needs to be amenable to wire bonding or solder bonding. Au (gold) is frequently used as top layer due to its

inertness, resistance to corrosion and oxidation, and its good electrical conductivity.

Intermediate layers (between the bottom and top layer) may serve a diffusion barrier that are intended to prevent intermixing and interdiffusion of metals. Noble metals (such as Pt, Pd, and Re) as well as refractory metals (such as Ti, W, Ta, and Mo) are frequently used as barrier metals. For example, a multilayer ohmic contact to n-type GaN employing an Al bottom layer, and Au top layer, and two barrier layers (Ti and Pt) has been reported by Brown *et al.* (2003). Barrier layers may also completely enclose a metal layer, particularly a metal that has the propensity to diffuse and electro-migrate. Ag (silver) has such propensity. Accordingly, it needs to be confined by barrier layers (Slater *et al.*, 2007; 2009).

Metal layers that prevent a contact from “balling up” and from “islanding” may also be included in a metal layer stack. The phenomenon of “balling up” and from “islanding” describes the possibility that a metal contact loses its contiguousness (i.e. “balls up” and “islands”), which can happen during contact annealing.

Yet another requirement of contacts is the mechanical adhesion of the contact metal to the semiconductor. Some metals, such as Al, have lower adhesion than, e.g., Ti or Cr. As a result, metals with strong adhesive properties, such as Ti, may be included in a contact stack (Luther *et al.*, 1997)

Finally, ohmic contacts may have an optical functionality: Whereas the primary function is to serve as an ohmic contact, the secondary function may be to serve as an optical reflector. Ag and Al are particularly reflective metals with Al being highly reflective in the UV spectral range and Ag being highly reflective in the visible spectral range.

7.10 References

- Berger H. H. “Models for contacts to planar devices” *Solid-State Electronics* **15**, 145 (1972)
- Brown M. G., Eliashevich I., Ouyang K., and Venugopalan H. “Contact to n-GaN with Au termination” US patent 6,653,215 (2003)
- Cao X. A., Stokes E. B., Sandvik P. M., LeBoeuf S. F., Kretchmer J., and Walker D. “Diffusion and tunneling currents in GaN / InGaN multiple quantum well light-emitting diodes” *IEEE Electron Device Letters* **23**, 535 (2002)
- Casey H. C., Jr., Muth J., Krishnankutty S., and Zavada J. M. “Dominance of tunneling current and band filling in InGaN/AlGaN double heterostructure blue light-emitting diodes” *Applied Physics Letters* **68**, 2867 (1996)
- Chen N. C., Chang P. H., Chiu A. P., Wang M. C., Feng W. S., Wu G. M., Shih C. F., and K. S. Liu “Modified transmission line model and its application to aluminum ohmic contacts with n-type GaN” *Applied Physics Letters* **84**, 2584 (2004)
- Emerson D., Abare A., Bergmann M., Slater D., and Edmond J. “Development of deep UV III–N optical

- sources” *7th International Workshop on Wide-Bandgap III–Nitrides*, Richmond VA, March (2002)
- Jang J.-S. and Seong T.-Y. “Electronic transport mechanisms of non-alloyed Pt ohmic contacts to p-GaN” *Applied Physics Letters* **76**, 2743 (2000)
- Krames M. R. et al. “High-brightness AlGaN light-emitting diodes” *Proceedings of SPIE* **3938**, 2 (2000)
- Li T., Lambert D. J. H., Wong M. M., Collins C. J., Yang B., Beck A. L., Chowdhury U., Dupuis R. D., and Campbell J. C. “Low-noise back-illuminated Al_xGa_{1-x}N-based p-i-n solar-blind ultraviolet photodetectors” *IEEE Journal of Quantum Electronics* **37**, 538 (2001)
- Lorenz M. R. and Pilkuhn M. “Preparation and properties of solution-grown epitaxial p-n junctions in GaP” *Journal of Applied Physics* **37**, 4094 (1966)
- Luther B. P., Mohney S. E., Jackson T. N., Khan M. A., Chen Q., and Yang J. W. “Investigation of the mechanism for ohmic contact formation in Al and Ti/Al contacts to n-type GaN” *Applied Physics Letters* **70**, 57 (1997)
- Mayes K., Yasan A., McClintock R., Shiell D., Darvish S. R., Kung P., and Razeghi M. “High-power 280 nm AlGaN light-emitting diodes based on an asymmetric single-quantum well” *Applied Physics Letters* **84**, 1046 (2004)
- Narukawa Y., Sano M., Ichikawa M., Minato S., Sakamoto T., Yamada T., and Mukai T. “Improvement of luminous efficiency in white light emitting diodes by reducing a forward-bias voltage” *Japanese Journal of Applied Physics* **46**, L963 (2007)
- Perlin P., Osinski M., Eliseev P. G., Smagley V. A., Mu J., Banas M., and Sartori P. “Low-temperature study of current and electroluminescence in InGaN / AlGaN / GaN double-heterostructure blue light-emitting diodes” *Applied Physics Letters* **69**, 1680 (1996)
- Rhoderick E. H. and Williams R. H. *Metal–Semiconductor Contacts* (Clarendon Press, Oxford, UK, 1988)
- Rideout V. L. “A review of the theory and technology for ohmic contacts to group III–V compound semiconductors” *Solid-State Electronics* **18**, 541 (1975)
- Sah C.-T., Noyce R. N., and Shockley W. “Carrier generation and recombination in p-n junctions and p-n junction characteristics” *Proceedings of the IRE* **45**, 1228 (1957)
- Schottky W. “Halbleitertheorie der Sperrsicht” (“Semiconductor theory of the blocking layer”) *Naturwissenschaften* **26**, 843 (1938)
- Schottky W. “Zur Halbleitertheorie der Sperrsicht- und Spitzengleichrichter” (“On the semiconductor theory of blocking layer rectifiers and point rectifiers”) *Zeitschrift für Physik* **113**, 367 (1939)
- Schottky W. “Vereinfachte und erweiterte Theorie der Randschicht-gleichrichter” (“Simplified and expanded theory of the boundary-layer rectifiers”) *Zeitschrift für Physik A* **118**, 539 (1942)
- Schroder D. K. “Semiconductor material and device characterization” 3rd edition (Wiley-IEEE Press, New York, 2006)
- Shah J. M., Li Y.-L., Gessmann Th., and Schubert E. F. “Experimental analysis and theoretical model for anomalously high ideality factors ($n >> 2.0$) in AlGaN/GaN p-n junction diodes” *Journal of Applied Physics* **94**, 2627 (2003)
- Slater Jr. D. B., Williams B. E., Andrews P. S., Edmond J. A., Allen S. T. “Light emitting diodes including barrier layers / sublayers” US patent 7,211,833 (2007)
- Slater Jr. D. B., Williams B. E., Andrews P. S., Edmond J. A., Allen S. T. “Methods of manufacturing light emitting diodes including barrier layers / sublayers” US patent 7,611,915 (2009)
- Zhu D., Xu J., Noemaun A. N., Kim J. K., Schubert E. F., Crawford M. H., and Koleske D. D. “The origin of the high diode-ideality factors in GaInN/GaN multiple quantum well light-emitting diodes” *Applied Physics Letters* **94**, 081113 (2009)

8

8 – Carrier transport effects in the active region

In this chapter, we will go beyond the Shockley theory of p-n junctions. We will consider additional effects occurring in the active region that are not covered by the Shockley theory. These additional effects can arise from the use of heterojunctions.

8.1 Carrier distribution in p-n homojunctions

The carrier distribution in p-n homojunctions, i.e. p-n junctions consisting of a single material, depends on the *diffusion constant* of the carriers. The diffusion constant of carriers is not easily measured. Much more common is the measurement of the *carrier mobility*; for example, by the Hall effect. The diffusion constant can be inferred from the carrier mobility by the ***Einstein relation***, which, for non-degenerate semiconductors, is given by

$$D_n = \frac{kT}{e} \mu_n \quad \text{and} \quad D_p = \frac{kT}{e} \mu_p . \quad (8.1)$$

Carriers injected into a neutral semiconductor, with no external electric field applied, propagate by diffusion. If carriers are injected into a region with opposite conductivity type, the minority carriers will eventually recombine. The mean distance a minority carrier diffuses before recombination is the ***diffusion length***. Electrons injected into the p-type region will, on average, diffuse over the diffusion length L_n , before recombining with holes. The diffusion length is given by

$$L_n = \sqrt{D_n \tau_n} \quad \text{and} \quad L_p = \sqrt{D_p \tau_p} \quad (8.2)$$

where τ_n and τ_p is the electron and hole minority carrier lifetimes, respectively. In typical semiconductors, the diffusion length is of the order of a several micrometers. For example, the diffusion length of electrons in p-type GaAs is given by $L_n = (220 \text{ cm}^2/\text{s} \times 10^{-8} \text{ s})^{1/2} \approx 15 \mu\text{m}$. Thus, minority carriers are distributed over a region several micrometers thick.

The distribution of carriers in a p-n junction under zero bias and under forward bias is shown

in *Figures 8.1* (a) and (b), respectively. Note that minority carriers are distributed over quite a large distance. Furthermore, the minority carrier concentration decreases as these carriers diffuse further into the adjacent region. Thus, recombination occurs over a large region, with a strongly changing minority carrier concentration. As will be shown below, the *large recombination region in homojunctions* is not beneficial for efficient recombination.

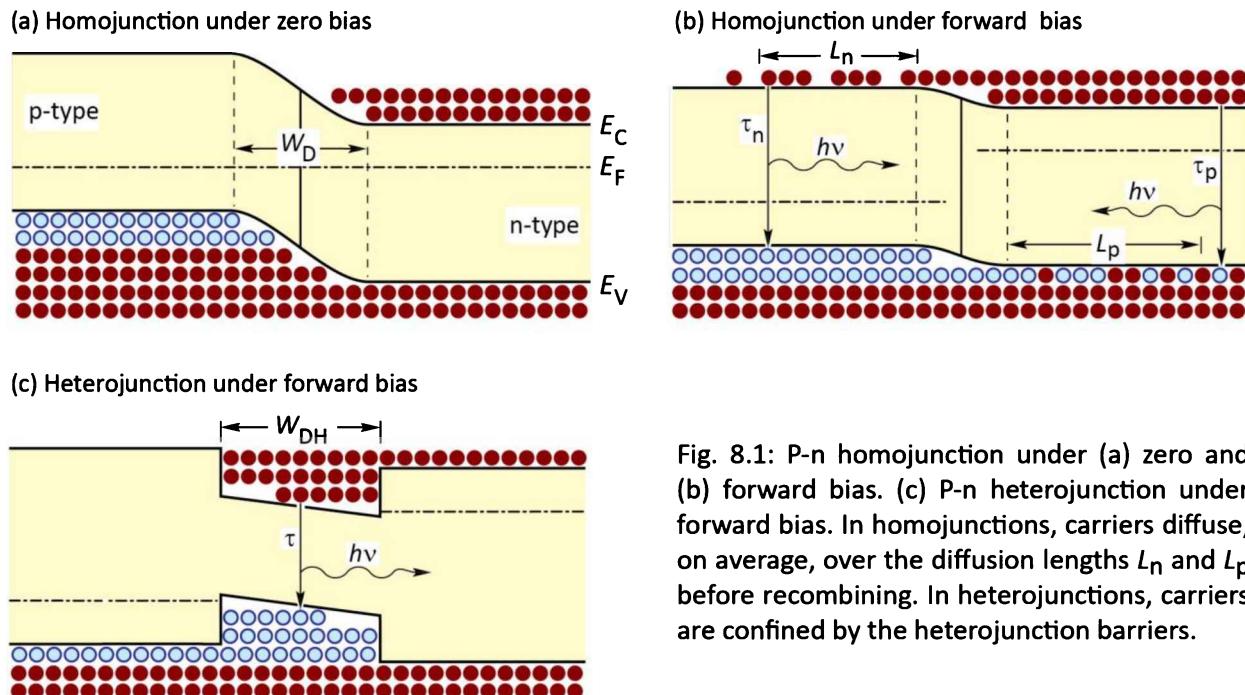


Fig. 8.1: P-n homojunction under (a) zero and (b) forward bias. (c) P-n heterojunction under forward bias. In homojunctions, carriers diffuse, on average, over the diffusion lengths L_n and L_p before recombining. In heterojunctions, carriers are confined by the heterojunction barriers.

8.2 Carrier distribution in p-n heterojunctions

All high-power light-emitting diodes do *not* use the homojunction design but rather employ the heterojunction design, which has several advantages over homojunction designs. Heterojunction devices employ two types of semiconductors, namely a small-bandgap active region and a large-bandgap barrier region. If a structure consists of two barriers, i.e. two large-bandgap semiconductors, then the structure is called a **double heterostructure** (frequently abbreviated as **DH**).

The effect of heterojunctions on the carrier distribution is shown in *Figure 8.1(c)*. Carriers injected into the active region of the double heterostructure are confined to the active region by means of the barriers. As a result, the thickness of the region in which carriers recombine *is given by the thickness of the active region rather than the diffusion length*. The consequences of this change are significant.

Firstly, the thickness of the active region is typically much smaller than the carrier-diffusion length. Diffusion lengths may range from about 1 to 20 μm . The active region of double heterojunctions may range from about 0.01 to 1.0 μm . Thus, carriers in the active region of a double heterostructure have a much higher concentration than carriers in homojunctions, which are distributed over several diffusion lengths. Recalling that the radiative recombination rate is given by the bimolecular recombination equation, i.e.,

$$R = B n p \quad (8.3)$$

it is clear that a high concentration of carriers in the active region *increases* the radiative recombination rate and decreases the recombination lifetime. For this reason, all high-efficiency LED designs employ double heterostructure or quantum well designs.

Secondly, by choosing a double-heterostructure design, the exposure of carriers to deep levels (luminescence killers) is reduced. The recombination of carriers through deep levels requires the presence of both, electrons and holes. Since the recombination volume is smaller for the double-heterostructure design, the absolute number of deep levels within the recombination volume is reduced, and, as a consequence, recombination through deep levels is reduced as well.

8.3 Effect of heterojunctions on device resistance

The employment of heterostructures allows one to improve the efficiency of LEDs by confining carriers to the active region, thereby avoiding diffusion of minority carriers over long distances. Heterostructures can also be used to confine light to waveguide regions; in particular, in edge-emitting LEDs. Generally, modern semiconductor LEDs and lasers have many heterojunctions, e.g. for contact layers, active regions, and waveguiding regions. Although heterostructures allow for improved LED designs, there are also problems associated with heterojunctions.

One of the problems introduced by heterostructures is the resistance caused by the heterointerface. The origin of the resistance is illustrated in **Figure 8.2 (a)**, which shows the band diagram of a heterostructure. The heterostructure consists of two semiconductors with different bandgap energy and it is assumed that both sides of the heterostructure are of n-type conductivity. Carriers in the large-bandgap material will diffuse over to the small-bandgap material where they occupy conduction band states of lower energy. As a result of the electron

transfer, an electrostatic dipole forms, consisting of a positively charged depletion layer with ionized donors in the large-bandgap material, and a negatively charged electron accumulation layer in the small-bandgap material. The charge transfer leads to the band bending illustrated in **Figure 8.2 (a)**. Carriers transferring from one semiconductor to the other must overcome this barrier by either tunneling or by thermal emission over the barrier. The resistance caused by heterojunctions can have a strong deleterious effect on device performance, especially in high-power devices. The thermal power produced by heterostructure resistances leads to heating of the active region, thereby decreasing the radiative efficiency.

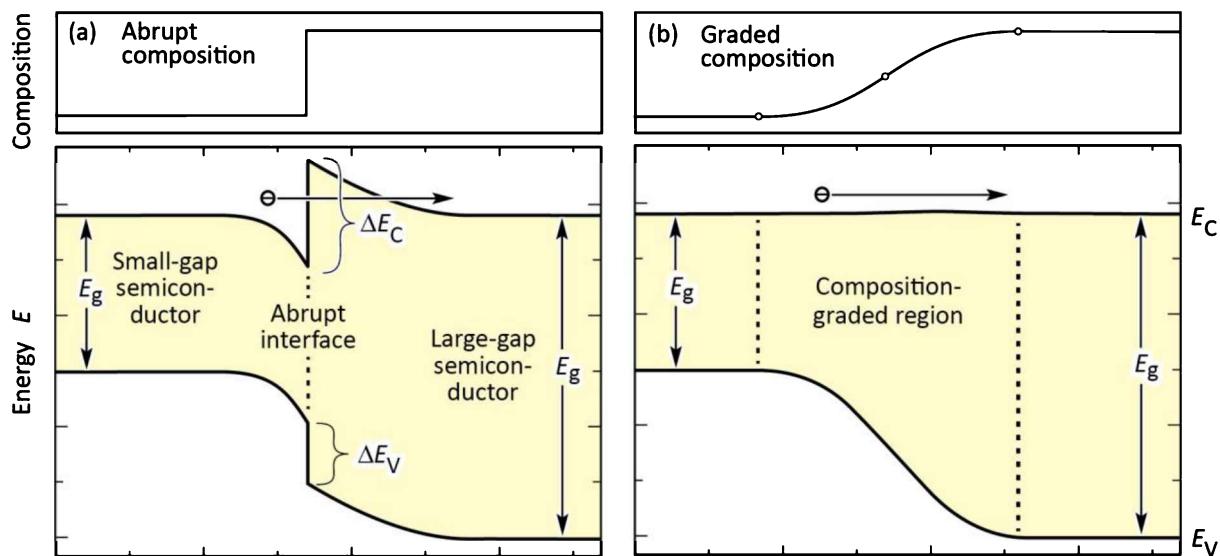


Fig. 8.2: Band diagram of (a) an abrupt n-type–n-type heterojunction and (b) a graded heterojunction of two semiconductors with different bandgap energy. The abrupt junction is more resistive than the graded junction due to the electron barrier forming at the abrupt junctions (after Schubert *et al.*, 1992).

It has been shown that heterostructure band discontinuities can be completely eliminated by **grading** of the chemical composition of the semiconductor in the vicinity of the heterostructure (Schubert *et al.*, 1992). The band diagram of a graded heterostructure is shown in **Figure 8.2 (b)**. Inspection of the figure reveals that there is no longer a spike in the conduction band which hinders the electron flow. It has been shown that the resistance of parabolically graded heterostructures is comparable to bulk material resistance. Thus, the additional resistance introduced by abrupt heterostructures can be completely eliminated by parabolic grading.

The shape of the graded region should be parabolic for the following reason. The large-bandgap material will be depleted of free carriers due to electron transfer to the small-bandgap

material. Thus, the charge concentration in the large-bandgap material will be the donor concentration. Assuming that the donor concentration N_D is a constant throughout the heterostructure, the solution of the Poisson equation yields the electrostatic potential

$$\Phi = \frac{e N_D}{2 \epsilon} x^2 . \quad (8.4)$$

The equation reveals that the potential depends *quadratically* on the spatial coordinate x , i.e. the potential has a *parabolic* shape. In order to compensate for the parabolic shape of the depletion potential, the composition of the semiconductor *is varied parabolically as well*, so that an overall *flat potential* results. It is assumed here that the parabolic variation of the chemical composition results in a parabolic change of the bandgap energy, i.e. that the bandgap energy depends linearly on the chemical composition and that bandgap *bowing* can be neglected.

Next, an approximate design rule for the grading of a heterostructure is given. Assume that the conduction band discontinuity of an abrupt heterojunction is given by ΔE_C and that the structure is uniformly doped with doping concentration N_D . Let us assume that carriers have transferred to the small-bandgap semiconductor, thus causing a depletion region of thickness W_D in the large-bandgap semiconductor. If the potential created in the depletion region is equal to $\Delta E_C/e$, then electrons will no longer transfer to the small-bandgap material. The thickness of the depletion region can be inferred from Eqn. (8.4) to be

$$W_D = \sqrt{\frac{2 \epsilon \Delta E_C}{e^2 N_D}} . \quad (8.5)$$

A heterostructure interface should be graded over the distance W_D in order to minimize the resistance introduced by an abrupt heterostructure. Although the result of Eqn. (8.5) is an approximation it provides excellent guidance for device design. Steps can be taken to refine the calculation. For example, the potential change due to the electron accumulation layer in the small-bandgap material can be taken into account. Several software packages are available that allow for the numerical calculation of semiconductor heterostructures, for example the software package *Atlas* by the Silvaco Corporation.

Exercise: Compositional grading of heterostructures. Assume that the conduction band discontinuity of an AlGaAs/GaAs heterostructure is given by $\Delta E_C = 300$ meV and that the structure is uniformly doped with donors of concentration $N_D = 5 \times 10^{17}$ cm⁻³. Over what distance should the interface

be graded in order to minimize the resistance occurring in abrupt heterostructures?

Solution: Calculating the depletion layer thickness from Eqn. (8.5) yields $W_D = 30$ nm. Thus, the Al composition in the heterostructure should be graded over 30 nm to minimize the heterostructure resistance. The composition graded region should have two parabolic regions as shown in **Figure 8.2 (b)**.

Grading is useful for all heterostructures, including the heterostructures adjoining the active region. The effect of grading in a double heterostructure is shown in **Figure 8.3**. The composition and the band diagram of an ungraded structure are shown in **Figure 8.3(a)**. At both heterointerfaces, barriers develop as a result of free charge transferring to the active region. These barriers increase the device resistance under forward-bias conditions.

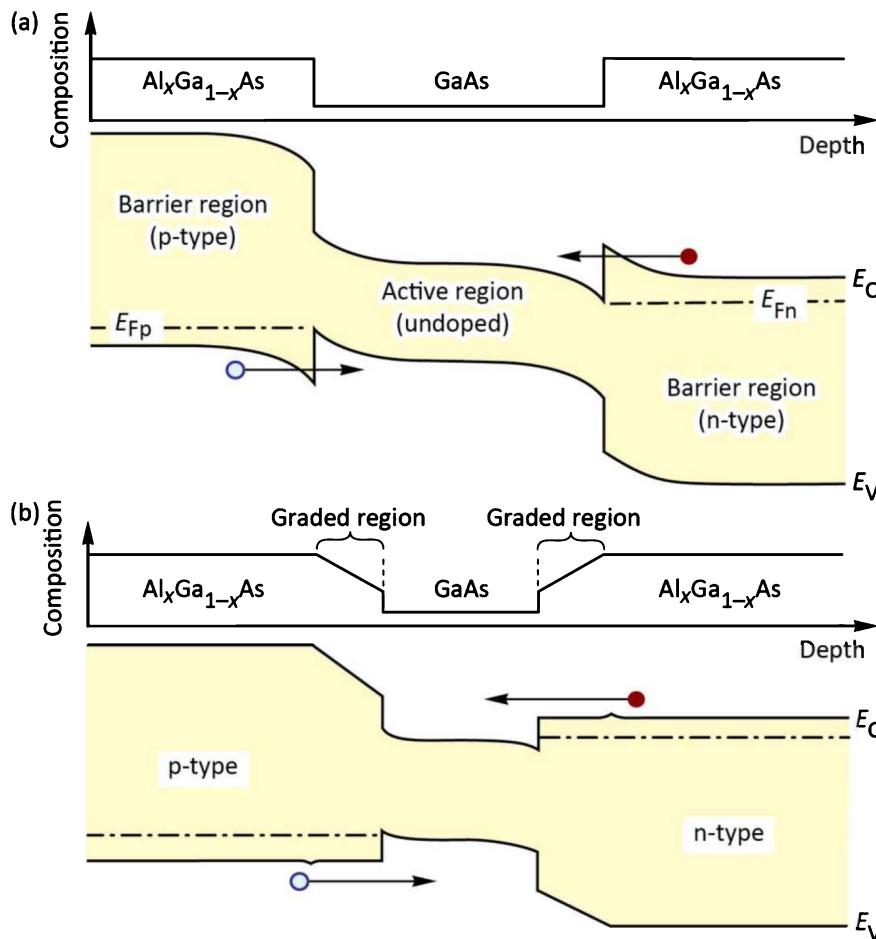


Fig. 8.3: Band diagram of (a) an abrupt double heterostructure and (b) a graded double heterostructure. The barrier–well interface of the abrupt junction is more resistive than for the graded junction due to barriers forming at the interfaces.

The case of graded heterointerfaces is shown in **Figure 8.3 (b)**. The figure shows two linearly graded regions cladding the active region. The band diagram illustrates that barriers at the heterointerfaces can be effectively reduced or completely eliminated by grading. Note that the

linear grading shown in **Figure 8.3(b)** results in small “spikes” at the interfaces between the linearly graded and the non-graded regions. The “spikes” are a result of the linear grading and the spikes would not occur for parabolic grading.

Generally, the transport of carriers in heterostructures should be as **adiabatic** as possible, i.e. the carrier transport within the semiconductor device should not generate unnecessary heat. This is particularly true for high-power devices where additional heat generated inside the device leads to a performance loss due to increased operating temperature.

Finally, note that *lattice matching* is desirable in all heterostructure devices. It is also desirable in graded structures in order to minimize the number of misfit dislocations that act as non-radiative recombination centers.

8.4 Carrier loss in double heterostructures

In an ideal LED, injected carriers are confined to the active region by the barrier layers adjoining the active region. By confining the carriers to the active region, a high carrier concentration is attained resulting in a high radiative efficiency of the recombination process.

The energy barriers confining the carriers to the active region are typically of the order of several hundred meV, i.e. much larger than kT . Nevertheless, some carriers will succeed in escaping from the active region into the barrier layers. The concentration of the escaping carriers in the barrier layers will be rather low, resulting in a low radiative efficiency of carriers in the barrier layers.

Free carriers in the active region are distributed according to the Fermi–Dirac distribution and, as a result, some carriers have a higher energy than the height of the confining barrier. Thus, some of the carriers escape from the active region into the barrier regions as illustrated in **Figure 8.4**.

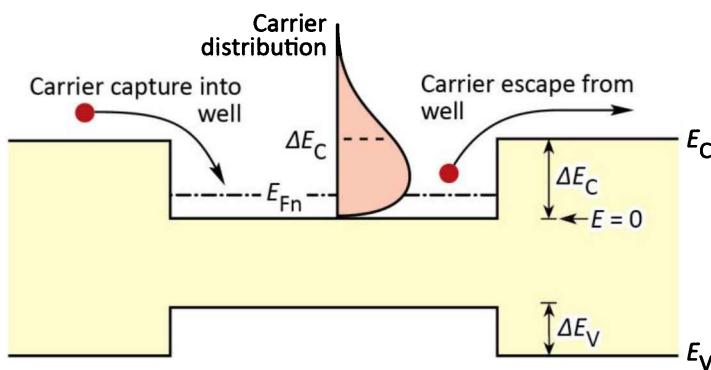


Fig. 8.4: Carrier capture and escape in a double hetero-structure. Also shown is the carrier distribution in the active layer.

Consider electrons in the active region of a double heterostructure and assume that the active region is clad by a barrier with height ΔE_C , as shown in **Figure 8.4**. The energy distribution of carriers is given by the Fermi–Dirac distribution. Thus, a certain fraction of the carriers residing in the active region have an energy that is higher than the energy of the barrier. The concentration of electrons with energy higher than the barrier is given by

$$n_B = \int_{E_B}^{\infty} \rho_{DOS}(E) f_{FD}(E) dE \quad (8.6)$$

where ρ_{DOS} is the density of states, f_{FD} is the Fermi–Dirac distribution function, and E_B is the height of the barrier. For a bulk-type density of states, the concentration of carriers with energy higher than E_B is given by

$$n_B = \frac{1}{2\pi^2} \left(\frac{2m^*}{\hbar^2} \right)^{3/2} \int_{E_B}^{\infty} \frac{\sqrt{E - E_C}}{1 + e^{(E - E_{Fn})/kT}} dE . \quad (8.7)$$

Taking into account that we are interested in carriers with energies much higher than the Fermi energy, the Fermi–Dirac distribution can be approximated by the Boltzmann distribution. One obtains

$$n_B = N_c e^{(E_{Fn} - E_B)/kT} \quad (8.8)$$

where N_c is the effective density of states in the active region. Equation (8.25) gives the concentration of free carriers at the active-region–cladding-region interface. Minority carriers at the edge of the cladding layer will diffuse into the cladding layer. The diffusion process is governed by the initial concentration n_B and the electron diffusion length L_n . Taking the location of the origin ($x = 0$) at the edge of the barrier, the carrier distribution can be written as

$$n_B(x) = n_B(0) e^{-x/L_n} = N_c e^{-(E_B - E_{Fn})/kT} e^{-x/L_n} \quad (8.9)$$

where $L_n = (D_n \tau_n)^{1/2}$ is the diffusion length, τ_n is the minority carrier lifetime, and D_n is the diffusion constant. The diffusion constant can be inferred from the mobility using the Einstein relation $D = \mu kT/e$.

The diffusion current density of electrons leaking over the barrier can be obtained from the carrier concentration gradient at $x = 0$, i.e.

$$J_n|_{x=0} = -e D_n \frac{dn_B(x)}{dx} \Big|_{x=0} = -e D_n \frac{n_B(0)}{L_n} . \quad (8.10)$$

The leakage current depends on the carrier concentration at the edge of the barrier. Thus, a high barrier height is required to minimize the leakage current. Clearly, barriers must be much larger than kT for efficient confinement of carriers. Some material systems such as AlGaN/GaN or AlGaAs/GaAs have relatively high barriers and therefore lower leakage currents over the barrier. Other material systems such as AlGalnP/AlGalnP emitting at 600~650 nm have lower barriers and therefore have a stronger carrier leakage over the barriers.

Note that the leakage increases exponentially with temperature. Thus, a decrease of the radiative efficiency of LEDs results as the temperature increases. To reduce the temperature dependence of the emission, high barriers are required. In addition to carrier leakage, other effects, such as Shockley–Read recombination also contribute to the lower radiative efficiency at high temperatures.

Exercise: Carrier leakage over a barrier. Electrons in the active region of a GaAs structure have a concentration of $2 \times 10^{18} \text{ cm}^{-3}$. Calculate the current density of the carrier loss over the barrier for barrier heights of 200 and 300 meV, assuming an electron mobility of $2000 \text{ cm}^2/(\text{V s})$ and a minority carrier lifetime of 5 ns. Compare the calculated leakage current to LED injection currents of $0.1\text{--}1.0 \text{ kA/cm}^2$.

Solution: The Fermi level in GaAs with electron density of $2 \times 10^{18} \text{ cm}^{-3}$ is 77 meV above the conduction band edge. Assuming that the effective density of states in the barrier is the same as in the GaAs active region, the carrier concentrations at the edge of the barrier are $3.9 \times 10^{15} \text{ cm}^{-3}$ for a 200 meV barrier and $8.3 \times 10^{13} \text{ cm}^{-3}$ for a 300 meV barrier. The diffusion constant, as inferred from the Einstein relation, is $D_n = 51.7 \text{ cm}^2/\text{s}$. The diffusion length is then given by $L_n = (D_n \tau_n)^{1/2} = 5.1 \mu\text{m}$. The leakage current is calculated using Eqn. (8.10), and one obtains 63 A/cm² for the 200 meV barrier and 1.3 A/cm² for the 300 meV barrier. Comparison with diode current densities of $0.1\text{--}1.0 \text{ kA/cm}^2$ suggests that leakage currents can be a significant loss mechanism, particularly for small barrier heights.

In the consideration above, we have assumed that electrons diffuse in the p-type region and have neglected any drift. However, if the p-type region has a substantial resistance, electron drift cannot be neglected. This drift will enhance the electron current. Furthermore, electrical contacts have been neglected. The minority carrier concentration at the contact–semiconductor interface

can be assumed to be zero due to the high surface recombination velocity of such interfaces. Taking these effects into account, the leakage current was calculated by Ebeling (1993). If the distance of the contact from the active–barrier interface is denoted by x_p , the leakage current is given by

$$J_n = -eD_n n_B(0) \left[\sqrt{\frac{1}{L_n^2} + \frac{1}{L_{nf}^2}} \coth\left(\sqrt{\frac{1}{L_n^2} + \frac{1}{L_{nf}^2}} x_p\right) + \frac{1}{L_{nf}} \right] \quad (8.11)$$

where

$$L_{nf} = \frac{kT}{e} \frac{\sigma_p}{J_{tot}}, \quad (8.12)$$

σ_p is the conductivity of the p-type cladding region, and J_{tot} is the total diode current density.

8.5 Carrier overflow in double heterostructures

The **overflow of carriers** from the active region into the confinement regions is another loss mechanism. Carrier overflow occurs at high injection current densities. As the injection current increases, the carrier concentration in the active region increases and the Fermi energy rises. For sufficiently high current densities, the Fermi energy will rise to the top of the barrier. The active region is flooded with carriers and a further increase in injection current density will *not* increase the carrier concentration in the active region. As a result, the optical intensity saturates. At high injection current densities, carrier overflow occurs, even if the barriers are sufficiently high, so that carrier leakage over the barriers at low injection current densities can be neglected.

Consider a double heterostructure LED with an active region of thickness W_{DH} , as shown in **Figure 8.5**. The rate equation of carrier supply to (by injection) and removal from (by recombination) the active region is given by

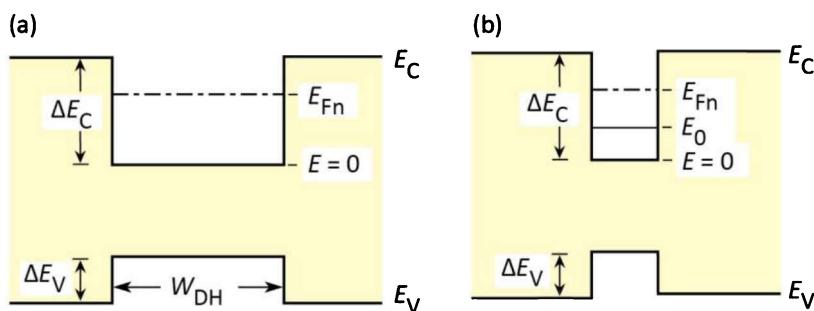


Fig. 8.5: Fermi level (E_{Fn}) and subband level (E_0) in (a) a double heterostructure and (b) a quantum well structure.

$$\frac{dn}{dt} = \frac{J}{e W_{DH}} - B n p \quad (8.13)$$

where B is the bimolecular recombination coefficient. For high injection densities, it is $n = p$. Solving Eqn. (8.13) for n under steady-state conditions ($dn/dt = 0$), yields

$$n = \sqrt{\frac{J}{e B W_{DH}}} . \quad (8.14)$$

The carrier density increases with the current injected into the device. As a result, the Fermi energy rises. In the high-density approximation, the Fermi energy is given by

$$\frac{E_F - E_C}{kT} = \left(\frac{3\sqrt{\pi}}{4} \frac{n}{N_c} \right)^{2/3} . \quad (8.15)$$

At high injection levels, the Fermi energy rises and will eventually reach the top of the barrier. At that point, it is $E_F - E_C = \Delta E_C$. Using this value, the current density at which the active region overflows can be calculated from Eqns. (8.14) and (8.15) and one obtains

$$J_{overflow} = \left(\frac{4 N_c}{3\sqrt{\pi}} \right)^2 \left(\frac{\Delta E_C}{kT} \right)^3 e B W_{DH} . \quad (8.16)$$

Either the conduction band or the valence band well may overflow first depending on the effective density of states (N_c, N_v) and the band discontinuities ($\Delta E_C, \Delta E_V$).

Exercise: *Carrier overflow in a double heterostructure.* Consider electrons in a GaAs double heterostructure with a barrier height of $\Delta E_C = 200$ meV and an active region thickness of $W_{DH} = 500$ Å. Calculate the electron-overflow current density.

Solution: Using $N_c = 4.4 \times 10^{17}$ cm⁻³ and $B = 10^{-10}$ cm³/s, one obtains from Eqn. (8.16) a current density of $J_{overflow} = 3990$ A/cm².

Generally, the problem of carrier overflow is more severe in structures with a small active-region volume. In particular, single-quantum-well structures and quantum-dot active regions have an inherently small volume. At a certain current density, the active region is filled with carriers, and the injection of additional carriers will not lead to an increase in the emitted light intensity.

Experimental results of an LED structure with one, four, six, and eight quantum wells (QWs) are shown in *Figure 8.6* (Hunt *et al.*, 1992). The light intensity for the single QW structure saturates at a low current level. As the number of quantum wells is increased, the current level at which saturation occurs increases, and the optical saturation intensity increases as well. The saturation of the light intensity displayed in *Figure 8.6* is caused by the overflow of carriers.

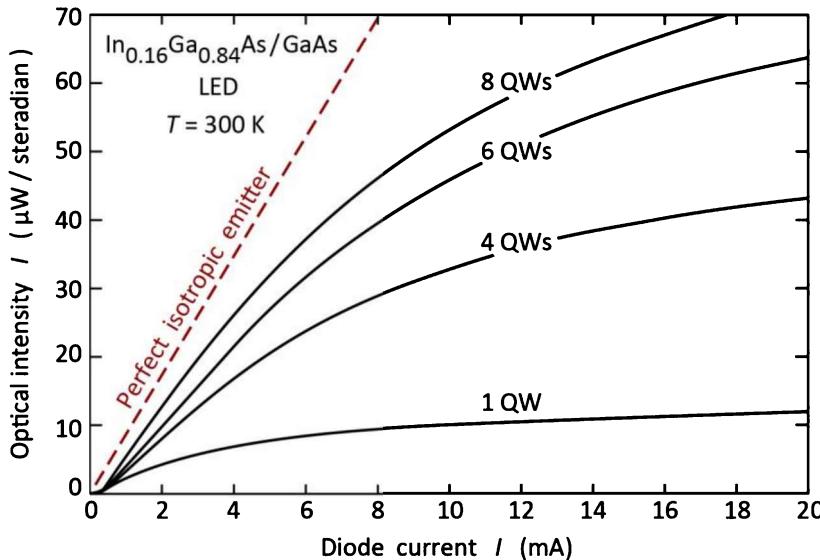


Fig. 8.6: Optical intensity emitted by $\text{In}_{0.16}\text{Ga}_{0.84}\text{As} / \text{GaAs}$ LEDs with active regions consisting of one, four, six, and eight quantum wells and theoretical intensity of a perfect isotropic emitter (dashed line) (after Hunt *et al.*, 1992).

The calculation of the overflow current level is different for quantum well structures and bulk active regions. For quantum well structures, we must employ the two-dimensional (2D) density of states, rather than the 3D density of states that was used in the above calculation. The Fermi level in a QW with one quantized state with energy E_0 is given by

$$\frac{E_F - E_0}{kT} = \ln \left[\exp \left(\frac{n^{2D}}{N_c^{2D}} \right) - 1 \right] \quad (8.17)$$

where n^{2D} is the 2D carrier density per cm^2 and N_c^{2D} is the effective 2D density of states given by

$$N_c^{2D} = \frac{m^*}{\pi \hbar^2} kT . \quad (8.18)$$

Because we are dealing with high carrier densities, the high-degeneracy approximation can be employed and one obtains

$$E_F - E_0 = \frac{\pi \hbar^2}{m^*} n^{2D} . \quad (8.19)$$

Next, we write the rate equation for the quantum well. The rate equation of carrier supply to (by injection) and removal from (by recombination) the active region is given by

$$\frac{dn^{2D}}{dt} = \frac{J}{e} - B^{2D} n^{2D} p^{2D} \quad (8.20)$$

where $B^{2D} \approx B/W_{QW}$ is the bimolecular recombination coefficient for a 2D structure. For high injection densities, it is $n^{2D} = p^{2D}$. Solving Eqn. (8.20) for n^{2D} under steady-state conditions ($dn^{2D}/dt = 0$), yields

$$n^{2D} = \sqrt{\frac{J}{e B^{2D}}} = \sqrt{\frac{J W_{QW}}{e B}} . \quad (8.21)$$

At high injection levels, the Fermi energy will reach the top of the barrier. At that point, $E_F - E_0 = \Delta E_C - E_0$. The use of this value in Eqn. (8.19) and subsequent elimination of n^{2D} from Eqns. (8.19) and (8.21) yields the current density at which the active region overflows

$$J_{overflow} = \left[\frac{m^*}{\pi \hbar^2} (\Delta E_C - E_0) \right]^2 \frac{e B}{W_{QW}} . \quad (8.22)$$

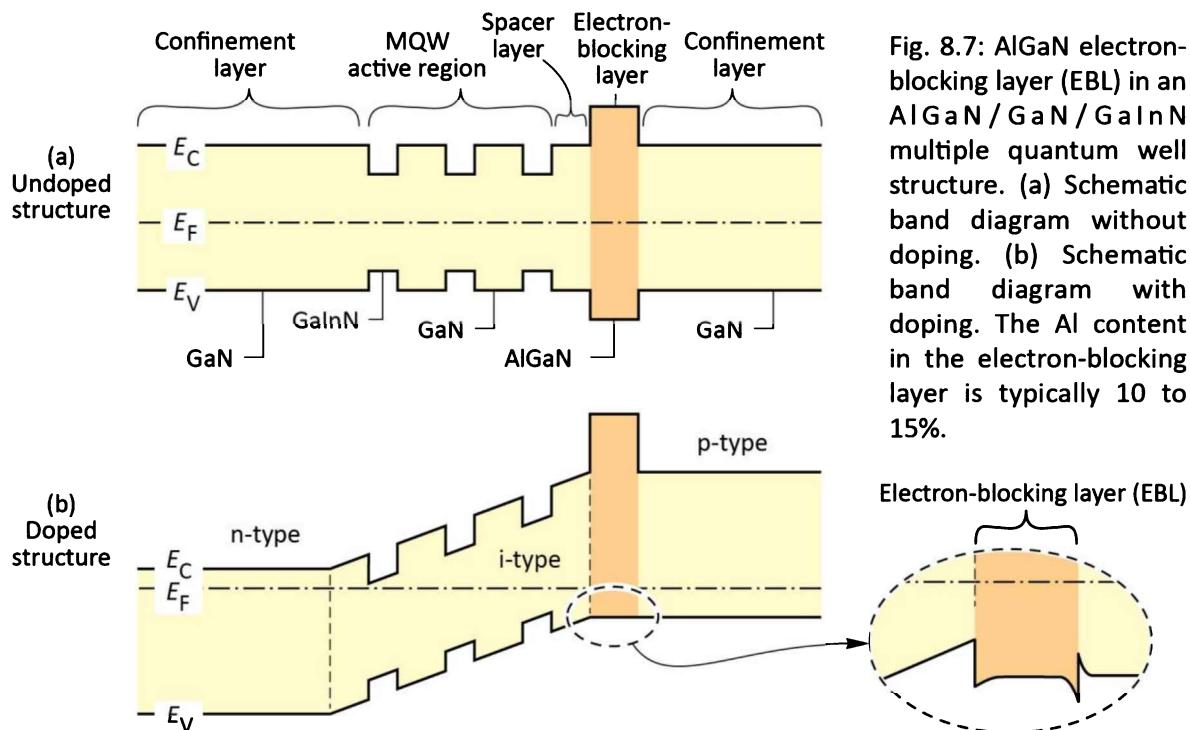
Thus, overflow of the active region is a potential problem in double heterostructures as well as quantum well structures. In order to avoid this problem, high-current LEDs must employ *thick* double heterostructure active regions, or *many* QWs of multiple QW (MQW) active regions, or a large injection (contact) area. By choosing these parameters, the volume of the active region can be designed in such a way that at the intended current density of operation, carrier overflow does not occur.

8.6 Electron-blocking layers

Carriers tend to escape from the active layer of an LED into the confinement layers. The carrier escape can be substantial in double heterostructures with a low barrier height at the active–confinement interface. In addition, high temperatures promote carrier loss out of the active region due to the increase in carrier thermal energy.

The electron leakage current is larger than the hole leakage current due to the usually larger diffusion constant of electrons compared with holes in III–V semiconductors. To reduce carrier leakage out of the active region, carrier-blocking layers are used. In particular, **electron-blocking layers** (EBLs) or **electron blockers** are used in GaN-based LED structures to reduce electron escape out of the active region. Such electron-blocking layers are regions with a high bandgap energy located at the confinement–active interface.

The band diagram of a GaInN LED with an electron-blocking layer is shown in **Figure 8.7**. The LED has GaN confinement layers and a GaInN/GaN multiple quantum well (MQW) active region. An AlGaN electron-blocking layer is included in the p-type confinement layer at the confinement–active interface. **Figure 8.7 (a)**, showing the undoped structure, illustrates that the AlGaN electron-blocking layer creates a barrier to current flow in both the conduction band as well as the valence band. The figure also shows a “spacer layer”. The spacer layer is a thin undoped GaN layer that separates the electron-blocking layer from the last-grown QW. The spacer layer has the purpose of preventing p-type dopants from diffusing into the MQW active region.



However, **Figure 8.7 (b)**, showing the doped structure, illustrates that the barrier in the valence band is screened by free carriers so that there is *no barrier* to the flow of *holes* in the p-type confinement layer. That is, the entire band discontinuity is located in the conduction band, i.e.

$$\text{Barrier height for electrons} = E_{C,\text{confinement}} - E_{C,\text{active}} + \Delta E_g \quad (8.23)$$

where ΔE_g is the difference in bandgap energy between the confinement and the electron-blocking layer.

The inset of **Figure 8.7** shows the valence band edge of the electron-blocking layer in greater detail. A potential spike (hole depletion layer in the electron blocker) and notch (hole accumulation layer in the p-type confinement layer) occur at the confinement–blocking layer interface. Holes must tunnel through the potential spike when propagating towards the active region. Note that the valence band edge can be completely smoothed out by compositional grading at the confinement–blocking layer interface so that the electron blocking layer does not impede the hole flow at all.

8.7 Effect of heterojunctions on diode voltage

The energy of an injected electron is converted into optical energy upon electron–hole recombination. Thus, conservation of energy requires that the **drive voltage** or **forward voltage** of a light-emitting device is equal to (or larger than) the bandgap energy divided by the elementary charge. The diode voltage is thus given by

$$V = h\nu/e \approx E_g/e . \quad (8.24)$$

There are several mechanisms causing the drive voltage to be slightly different from this value and these mechanisms will be discussed below.

Firstly, if the diode has a significant series resistance, an additional voltage drop occurs. The additional resistance can be caused by (i) *contact resistance*, (ii) *resistances caused by abrupt heterostructures*, and (iii) *bulk resistance* occurring particularly in materials with low carrier concentrations or low carrier mobilities. A voltage drop of magnitude $I R_s$ occurs at the series resistance thereby increasing the drive voltage.

Secondly, carrier energy may be lost upon injection into a quantum well structure or double heterostructure. An example of non-adiabatic injection is shown in **Figure 8.8**, which shows a thin quantum well under forward-bias conditions. The figure illustrates that upon injection into the quantum well, the electron loses energy $\Delta E_C - E_0$, where ΔE_C is the band discontinuity and E_0 is the energy of the lowest quantized state in the conduction-band quantum well. Similarly, the

energy lost by holes is given by $\Delta E_V - E_0$, where ΔE_V is the band discontinuity and E_0 is the energy of the lowest state in the valence-band quantum well. Upon injection of carriers into the well, the carrier energy is dissipated by **phonon emission**, i.e. by conversion of the carrier energy to heat. The energy loss due to **non-adiabatic injection** of carriers is relevant in semiconductors with large band discontinuities, for example GaN and other group-III nitride materials.

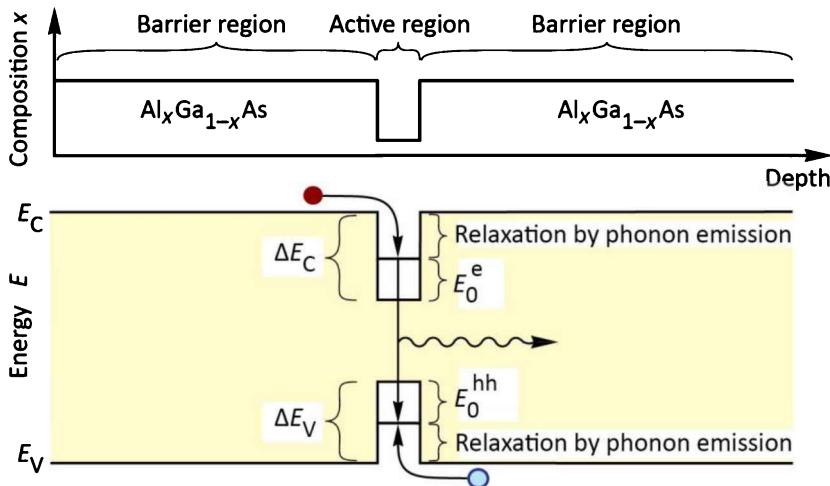


Fig. 8.8: Chemical composition and energy band diagram of a quantum well structure, illustrating the energy loss of carriers as they are captured into the quantum well.

Thus, the total voltage drop across a forward-biased LED is given by

$$V = \frac{E_g}{e} + IR_S + \frac{\Delta E_C - E_0}{e} + \frac{\Delta E_V - E_0}{e} \quad (8.25)$$

where the first summand on the right-hand side of the equation is the theoretical voltage minimum, the second summand is due to the series resistance in the device, and the third and fourth summands are due to non-adiabatic injection of carriers into the active region.

One finds experimentally that the diode voltage can be *slightly lower than the minimum value* predicted by Eqn. (8.25), i.e. can be slightly lower than $E_g/e \approx h\nu/e$. Both electrons and holes carry, on average, the thermal energy kT . In a forward-biased p-n junction, high-energy carriers are more likely than low-energy carriers to diffuse over to the side of opposite conductivity type where they recombine. At room temperature, $4kT/e$ amounts to a voltage of about 100 mV. In low-resistance devices, the diode voltage can be 100~200 mV *lower* than $h\nu/e$. For example, in forward-biased GaAs LEDs ($E_g = 1.42$ eV), some photon emission with $h\nu = 1.42$ eV is observed at diode voltages of about 1.32 V, i.e. lower than the photon energy.

8.8 References

- Ebeling K. J. *Integrated Opto-Electronics* Chapter 9 (Springer, Berlin, 1993)
- Hunt N. E. J., Schubert E. F., Sivco D. L., Cho A. Y., and Zydzik G. J. "Power and efficiency limits in single-mirror light-emitting diodes with enhanced intensity" *Electronics Letters* **28**, 2169 (1992)
- Schubert E. F., Tu L.-W., Zydzik G. J., Kopf R. F., Benvenuti A., and Pinto M. R. "Elimination of heterojunction band discontinuities by modulation doping" *Applied Physics Letters* **60**, 466 (1992)

9

9 – Design of current flow

LEDs may be grown on conductive and insulating substrates. Whereas the current flow is mostly vertical (normal to the substrate plane) in structures grown on conductive substrates, it is mostly lateral (horizontal) in devices grown on insulating substrates. The location and size of ohmic contacts are relevant to light extraction, because metal contacts are opaque. The current chapter discusses the current flow patterns of different device structures aimed at high extraction efficiency.

9.1 Current-spreading layer

In LEDs with thin top confinement layers, the current is injected into the active region mostly under the top electrode. Thus, light is generated under an opaque metal electrode. This results in a low extraction efficiency. This problem can be avoided with a ***current-spreading layer*** that spreads the current under the top electrode to regions not covered by the opaque top electrode.

The ***current-spreading layer*** is synonymous with the ***window layer***. The term window layer is occasionally used to emphasize the *transparent character* of this layer and its ability to enhance the extraction efficiency. It is understood that a current-spreading layer should be optically transparent so as to not absorb the light emitted by the active region.

The usefulness of current-spreading layers was realized during the infancy of LEDs. Nuese *et al.* (1969) demonstrated a substantial improvement of the optical output power in GaAsP LEDs by employing a current-spreading or window layer. The window layer is the top semiconductor layer located between the upper cladding layer and the top ohmic contact. The effect of a current-spreading layer is illustrated in ***Figure 9.1***. Light is emitted only around the perimeter of the top contact for LEDs without a current-spreading layer, as shown in ***Figure 9.1 (a)***. The addition of a current-spreading layer results in more uniform and brighter surface emission as shown in ***Figure 9.1 (b)***.

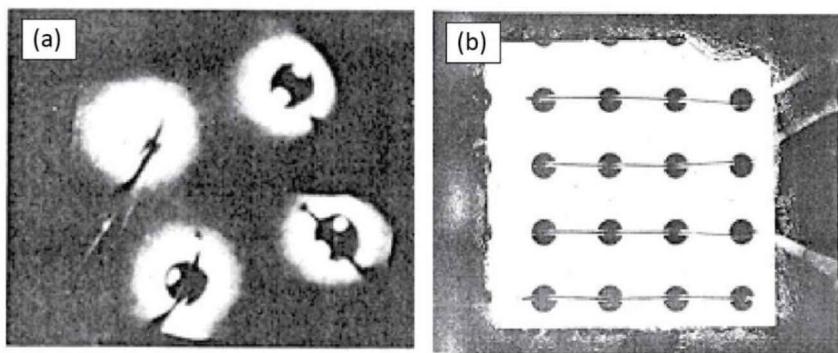


Fig. 9.1: Effect of the current-spreading layer on LED output. (a) Top view without a current-spreading layer. Emission occurs only near the perimeter of the contact. (b) Top view with a current-spreading layer (after Nuese *et al.*, 1969).

Nuese *et al.* (1969) demonstrated current-spreading layers composed of the ternary GaAsP and the binary GaP and discussed the requirements of the current-spreading layer. These requirements include low resistivity and large thickness for current spreading, and transparency to minimize absorption losses. To reduce absorption losses, Nuese *et al.* (1969) employed a high P mole fraction in the $\text{GaAs}_{1-x}\text{P}_x$ current-spreading layer, namely $0.45 < x \leq 1.0$, higher than the P mole fraction in the $\text{GaAs}_{1-x}\text{P}_x$ active region, where $x = 0.45$. Thus, the bandgap energy of the current-spreading layer is higher than the bandgap of the active region.

Although Nuese *et al.* discussed the properties of current-spreading layers *qualitatively*, the authors did not provide a *quantitative* theoretical framework of current spreading. The theoretical foundation of current-spreading layers in devices with linear contact geometry (i.e. a strip-shaped contact) was given by Thompson, as discussed below. The use of a current-spreading layer was adopted in most top-emitting LED designs, including AlGaAs LEDs (Nishizawa *et al.*, 1983; Moyer 1988), GaP LEDs (Groves *et al.*, 1977, 1978a, 1978b), and AlGaN_xP LEDs (Kuo *et al.*, 1990; Sugawara *et al.*, 1991, 1992a, 1992b).

The effect of the current-spreading layer is schematically illustrated in **Figure 9.2**. Without a current-spreading layer, the current-injected area of the active region is limited to approximately the contact size, as indicated in **Figure 9.2 (a)**. As a consequence, the light is emitted mostly below the opaque top contact which in turn results in poor light extraction. The addition of the current-spreading layer results in a larger current-injected area, as shown in **Figure 9.2 (b)**, as well as good light extraction.

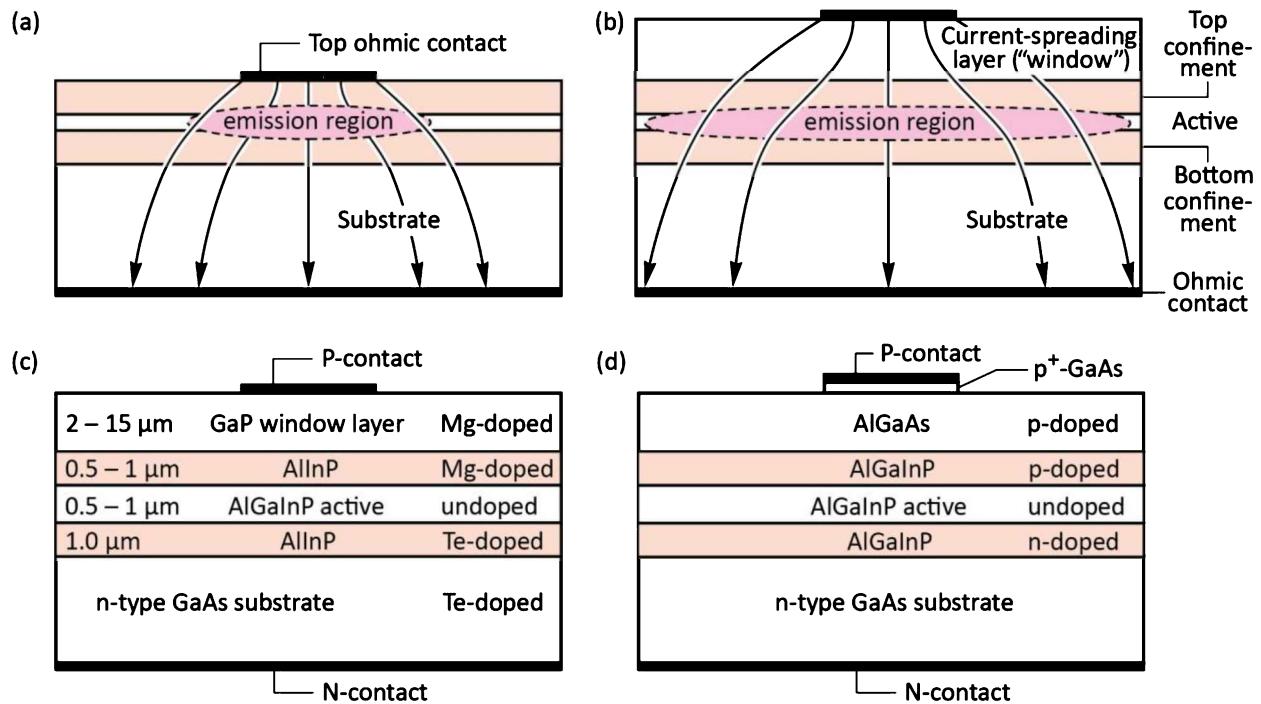


Fig. 9.2: Current-spreading structures in high-brightness AlGalnP LEDs. Illustration of the effect of a current-spreading layer for LEDs (a) without and (b) with a spreading layer on the light extraction efficiency. (c) GaP current-spreading structure (Fletcher *et al.*, 1991a, 1991b). (d) AlGaAs current-spreading structure (Sugawara *et al.*, 1992a, 1992b).

Current-spreading layers are predominantly employed in top-emitting LEDs. Two different approaches for AlGalnP visible LEDs, grown on GaAs substrates, are shown in *Figures 9.2 (c)* and (d). A GaP current-spreading layer was reported by Kuo *et al.* (1990) and Fletcher *et al.*, (1991a, 1991b). GaP has a bandgap of $E_{g,GaP} = 2.26$ eV and is thus transparent for red, orange, yellow, and part of the green spectrum. AlGalnP LEDs with emission wavelengths as short as 550 nm have been fabricated. GaP, as a binary compound semiconductor, is very transparent for energies below the bandgap, i.e. the Urbach tail energy of GaP is small. Furthermore, GaP is an indirect-gap semiconductor, which is inherently less absorbing compared with direct-gap semiconductors. Thus, little light is absorbed even in thick GaP current-spreading layers.

However, GaP is lattice mismatched to the underlying epitaxial layers. The lower confinement, active, and upper confinement layers are lattice matched to the GaAs substrate. Since GaP has a lattice constant that is about 3.6% smaller than that of GaAs, a high density of threading dislocations and stacking faults is expected at the upper-cladding-layer-to-GaP interface. One could assume that these dislocations, which will act as non-radiative

recombination centers, will not degrade the internal quantum efficiency of the LED because they are located at the confinement–window layer interface and in the window layer, far away from the active region. However, if the dislocations propagated downward towards the active region during device operation, the efficiency and the reliability of the LEDs would be affected. The issues associated with the confinement–window interface have apparently been resolved since AlGaN_P/GaAs LEDs with GaP window layers have excellent reliability and efficiency.

An alternative approach for increasing the extraction efficiency in AlGaN_P/GaAs LEDs uses AlGaAs current-spreading layers (Sugawara *et al.*, 1991, 1992a, 1992b). Al_xGa_{1-x}As is lattice-matched to GaAs for all chemical compositions $0 \leq x \leq 1$. AlAs has a bandgap energy of $E_{g,AlAs} = 2.9$ eV. For $x > 0.45$, Al_xGa_{1-x}As becomes an indirect semiconductor. The absorption coefficient of indirect semiconductors is much smaller than that of direct-gap semiconductors. The AlGaAs current-spreading layer is lattice matched to the underlying confinement layer and thus misfit dislocations, as in the case of GaP spreading layers, do not arise for AlGaAs current-spreading layers. However, the absorption of light in the AlGaAs layers is higher than the absorption in the GaP spreading layers. The AlGaAs is a ternary alloy and fluctuations of the cation concentration (Al and Ga) lead to a local variation of the bandgap energy. The compositional fluctuations lead to an absorption tail for energies below the AlGaAs bandgap so that AlGaAs has a larger Urbach energy than GaP.

It is well known that Al-containing compounds are difficult to grow by organo-metallic vapor-phase epitaxy (OMVPE), the common epitaxial crystal growth technique for LEDs. Aluminum is a very reactive element and OMVPE cleanliness is essential. Even minor leaks in the growth system will result in the degradation of Al-containing films. This applies, in particular, to compounds containing a high percentage of Al such as AlAs. The optical properties of AlGaAs current-spreading layers are therefore usually inferior to the properties of GaP spreading layers. The electrical properties of AlAs or AlGaAs with high Al content are also inferior to GaP. Finally, AlAs tends to oxidize over time when exposed to water or humid air (Choquette *et al.*, 1997). Despite these difficulties, viable AlGaN_P LEDs with AlGaAs current-spreading layers have been developed and are commercially available.

The increase in light-extraction efficiency is illustrated in *Figure 9.3* for p-type GaP current-spreading layer with a resistivity of $0.05 \Omega \text{ cm}$ and thicknesses ranging from 2 to $15 \mu\text{m}$ (Fletcher *et al.*, 1991a). The data shown in *Figure 9.3* were obtained using a near-field microscope and a video analyzer. To obtain the light intensity profile, a single-line scan was taken across the chip, including the center p-type contact pad, as shown in the inset. Since the intensity is directly proportional to the p-n current density at any given point, the current-spreading characteristics are obtained by this measurement. For a window thickness of $2 \mu\text{m}$, current spreading is limited. As the window layer thickness is increased to $15 \mu\text{m}$, the current spreads well beyond the contact, reaching almost the edge of the chip. An even larger thickness of the window layer would spread the current to the edges of the chip. Such strong current spreading is not desirable due to surface recombination.

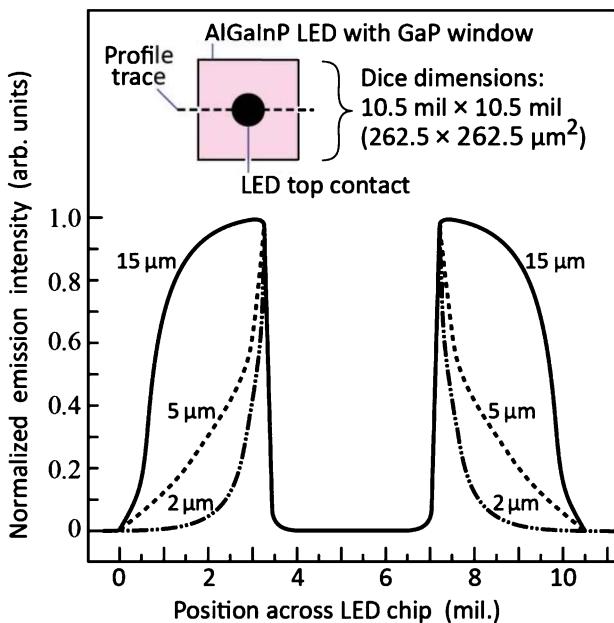


Fig. 9.3: The effect of GaP window thickness on current spreading is illustrated by surface light emission intensity profiles for three different AlGaNp LED chips with window layer thicknesses of 2, 5, and 15 μm . The profile is indicated by the dashed line in the inset. The dip in the middle of the profiles is due to the opaque ohmic contact pad. A microscope fitted with a video camera was used in the measurements (after Fletcher *et al.*, 1991a). Note: 1 mil = $25.4 \mu\text{m}$.

The effect of current spreading on the efficiency of an AlGaNp/GaAs LED with a GaP current-spreading layer is shown in *Figure 9.4*. For as sufficiently thick window layer the extraction efficiency is increased by a factor of approximately 8. The comparison of pulsed with direct current (DC) measurements shown in *Figure 9.4* shows that the efficiency drop occurring at high currents is caused by heating of the device.

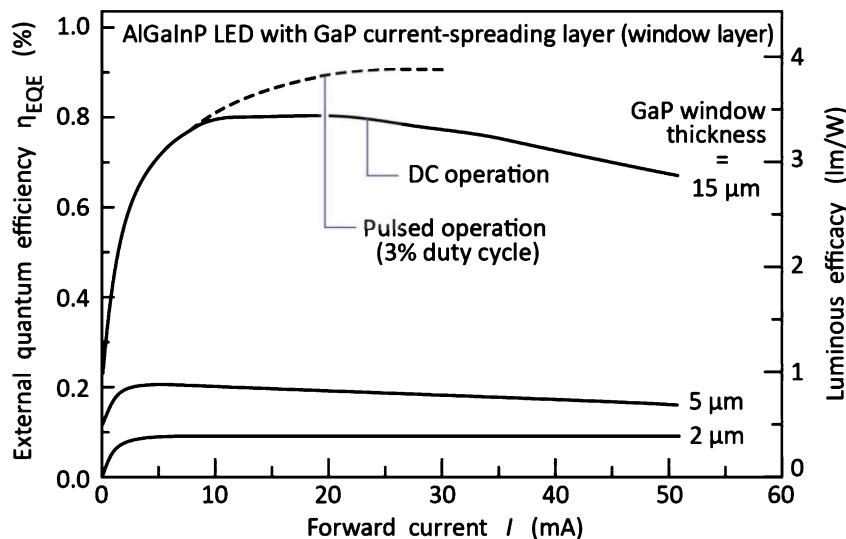


Fig. 9.4: Bare chip external quantum efficiency and luminous efficacy versus forward current for AlGaNp LEDs with GaP current-spreading layer thicknesses of 2, 5, and 15 μm. Solid curves are under DC conditions. Dashed curve is under pulsed condition using 400 ns pulses and a 3% duty cycle. Heating is essentially eliminated for pulsed condition (after Fletcher *et al.*, 1991a).

The optimum thickness range of current-spreading layers in AlGaNp/GaAs LEDs with $\text{Al}_{0.70}\text{Ga}_{0.30}\text{As}$ current-spreading layers was investigated by Sugawara *et al.* (1991, 1992a, 1992b). The p-type doping concentration of the $\text{Al}_{0.70}\text{Ga}_{0.30}\text{As}$ spreading layer was $3 \times 10^{18} \text{ cm}^{-3}$. The luminous efficiency of the LED versus current-spreading layer thickness is shown in **Figure 9.5**. Inspection of the figure reveals that the optimum thickness of the current-spreading layer is between 5 and 30 μm. For a current-spreading layer thickness of 15 μm, the efficiency of the device increases by a factor of 30 compared with a device with no current-spreading layer at all. The optimum doping concentration in the p-type current-spreading layer was found in the low 10^{18} cm^{-3} range.

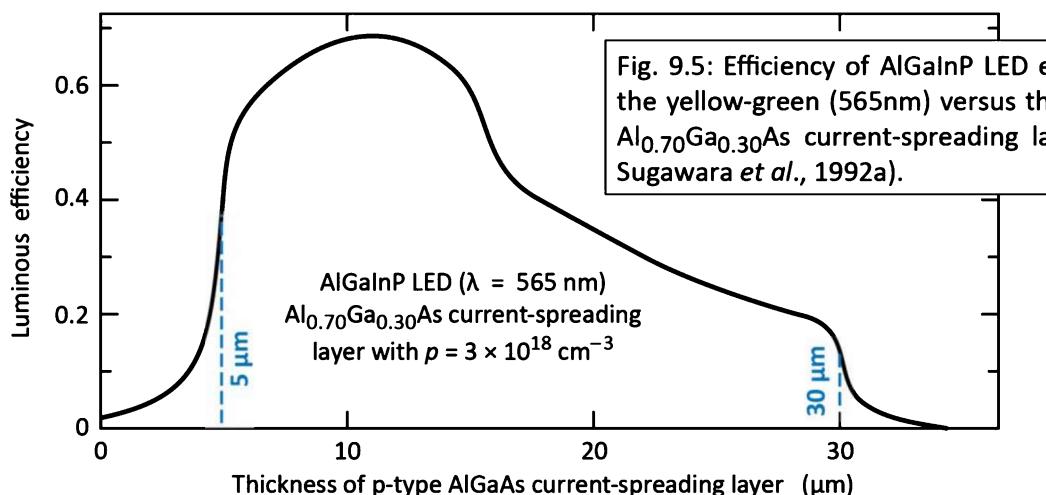


Fig. 9.5: Efficiency of AlGaNp LED emitting in the yellow-green (565 nm) versus thickness of Al_{0.70}Ga_{0.30}As current-spreading layer (after Sugawara *et al.*, 1992a).

The disadvantage of no or a very thin current-spreading layer is that most of the light is generated under the opaque metal contact pad, thereby hindering the escape of light from the LED die. A

very thick window layer is equally disadvantageous. Firstly, a thick window layer spreads the current all the way to the edge of the LED dies, which leads to increased surface recombination and thus lowers the efficiency of the LED. Secondly, light absorption increases with the thickness of the window layer due to the absorption of below-bandgap light in the window layer. Thirdly, a thick window layer will increase the ohmic resistance of the device thereby lowering the overall efficiency. Fourthly, the long growth times required for thick current-spreading layers may result in the diffusion of dopants from the confinement layers into the active region, thereby lowering the internal quantum efficiency.

Current spreading is an important issue in many LED materials, in particular in those materials that possess low conductivity. Current spreading in the top p-type layer of GaN/GaInN LEDs is very weak due to the high resistivity of the p-type top cladding layer. The hole mobility in III–V nitrides is typically $1\sim20 \text{ cm}^2/(\text{V s})$ and the hole concentrations are in the 10^{17} cm^{-3} range, resulting in resistivities $> 1 \Omega \text{ cm}$. To address this problem, Jeon *et al.* (2001) demonstrated an LED with a **tunnel junction** adjoining the p-type confinement region above the active region. An n-type layer on top of the tunnel junction allows for lateral current spreading under the top contact. Owing to the employment of the tunnel junction, the LED has *two* n-type but *no* p-type ohmic contacts.

9.2 Theory of current spreading

The theory of current spreading under a **linear stripe top contact** geometry has been reported by Thompson (1980). Such a stripe-like geometry is typical for semiconductor lasers. **Figure 9.6(a)** shows the schematic cross section of a stripe-geometry semiconductor laser. The laser has a current-spreading layer located above the p-n junction. Because of the symmetry of the laser, only the right half of the laser is shown, so that the left edge of the contact shown in the diagram is actually the center of the laser stripe. The model assumes a constant potential and current density (J_0) under the metal ($x < r_c$). The potential throughout the substrate is assumed to be constant. The current density $J(x)$ extending away from the contact is given by

$$J(x) = \frac{2 J_0}{[(x - r_c) / L_s + \sqrt{2}]^2} \quad (x \geq r_c) \quad (9.1)$$

where L_s is the current spreading length given by

$$L_s = \sqrt{\frac{t n_{ideal} kT}{\rho J_0 e}} \quad (9.2)$$

where ρ is the resistivity of the current spreading-layer, t is the thickness of the current-spreading layer, and n_{ideal} is the diode ideality factor. The diode ideality factor has typical values of $1.05 < n_{ideal} < 1.35$.

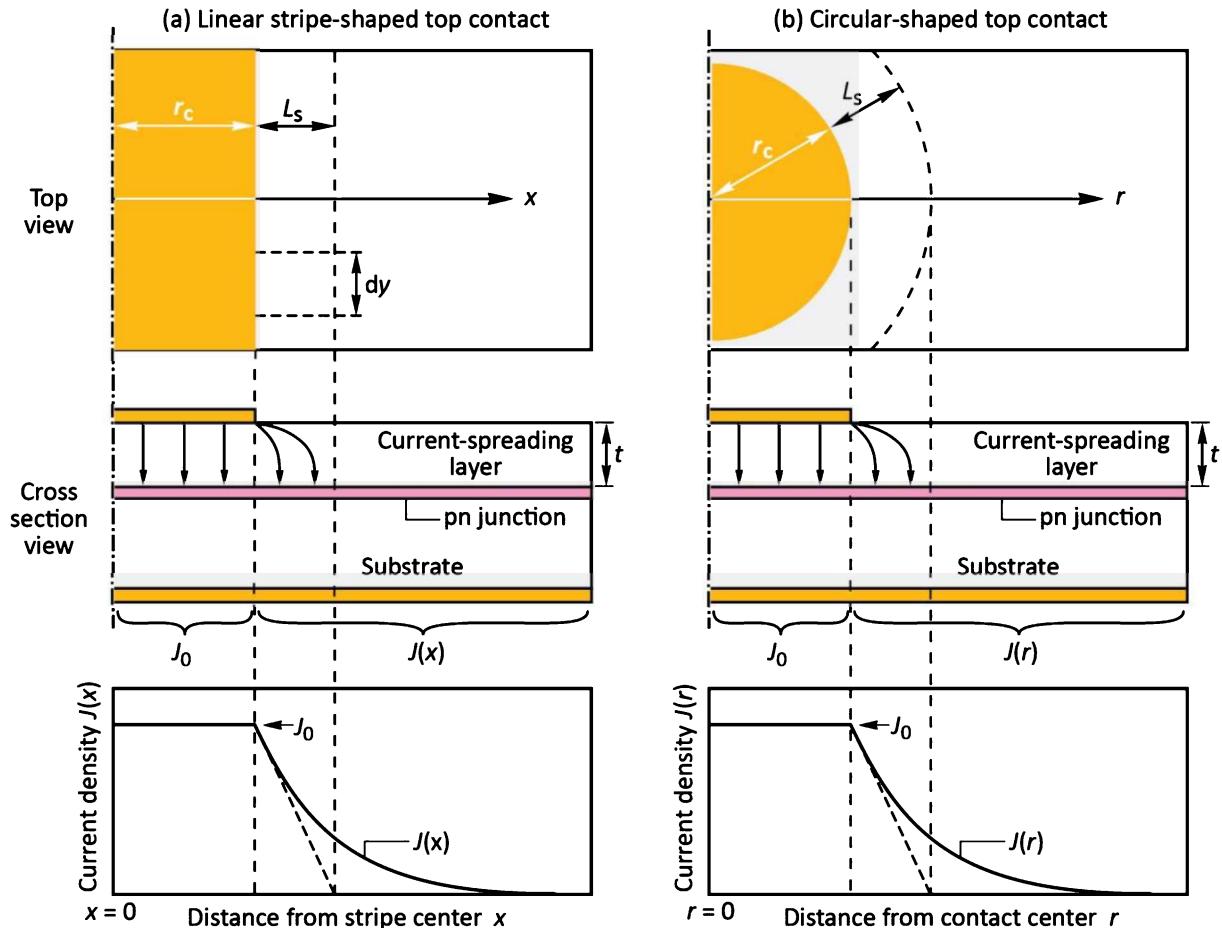


Fig. 9.6: Schematic illustration of current spreading in structures with different top contact geometries. (a) Linear stripe contact geometry. (b) Circular contact geometry.

We next develop a theoretical model that can be applied to *linear stripe* (see *Figure 9.6 (a)*) as well as *circular contact* (see *Figure 9.6 (b)*) shapes. We first consider the ***linear stripe contact geometry***. We assume that the current at the edge of the spreading region ($x = r_c + L_s$) is a factor of e^{-1} lower than under the metal contact. Then the voltage drop across the junction at the edge of the current-spreading layer is $n_{ideal} kT/e$ lower than under the metal contact. This

voltage drops within the current-spreading region. The resistance of the current-spreading region along the lateral direction per unit stripe length dy is given by

$$R = \rho \frac{L_s}{t dy} . \quad (9.3)$$

The current flowing vertically through the junction in the current spreading region is given by

$$I = J_0 L_s dy . \quad (9.4)$$

Using Ohm's law, one obtains

$$\rho \frac{L_s}{t dy} J_0 L_s dy = \frac{n_{ideal} kT}{e} . \quad (9.5)$$

Solving this equation for t yields

$$t = \rho L_s^2 J_0 \frac{e}{n_{ideal} kT} . \quad (9.6)$$

Comparison of this equation with Eqn. (9.2) yields that the two equations are identical. Equation (9.6) allows one to calculate the required current-spreading layer thickness t for a given resistivity of this layer and the desired current-spreading length L_s .

We next consider the **circular contact geometry** shown in **Figure 9.6(b)**. The circular geometry is relevant to LEDs with a circular top contact. Proceeding in a similar way, we write the lateral resistance from the edge of the contact to the edge of the current-spreading region. This resistance is given by

$$R = \int_{r_c}^{r_c+L_s} \rho \frac{1}{A} dr = \int_{r_c}^{r_c+L_s} \rho \frac{1}{t 2\pi r} dr = \frac{\rho}{2\pi t} \ln\left(1 + \frac{L_s}{r_c}\right) \quad (9.7)$$

The current flowing vertically through the junction in the current spreading region is given by

$$I = J_0 [\pi (L_s + r_c)^2 - \pi r_c^2] = J_0 \pi L_s (L_s + 2r_c) . \quad (9.8)$$

Using Ohm's law, one obtains

$$\frac{\rho}{2\pi t} \ln\left(1 + \frac{L_s}{r_c}\right) J_0 \pi L_s (L_s + 2r_c) = \frac{n_{ideal} kT}{e} . \quad (9.9)$$

Solving this equation for t yields

$$t = \rho L_s \left(r_c + \frac{L_s}{2} \right) \left(J_0 \frac{e}{n_{ideal} kT} \right) \ln \left(1 + \frac{L_s}{r_c} \right). \quad (9.10)$$

Equation (9.10) allows one to calculate the required current-spreading layer thickness t for a given resistivity of this layer and the desired current-spreading length L_s . Note that for large values of r_c , we can simplify Eqn. (9.10) using the approximation $\ln(1 + x) \approx x$, valid for $x \ll 1$. Thus, in the limit of large values of r_c (e.g. $r_c \rightarrow \infty$), Eqn. (9.10) and Eqn. (9.6) become identical, as expected.

Exercise: Current crowding occurring at very high current levels in devices with current-spreading layer. In device structures with vertical current flow (current flowing from top to bottom of chip), the current-spreading layer ensures that the current spreads out over the entire p-n junction area. However, as the current increases to very high levels, the current tends to crowd under the top contact. This is illustrated in **Figure 9.7(a) and (b)**. Explain the phenomenon of current crowding occurring at very high current densities.

Solution 1: The equation for the current-spreading length has the dependence $L_s \propto J_0^{-1/2}$. Thus, as the current density increases, L_s decreases, and the current “bunches” under the top contact.

Solution 2: An intuitive explanation for current crowding can be obtained from the equivalent circuit shown in **Figure 9.7(c)**. At very high current densities, the resistors that represent the p-n junction decrease (whereas the resistors representing the current-spreading layer remain constant), thereby causing the current to flow directly downward from the top contact.

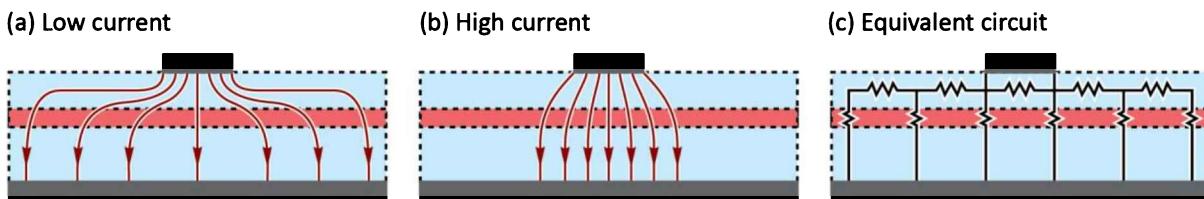


Fig. 9.7: Schematic current flow in device with current-spreading layer at (a) low and (b) high current. Current spreading decreases at very high current densities which results in current “bunching” under the top contact as shown in (b). (c) Equivalent circuit.

9.3 Current crowding in LEDs on insulating substrates

Current crowding also occurs in mesa-structure LEDs grown on insulating substrates. This type of LED includes GaInN/GaN LEDs grown on sapphire substrates. In these LEDs, the p-type contact is usually located on the top of the mesa, and the n-type contact is located on an n-type buffer layer at the bottom of the mesa. As a result, the current tends to *crowd* at the edge of the mesa contact

adjoining the n-type contact.

A lateral p-side-up mesa LED grown on an insulating substrate is shown in **Figure 9.8 (a)**. It is intuitively clear that the p-n junction current *crowds* near the edge of the mesa as indicated in the figure. An equivalent circuit model is shown in **Figure 9.8 (b)** and includes the p-type contact resistance and the resistances of the n-type and p-type cladding layers. The p-n junction is approximated by an ideal diode. The circuit model also shows several nodes separated by a distance dx . Assuming that V is the voltage in the n-type layer along the x -direction, then dV is the voltage drop across the n-layer resistance of length dx . The incremental current flowing downward through one of the diodes is given by $dI = J_0 [\exp(eV_j/kT) - 1] w dx$, where J_0 is the saturation current density of the p-n junction. Calculating the *difference* in voltage drop between two adjacent resistors and applying Kirchhoff's current law to the node located between the two resistors yields the differential equation

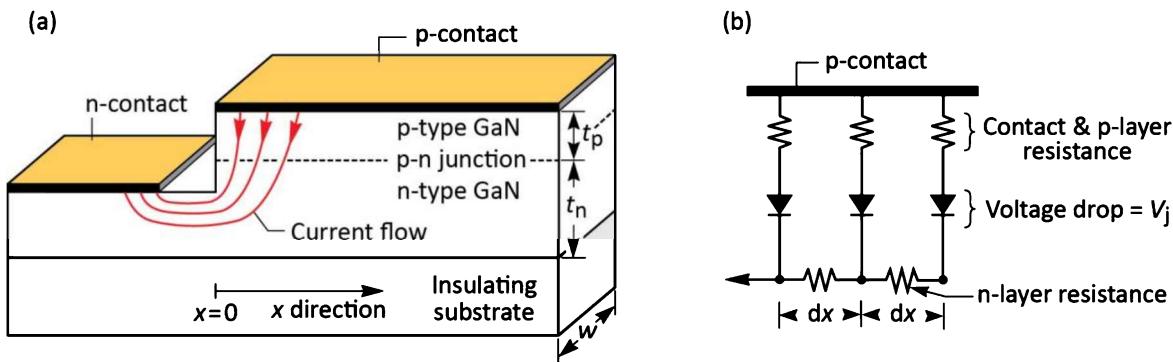


Fig. 9.8: (a) Current crowding in a mesa-structure GaN-based LED grown on an insulating substrate. (b) Equivalent circuit consisting of n-type and p-type layer resistances, p-type contact resistance, and ideal diodes representing the p-n junction.

$$\frac{d^2V}{dx^2} = \frac{\rho_n}{t_n} J_0 \left[\exp\left(\frac{e V_j}{kT}\right) - 1 \right]. \quad (9.11)$$

In the case of zero or negligible resistance of the p-type layer, it is $dV = dV_j$. In this case, Eqn. (9.11) can be easily solved and an analytic solution was given by Thompson (1980) who calculated the spreading length in p-n junction diodes grown on conductive substrates. In Thompson's study, the material resistivity of the top p-type cladding layer was considered, but the lower n-type cladding layer resistivity was neglected. However, in GaN/InGaN LEDs with the p-type layer on top and the n-type layer underneath, the resistive n-type layer causes current crowding and cannot be neglected. Furthermore, p-type resistances can be high, so that they

should not be neglected either. As will be shown in the following calculation, both types of resistances play peculiar roles in the current-crowding problem.

Next, we take into account the resistance of the n-type layer, p-type layer, and the p-type contact resistance. The voltage drop across the p-n junction and the p-type resistors is given by

$$V = R_v I_0 \left[\exp\left(\frac{e V_j}{kT}\right) - 1 \right] + V_j \quad (9.12)$$

where R_v ("vertical resistance") is the sum of the p-type layer resistance and the p-type contact resistance of the area element $w dx$, that is

$$R_v = \rho_p \frac{t_p}{w dx} + \rho_c \frac{1}{w dx} \quad (9.13)$$

where ρ_p is the resistivity of the p-type layer and ρ_c is the p-type specific contact resistance. Forming the second derivative of V with respect to x in Eqn. (9.12) and inserting the result into Eqn. (9.11) yields the differential equation

$$\begin{aligned} \frac{e}{kT} (\rho_c + \rho_p t_p) J_0 \exp\left(\frac{e V_j}{kT}\right) \left[\frac{d^2 V_j}{dx^2} + \frac{e}{kT} \left(\frac{dV_j}{dx} \right)^2 \right] + \frac{d^2 V_j}{dx^2} &= \\ = \frac{\rho_n}{t_n} J_0 \left[\exp\left(\frac{e V_j}{kT}\right) - 1 \right]. \end{aligned} \quad (9.14)$$

In order to solve the differential equation, we restrict ourselves to the forward-bias operation of the diode. In this case, the junction voltage is much larger than kT/e , that is

$$V_j \gg kT/e \quad \text{and} \quad \exp(eV_j/kT) \gg 1. \quad (9.15)$$

Furthermore, we assume that the voltage drop across the p-type series resistance and contact resistance is much larger than kT/e

$$(\rho_c + \rho_p t_p) J_0 \exp(eV_j/kT) \gg kT/e. \quad (9.16)$$

This condition applies to typical GaN/GaN LEDs. Using the approximations of Eqns. (9.15) and (9.16), Eqn. (9.14) can be simplified to

$$\frac{d^2V_j}{dx^2} + \frac{e}{kT} \left(\frac{dV_j}{dx} \right)^2 = -\frac{\rho_n}{(\rho_c + \rho_p t_p) t_n} \frac{kT}{e}. \quad (9.17)$$

Solving Eqn. (9.17) for V_j yields $V_j(x) = V_j(0) - (kT/e)(x/L_s)$. Inserting V_j into the equation $J = J_0 \exp(eV_j/kT)$ yields the solution of the differential equation as

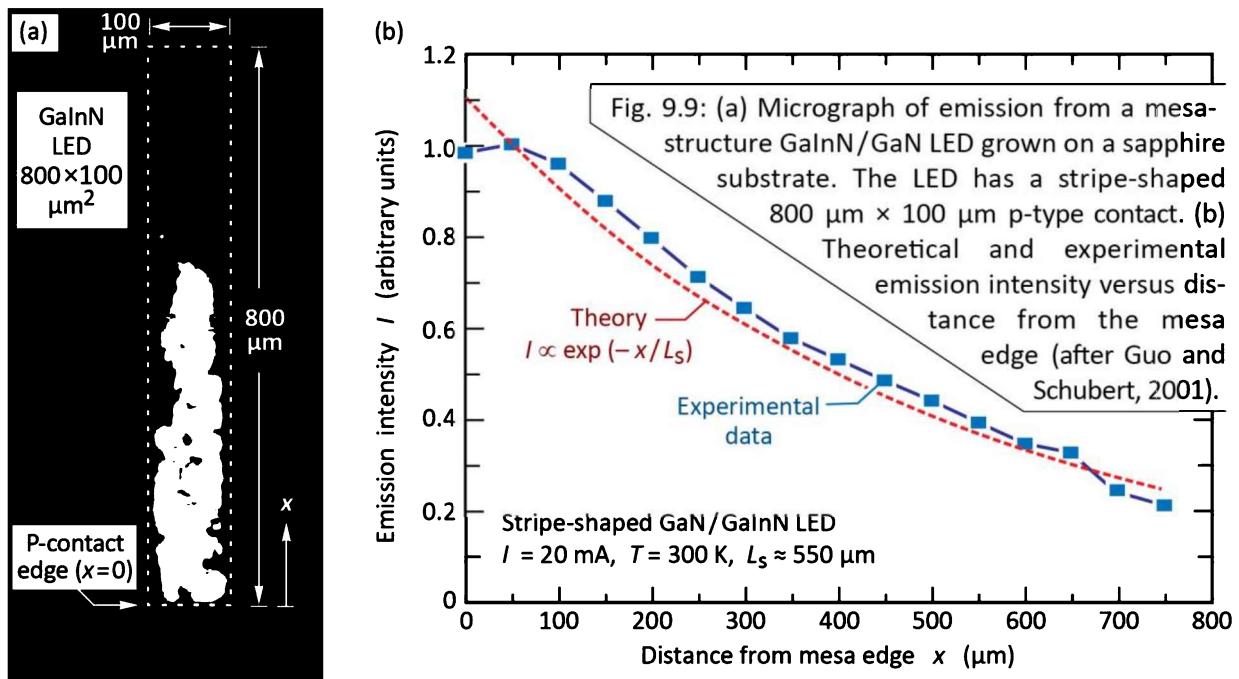
$$J(x) = J(0) \exp(-x/L_s) \quad (9.18)$$

where $J(0)$ is the current density at the p-type mesa edge and L_s is denoted as the ***current-spreading length***, that is, as the length where the current density has dropped to the $1/e$ value of the current density at the edge, so that $J(L_s)/J(0) = 1/e$. The current-spreading length is given by

$$L_s = \sqrt{(\rho_c + \rho_p t_p) t_n / \rho_n}. \quad (9.19)$$

Equation (9.19) shows that the current distribution depends on epitaxial layer thicknesses and material properties. A thick low-resistivity n-type buffer layer is needed to ensure that current crowding is minimized. Equation (9.19) also illustrates a somewhat surprising result; namely that the *decrease* of p-type specific contact resistance or p-type layer resistivity *enhances* the current-crowding effect. For *low* p-type contact and confinement resistances, strong current crowding results, unless the n-type buffer layer is very conductive so that t_n/ρ_n is very large. In GaN/GaN devices, the sum of p-type contact and p-type layer resistances can be larger than the n-type cladding resistance, especially if t_n is small.

An experimental result on the current-crowding effect in GaInN/GaN LEDs grown on a sapphire substrate is shown in **Figure 9.9** (Guo and Schubert, 2001). A micrograph of the optical emission from a GaInN LED is shown in **Figure 9.9 (a)**. The picture was taken from the sapphire substrate side of the LED and shows the intensity of blue light emission. The micrograph clearly reveals that the emission intensity decreases with increasing distance from the mesa edge. **Figure 9.9 (b)** shows the experimental intensity as a function of the distance from the mesa edge. Also shown is a theoretical fit to the experimental data using the exponential decrease in current density derived above. The experimental and the theoretical data exhibit very good agreement if a current-spreading length of 550 μm is used in the calculation.



High contact resistances and high p-GaN resistivity are not desirable for high-power devices since these resistances generate heat. On the other hand, these resistances alleviate the current crowding effect. Note that with the expected future improvement of the contact and p-type doping in GaN devices, and larger device and contact sizes, current crowding will become increasingly severe unless novel contact geometries are introduced to alleviate the problem. Such novel contact geometries can include interdigitated structures (Guo *et al.*, 2001; Steigerwald *et al.*, 2001) with p-type finger widths of less than L_s . For device dimensions much smaller than L_s , the current-crowding effect becomes irrelevant.

The schematic structure and a photograph of an interdigitated stripe-contact geometry are shown in **Figure 9.10(a)~(c)**. Uniform current injection into the active region is achieved by the p-type contact width ($W_{\text{p-contact}}$) being smaller than the current spreading length. The width of the n-type contact ($W_{\text{n-contact}}$) must be at least equal to the contact transfer length to ensure low contact resistance. The contact transfer length follows from the transmission line model (TLM) used for characterization of ohmic contacts (see, for example, Schroder, 1998).

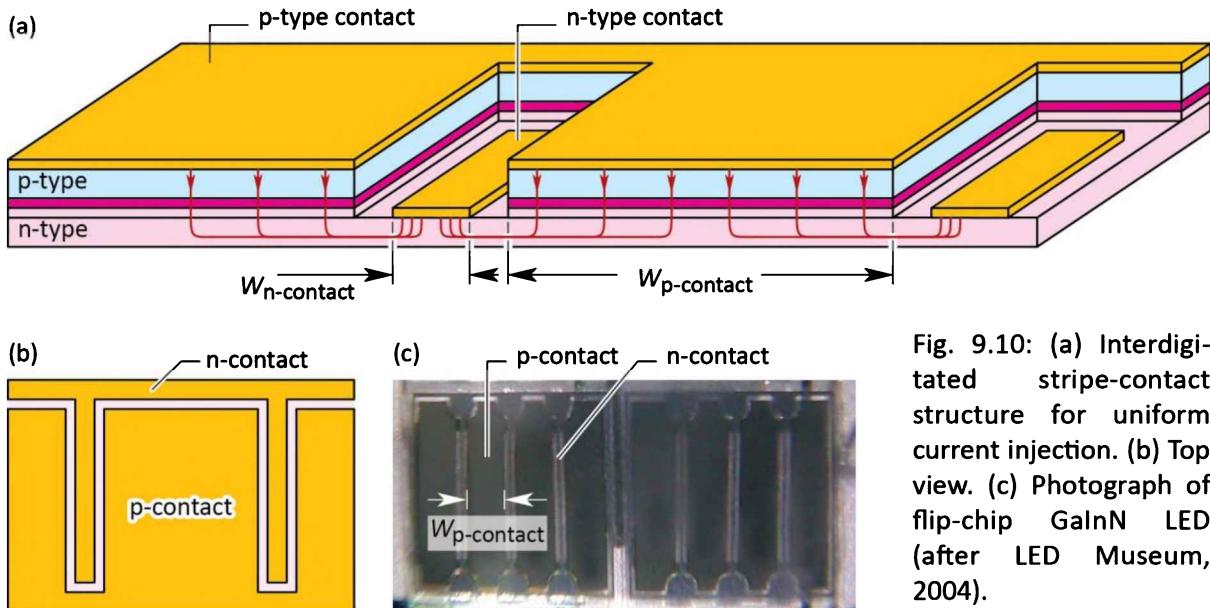


Fig. 9.10: (a) Interdigitated stripe-contact structure for uniform current injection. (b) Top view. (c) Photograph of flip-chip GaInN LED (after LED Museum, 2004).

9.4 Buried contact design

There is a fundamental desire in LED design for the emitted light to not be obstructed by opaque ohmic metal contacts. That is, if a structure is emitting light through the top surface of the device, then the top surface should be free of opaque ohmic metal contacts. This can be accomplished by the buried contact structure shown in *Figure 9.11*. The structure consists of a large-area p-type contact and an array of small circular n-type contacts, both of which are located on the bottom side of the LED.

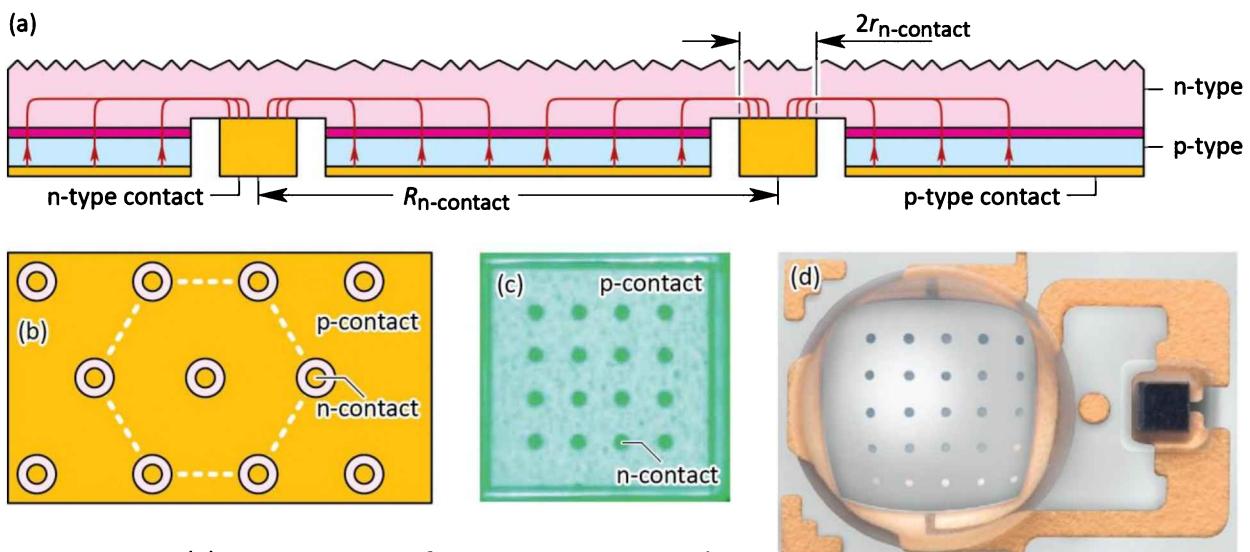


Fig. 9.11: (a) Cross section of top-emitting LED with buried p- and n-type contacts that are both located on the bottom side. The p-type contact is optically reflective. Light emission occurs through the roughened top surface. (b) Schematic bottom view of a contact structure. (c) Photograph of GaInN green LED with buried contacts (after Chen *et al.* 2008). (d) Commercial GaInN blue LED employing buried contacts (after Philips, 2011).

The buried contact structure has the advantage of minimizing the n-type contact area while maximizing the area of the light-emitting region (i.e. pn-junction area). For large-area devices, the n-type contacts are advantageously arranged in a hexagonal close-packed pattern, as shown in *Figure 9.11 (b)*. Such hexagonal close-packed pattern minimizes the distance that current is flowing laterally in the n-type layer.

Buried-contact LED structures were demonstrated in the GaN material system by Chen *et al.*, 2008. The authors reported a chip layout that employed an array of n-contact vias fabricated through a continuous p-contact area, with an interlayer dielectric that passivates the mesa walls and sits between overlapping metal layers to enable uniform current distribution. Photographs of devices showing the n-type contact array are shown in *Figure 9.11 (c)* and (d) (after Chen *et al.*, 2008; and Philips, 2011).

9.5 Lateral injection schemes

A device structure with a lateral current-injection scheme is shown in *Figure 9.12 (a)*. The current is transported laterally in both the n-type and p-type cladding layers. Ideally, the light would be generated in the region between the contacts where they would not hinder the extraction of light. If the n-type sheet resistance ρ_n/t_n (where ρ_n and t_n are the resistivity and layer thickness of the n-type material, respectively) is much lower than the p-type sheet resistance ρ_p/t_p , the current prefers to flow laterally in the low-resistance n-layer rather than the p-layer. As a result, the junction current crowds near the p-type contact.

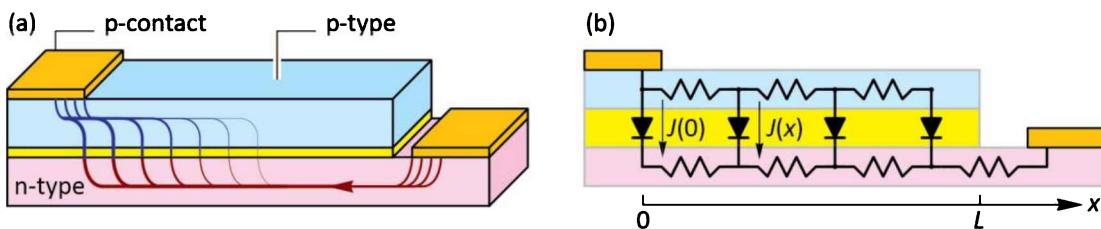


Fig. 9.12: (a) Lateral injection geometry and schematic current distribution for $\rho_n \ll \rho_p$.
 (b) Corresponding equivalent circuit.

A schematic equivalent circuit suitable for the quantitative analysis is shown in *Figure 9.12 (b)* where a p-n junction current density of $J(0)$ is assumed at the edge of the p-type contact. The analytic solution of the equivalent circuit is an exponential given by (Joyce and Wemple, 1970; Rattier *et al.*, 2002)

$$J(x) = J(0) \exp(-x/L_s) \quad (9.20)$$

where

$$L_s = \sqrt{\frac{2 V_a}{J(0) [(\rho_p / t_p) + (\rho_n / t_n)]}} \quad (9.21)$$

$J(x = 0) = J(0)$ is the current density at the edge of the contact. Rattier *et al.* (2002) stated that the voltage V_a would be an activation voltage with magnitude of a few times kT/e , e.g. 50~75 mV. The analytic solution is shown in **Figure 9.13**.

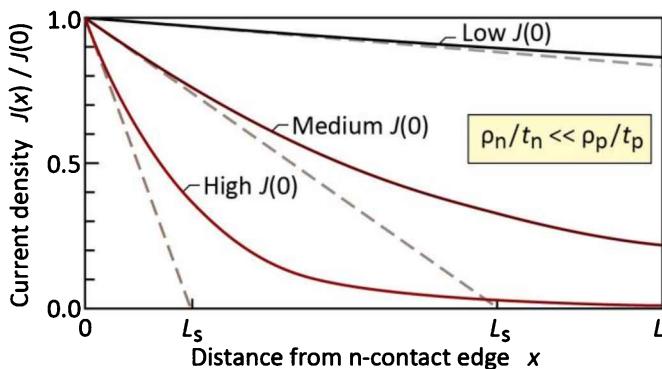


Fig. 9.13: Current density distribution at high, medium, and low current normalized to the initial current density for device with lateral injection geometry. The n-type sheet resistance is assumed to be lower than the p-type sheet resistance.

For uniform light generation across the gap between the contacts, it is desirable to have a long exponential decay length L_s . This can be achieved by high doping or thick confinement layers. To attain high powers, one may be tempted to scale the device structure in size. However, for large contact separations L , the device becomes generally more resistive unless very thick confinement layers are being used (which may be unpractical). Scaling such device structures can be accomplished by employing arrays of many small devices rather than scaling up a single device.

9.6 Current-blocking layers

In conventional DH LEDs with small top contacts and large backside contacts, the current injected by the top contact enters the active region predominantly under the top contact. The extraction of light generated in the active region is thus strongly hindered by the opaque metal contact. One possibility to alleviate this problem is the use of a thick current-spreading layer. Another possibility is the use of a current-blocking layer. This layer blocks the current from entering the active region below the top contact. The current is deflected away from the top contact, thereby allowing for higher light-extraction efficiency.

The schematic structure of an LED employing a current-blocking layer is shown in **Figure 9.14**. The blocking layer is located on top of the upper confinement layer and has approximately the same size as the top metal contact. The current-blocking layer has n-type conductivity and is embedded in material with p-type conductivity. Owing to the p-n junction surrounding the current-blocking layer, the current flows around the current-blocking layer as indicated in **Figure 9.14**.

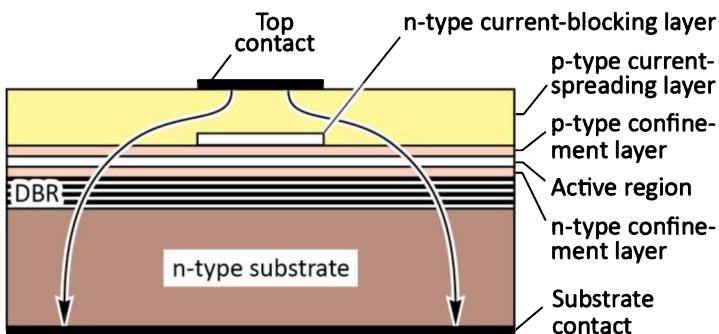


Fig. 9.14: LED with n-type current-blocking layer on upper confinement layer. Light emission occurs in the regions not covered by the opaque top contact. The LED is fabricated by *epitaxial regrowth* of the current-spreading layer after having etched the current-blocking layer.

Current-blocking layers can be fabricated by **epitaxial regrowth**. In this process, the DH and the thin n-type current-blocking layer are grown over the entire wafer surface. Subsequently, the wafer is taken out of the growth system for etching. The regions to be etched are defined by photolithography. The entire blocking layer is etched away except the region where the top ohmic contact is going to be located, as shown in **Figure 9.14**. Frequently, the current-blocking-layer etch is *selective* so that it does not etch into the upper confinement layer. Subsequently, the wafer is returned to the growth system for resumption of epitaxial growth, i.e. for the growth of the current-spreading layer.

Regrowth processes are expensive due to the reduction of the device and wafer yield that usually accompanies regrowth processes. Cleaning of the wafer surface after the etching, directly before regrowth, is critical. Defects occurring at the regrowth interface can lead to a reduction in yield. Therefore, processes requiring epitaxial regrowth are more expensive and are not suited for low-cost devices. However, for more expensive devices, such as communication LEDs, epitaxial regrowth processes are used.

In AlGaInP LEDs, n-type GaAs has been used as the current-blocking layer. The n-type GaAs layer is located on top of the AlGaInP upper confinement layer. The GaAs current-blocking layer is lattice matched to the underlying AlGaInP confinement layer. Selective wet chemical etches are available that etch the GaAs but do not etch the AlGaInP (Adachi and Oe, 1983).

Vertical-cavity surface-emitting lasers (VCSELs) also employ current-blocking layers to funnel the current to the active region located between the mirrors of the laser. However, oxygen or hydrogen implantation rather than epitaxial regrowth is used in VCSELs to form current-blocking layers. The implantation depth is limited so that the lateral resistance can become substantial for devices with large-area contacts.

9.7 Surface-leakage current

The leakage current of a pn-junction device has two components: (i) Bulk-leakage current and (ii) Surface-leakage current. The *bulk-leakage current* is usually caused by carrier recombination due to defects (point defects, such as unwanted impurities, and extended defects, such as threading dislocations). The bulk-leakage current due to minority carriers (as derived by Shockley and expressed in the Shockley equation) also occurs, but is secondary in magnitude, particularly in wide-bandgap semiconductors.

The *surface-leakage current*, schematically illustrated in **Figure 9.15**, occurs at the perimeter of the pn-junction plane, i.e., where the pn-junction plane meets the semiconductor surface. The surface-leakage current is caused by surface recombination (Cutler and Bath, 1957; Sah *et al.*, 1957).

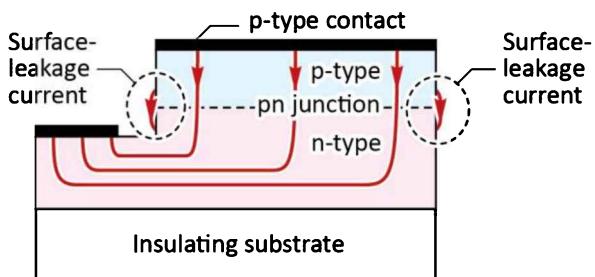


Fig. 9.15: Surface-leakage current that occurs at the perimeter of the pn-junction plane, i.e. where the pn-junction plane meets the semiconductor surface. The surface-leakage current is non-radiative and caused by surface recombination.

Consider an undoped active region that is injected with carriers so that $\Delta n_{surface} = \Delta p_{surface} = n_{surface} = p_{surface}$. The surface recombination rate (per unit area) is given by $R_{surface} = S n_{surface}$ where S is the surface recombination velocity. The surface-leakage current is given by

$$I_{surface-leakage} = e R_{surface} A = e S n_{surface} A \quad (9.22)$$

where e is the elementary charge and A is the area of the surface leakage, i.e. the length of the pn-junction-plane's perimeter multiplied by the thickness of the active region.

The relative contribution of the surface-leakage current is proportional to the pn-junction plane's perimeter-to-area ratio. For a square-shaped mesa with perimeter $4 L_{\text{LED}}$ (L_{LED} being the edge-length of the LED) and area L_{LED}^2 , the perimeter-to-area ratio is given by $4 / L_{\text{LED}}$ illustrating that the surface-leakage current becomes increasingly relevant for small-area LEDs. The contribution of the surface-leakage current has been estimated for GaInN blue LEDs to be 3.8% of the total leakage current (Ferdous *et al.*, 2007). Although the surface-leakage current is relatively small, its relevance with respect to device aging and reliability is significant. That is, the high current density occurring at the surface, if unmitigated, results in rapid degradation of the device.

Therefore, surface passivation and surface protection is commonly employed. This is accomplished by covering the exposed perimeter of the pn junction with a passivation layer, made of, e.g., SiO_2 or Si_3N_4 . The passivation layer (i) passivates surface states, (ii) aids in the dissipation of the thermal power generated at the surface, and (iii) slows down the structural degradation occurring at the exposed surface. A systematic investigation of different passivation films on a GaN-based LED was reported by Yang *et al.* (2005). The authors concluded that SiO_2 or SiN_x deposited by plasma-enhanced chemical-vapor deposition (PECVD), and Sc_2O_3 or MgO deposited by radio-frequency plasma-enhanced molecular-beam epitaxy, all show excellent passivation qualities. PECVD-films are favored in manufacturing applications since they are easily deposited and the cost per die is relatively small (Yang *et al.*, 2005).

9.8 References

- Adachi S. and Oe K. "Chemical etching characteristics of (001) GaAs" *Journal of the Electrochemical Society* **130**, 2427 (1983)
- Chen G., Craven M., Kim A., Munkholm A., Watanabe S., Camras M., Götz W., and Steranka F. "Performance of high-power III-nitride light emitting diodes" *Physica Status Solidi* **205**, 1086 (2008)
- Choquette K. D., Geib K. M., Ashby C. I. H., Tweten R. D., Blum O., Hou H. Q., Follstaedt D. M., Hammons B. E., Mathes D., and Hull R. "Advances in selective wet oxidation of AlGaAs alloys" *IEEE Journal of Selected Topics in Quantum Electronics* **3**, 916 (1997)
- Cutler M. and Bath H. M. "Surface leakage current in silicon fused junction diodes" *Proceedings of the IRE* **45**, 39 (1957)
- Ferdous M. S., Wang X., Fairchild M. N., and Hersee S. D. "Effect of threading defects on InGaN/GaN multiple quantum well light emitting diodes" *Applied Physics Letters* **91**, 231107 (2007)
- Fletcher R. M., Kuo C. P., Osentowski T. D., Huang K. H., and Crawford M. G. "The growth and properties of high performance AlInGaP emitters using lattice mismatched GaP window layers" *Journal of Electronic Materials* **20**, 1125 (1991a)
- Fletcher R. M., Kuo C. P., Osentowski T. D., and Robbins V. M. "Light-emitting diode with an electrically conductive window" US Patent 5,008,718 (1991b)

- Groves W. O. and Epstein A. S. "Epitaxial deposition of III–V compounds containing isoelectronic impurities" US Patent 4,001,056 (1977)
- Groves W. O., Herzog A. H., and Craford M. G. "Process for the preparation of electroluminescent III–V materials containing isoelectronic impurities" US Patent Re. 29,648 (1978a)
- Groves W. O., Herzog A. H., and Craford M. G. "GaAsP electroluminescent device doped with isoelectronic impurities" US Patent Re. 29,845 (1978b)
- Guo X. and Schubert E. F. "Current crowding and optical saturation effects in GaInN/GaN light-emitting diodes grown on insulating substrates" *Applied Physics Letters* **78**, 3337 (2001)
- Guo X., Li Y.-L., and Schubert E. F. "Efficiency of GaN/GaN light-emitting diodes with interdigitated mesa geometry" *Applied Physics Letters* **79**, 1936 (2001)
- Jeon S.-R., Song Y.-H., Jang H.-J., Yang G. M., Hwang S. W., and Son S. J. "Lateral current spreading in GaN-based light-emitting diodes utilizing tunnel contact junctions" *Applied Physics Letters* **78**, 3265 (2001)
- Joyce W. B. and Wemple S. H. "Steady-state junction-current distributions in thin resistive films on semiconductor junctions (solutions of $\nabla^2\nu = \pm e^\nu$)" *Journal of Applied Physics* **41**, 3818 (1970)
- Kuo C. P., Fletcher R. M., Osentowski T. D., Lardizabal M. C., Craford M. G., and Robins V. M. "High performance AlGaN/P visible light emitting diodes" *Applied Physics Letters* **57**, 2937 (1990)
- LED Museum, www.ledmuseum.org (2004)
- Moyer C. D. "Photon recycling light emitting diode" US Patent 4,775,876 (1988)
- Nishizawa J., Koike M., and Jin C. C. "Efficiency of GaAlAs heterostructure red light-emitting diodes" *Journal of Applied Physics* **54**, 2807 (1983)
- Nuese C. J., Tietjen J. J., Gannon J. J., and Gossenberger H. F. "Optimization of electroluminescent efficiencies for vapor-grown GaAsP diodes" *Journal of the Electrochemical Society: Solid State Science* **116**, 248 (1969)
- Philips Lumileds Lighting Company, Luxeon Rebel royal-blue LED product (2011)
- Rattier M., Bensity H., Stanley R. P., Carlin J.-F., Houdre R., Oesterle U., Smith C. J. M., Weisbuch C., and Krauss T. F. "Toward ultra-efficient aluminum oxide microcavity light-emitting diodes: Guided mode extraction by photonic crystals" *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 238 (2002)
- Sah C.-T., Noyce R. N., and Shockley W. "Carrier generation and recombination in p-n junctions and p-n junction characteristics" *Proceedings of the IRE* **45**, 1228 (1957)
- Schroder D. K. "Semiconductor material and device characterization" (John Wiley and Sons, New York, 1998)
- Sugawara H., Ishikawa M., and Hatakoshi G. "High-efficiency InGaAlP/GaAs visible light-emitting diodes" *Applied Physics Letters* **58**, 1010 (1991)
- Sugawara H., Ishikawa M., Kokubun Y., Nishikawa Y., Naritsuka S., Itaya T., Hatakoshi G., Suzuki M. "Semiconductor light-emitting device" US Patent 5,153,889, issued Oct. 6 (1992a)
- Sugawara H., Itaya T., Nozaki H., and Hatakoshi G. "High-brightness InGaAlP green light-emitting diodes" *Applied Physics Letters* **61**, 1775 (1992b)
- Steigerwald D. A., Rudaz S. L., Thomas K. J., Lester S. D., Martin P. S., Imler W. R., Fletcher R. M., Kish Jr. F. A., Maranowski S. A. "Electrode structures for light-emitting devices" US Patent 6,307,218 (2001)
- Thompson G. H. B. *Physics of Semiconductor Laser Devices* (John Wiley and Sons, New York, 1980)
- Yang H. S., Han S. Y., Hlad M., Gila B. P., Baik K. H., Pearton S. J., Jang S., Kang B. S., and Ren F. "Comparison of Surface Passivation Layers on InGaN/GaN MQW LEDs" *Journal of Semiconductor Technology and Science* **5**, 131 (2005)

10

10 – Alternating-current (AC) and high-voltage (HV) LEDs

A power supply driving a single LED needs to provide a direct-current (DC) voltage on the order of 2~3 V. This voltage is much lower than common alternating-current (AC) line voltages of 120~240 V. To overcome the need for an AC-to-DC conversion with a high AC-to-DC-voltage ratio, alternating-current LEDs or **AC LEDs** are employed.

- First, alternating-current LEDs, **AC LEDs**, are LEDs, connected in series, which are optimized for the use with an AC voltage supply without the need for AC-to-DC conversion.
- Second, high-voltage LEDs, **HV LEDs**, are LEDs, connected in series, to allow higher drive voltages.

This chapter will first discuss AC LEDs and then discuss HV LEDs. AC LEDs and HV LEDs can be (*i*) multiple discrete LED devices that are interconnected in the same package or (*ii*) multiple LED devices that are integrated on the same chip. Such monolithic integration of optoelectronic devices is reminiscent of the microelectronic integrated circuits made in Si, pioneered by Jack Kilby and Robert Noyce in 1958 (Noyce, 1961 and Kilby, 1964). Monolithic integration was also demonstrated in compound semiconductor devices (see, for example, Spitzer *et al.*, 1975; Bar-Chaim *et al.*, 1982; Dagenais *et al.*, 1994; Ebeling, 1993; and Liu *et al.*, 1995). Particularly noteworthy is the work of Spitzer *et al.* (1975) who demonstrated a monolithically integrated GaP diode array that the authors analyzed and optimized.

10.1 Simple circuits for AC-driven LEDs

The electrical drive of LEDs can be direct current (DC) or alternating current (AC). While DC drive is very common for LEDs, there has been some interest in AC drive for LEDs (Ao *et al.*, 2002; Jiang *et al.*, 2005; Cho *et al.*, 2007; Yen *et al.*, 2007; Lee *et al.*, 2008; Yen *et al.*, 2008). This interest is motivated by the fact that no AC-to-DC power conversion is needed for AC-driven LEDs. That is, when configured in series-type of circuit, the string of LEDs can be driven directly by common 120 V or 240 V AC line voltages.

A common AC drive voltage can be described by a sinusoidal function, that is

$$V(t) = V_0 \cos \omega t \quad (10.1)$$

where V_0 is the peak voltage of the AC line voltage. The peak voltage is related to the effective voltage (root-mean-square voltage, V_{rms}) by

$$V_0 = \sqrt{2} V_{rms} . \quad (10.2)$$

That is, for a 120 V line voltage, $V_{rms} = 120$ V and $V_0 = 2^{1/2} \times 120$ V ≈ 170 V; for a 240 V supply voltage, $V_{rms} = 240$ V and $V_0 = 2^{1/2} \times 240$ V ≈ 339 V.

A **unidirectional** AC-driven LED circuit is shown in *Figure 10.1(a)*. The circuit has one branch and contains N LEDs. The voltage drop across one LED is given by $V_{LED}(t) = V(t)/N$. The number of LEDs, N , is selected in such a way that V_{LED} , at its peak value, significantly exceeds the common DC operating voltage of the LED. The unidirectional circuit emits light during one half-cycle of the sinusoidal drive voltage. The intermittent emission may cause flickering effects (Lee *et al.*, 2008).

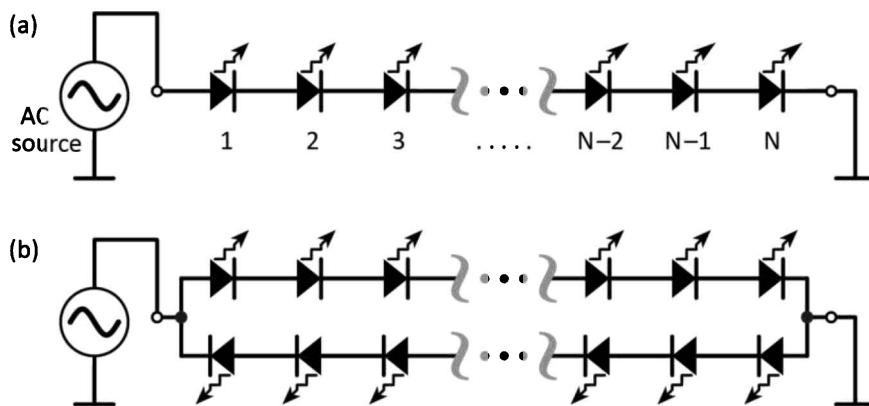


Fig. 10.1: (a) One-branch and (b) two-branch drive circuit for alternating-current (AC) LEDs. The two-branch drive circuit reduces visual flickering effects due to its higher frequency of emitted light pulses.

A **bidirectional** AC-driven LED circuit is shown in *Figure 10.1(b)*. The circuit has two branches and contains $2N$ LEDs. The bidirectional circuit emits light during the positive and negative half-cycle of the sinusoidal drive voltage. The advantage of the bidirectional circuit includes a higher pulse frequency and thus reduced flickering effects.

A photograph of a monolithically integrated bi-directional AC-driven LED circuit with an area of $1.1\text{ mm} \times 1.0\text{ mm}$ is shown in *Figure 10.2* for a positive and negative bias (Ao *et al.*, 2002). Different elements light up during the positive and negative half cycle. The AC LED consisted of 20 LED elements in each circuit branch. Under an AC voltage of 72 V, the optical output power was 40 mW.

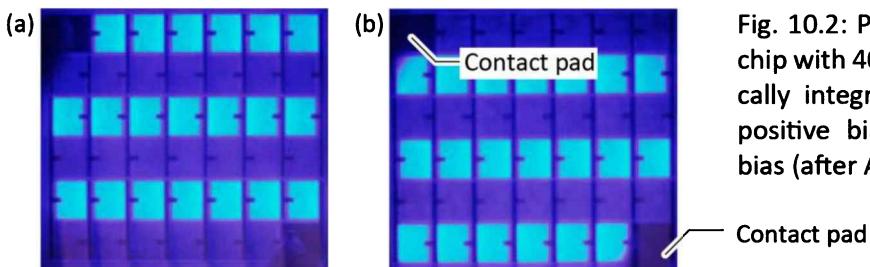


Fig. 10.2: Photograph of an LED chip with 40 (20 + 20) monolithically integrated devices for (a) positive bias and (b) negative bias (after Ao *et al.*, 2002).

10.2 LED bridge rectifier configurations for AC-driven LEDs

It is desirable that the diodes that are located on the chip are used as efficiently as possible. By including a bridge rectifier, the LED load is emitting light during each half-cycle of the AC signal (Cho *et al.*, 2007). **Figure 10.3** shows such a bridge rectifier circuit. As indicated in the figure, the bridge rectifier itself can be made of LEDs thereby monolithically integrating an LED bridge rectifier and the LED load. Such integration makes efficient use of the chip area and minimizes the need for external rectifier parts.

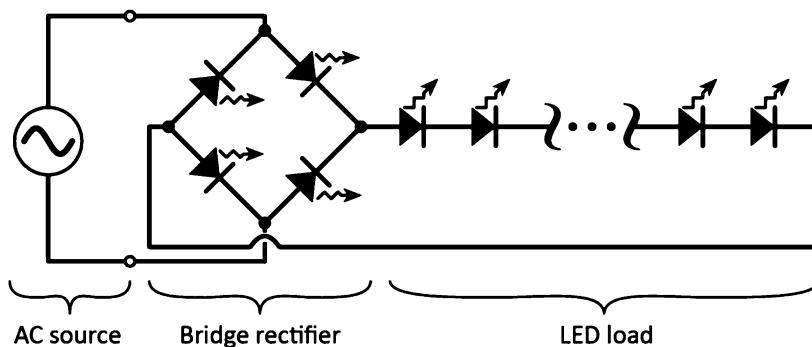


Fig. 10.3: AC drive circuit using a bridge rectifier composed of LEDs. The bridge-rectifier circuit increases the time LEDs are used from one half to one full AC cycle.

Figure 10.4 shows the implementation of an AC-driven LED chip that monolithically integrates two LED bridge rectifiers and associated LED loads (Yen *et al.*, 2008). The photographs shown in the figure reveal that the majority of LEDs are emitting light for a positive as well as a negative drive voltage.

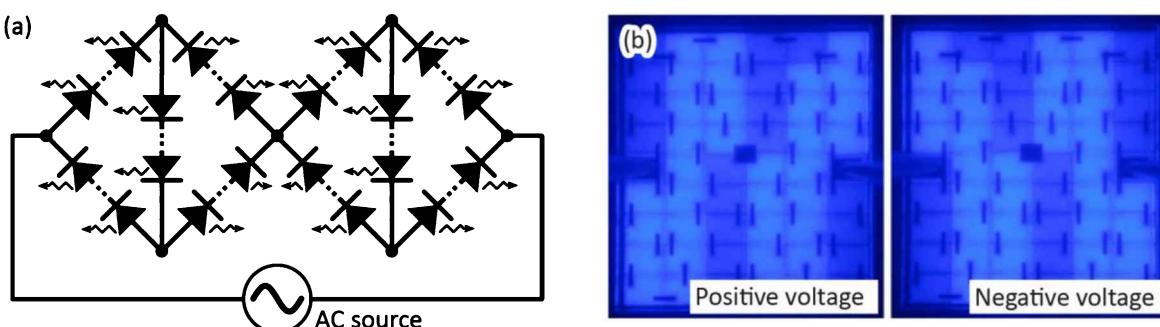


Fig. 10.4: (a) Circuit diagram of two LED bridge rectifiers and associated LED loads. (b) Photograph of single AC LED chip containing two bridge rectifiers and LED loads. The chip is driven by a positive and a negative DC voltage (after Yen *et al.*, 2008).

10.3 Analysis of pulsed drive current in AC-driven LEDs

Consider an LED that is driven by a sinusoidal voltage. Since the current through the LED depends exponentially on the voltage, the LED is emitting light only for a short time during a half-cycle of the drive voltage. This is schematically shown in **Figure 10.5**.

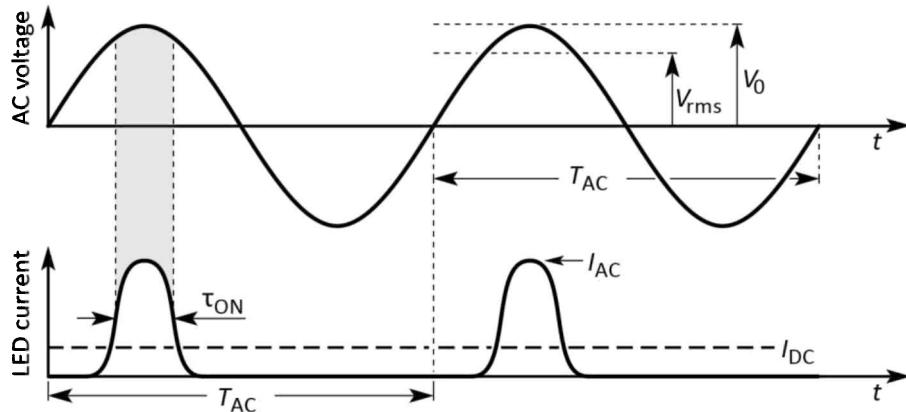


Fig. 10.5: Illustration of AC drive voltage with period T_{AC} and associated LED current with effective ON time τ_{ON} .

Next, we proceed to calculate the time, τ_{ON} , during which the LED emits light. Assume that one LED is driven by a sinusoidal voltage with peak voltage V_0 and angular frequency ω , that is

$$V(t) = V_0 \cos \omega t \quad (10.3)$$

Since the current of a diode depends exponentially on the forward voltage, a significant current will flow through the LED only when the AC voltage is close to its peak value, as illustrated in **Figure 10.5**. We therefore expand the cosine function into a power series:

$$V(t) = V_0 \cos \omega t = V_0 \left(1 - \frac{(\omega t)^2}{2!} + \frac{(\omega t)^4}{4!} - \frac{(\omega t)^6}{6!} + \dots - \dots \right). \quad (10.4)$$

Near $t = 0$, we write

$$V(t) \approx V_0 \left(1 - \frac{1}{2} \omega^2 t^2 \right). \quad (10.5)$$

For a diode-forward voltage with $V \gg kT/e$, we can write the Shockley-diode equation as

$$I(t) \approx I_S e^{eV(t)/kT}. \quad (10.6)$$

Inserting Eqn. (10.5) into Eqn. (10.6) yields

$$I(t) \approx I_S \exp \frac{eV_0(1 - \frac{1}{2}\omega^2 t^2)}{kT} = I_S \exp \left(\frac{eV_0}{kT} \right) \exp \left(\frac{-\frac{1}{2}eV_0 \omega^2 t^2}{kT} \right) \quad (10.7)$$

which is a pulse with the shape of a gaussian function (the gaussian function is also called normal distribution function). The gaussian function is given by

$$y(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right) \quad (10.8)$$

where σ is the standard deviation and μ is the average. Using $\mu = 0$ and comparing the exponents of the exponential function in Eqn. (10.7) with Eqn. (10.8), we can identify the temporal standard deviation σ_t as

$$\sigma_t = \sqrt{\frac{kT}{\omega^2 e V_0}} = \sqrt{\frac{kT}{e V_0}} \frac{T_{AC}}{2\pi} \quad (10.9)$$

where T_{AC} is the period of the sinusoidal AC drive voltage. The full-width-at-half-maximum ($FWHM$) of a gaussian function is related to its standard deviation by $FWHM_t = 2(2 \ln 2)^{1/2} \sigma_t \approx 2.355 \sigma_t$. Denoting the time that the LED emits light as τ_{ON} (see **Figure 10.5**), and using $FWHM_t = \tau_{ON}$, we obtain

$$\tau_{ON} = 2\sqrt{2 \ln 2} \sigma_t = 2.355 \sqrt{\frac{kT}{e V_0}} \frac{T_{AC}}{2\pi} \quad (10.10)$$

As an example, we consider an LED driven by an AC voltage with $V_0 = 3.0$ V. Using $kT = 26$ meV and assuming a 50 Hz line voltage ($T_{AC} = 20$ ms), we obtain

$$\tau_{ON} = 0.698 \text{ ms} \approx 1 \text{ ms.} \quad (10.11)$$

The result shows that an AC-driven LED is emitting light during only about 5% of the full cycle of the sinusoidal drive voltage (about 1 ms out of 20 ms) or only about 10% of the half-cycle of the sinusoidal drive voltage (about 1 ms out of 10 ms).

The LED ON-time is longer than the value calculated above when taking into account the unavoidable series resistance of an LED or when deliberately adding a series resistance to the circuit. The series resistance makes the current–voltage curve less exponential and more linear, so that $\tau_{ON}/T_{AC} > 5\%$.

In order for an AC LED to deliver about the same power as a DC LED, the AC drive current, I_{AC} , would need to be higher by a factor equal to the inverse duty cycle, that is,

$$\frac{I_{AC}}{I_{DC}} = \frac{T_{AC}}{\tau_{ON}} \quad (10.12)$$

where I_{AC} is illustrated in **Figure 10.5**. Since $\tau_{ON} \ll T_{AC}$, it is $I_{AC} \gg I_{DC}$. That is, AC LEDs need to use AC peak currents that are generally much higher than equivalent DC currents.

10.4 Challenges associated with AC-driven LEDs

The analysis in the previous section shows that the AC peak current of an LED would need to be significantly higher, by about a factor of 10, than the DC current of a LED emitting about the same optical power. The reason for the much higher AC current is found in the inherently short duty cycle of the AC LED which in turn is caused by the exponential dependence of the LED current on forward voltage. The high AC peak current poses two formidable challenges:

Firstly, AC LEDs suffer from larger **resistive losses** than DC-driven LEDs. Resistive power losses are given by $I^2 R_s$ where R_s is the series resistance of the LED. For example, a factor of two (2) higher AC peak current results in a factor of $2^2 = 4$ higher resistive losses.

The fraction of power consumed in the series resistance is given by $I^2 R_s / (I V_{f,junction} + I^2 R_s) \approx I^2 R_s / (I h\nu/e + I^2 R_s)$. As the drive current increases, so does the fraction of power consumed in the series resistance of the diode.

Secondly, AC LEDs suffer from much more strongly from the **efficiency droop** than DC-driven LEDs. A typical efficiency-versus-current curve for III–V nitride LEDs is shown in **Figure 10.6**. Inspection of the figure reveals that AC LEDs necessarily operate in a lower efficiency regime than DC LEDs. Although the efficiency droop may vary from material to material and from device to device, the efficiency droop cannot be completely eliminated. This puts AC LEDs at an inherent disadvantage over their DC counterparts.

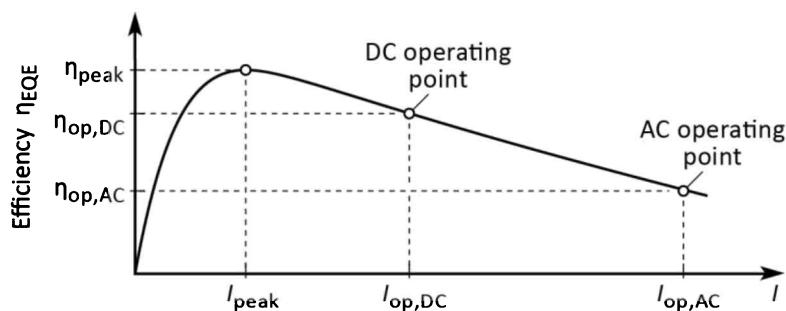


Fig. 10.6: Illustration of typical operating points for LEDs driven by a DC and an AC drive voltage. The AC drive current is about a factor $(\eta_{op,DC}/\eta_{op,AC}) \times (T_{AC}/\tau_{ON})$ higher than the DC drive current to reach the same light-output power.

In order to overcome the above challenges of AC-driven LEDs, Lee *et al.* (2008) proposed a switching matrix that switches the AC signal between different groups of LEDs. As a result, the time during which AC-driven LEDs emit light, is prolonged. However, compensatory measures, such as a switching matrix, intended to overcome the inherent challenges of AC-driven LEDs, add complexity to an AC-LED system.

Despite the above drawbacks, reasonably efficient AC LEDs have been demonstrated. Onushkin *et al.* (2009) reported a blue and white LED chips that was designed and fabricated as an integrated circuit with several isolated LED structures grown on a single sapphire substrate. The operation of the white LED directly from a 220 V AC power source with a drive power of 4 W resulted in an average luminous flux 320 lm and a luminous efficacy of 80 lm/W.

10.5 High-voltage (HV) LEDs

LEDs that are operated by a rectified high-voltage AC signal can be referred to as **HV-AC LEDs** or simply **HV-LEDs**. They can be operated by the common household AC line voltage (110~240 V depending on the geographical region). The power supply of a HV LED is a bridge rectifier, most commonly Si pn-junction bridge rectifier. The voltage drop across a Si bridge rectifier is $V_{f,Si} \approx 0.8\text{ V}$. The electrical power lost in the bridge rectifier is about $V_{f,Si} I_f$. The power-conversion efficiency (*PCE*) of such supply is $PCE = (V_{supply} I_f - V_{f,Si} I_f) / (V_{supply} I_f) > 99\%$ which is an excellent value. Furthermore, the cost of a Si bridge rectifier is very low. Indeed the high efficiency of such a power supply and its low cost are clear advantages of HV LEDs. On the other hand, the driving of LEDs by a rectified “raw” sinusoidal voltage suffers from the same drawbacks as AC LEDs. How can the efficiency of HV LED systems be improved?

Series resistance: A series resistance enables longer “ON” times and thus better efficiencies. However, if the series resistance is too large, it will consume more power thereby decreasing the efficiency. Therefore, there is an optimum resistance value that maximizes the power-conversion efficiency of a HV LED.

Number of LEDs: LEDs configured in a series circuit, consisting of a very *small* number of LEDs, when driven by a rectified AC signal, are subjected to voltages much greater than their turn-on voltage. As a result, most of the voltage drops across the series resistance and the efficiency of the circuit is low. If the circuit consists of a very *large* number of LEDs, the voltage may not exceed the turn-on voltage. Therefore, there is an optimum

number of LEDs for which the efficiency is highest. The problem was analyzed by Spitzer *et al.* (1975) and the reader is referred to this literature source for a detailed discussion.

Capacitor: Inclusion of a capacitor will smoothen the ripple of the rectified AC signal thereby operating the HV LEDs in a more DC-like mode.

A photograph of an HV LED is shown in *Figure 10.7*. HV LEDs are particularly suited for applications that are space-constrained (such as small light bulbs) if high efficiency is not a priority.

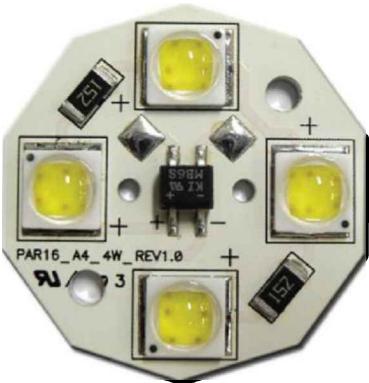


Fig. 10.7: Photograph of AC LED circuit board having a Si bridge rectifier in the center of the board. Two series resistors are included on the board. Each of the four LED packages contains multiple pn-junction LEDs. The module can be driven by 120 V AC, emits white light with color temperature of 3000 K (“warm white”), has an efficacy of 75 lm/W, and a color rendering index of CRI = 85 (after Seoul, 2009).

HV-LEDs can be connected in series-circuit configurations, parallel-circuit configurations, or a mixture thereof. This is done to allow for greater flexibility for the drive voltage and drive current. For example, series circuits are desired for higher drive voltages. Parallel circuits can enhance reliability, since the failure of a device on one branch of the parallel circuit does not affect the other branch. Several series and parallel circuits used for LEDs are shown in *Figure 10.8*.

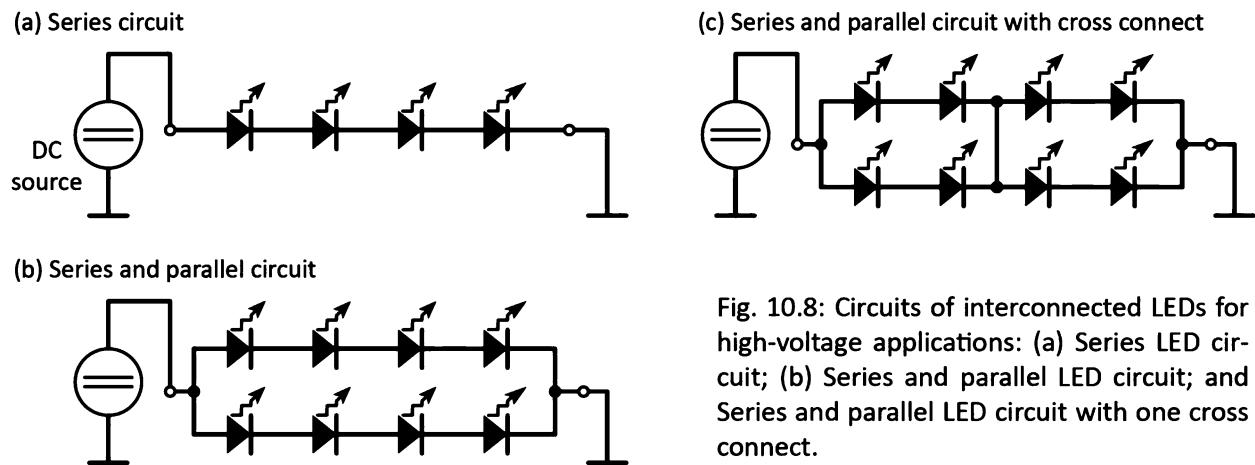


Fig. 10.8: Circuits of interconnected LEDs for high-voltage applications: (a) Series LED circuit; (b) Series and parallel LED circuit; and Series and parallel LED circuit with one cross connect.

Figure 10.9 shows a two-LED series circuit and a two-LED parallel circuit. A series circuit configuration ensures that both LEDs are driven with the same current. However, this is not the case for a parallel circuit configuration: The forward voltages of LEDs, even if the LEDs are of the same type, can be slightly different. Since the diode current depends exponentially on the drive voltage, the same drive voltage can result in significantly different drive currents. As a result, the two LEDs in the parallel circuit can have a significantly different brightness. For this reason, a caution is frequently expressed in conjunction with parallel circuits, as illustrated in **Figure 10.9**.

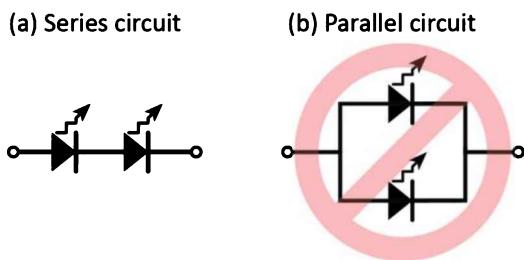


Fig. 10.9: (a) In a *series* circuit, both LEDs are driven with the same current. (b) A *parallel* circuit may result in different drive currents (current depends exponentially on voltage). For this reason, parallel LED circuits are rarely used.

An LED lamp using multiple LEDs to allow for higher drive voltages was disclosed by Uchida (1988). The author demonstrated a glass light bulb containing a printed circuit board holding several LEDs connected in series. The circuit also contained a series resistor. The lamp was shown to be suitable for power supplies of 100 V.

Collins *et al.* (2003) disclosed a monolithically integrated LED circuit. The individual LEDs were connected through metal interconnects allowing for series- and parallel-circuit configurations. Two LEDs located on the same substrate and interconnected in series by a metal interconnect are shown in **Figure 10.10** (Collins *et al.*, 2003). HV LEDs emitting white light are commercially available and have been offered for various DC supply voltages including about 20 V and 48 V.

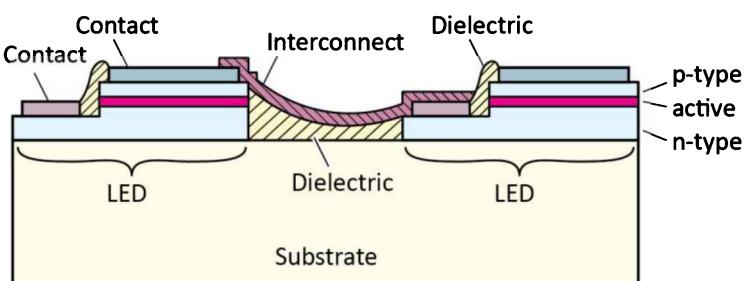


Fig. 10.10: Monolithic integration of two LEDs by a metal interconnect. The interconnect line connects a cathode with an anode resulting in a series circuit (after Collins *et al.*, 2003).

10.6 References

- Ao J.-P., Sato H., Mizobuchi T., Marioka K., Kawano S., Muramoto Y., Lee Y.-B., Sato D., Ohno Y., and Sakai S. "Monolithic blue LED series arrays for high-voltage AC operation" *Physica Status Solidi (a)* **194**, 376 (2002)
- Bar-Chaim N., Margalit S., Yariv A. "GaAs Integrated Optoelectronics" *IEEE Transactions on Electron Devices* **ED-29**, 1372 (1982)
- Cho J., Jung J., Chae J. H., Kim H., Kim H., Lee J. W., Yoon S., Sone C., Jang T., Park Y., and Yoon E. "Alternating-current light emitting diodes with a diode bridge circuitry" *Japanese Journal of Applied Physics* **46**, L1194 (2007)
- Collins W. D., Bhat J. C., and Steigerwald D. A. "Monolithic series / parallel LED array formed on highly resistive substrates" US patent 6,547,249; filed 2001 (issued 2003)
- Dagenais M., Leheny R. F. and Crow J. "Integrated optoelectronics (Quantum electronics – Principles and applications)" 1994 Dagenais and Leheny (Academic Press, Elsevier Science and Technology, Maryland Heights, Missouri, 1994)
- Ebeling K. J. "Integrated optoelectronics: Waveguide optics, photonics, semiconductors" (Springer-Verlag, Berlin, 1993)
- Jiang H., Lin J., Manhattan, and Jin S. "Light emitting diodes for high AC voltage operation and general lighting" United States Patent Application Publication US 2005 / 0185401 A1 (2005)
- Kilby J. S. "Miniaturized electronic circuits" US patent 3,138,743; filed 1959 (issued 1964)
- Lee C. H., Speck J. S., Kim H. S., Kim J. J., Kim S. H., and Lee J. H. "Light-emitting device for AC power operation" United States Patent Application Publication US 2008 / 0211421 A1 (2008)
- Liu H. C., Li J., Wasilewski Z. R. and Buchanan M. "Integrated quantum well intersub-band photodetector and light emitting diode" *Electronics Letters* **31**, 832 (1995)
- Noyce R. N. "Semiconductor device-and-lead structure" US patent 2,981,877; filed 1959 (issued 1961)
- Onushkin G. A., Lee Y.-J., Yang J.-J., Kim H.-K., Son J.-K., Park G.-H., and Park Y.-J. "Efficient alternating current operated white light-emitting diode chip" *IEEE Photonics Technology Letters* **21**, 33 (2009)
- Seoul Semiconductor Company, Seoul, South Korea, Model "A4-AC Acriche LED" (2009)
- Spitzer S. M., Schumaker N. E., Koszi L. A., and North J. C. "Multijunction AC or DC integrated GaP light emitting diode array" *IEEE Transactions on Electron Devices* **ED-22**, 701 (1975)
- Uchida A. "LED lamp" US patent 4,727,289; filed 1986 (issued 1988)
- Yen H.-H., Yeh W.-Y., and Kuo H.-C. "GaN alternating current light-emitting device" *Physica Status Solidi (a)* **204**, 2077 (2007)
- Yen H.-H., Kuo H.-C., and Yeh W.-Y. "Characteristics of single-chip GaN-based alternating current light-emitting diode" *Japanese Journal of Applied Physics* **47**, 8808 (2008)

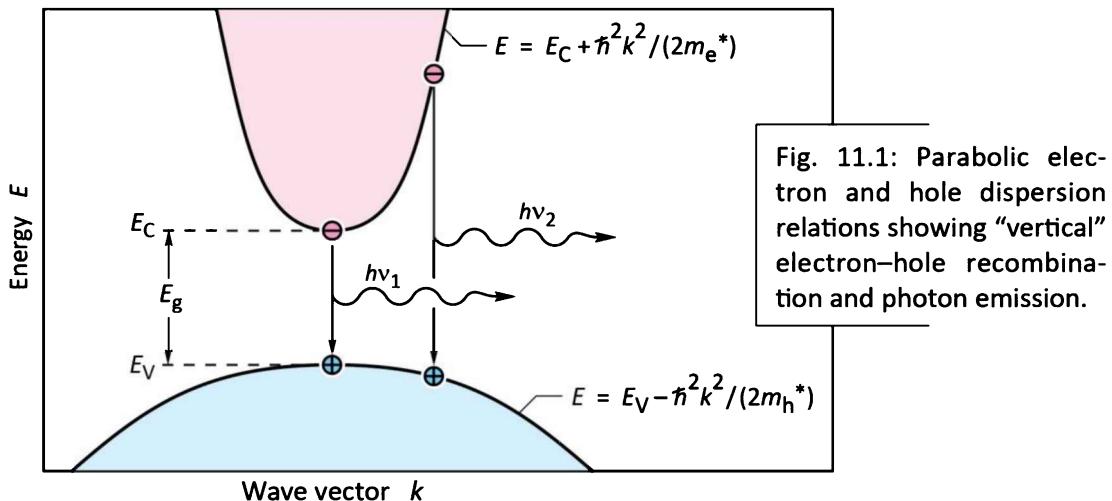
11

11 – Basic optical properties

11.1 Emission spectrum

The physical mechanism by which semiconductor LEDs emit light is spontaneous recombination of electron–hole pairs and simultaneous emission of photons. The spontaneous emission process is fundamentally different from the stimulated emission process occurring in semiconductor lasers and superluminescent LEDs. Spontaneous recombination has certain characteristics that determine the optical properties of LEDs. The properties of spontaneous emission in LEDs will be discussed in this section.

An electron–hole recombination process is illustrated schematically in **Figure 11.1**. Electrons in the conduction band and holes in the valence band are assumed to have the parabolic dispersion relations



$$E = E_C + \frac{\hbar^2 k^2}{2 m_e^*} \quad (\text{for electrons}) \quad (11.1)$$

and

$$E = E_V - \frac{\hbar^2 k^2}{2 m_h^*} \quad (\text{for holes}) \quad (11.2)$$

where m_e^* and m_h^* are the electron and hole effective masses, \hbar is Planck’s constant divided by

2π , k is the carrier wave number, and E_C and E_V are the conduction and valence band-edge energies, respectively.

The requirement of energy and momentum conservation leads to further insight into the radiative recombination mechanism. It follows from the Boltzmann distribution that electrons and holes have an average kinetic energy of kT . Energy conservation requires that the photon energy is given by the difference between the electron energy, E_e , and the hole energy, E_h , i.e.

$$h\nu = E_e - E_h \approx E_g . \quad (11.3)$$

The photon energy is approximately equal to the bandgap energy, E_g , if the thermal energy is small compared with the bandgap energy $kT \ll E_g$. Thus, the desired emission wavelength of an LED can be attained by choosing a semiconductor material with appropriate bandgap energy. For example, GaAs has a bandgap energy of 1.42 eV at room temperature and thus GaAs LEDs emit at the infrared wavelength of 870 nm.

It is helpful to compare the average carrier momentum with the photon momentum. A carrier with kinetic energy kT and effective mass m^* has the momentum

$$p = m^*v = \sqrt{2m^*\frac{1}{2}m^*v^2} = \sqrt{2m^*kT} . \quad (11.4)$$

The momentum of a photon with energy E_g can be derived from the de Broglie relation

$$p = \hbar k = \frac{h\nu}{c} = \frac{E_g}{c} . \quad (11.5)$$

Calculation of the carrier momentum (using Eqn. 11.4) and the photon momentum (using Eqn. 11.5) yields that the carrier momentum is *orders of magnitude larger* than the photon momentum. Therefore, the electron momentum cannot change significantly during the transition from the conduction to the valence band. The transitions are therefore “vertical” as shown in **Figure 11.1**, i.e. electrons only recombine with holes that have the same momentum or k value.

Using the requirement that electron and hole momenta are the same, the photon energy can be written as the *joint dispersion relation*

$$h\nu = E_C + \frac{\hbar^2 k^2}{2m_e^*} - E_V + \frac{\hbar^2 k^2}{2m_h^*} = E_g + \frac{\hbar^2 k^2}{2m_r^*} \quad (11.6)$$

where m_r^* is the reduced mass given by

$$\frac{1}{m_r^*} = \frac{1}{m_e^*} + \frac{1}{m_h^*} . \quad (11.7)$$

Using the joint dispersion relation, the joint density of states can be calculated and one obtains

$$\rho(E) = \frac{1}{2\pi^2} \left(\frac{2 m_r^*}{\hbar^2} \right)^{3/2} \sqrt{E - E_g} . \quad (11.8)$$

The distribution of carriers in the allowed bands is given by the Boltzmann distribution, i.e.

$$f_B(E) = e^{-E/kT} . \quad (11.9)$$

The *emission intensity* as a function of energy is proportional to the product of Eqns. (11.8) and (11.9),

$$I(E) \propto \sqrt{E - E_g} e^{-E/kT} . \quad (11.10)$$

The lineshape of an LED, as given by Eqn. (11.10), is shown in **Figure 11.2**. The maximum emission intensity occurs at

$$E = E_g + \frac{1}{2} kT . \quad (11.11)$$

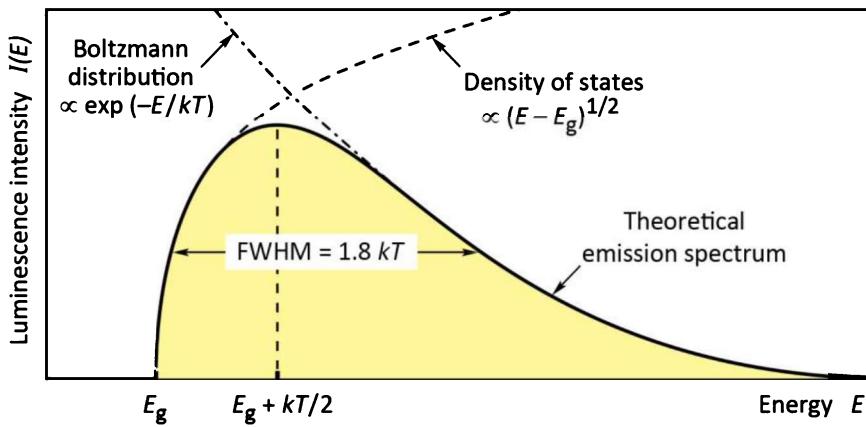


Fig. 11.2: Theoretical emission spectrum of an LED. The full-width at half-maximum (FWHM) of the emission line is equal to $1.8kT$.

The full-width at half-maximum of the emission is

$$\Delta E = 1.8 kT \quad \text{or} \quad \Delta\lambda = \frac{1.8 kT \lambda^2}{h c} . \quad (11.12)$$

For example, the theoretical room-temperature linewidth of a GaAs LED emitting at 870 nm is $\Delta E = 46 \text{ meV}$ or $\Delta\lambda = 28 \text{ nm}$.

The spectral linewidth of LED emission is important in several respects. Firstly, the linewidth of an LED emitting in the visible range is relatively narrow compared with the range of the entire visible spectrum. The LED emission is even narrower than the spectral width of a single color as perceived by the human eye. For example, *red* colors range in wavelength from 625 nm to 730 nm, which is much wider than the typical emission spectrum of an LED. Therefore, LED emission is perceived by the human eye as *monochromatic*.

Secondly, optical fibers are dispersive, which leads to a range of propagation velocities for a light pulse consisting of a range of wavelengths. The material dispersion in optical fibers limits the “*bit rate × distance product*” achievable with LEDs.

The spontaneous lifetime of carriers in LEDs in direct-gap semiconductors is of the order of 1~100 ns depending on the active region doping concentration (or carrier concentrations) and the material quality. Thus, modulation speeds up to 1 Gbit/s are attainable with LEDs.

11.2 The light escape cone

Light generated inside a semiconductor cannot escape from the semiconductor if it is totally internally reflected at the semiconductor–air interface. If the angle of incidence of a light ray is close to normal incidence, light can escape from the semiconductor. However, total internal reflection occurs for light rays with oblique and grazing-angle incidence. Total internal reflection reduces the external efficiency significantly, in particular for LEDs consisting of high-refractive-index materials.

Assume that the angle of incidence inside the semiconductor at the semiconductor–air interface is given by ϕ . Then the angle of incidence of the refracted ray, Φ , can be inferred from Snell’s law of refraction

$$\bar{n}_s \sin \phi = \bar{n}_{air} \sin \Phi \quad (11.13)$$

where \bar{n}_s and \bar{n}_{air} are the refractive indices of the semiconductor and air, respectively. The ***critical angle for total internal reflection*** is obtained using $\Phi = 90^\circ$, as illustrated in ***Figure 11.3 (a)***. Using Snell’s law, one obtains

$$\sin \phi_c = \frac{\bar{n}_{air}}{\bar{n}_s} \sin 90^\circ = \frac{\bar{n}_{air}}{\bar{n}_s}$$
(11.14a)

and

$$\phi_c = \arcsin \frac{\bar{n}_{air}}{\bar{n}_s}.$$
(11.14b)

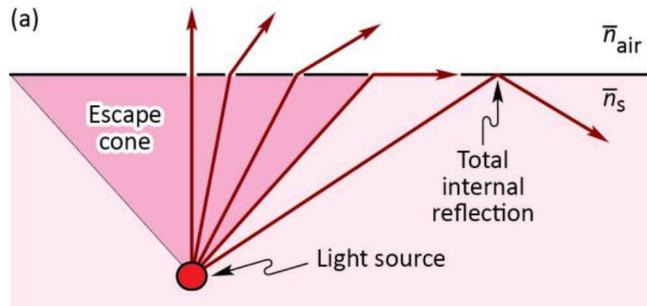
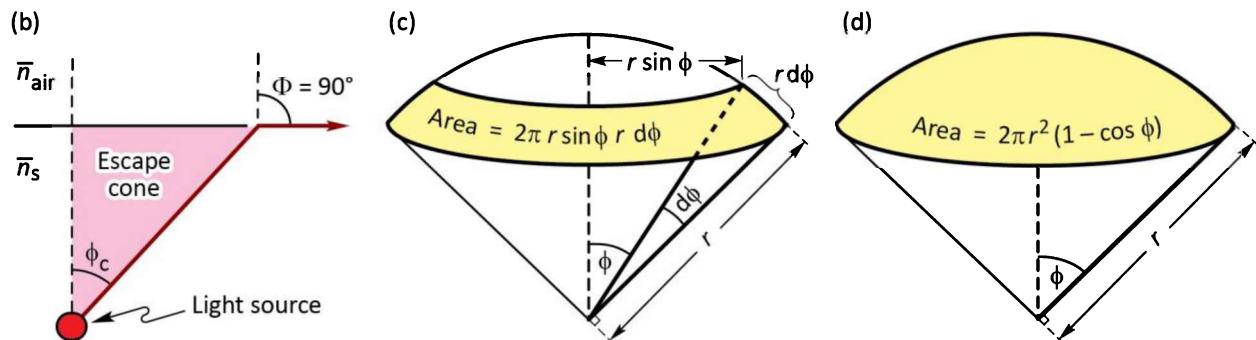


Fig. 11.3: (a) Illustration of refraction and total internal reflection at a semiconductor-air interface. (b) Definition of the escape cone by the critical angle ϕ_c . (c) Area element dA . (d) Area of calotte-shaped section of the sphere defined by radius r and angle ϕ_c .



The refractive indices of semiconductors are usually quite high. For example, GaAs has a refractive index of 3.4. Thus, according to Eqn. (11.14), the critical angle for total internal reflection is quite small. In this case, we can use the approximation $\sin \phi_c \approx \phi_c$. The critical angle for total internal reflection is then given by

$$\phi_c \approx \frac{\bar{n}_{air}}{\bar{n}_s}.$$
(11.15)

The angle of total internal reflection defines the light escape cone. Light emitted into the cone can escape from the semiconductor, whereas light emitted outside the cone is subject to total internal reflection.

Next, we calculate the surface area of the spherical cone with radius r in order to determine the total fraction of light that is emitted into the light escape cone. The surface area of the calotte-shaped surface shown in **Figures** 11.3 (b) and (c) is given by the integral

$$A = \int dA = \int_{\phi=0}^{\phi_c} 2\pi r \sin \phi \ r \ d\phi = 2\pi r^2 (1 - \cos \phi_c) . \quad (11.16)$$

Let us assume that light is emitted from a point-like source in the semiconductor with a total power of P_{source} . Then the power that can escape from the semiconductor is given by

$$P_{escape} = P_{source} \frac{2\pi r^2 (1 - \cos \phi_c)}{4\pi r^2} \quad (11.17)$$

where $4\pi r^2$ is the entire surface area of the sphere with radius r .

The calculation indicates that only a fraction of the light emitted inside a semiconductor can escape from the semiconductor. The fraction is given by

$$\frac{P_{escape}}{P_{source}} = \frac{1}{2} (1 - \cos \phi_c) . \quad (11.18)$$

Because the critical angle of total internal reflection for high-index materials is relatively small, the cosine term can be expanded into a power series. Neglecting higher-than-second-order terms yields

$$\frac{P_{escape}}{P_{source}} \approx \frac{1}{2} \left[1 - \left(1 - \frac{\phi_c^2}{2} \right) \right] = \frac{1}{4} \phi_c^2 . \quad (11.19)$$

Using the approximation of Eqn. (11.15), one obtains

$$\frac{P_{escape}}{P_{source}} \approx \frac{1}{4} \frac{\bar{n}_{air}^2}{\bar{n}_s^2} . \quad (11.20)$$

The escape problem is a significant problem for high-efficiency LEDs. In most semiconductors, the refractive index is quite high ($\bar{n} > 2.5$) and thus only a small percentage of the light generated in the semiconductor can escape from a planar LED. The problem is less significant in semiconductors with a small refractive index and for polymers, which have refractive indices of the order of 1.5.

Exercise: Light escape from planar GaAs, GaN, and polymer LED structures. The refractive indices of GaAs, GaN, and light-emitting polymers are 3.4, 2.5, and 1.5, respectively. Calculate the critical angle of total internal reflection for GaAs, GaN, and for polymers. Also calculate the fraction of light power that can escape from a planar GaAs and GaN semiconductor structures and a polymer LED structure.

What improvement can be attained if a planar GaAs LED is encapsulated in a transparent polymer of refractive index 1.5, if the reflection at the polymer–air interface is neglected?

Solution:

Critical angle for total internal reflection:

$$\text{GaAs: } \phi_c = 17.1^\circ \quad \text{GaN: } \phi_c = 23.6^\circ \quad \text{Polymer: } \phi_c = 41.8^\circ.$$

Fraction of light that can escape:

$$\text{GaAs: } 2.21\% \quad \text{GaN: } 4.18\% \quad \text{Polymer: } 12.7\%.$$

Improvement of the GaAs planar LED due to polymer encapsulation: 232%.

11.3 Emission pattern

All LEDs have a certain **emission pattern, radiation pattern** or **far-field pattern**. The intensity, measured in W/cm², depends on the longitudinal and azimuth angle and the distance from the LED. The total optical power emitted by the LED is obtained by integration over the area of a sphere.

$$P = \int_A \int_{\lambda} I(\lambda) d\lambda dA \quad (11.21)$$

where $I(\lambda)$ is the **spectral light intensity** (measured in W per nm per cm²) and A is the surface area of the sphere. The integration is carried out over the entire surface area.

11.4 The lambertian emission pattern

The index contrast between the light-emitting material and the surrounding material leads to a non-isotropic emission pattern. For high-index light-emitting materials with a planar surface, a lambertian emission pattern is obtained. **Figure 11.4** illustrates a point-like light source located a short distance below a semiconductor–air interface. Consider a light ray emitted from the source at an angle ϕ with respect to the surface normal. The light ray is refracted at the semiconductor–air interface and the refracted light ray has an angle Φ with respect to the surface normal. The two angles are related by Snell's law, which, for small angles of ϕ (for which $\sin \phi \approx \phi$), can be written as

$$\bar{n}_s \phi = \bar{n}_{air} \sin \Phi \quad (11.22)$$

Light emitted into the angle $d\phi$ in the semiconductor is emitted into the angle $d\Phi$ in air as shown in **Figure 11.4** (a). Differentiating the equation with respect to Φ , and solving the resulting equation for $d\Phi$, yields

$$d\Phi = \frac{\bar{n}_s}{\bar{n}_{air}} \frac{1}{\cos \Phi} d\phi . \quad (11.23)$$

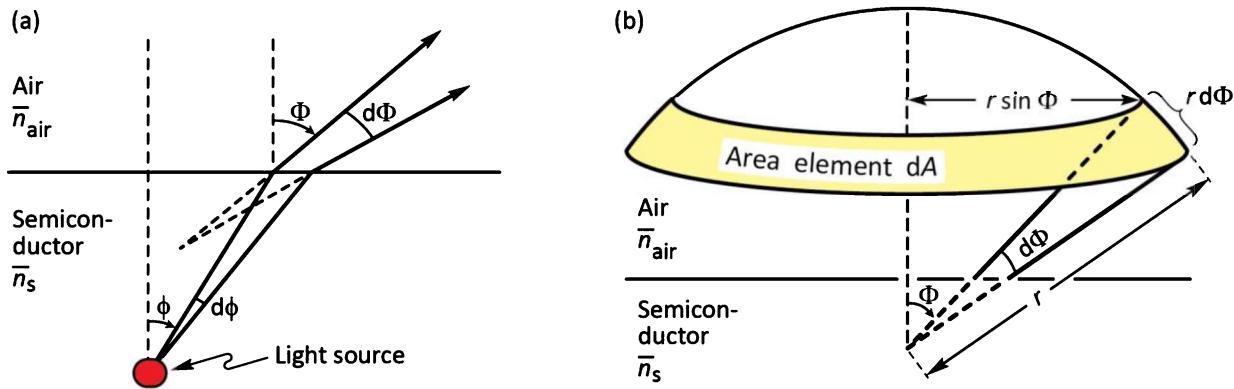


Fig. 11.4: Geometrical model used to derive the lambertian emission pattern. (a) The light emitted into angle $d\phi$ inside the semiconductor is emitted into the angle $d\Phi$ in air.
(b) Illustration of the area element dA of the calotte-shaped section of the sphere.

Power conservation requires that the optical power emitted into the angle $d\phi$ in the semiconductor be equal to the optical power emitted into the angle $d\Phi$ in air, i.e.

$$I_s dA_s = I_{air} dA_{air} \quad (11.24)$$

where I_s and I_{air} are the light intensities (measured in units of W/m^2) in the semiconductor and air, respectively. Owing to the cylindrical symmetry of the emission pattern we choose the area element shown in **Figure 11.4 (b)**. The area element in air is given by

$$dA_{air} = 2\pi r \sin \Phi r d\Phi . \quad (11.25)$$

Using Eqns. (11.23) and (11.24) yields

$$dA_{air} = 2\pi r^2 \frac{\bar{n}_s^2}{\bar{n}_{air}^2} \frac{1}{\cos \Phi} \phi d\phi \quad (11.26)$$

Similarly, the surface element in the semiconductor is given by

$$dA_s = 2\pi r \sin \Phi r d\phi \approx 2\pi r^2 \phi d\phi . \quad (11.27)$$

The light intensity in the semiconductor at a distance r from the light source is given by the total source power divided by the surface area of a sphere with radius r , i.e.

$$I_s = \frac{P_{source}}{4\pi r^2} . \quad (11.28)$$

The light intensity in air can then be inferred from Eqns. (11.24), (11.26), (11.27), and (11.28).

One obtains the ***lambertian emission pattern*** given by

$$I_{air} = \frac{P_{source}}{4\pi r^2} \frac{\bar{n}_{air}^2}{\bar{n}_s^2} \cos \Phi . \quad (11.29)$$

The lambertian emission pattern follows a cosine dependence on the angle Φ (**Lambert's cosine law**). The intensity is highest for emission normal to the semiconductor surface, i.e. for $\Phi = 0^\circ$. At an angle of $\Phi = 60^\circ$, the intensity decreases to half of its maximum value. The lambertian emission pattern is shown schematically in **Figure 11.5**.

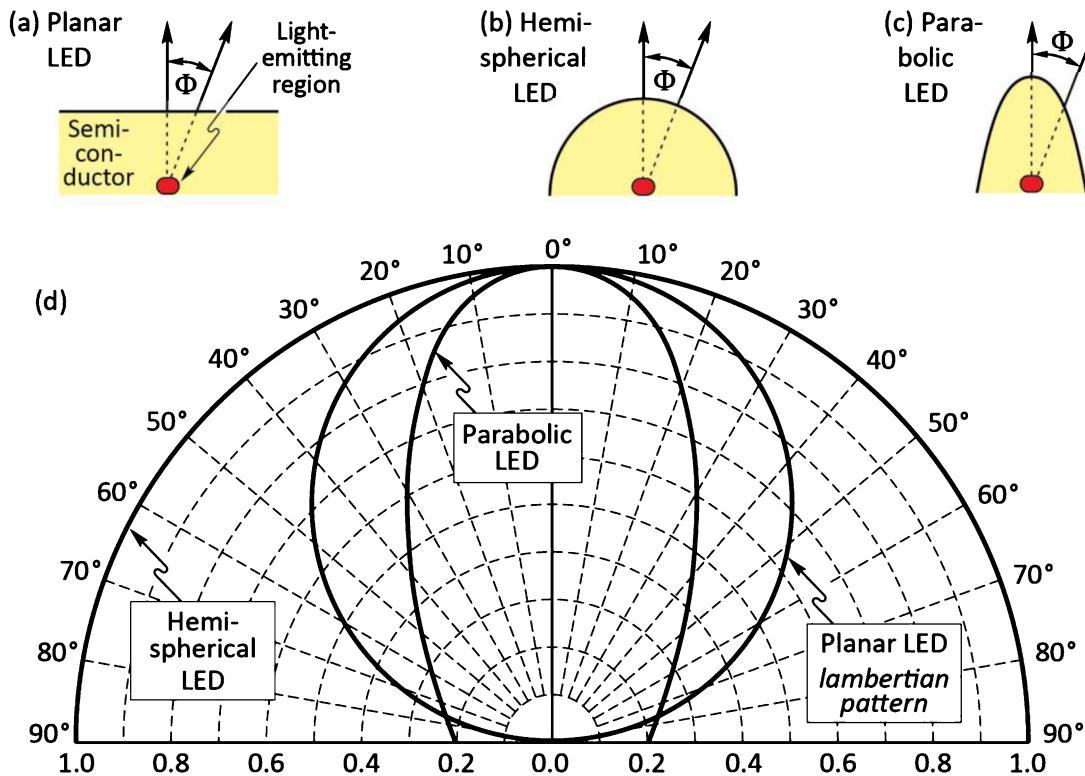


Fig. 11.5: Light-emitting diodes with (a) planar, (b) hemispherical, and (c) parabolic surfaces. (d) Far-field patterns of the different types of LEDs. At an angle of $\Phi = 60^\circ$, the lambertian emission pattern decreases to 50% of its maximum value occurring at $\Phi = 0^\circ$. The three emission patterns are normalized to unity intensity at $\Phi = 0^\circ$.

Several other surface shapes are also shown in **Figure 11.5**. These non-planar surfaces exhibit various emission patterns. An isotropic emission pattern is obtained for hemispherically shaped

LEDs, which have the light-emitting region in the center of the sphere. A strongly directed emission pattern can be obtained in LEDs with parabolically shaped surfaces. However, both hemispherical as well as parabolic surfaces are difficult to fabricate.

The total power emitted into air can be calculated by integrating the intensity over the entire hemisphere. The total power is then given by

$$P_{air} = \int_{\Phi=0^\circ}^{90^\circ} I_{air} 2\pi r \sin \Phi \ r d\Phi . \quad (11.30)$$

By using the Lambertian emission pattern for I_{air} in Eqn. (11.30) and using $\cos \Phi \sin \Phi = \frac{1}{2}\sin(2\Phi)$, the integral can be calculated to yield

$$P_{air} = \frac{P_{source}}{4} \frac{\bar{n}_{air}^2}{\bar{n}_s^2} . \quad (11.31)$$

This result is identical to Eqn. (11.20). This is not surprising because the light power that escapes from the semiconductor (P_{escape}) must be identical to the power in air (P_{air}).

In the calculation above, **Fresnel reflection** at the semiconductor–air interface has been neglected. At normal incidence, the Fresnel power transmittance is given by

$$T = 1 - R = 1 - \left(\frac{\bar{n}_s - \bar{n}_{air}}{\bar{n}_s + \bar{n}_{air}} \right)^2 = \frac{4 \bar{n}_s \bar{n}_{air}}{(\bar{n}_s + \bar{n}_{air})^2} . \quad (11.32)$$

Fresnel reflection losses must be taken into account in a rigorous calculation.

Exercise: LED-to-fiber coupling efficiency. Consider a GaAs LED with a point-like light-emitting region located in close proximity to the planar GaAs LED surface. An optical fiber has an acceptance angle of 12° in air. What fraction of the light emitted by the active region can be coupled into the fiber? Assume a GaAs refractive index of 3.4. Neglect Fresnel reflection losses at the semiconductor–air and air–fiber interfaces.

Solution: The acceptance angle in the semiconductor is obtained from Snell's law and is 3.5° . Thus, 0.093% of the power emitted by the active region can be coupled into the fiber.

11.5 Encapsulants

The light-extraction efficiency can be enhanced by using dome-shaped encapsulants with a large refractive index. As a result of the encapsulation, the angle of total internal reflection through the top surface of the semiconductor is increased (Nuese *et al.*, 1969). It follows from Eqn. (11.18) that the ratio of extraction efficiency with and without encapsulant is given by

$$\frac{\eta_{\text{encapsulant}}}{\eta_{\text{air}}} = \frac{1 - \cos \phi_{c,\text{encapsulant}}}{1 - \cos \phi_{c,\text{air}}} \quad (11.33)$$

where $\phi_{c,\text{encapsulant}}$ and $\phi_{c,\text{air}}$ are the critical angles for total internal reflection at the semiconductor–encapsulant and semiconductor–air interface, respectively. **Figure 11.6** shows the calculated ratio of the extraction efficiency with and without an encapsulant dome. Inspection of the figure yields that the efficiency of a typical semiconductor LED increases by a factor of 2~3 upon encapsulation. Typical encapsulation materials, such as epoxy and silicone, have a refractive index of about 1.5.

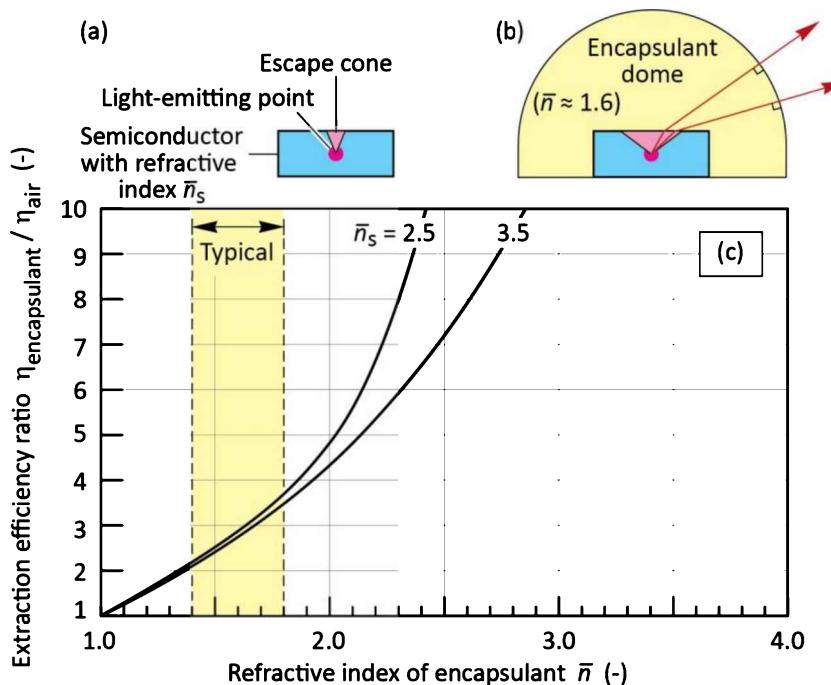


Fig. 11.6: LED (a) without and (b) with dome-shaped encapsulant. A larger escape angle is obtained for the LED with encapsulant dome. (c) Calculated ratio of light extraction efficiency emitted through the top surface of a planar LED with and without encapsulant dome. The refractive indices of typical encapsulant materials range from 1.4 to 1.8 (adapted from Nuese *et al.*, 1969).

The inset of **Figure 11.6** shows that light is incident at an angle of approximately 90° at the encapsulation–air interface due to the dome-shape of the encapsulant. Thus, total internal reflection losses do not occur at the encapsulant–air interface. Besides improving the external efficiency of an LED, the encapsulant can also be used as a spherical lens for applications requiring a directed emission pattern. Note that in organic polymer LEDs, encapsulants increase the extraction efficiency by only a small amount due to the inherently small refractive index of polymers.

Advanced encapsulants including graded-index encapsulants, encapsulants with a high refractive index ($\bar{n} > 2.0$), and encapsulants containing mineral diffusers will be discussed in the

chapter on packaging.

11.6 Temperature dependence of emission intensity

The emission intensity of LEDs decreases with increasing temperature. This decrease of the emission intensity is due to several temperature-dependent factors including (i) non-radiative recombination via deep levels, (ii) surface recombination, and (iii) carrier loss over heterostructure barriers.

Near room temperature, the temperature dependence of the LED emission intensity is frequently described by the phenomenological equation

$$I = I|_{300\text{ K}} \exp -\frac{T - 300\text{ K}}{T_1} \quad (11.34)$$

where T_1 is the ***characteristic temperature***. A *high* characteristic temperature, implying *weak* temperature dependence, is desirable.

It is interesting to note that both LEDs as well as semiconductor lasers have a distinct temperature dependence of the emission intensity. In LEDs, the decrease is expressed in terms of the “ T_1 equation”. In semiconductor lasers, the threshold current, i.e. the electrical current required for the onset of lasing, increases. In lasers the increase in threshold current is expressed in terms of the well-known “ T_0 equation”. This equation is given by

$$I_{th} = I_{th}|_{300\text{ K}} \exp \frac{T - 300\text{ K}}{T_0} \quad (11.35)$$

where I_{th} is the threshold current of the laser. Note the formal similarity of the “ T_1 equation” (Eqn. 11.34) and the “ T_0 equation” (Eqn. 11.35). Both equations are purely phenomenological equations intended to describe the experimental results without a strong theoretical framework allowing the derivation of the equations from basic principles.

Figure 11.7 shows experimental results for the temperature dependence of the emission intensity (Toyoda Gosei Corporation, 2000), specifically the temperature dependence of the emission intensity at a constant current for a GaInN/GaN blue, a GaInN/GaN green, and a AlGaInP/GaAs red LED. Inspection of the figure reveals that the blue LED has the highest T_1 and the red LED has the lowest T_1 . III–V nitride LEDs have deeper wells so that carrier confinement is more effective in III–V nitride structures than in the III–V phosphide structures.

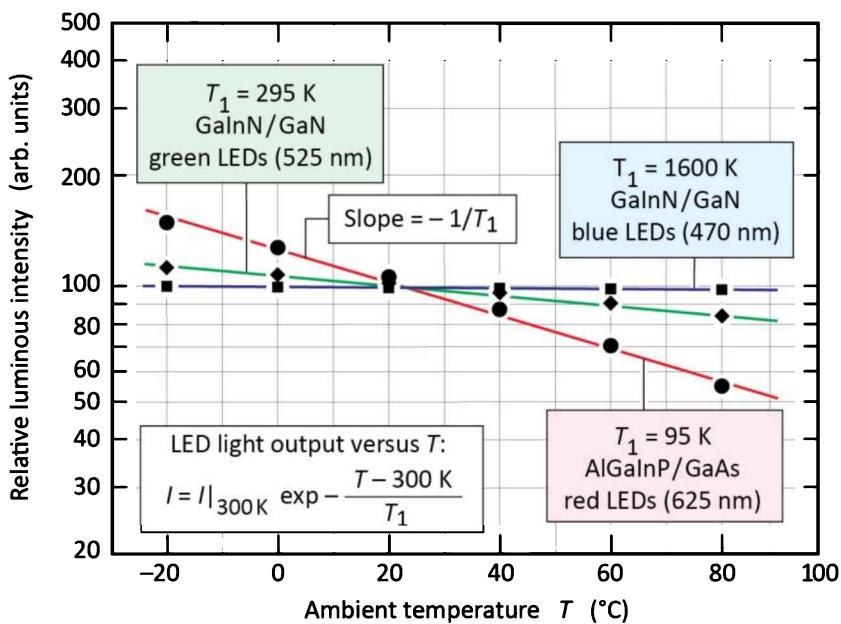


Fig. 11.7: Characteristic temperature T_1 of GaInN/GaN blue, GaInN/GaN green, and AlGaNp/GaAs red LEDs near room temperature (after data from Toyoda Gosei Company, 2000). Additional data by Toyoda Gosei Company (2004) show the following values for T_1 : Blue GaInN LED, 460 nm, $T_1 = 1600$ K; Cyan GaInN LED, 505 nm, $T_1 = 832$ K; Green GaInN LED, 525 nm, $T_1 = 341$ K; Red AlGaNp LED, 625 nm, $T_1 = 199$ K.

Note that the characteristic temperature can strongly depend on the layer structure of the active region, particularly the presence or absence of an electron-blocking layer. For blue GaInN LEDs emitting at 475 nm, Huh and Park (2004) and Huh *et al.* (2004) showed that the characteristic temperature can be as low as 62.5 K (without an electron-blocking layer) and as high as 2 200 K (with an electron-blocking layer). Characteristic temperatures in GaInN LEDs ranging from about 1 260 K to 1 640 K were reported by Lee *et al.* (2010).

11.7 References

- Huh C. and Park S.-J. "Effects of temperature on InGaN / GaN LEDs with different MQW structures" *Electrochemical and Solid-State Letters* **7**, G266 (2004)
- Huh C., Schaff W. J., Eastman L. F., and Park S.-J. "Temperature dependence of light-output performance of InGaN/GaN multiple-quantum well light-emitting diodes with various In compositions" *Proceedings of SPIE* **5187**, 330 (2004)
- Lee H. K., Yu J. S., and Lee Y. T. "Thermal analysis and characterization of the effect of substrate thinning on the performances of GaN-based light emitting diodes" *Physica Status Solidi A* **207**, 1497 (2010)
- Nuese C. J., Tietjen J. J., Gannon J. J., and Gossenberger H. F. "Optimization of electroluminescent efficiencies for vapor-grown GaAsP diodes" *Journal of the Electrochemical Society: Solid State Science* **116**, 248 (1969)
- Toyoda Gosei Corporation, Japan, General LED catalogue (2000)
- Toyoda Gosei Corporation, Japan, General LED catalogue (2004)

12

12 – Reflectors

Ideal reflectors incorporated into a device structure should have (i) high reflectivity, (ii) a sufficiently broad spectral range of the high-reflectivity band, (iii) omnidirectional characteristics, and (iv) low resistivity (provided current flows through the reflector). It is a question of great practical and intellectual interest what type of reflector best meets these multiple requirements.

Different types of reflectors are shown in *Figure 12.1* including a metallic reflector, distributed Bragg reflector (DBR), hybrid metal–DBR, reflectors based on total internal reflection (TIR), and an omni-directional reflector (ODR). The characteristics of the different types of reflectors will be discussed below.

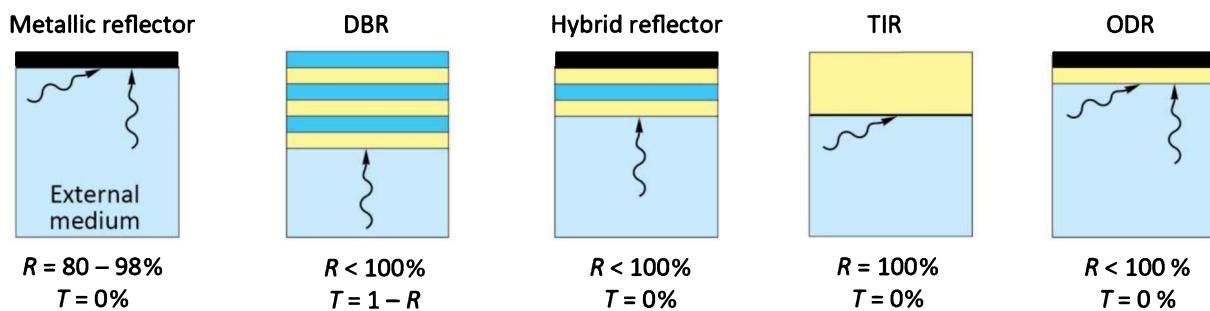


Fig. 12.1: Metallic reflector, distributed Bragg reflector (DBR), hybrid reflector, total internal reflector (TIR), and a triple-layer omni-directional reflector (ODR). Also given are angles of incidence for high reflectivity and typical reflectances and transmittances.

An **external medium** is indicated in the figure. In an LED structure with a reflector, the external medium is a semiconductor. The external medium has a significant influence on the reflector properties. For example, the reflectivity of a metal–semiconductor reflector is lower compared to a metal–air reflector.

12.1 Metallic reflectors, reflective contacts, and transparent contacts

Having been used by humankind for several millenniums, metal–air reflectors are the oldest type of reflector with high reflectivity. Metal reflectors are characterized by a broad spectral

reflectivity band and a weak angular dependence of the reflectivity. The first high-quality metallic reflectors were used in reflection telescopes for astronomy applications (Bell, 1922).

Experimental reflectance spectra of metal–air reflectors at normal angle of incidence are shown in *Figure 12.2* (Spisz, 1969; Mochizuki, 2013). The metal reflectance spectra are characterized by a broad spectral band of high reflectivity. Inspection of the figure reveals that Al, Ag, and Au are particularly good reflectors at UV, visible, and IR wavelengths, respectively.

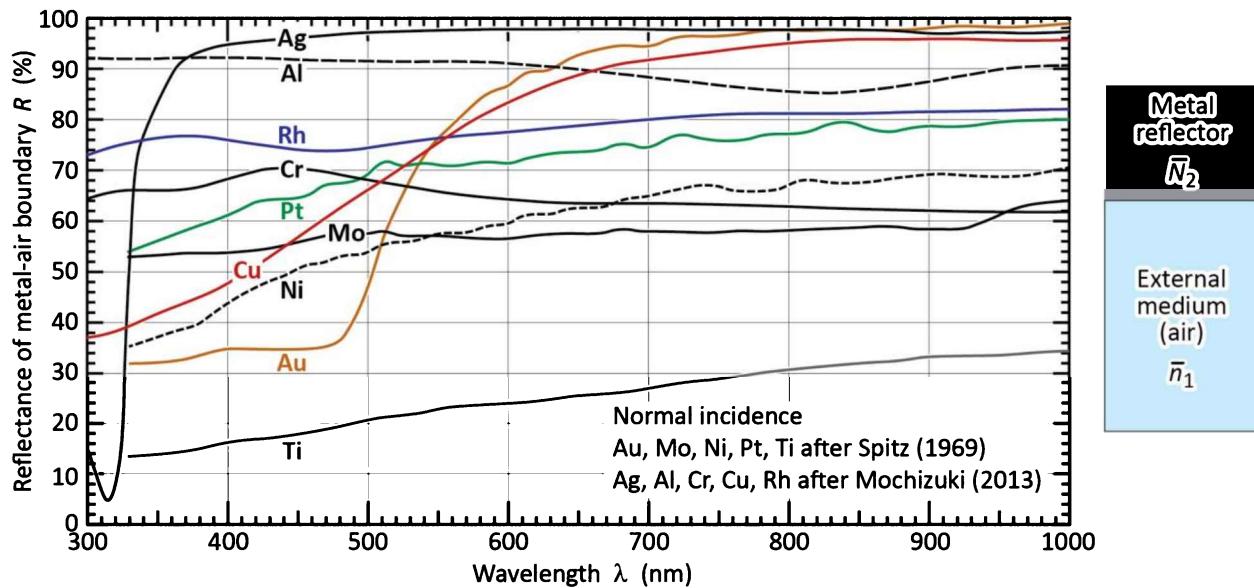


Fig. 12.2: Measured reflectance of various metals as a function of wavelength for a normal angle of incidence. Ag (silver), Al (aluminium), Au (gold), and Cu (copper) are good reflectors with $R > 90\%$ at some wavelengths (after Spisz, 1969, and Mochizuki, 2013).

The reflectance (amplitude reflection coefficient at normal incidence) of a metal reflector and an external medium can be calculated from the Fresnel equation, which for media with complex refractive indices is given by⁸

$$r = \frac{\vec{E}_r}{\vec{E}_i} = \frac{\bar{N}_1 - \bar{N}_2}{\bar{N}_1 + \bar{N}_2} \quad (12.1)$$

where \bar{N}_1 and \bar{N}_2 is the complex refractive index of a first medium and second medium (the metal), respectively. The complex refractive index has a real part, \bar{n} , (called refractive index) and an imaginary part, $i\bar{k}$, (called extinction coefficient) so that $\bar{N} = \bar{n} + i\bar{k}$.⁹ If the complex

⁸ The equation applies to normal incidence; for non-normal incidence, the reader is referred to the literature (Saleh and Teich, 2007).

⁹ A plane wave includes the oscillating factor $\exp(-i\bar{N}kx)$, where $k = 2\pi/\lambda$ and x is the propagation direction.

refractive index of a medium has only a real part, then an electromagnetic wave propagates unattenuated. If the complex refractive index of a medium has a non-zero imaginary part ($\bar{k} \neq 0$), then the propagating wave is attenuated. For dielectric materials, it is generally $\bar{k} \approx 0$; in contrast, metals generally have large values of \bar{k} , so that waves are rapidly attenuated.

The reflected and transmitted *intensities* are proportional to the square of the electric field. The **reflectivity** or **reflectance** (power reflection coefficient) is given by

$$R = \frac{|\vec{E}_r|^2}{|\vec{E}_i|^2} = |r|^2 = \left| \frac{\bar{N}_1 - \bar{N}_2}{\bar{N}_1 + \bar{N}_2} \right|^2 = \frac{|\bar{N}_1 - \bar{N}_2|^2}{|\bar{N}_1 + \bar{N}_2|^2}. \quad (12.2)$$

The energy conservation law requires that the power transmission coefficient in a lossless reflector is given by

$$T = 1 - R. \quad (12.3)$$

Metals are lossy media so that the transmittance $T \approx 0$ for thick metal films. The Fresnel equation can be used to calculate the reflectance of a metal–dielectric interface. Assuming that the complex refractive indices of the dielectric and the metal are $\bar{N}_1 = \bar{n}_1$ and $\bar{N}_2 = \bar{n}_2 + i\bar{k}_2$, respectively, the reflectance is given by

$$r = \frac{\bar{n}_1 - (\bar{n}_2 + i\bar{k}_2)}{\bar{n}_1 + (\bar{n}_2 + i\bar{k}_2)} \quad \text{and} \quad R = \frac{(\bar{n}_1 - \bar{n}_2)^2 + \bar{k}_2^2}{(\bar{n}_1 + \bar{n}_2)^2 + \bar{k}_2^2}. \quad (12.4)$$

An ideal metal has very high conductivity, i.e. $\sigma \rightarrow \infty$ and therefore $\bar{k} \rightarrow \infty$. Note that conductivity σ and the imaginary part of the refractive index \bar{k} are related by $\sigma = 2\bar{n}\omega\epsilon_0\bar{k}$. Also note that the absorption constant α and the imaginary part of the refractive index are related by $\alpha = 4\pi\bar{k}/\lambda_0$. We thus obtain for an ideal metal

$$|r| \approx 1 \quad \text{and} \quad R \approx 1 \quad \text{and} \quad \phi_r = \pi. \quad (12.5)$$

That is, the ideal metal has a unit reflection coefficient, unit reflectance, and a phase shift of π .

Real metals have high conductivity (but not an infinitely high conductivity) and the reflectivity of any real metal is therefore less than unity. Loss mechanisms in metals were first analyzed in terms of the Drude model (Drude, 1904). **Table 12.1** gives the \bar{n} and \bar{k} values of several metals and semiconductors.

Table 12.1: Real and imaginary part of the refractive index for different semiconductors and metals at 0.5 and 1.0 μm .

Material	GaP	GaP	Si	Ag	Ag	Au	Au	Al	Al
$\lambda =$	0.5 μm	1.0 μm	1.0 μm	0.5 μm	1.0 μm	0.5 μm	1.0 μm	0.5 μm	1.0 μm
\bar{n}	3.5	3.1	3.6	0.05	0.04	0.86	0.26	0.77	1.35
\bar{k}	≈ 0	≈ 0	≈ 0	3.1	7.1	1.90	6.82	6.08	10.7

Using Eqn. (12.4), the reflectance of a metal–air and metal–semiconductor boundary can be calculated; the results are given in *Table 12.2*. Inspection of the table reveals that the visible-spectrum reflectivity is generally lower for metal-semiconductor interfaces than that of metal-air interfaces. This is due to the lower refractive-index contrast of the metal-semiconductor boundary.

Table 12.2: Calculated reflectivity of metal–air and metal–semiconductor reflectors at 0.5 and 1.0 μm .

Material	R (%)	Material	R (%)	Material	R (%)
Ag / air (0.5 μm)	0.982	Al / air (0.5 μm)	0.923	Au / air (0.5 μm)	0.514
Ag / air (1.0 μm)	0.997	Al / air (1.0 μm)	0.955	Au / air (1.0 μm)	0.979
Ag / GaP (0.5 μm)	0.969	Al / GaP (0.5 μm)	0.805	Au / GaP (0.5 μm)	0.470
Ag / GaP (1.0 μm)	0.992	Al / GaP (1.0 μm)	0.876	Au / GaP (1.0 μm)	0.945
Ag / Si (1.0 μm)	0.991	Al / Si (1.0 μm)	0.861	Au / Si (1.0 μm)	0.939

Although metals are simple and viable reflectors, the reflection losses or **mirror losses** are quite high. The loss of one reflection event, $(1 - R)$, is about 5% for metal-semiconductor reflectors. Losses are particularly high for waveguided modes shown in *Figure 12.3*. The intensity of waveguided modes decays according to

$$\frac{I}{I_0} = R^N = (1 - L)^N \approx 1 - NL \quad (12.6)$$

where N is the number of reflection events, and the mirror loss is $L = 1 - R$ (where $L \ll 1.0$ and $R \approx 1.0$). The equation illustrates that a small difference (e.g. a few %) in R can make a large difference (e.g. a factor of 2) in the intensity of waveguided modes after N reflection events.

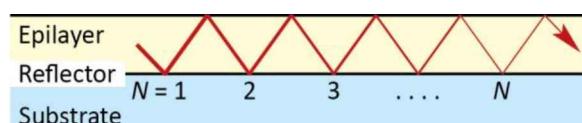


Fig. 12.3: Attenuation of waveguide mode due to lossy reflector.

Metal–semiconductor reflectors have been used in AlGaNp LEDs by Horng *et al.* (1999a, 1999b) to increase the light-extraction efficiency. The layer sequence of the finished device consisted of AlGaNp / AuBe / SiO₂ / Si. The p-type AlGaNp–AuBe interface served as a reflector and as a broad-area ohmic contact. The AuBe layer also served as a supply layer for Be acceptors to attain low contact resistance. The contacts were annealed at 450°C for 15 minutes. The LEDs were fabricated by a wafer-bonding process using a Si substrate. After the bonding process, the GaAs substrate, on which the epitaxial layers had been grown, was removed. Si has a higher thermal conductivity than GaAs thus allowing for lower junction temperatures and reducing the joule-heating-induced emission-wavelength shift. The emission from the metal-reflector AlGaNp LEDs exceeded those of reference LEDs with DBRs fabricated on GaAs substrates.

Annealing and alloying forms low-resistance ohmic contacts. Typical annealing temperatures for alloyed contacts are between 375 and 450°C for III–V arsenides and III–V phosphides and up to 600°C for III–V nitrides. During the annealing process, the metal surface changes from smooth to rough and a concomitant decrease in the optical reflectivity results.

In contrast, non-alloyed contacts are just deposited on the semiconductor without annealing. Highly doped semiconductor surface layers are needed to obtain good ohmic *I*–*V* characteristics for such non-alloyed ohmic contacts. However, even for highly doped semiconductors, the contact resistance of non-alloyed contacts is usually higher than for alloyed contacts.

Thick metallic reflectors and hybrid reflectors are absorbing reflectors that should not be used as light-exit reflectors in LEDs and vertical-cavity surface emitting lasers. Metal contacts become practically opaque for thicknesses > 50 nm. That is, the transmittance of hybrid reflectors is near zero, unless the thickness of the metal is very thin (Tu *et al.*, 1990).

Very thin metal contacts are semi-transparent. Most metal contacts have a transmittance of approximately 50% at a metal film thickness of 5~10 nm. The exact value of the transmittance needs to be calculated by taking into account the real and imaginary part of the refractive index (see, for example, Palik, 1998). However, very thin metallic contacts may form an islanded structure rather than a single continuous film. Furthermore, the electrical resistance of thin metal films can be large, in particular if an islanded structure is formed.

In LEDs with transparent substrates, e.g. AlGaNp LEDs on GaP, light emanating from the active region is incident on the substrate contact. To increase the reflectance of the backside, an ohmic contact geometry covering only a small fraction of the substrate surface can be used, such

as a multiple-stripe or a ring-shaped contact. Using Ag-loaded conductive epoxy to attach the LED die to the package provides a high-reflectivity material in the regions not covered by the ohmic contact.

The die-attach epoxy can also serve as a reflector in LEDs grown on a transparent material, e.g. GaInN LEDs grown on sapphire substrates. Conductive Ag-loaded epoxy has a high conductivity as well as high reflectance. Such a highly reflective epoxy can increase the extraction efficiency in LEDs grown on transparent substrates.

There are ohmic contacts that are transparent to visible light. Such *transparent ohmic contacts* include indium tin oxide, frequently referred to as ITO (Ray *et al.*, 1983; Sheu *et al.*, 1998; Margalith *et al.*, 1999; Mergel *et al.*, 2000; Shin *et al.*, 2001). The material can be considered as a tin oxide semiconductor that is doped with indium. Indium substitutes for tin and therefore acts as an acceptor. Generally, the specific contact resistance of ITO contacts is higher than the contact resistance of alloyed metal contacts.

12.2 Total internal reflectors

Total internal reflection is a fascinating phenomenon occurring at the boundary between two dielectric media with different refractive indices. Total internal reflection was first discovered by Johannes Kepler in the early 1600s (Kepler, 1611). Kepler attempted to explain the apparent bending of objects partially submersed in water. Kepler discovered that for rays near normal incidence, the ratio of the angles of incidence and refraction is proportional to the ratio of what are now known as the refractive indices of the media. Kepler's relationship can be expressed as

$$\bar{n}_1 \theta_1 = \bar{n}_2 \theta_2 \quad (12.7)$$

where the angles θ_1 and θ_2 are measured with respect to the surface normal. Comparison with the law discovered between 1621 and 1625 by Cornelius Willebrord Snell (*Snell's law*)

$$\bar{n}_1 \sin \theta_1 = \bar{n}_2 \sin \theta_2 \quad (12.8)$$

reveals that Kepler found the small-angle approximation of Snell's law. The angles used in Snell's law are shown in *Figure 12.4*.

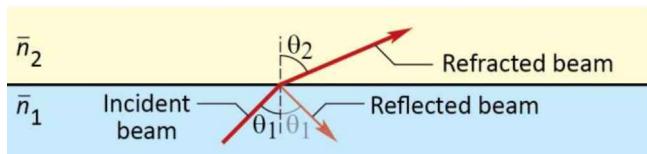


Fig. 12.4: Reflected and refracted light ray at the boundary between two media with refractive indices \bar{n}_1 and \bar{n}_2 , where $\bar{n}_1 > \bar{n}_2$.

Kepler also discovered that, with suitable materials and sufficiently shallow angles of incidence, the refracted angle can be made to exceed 90° , resulting in total internal reflection. The critical angle for total internal reflection can be derived from Snell's law using the condition $\theta_2 = 90^\circ$. Hence

$$\theta_{1,\text{crit}} = \arcsin\left(\frac{\bar{n}_2}{\bar{n}_1}\right). \quad (12.9)$$

Because the ratio of the refractive indices on the right-hand side of the equation, \bar{n}_2/\bar{n}_1 , must be ≤ 1.0 , total internal reflection can occur only in the optically denser material. Total internal reflection occurs for all angles of incidence $\theta_1 > \theta_{1,\text{crit}}$. For grazing angles of incidence and a sufficiently high index contrast, a light ray cannot leave a medium of high refractive index.

Isaac Newton later showed that, for most transparent media, the refractive index could be taken as unity plus a term proportional to the medium's mass density (measured in units of g/cm^3). Media with high refractive index are therefore frequently called ***optically dense*** materials.

The most fascinating application of total internal reflection is fiber optic communication in which light rays, by undergoing total internal reflection events, are guided over thousands of kilometers in the core of a silica fiber. In 1841, the guiding of light was first demonstrated by Daniel Colladon using a jet of water as the high-index material (Hecht, 2001). This phenomenon is used to the present day to enhance the appearance of water fountains at night. The apparatus built by Daniel Colladon, who is considered the father of light guiding, is shown in **Figure 12.5**.

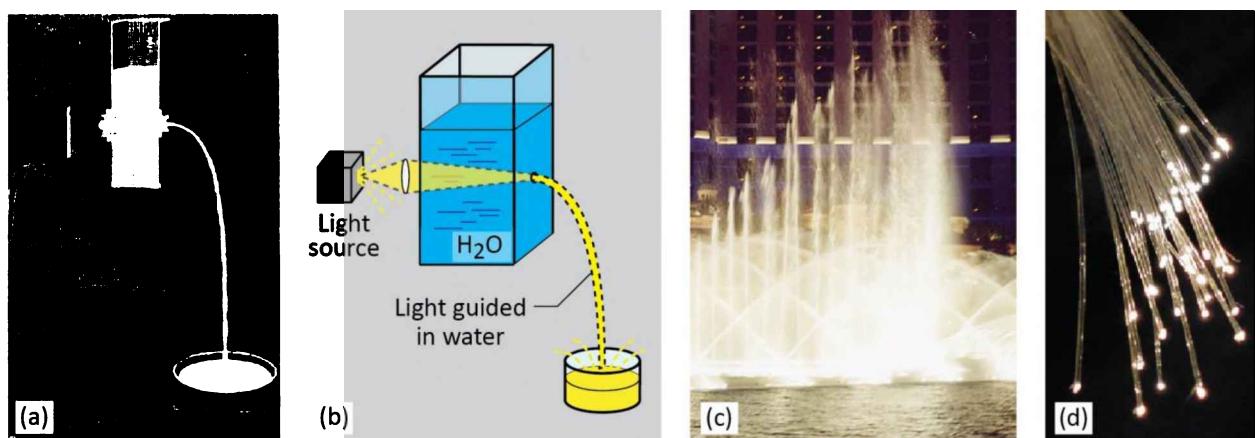


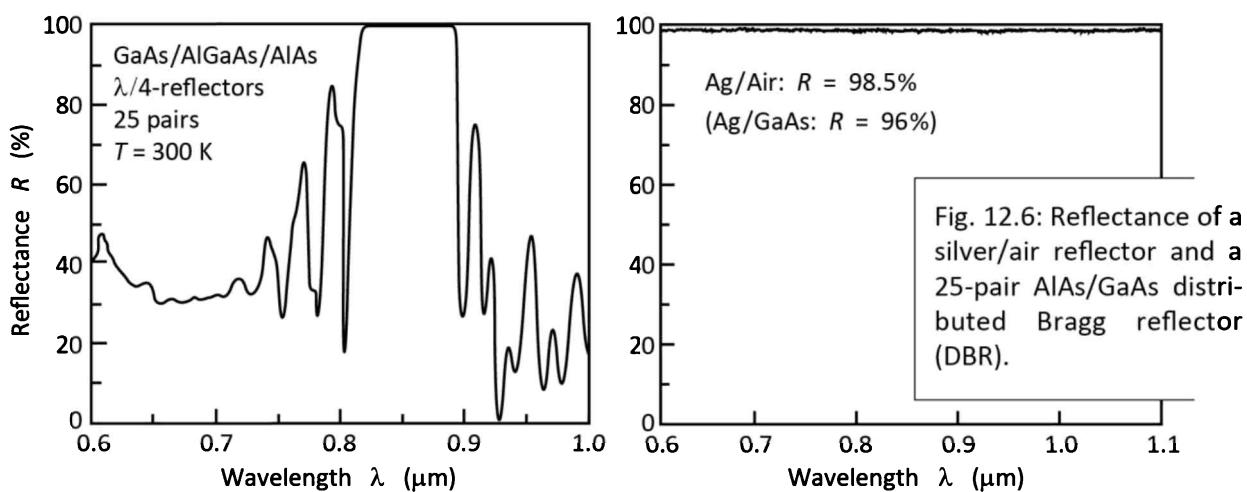
Fig. 12.5: (a) Historical drawing and (b) schematic illustration of apparatus used in 1841 by Swiss engineer Daniel Colladon to demonstrate the guiding of light by total internal reflection in a jet of water. (c) Light guiding used to beautify appearance of fountain in Las Vegas, Nevada. (d) Fiber-optic strands (after The Free Dictionary.com, 2005).

In light-emitting devices based on semiconductors, total internal reflection represents a major problem as it hinders the out-coupling of light out of the semiconductor die. As semiconductors have a large refractive index of typically $2.0 \sim 3.5$, the critical angle for total internal reflection is small. Whereas the problem is severe for III–V arsenides and phosphides ($\bar{n} \approx 3.0$) it is less severe for III–V nitrides ($\bar{n} \approx 2.0$). Total internal reflection is of little concern in *organic* light-emitting diodes due to the low refractive indices of organic materials.

One of the unique features of total internal reflection is that the magnitude of the reflection is $R = 1.0$ (“total”). This enables the demonstration of reflectors with zero mirror loss, a feature that has been employed advantageously in lasers. Edge-emitting lasers (Smith *et al.*, 1993) as well as microdisk lasers (“whispering gallery lasers”) (McCall *et al.*, 1992) have employed total internal reflection to demonstrate cavities with very high cavity-quality factors.

12.3 Distributed Bragg reflectors

The reflectance spectra of a metal mirror and a DBR are compared in **Figure 12.6**. Whereas metal reflectors exhibit a broad band with high reflectivity, DBRs display only a narrow band of high reflectivity denoted as the *stop band*. However, while the metal reflector has a certain reflectivity, the DBR’s reflectivity can be increased by increasing the number of reflector pairs so long as these reflector pairs are fully transparent.



For LED structures on light-absorbing substrates, about 50% of the light emitted by the active region is absorbed by the substrate. This represents a substantial loss. The absorption of light in

the substrate can be avoided by placing a reflector between the substrate and the LED active layers. Light emanating from the active region towards the substrate will then be reflected and can escape from the semiconductor through the top surface.

Distributed Bragg reflectors (DBRs) are well suited for inclusion between the substrate and the active layers. The schematic LED structure with a DBR is shown in **Figure 12.7**. LEDs with DBRs were first demonstrated by Kato *et al.* (1991) in the AlGaAs / GaAs material system. The LED had a 25-pair AlAs / GaAs or AlGaAs / GaAs DBR and emitted in the infrared at 870 nm.

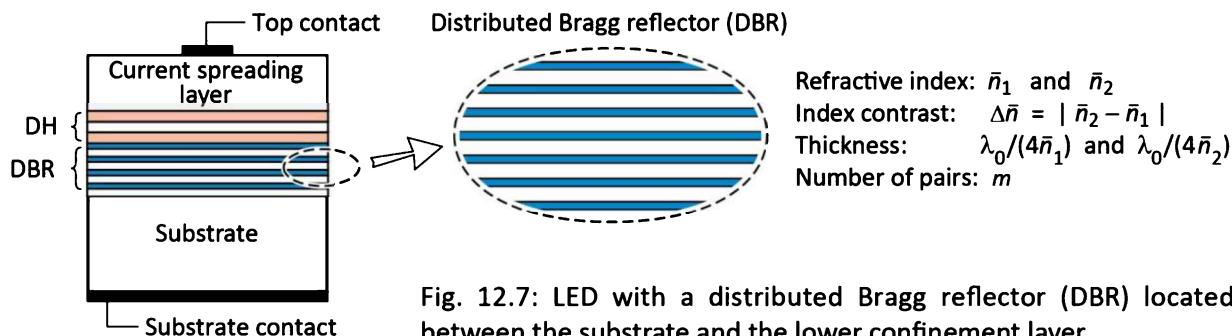


Fig. 12.7: LED with a distributed Bragg reflector (DBR) located between the substrate and the lower confinement layer.

A DBR is a multi-layer reflector consisting of typically 5~50 pairs of two materials with different refractive indices. As a result of the difference in refractive index, Fresnel reflection occurs at each of the interfaces. Usually the refractive index difference between the two materials is small so that the magnitude of the Fresnel reflection at one interface is also quite small. However, DBRs consist of many interfaces. More importantly, the thicknesses of the two materials are chosen in such a way that *all reflected waves* are in *constructive interference*. For normal incidence, this condition is fulfilled when both materials have a thickness of a *quarter wavelength* of the light, i.e.

$$t_{l,h} = \frac{\lambda_{l,h}}{4} = \frac{\lambda_0}{4\bar{n}_{l,h}} \quad (\text{normal incidence}) \quad (12.10)$$

where λ_0 is the vacuum Bragg wavelength of the light, $t_{l,h}$ is the thickness of the low-index (l) and high-index (h) material, and $\bar{n}_{l,h}$ is the refractive index of the low-index (l) and high-index

(h) material. The thickness of $t_{l,h}$ given in Eqn. (12.10) may not only be $\lambda/4$, but also an odd-numbered integer multiple of $\lambda/4$, i.e. $3\lambda/4$, $5\lambda/4$, $7\lambda/4$, and so on. These thicknesses will also result in constructive interference of the reflected waves. For layer thicknesses greater than $\lambda/4$, e.g. $3\lambda/4$, the reflector will, however, have a narrower high-reflectivity stop band.

For an oblique angle of incidence, the wave vector can be separated into a parallel and a normal component. As in the normal-incident case, the thickness of the DBR layers must be a quarter wavelength for the wave vector component normal to the DBR layers. For an oblique angle of incidence $\theta_{l,h}$, the optimum thicknesses for high reflectivity are given by

$$t_{l,h} = \frac{\lambda_{l,h}}{4 \cos \theta_{l,h}} = \frac{\lambda_0}{4 \bar{n}_{l,h} \cos \theta_{l,h}}$$

(oblique incidence). (12.11)

Again, the thickness $t_{l,h}$ given by the equation can also be an odd-numbered integer multiple of the value given by the equation. For DBRs with sufficiently many quarter-wave pairs, reflectivities near 100% can be obtained.

A DBR must fulfill several additional conditions. Firstly, since a DH is usually grown on top of the DBR, the DBR must be lattice matched to the DH in order to avoid misfit dislocations. Secondly, to attain high-reflectivity DBRs, the constituent DBR materials need to be transparent at the wavelength of operation unless the DBR has a high index contrast. High index-contrast DBRs (e.g. Si / SiO₂) yield high reflectivity, even if one of the materials is weakly absorbing at the wavelength of interest. Thirdly, if the DBR is in the current path, the DBR must be conductive. The reflectances of a Si / SiO₂ and an AlAs / GaAs DBR versus wavelength are shown in *Figure 12.8*. The Bragg wavelength is located in the center of the high-reflectivity band or ***stop band***. Inspection of the figure reveals that (i) the reflectivity of high-contrast DBRs (Si / SiO₂) is much higher than the reflectivity of low-index-contrast DBRs (AlAs / GaAs) for the same number of quarter-wave pairs and that (ii) the width of the stop band of the high-index-difference DBR is much wider than the stop band width of the low-contrast DBR.

Properties of DBRs can be calculated by the matrix method (Born and Wolf, 1989). The properties of DBRs have been analyzed in detail by Coldren and Corzine (1995), Yariv (1989), and Björk *et al.* (1995). Here, only a brief summary will be given.

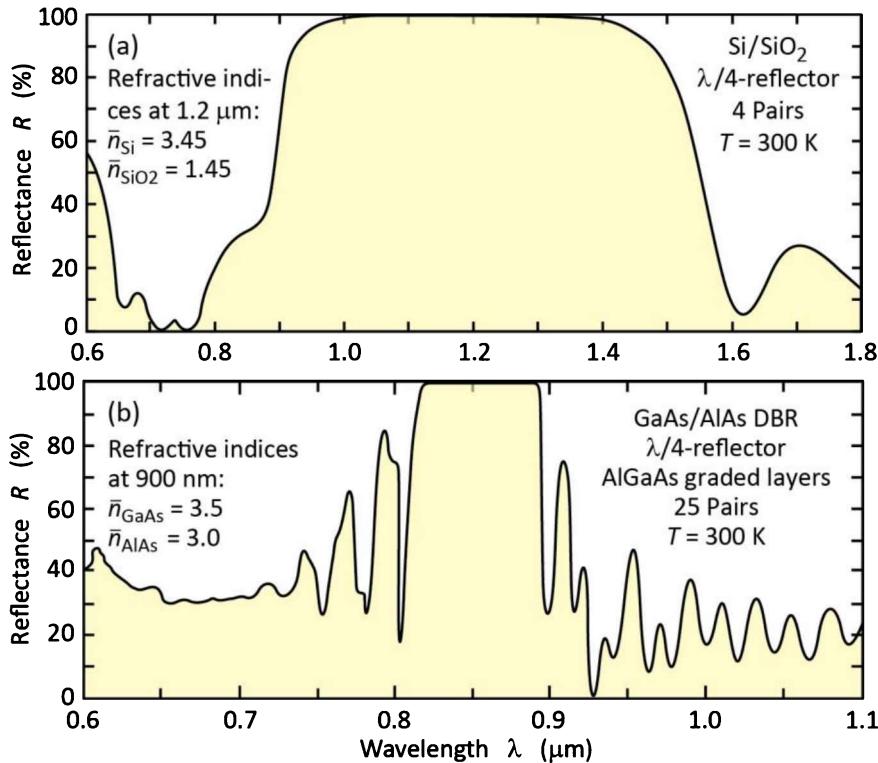


Fig. 12.8: Reflectance versus wavelength of two distributed Bragg reflectors (DBRs). (a) Four-pair Si/SiO_2 DBR with high index contrast. (b) 25-pair AlAs/GaAs DBR with low index contrast. The high-index-contrast DBR only needs four pairs to attain high reflectivity. Note that the stop band of the high-index-contrast DBR is wider compared with the low-index-contrast DBR.

Consider a distributed Bragg reflector consisting of m pairs of two dielectric, lossless materials with refractive indices \bar{n}_l and \bar{n}_h . The thicknesses of the layers are assumed to be a quarter wave, i.e. $L_l = \lambda_{\text{Bragg}}/(4\bar{n}_l)$, and $L_h = \lambda_{\text{Bragg}}/(4\bar{n}_h)$. The period of the DBR is $L_l + L_h$. The reflectivity of a single interface is given by Fresnel's equation for normal incidence

$$r = \frac{\bar{n}_h - \bar{n}_l}{\bar{n}_h + \bar{n}_l} . \quad (12.12)$$

Multiple reflections at the interfaces of the DBR and constructive interference of the multiple reflected waves increase the reflectivity with increasing numbers of pairs. The reflectivity has a maximum at the Bragg wavelength λ_{Bragg} . The reflectivity at the Bragg wavelength of a DBR with m quarter-wave pairs is given by (Coldren and Corzine, 1995)

$$R_{\text{DBR}} = |r_{\text{DBR}}|^2 = \left[\frac{1 - (\bar{n}_l / \bar{n}_h)^{2m}}{1 + (\bar{n}_l / \bar{n}_h)^{2m}} \right]^2 . \quad (12.13)$$

The stop band of a DBR depends on the difference in refractive index of the two constituent materials, $\bar{n}_h - \bar{n}_l = \Delta\bar{n}$. The spectral width of the stop band is given by (Yariv and Nakamura, 1977; Yariv, 1989; Karim, 2001)

$$\Delta\lambda_{stopband} = \frac{2 \lambda_{Bragg} \Delta\bar{n}}{\pi \bar{n}_{eff}}$$
(12.14)

where \bar{n}_{eff} is the effective refractive index of the DBR. For efficient operation of the LED, the stop band should be wider than the emission spectrum of the active region.

The effective refractive index of the DBR can be calculated by requiring the same optical path length normal to the layers for the DBR and an effective medium. The effective refractive index is then given by

$$\bar{n}_{eff} = 2 \left(\frac{1}{\bar{n}_l} + \frac{1}{\bar{n}_h} \right)^{-1}.$$
(12.15)

For small index differences, i.e. $\Delta\bar{n} \ll \bar{n}_l$, the effective refractive index can be approximated by

$$\bar{n}_{eff} = \frac{1}{2} (\bar{n}_l + \bar{n}_h).$$
(12.16)

The optical wave penetrates into the DBR only by a finite number of quarter-wave pairs. That is, a finite number out of the total number of quarter-wave pairs are effectively reflecting the wave. The effective number of pairs “seen” by the wave electric field is given by (Coldren and Corzine, 1995)

$$m_{eff} \approx \frac{1}{2} \frac{\bar{n}_h + \bar{n}_l}{\bar{n}_h - \bar{n}_l} \tanh \left(2m \frac{\bar{n}_h - \bar{n}_l}{\bar{n}_h + \bar{n}_l} \right).$$
(12.17)

For thick DBRs ($m \rightarrow \infty$), the tanh function approaches unity and one obtains

$$m_{eff} \approx \frac{1}{2} \frac{\bar{n}_h + \bar{n}_l}{\bar{n}_h - \bar{n}_l}.$$
(12.18)

At the Bragg wavelength ($\lambda = \lambda_{Bragg}$), the phase change of the reflected wave is zero. In the vicinity of the Bragg wavelength ($\lambda \approx \lambda_{Bragg}$), the phase of the reflected wave changes *linearly* with wavelength. It is therefore possible to approximate a DBR with a metal-like mirror located a distance L_{pen} behind the first dielectric interface, as shown in **Figure 12.9**. The reflection of the DBR can thus be expressed as

$$r_{DBR} \approx |r_{DBR}| e^{-2i(\beta - \beta_{Bragg})L_{pen}}$$
(12.19)

where $\beta = 2\pi/\lambda$ is the average phase constant of the wave. The phase change at $z=0$

(see **Figure 12.9**) of the wave reflected by the metal mirror is given by

$$r_{\text{metal}}|_{z=0} = |r_{\text{metal}}| e^{2i(2\pi/\lambda)L_{\text{pen}}} . \quad (12.20)$$

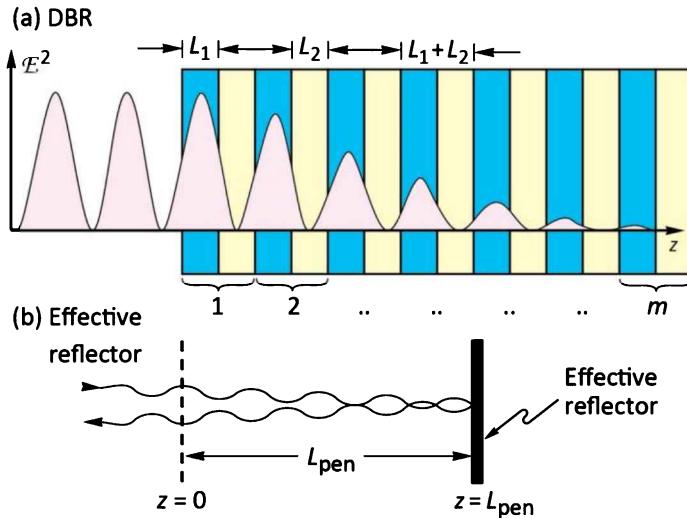


Fig. 12.9: Illustration of the DBR penetration depth. (a) DBR consisting of two materials with thickness L_1 and L_2 . (b) Ideal (metallic) reflector displaced from the DBR surface by the penetration depth.

Equating the phase changes given by Eqns. (12.19) and (12.20) and using the phase changes of a DBR (Coldren and Corzine, 1995), the penetration depth is given by

$$L_{\text{pen}} = \frac{L_1 + L_2}{4r} \tanh(2mr) . \quad (12.21)$$

For a large number of pairs ($m \rightarrow \infty$), the penetration depth is given by

$$L_{\text{pen}} \approx \frac{L_1 + L_2}{4r} = \frac{L_1 + L_2}{4} \frac{\bar{n}_h + \bar{n}_l}{\bar{n}_h - \bar{n}_l} . \quad (12.22)$$

Comparison of Eqns. (12.22) and (12.18) yields that

$$L_{\text{pen}} = \frac{1}{2} m_{\text{eff}} (L_1 + L_2) . \quad (12.23)$$

The factor of $(1/2)$ in Eqn. (12.23) is due to the fact that m_{eff} applies to the effective number of periods seen by the *electric field*, whereas L_{pen} applies to the optical power. The optical power is equal to the square of the electric field and hence it penetrates half as far into the mirror. The effective length of a cavity consisting of two DBRs is thus given by the sum of the thickness of the center region plus the two penetration depths into the DBRs. The effective length of a cavity with DBRs is thus longer than the effective length of a cavity with metal mirrors.

The reflectivity of the DBR depends on the polar angle of incidence and on the wavelength. Although an analytic result can be obtained for the reflectivity of a DBR at the Bragg wavelength for normal incidence, the reflectivity of a DBR at an arbitrary wavelength and an arbitrary angle of incidence can only be calculated numerically. If light from an isotropic source is reflected by a DBR, the reflected intensity can be obtained by integration over all angles. The angle-integrated reflectance at a certain wavelength λ is then given by

$$R_{int}(\lambda) = \frac{\int_0^{\pi/2} R(\lambda, \theta) 2\pi \sin \theta \, d\theta}{\int_0^{\pi/2} 2\pi \sin \theta \, d\theta} = \int_0^{\pi/2} R(\lambda, \theta) \sin \theta \, d\theta. \quad (12.24)$$

The total light intensity reflected by the DBR is calculated by

$$I_r = \int_{\lambda} I_i(\lambda) R_{int}(\lambda) \, d\lambda \quad (12.25)$$

where $I_i(\lambda)$ is the emission intensity spectrum of the active region incident on the DBR and I_r is the intensity reflected by the DBR. For an isotropic emitter such as the active region of an LED, it can be assumed that the emission spectrum incident on the DBR is independent of the emission angle.

An efficient DBR will be optimized in such a way that it maximizes the intensity of the reflected light. In addition, the escape of the light reflected by the DBR from the LED die must be taken into account. Maximizing the LED extraction efficiency using a DBR is a complicated problem that cannot be solved analytically. *Ray tracing computer software* is used to maximize the extraction efficiency in LED structures with DBRs.

A priori, it is not clear that the use of conventional DBRs provides the best extraction efficiency. DBRs with layer thicknesses thinner or thicker than a quarter-wave should also be considered for maximization of the extraction efficiency. Such variable-thickness DBRs have a *lower reflectivity* but a *wider stop-band width* compared with quarter-wave DBRs. For active regions with broad emission spectra, such variable-period DBRs can be advantageous.

Ideally, the layers comprising the DBR are transparent. Layers of **transparent DBRs** have negligible absorption losses. However, transparent materials may not always be available so that *absorbing* materials must be used. Such **absorbing DBRs** have a maximum reflectivity of less than 100%, even if an infinite number of pairs are used.

An example of a partially absorbing Si/SiO₂ DBR was shown in *Figure 12.8*. Silicon absorbs light for $\lambda < 1.1 \mu\text{m}$, i.e. for $h\nu > E_g$. However, the results shown in the figure demonstrate that very high reflectivities can be attained at $1.0 \mu\text{m}$, where Si is absorbing. This is due to the high index contrast between Si and SiO₂.

Transparent and absorbing DBRs, lattice matched to GaAs, are used in the AlGaInP/GaAs material system. This material system is used for high-efficiency visible LEDs emitting at $\lambda > 550 \text{ nm}$ (yellow-green, yellow, amber, orange, and red). The properties of transparent and absorbing DBRs used in this material system have been compiled by Kish and Fletcher (1997) and are summarized in *Table 12.3*. Inspection of the table reveals that the absorbing Al_{0.5}In_{0.5}P/GaAs DBRs have the advantage of a high refractive index contrast. However, the absorbing nature of the DBR imposes an upper limit on the maximum reflectivity. High-contrast DBRs have a *wider* stop-band width. The transparent Al_{0.5}In_{0.5}P/(AlGa)_{0.5}In_{0.5}P DBRs have the advantage of negligible optical losses. However, many pairs are needed to attain high reflectivity and the stop-band width is narrower than in high index-contrast DBRs.

Table 12.3: Properties of distributed Bragg reflector (DBR) materials used in LEDs for visible and infrared applications. The DBRs marked as “lossy” are absorbing at the Bragg wavelength (data after Adachi, 1990; Adachi *et al.*, 1994; Kish and Fletcher, 1997; Babic *et al.*, 1999; Palik, 1998).

Material system	Bragg wavelength	\bar{n}_{low}	\bar{n}_{high}	$\Delta\bar{n}$	Transparency range
Al _{0.5} In _{0.5} P / GaAs	590 nm	3.13	3.90	0.87	> 870 nm (lossy)
Al _{0.5} In _{0.5} P / Ga _{0.5} In _{0.5} P	590 nm	3.13	3.74	0.61	> 649 nm (lossy)
Al _{0.5} In _{0.5} P / (Al _{0.3} Ga _{0.7}) _{0.5} In _{0.5} P	615 nm	3.08	3.45	0.37	> 592 nm
Al _{0.5} In _{0.5} P / (Al _{0.4} Ga _{0.6}) _{0.5} In _{0.5} P	590 nm	3.13	3.47	0.34	> 576 nm
Al _{0.5} In _{0.5} P / (Al _{0.5} Ga _{0.5}) _{0.5} In _{0.5} P	570 nm	3.15	3.46	0.31	> 560 nm
AlAs / GaAs	900 nm	2.97	3.54	0.57	> 870 nm
SiO ₂ / Si	1300 nm	1.46	3.51	2.05	> 1106 nm

In practice, transparent layers are used at and near the top (epitaxial side) of the DBR whereas absorbing layers are used towards the bottom (substrate side) of the DBR. For DBRs used in AlGaInP/GaAs LEDs, each of the layers is different and optimized to keep the number of pairs low, the absorption low, and the reflectivity spectrum broad (Streubel, 2000).

Table 12.3 also shows properties of the AlAs/GaAs and the SiO₂/Si material systems. The SiO₂/Si material system is an example of a high index-contrast system. However, it cannot be

used for current conduction due to the insulating nature of the SiO_2 . The AlAs / GaAs material system is used in resonant-cavity LEDs and vertical-cavity surface-emitting lasers emitting in the range 880~980 nm.

A DBR that is resonant with the peak-emission wavelength, is not necessarily the optimum reflector for an absorbing substrate (AS) LED. Although a DBR has a high reflectivity for normal incidence, it rapidly decreases for off-normal angles of incidence. Assume that the angle of incidence is $\theta = 0^\circ$ at normal incidence. Since the solid angle (per angle interval $d\theta$) increases with angle θ according to a *sine* function, it is advantageous to shift the normal-incidence resonance wavelength of the DBR to wavelengths *longer* than the peak emission wavelength.

The calculated reflectivities of a transparent and an absorbing DBR versus wavelength and polar angle of incidence are shown in *Figure 12.10*. Whereas the DBR made of the transparent materials has a reflectivity close to 100% at the Bragg wavelength, the DBR that includes the absorbing GaAs layers has a maximum reflectivity of about 55%. This value cannot be increased by adding additional pairs to the DBR, because the limitation of the value lies in the absorptive nature of the DBR.

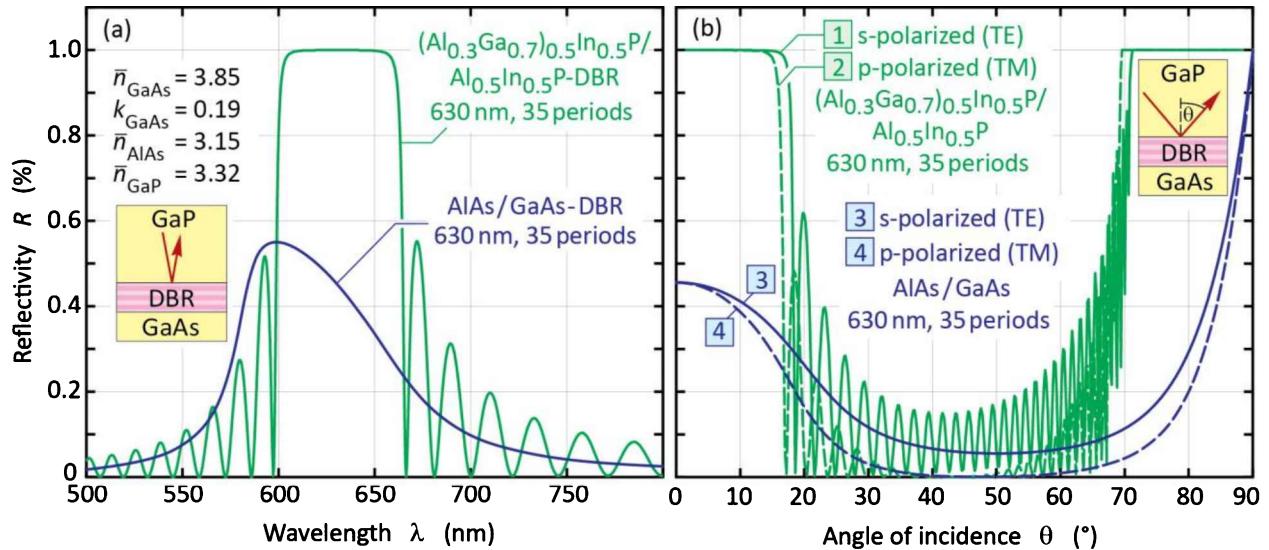


Fig. 12.10: Calculated reflectivity (inside the cladding GaP) versus (a) wavelength and (b) polar angle of incidence of a transparent AlGaN/P/AlInP DBR and an absorbing AlAs/GaAs DBR.

Figure 12.10 also shows the reflectivity of the DBRs versus the angle of incidence. Inspection of the figure reveals a major drawback of DBRs, namely the limitation of the high reflectivity band to small angles of incidence. For angles of incidence greater 20°, the reflectivity strongly

decreases to assume values close to zero. Thus, for oblique angles of incidence ($20^\circ < \theta < 70^\circ$), the DBR becomes non-reflective. The lack of reflectivity is a major loss mechanism in AlGaNp LEDs whose active layers are grown on top of a DBR that in turn is located on the absorbing GaAs substrate.

Next a formula will be derived that gives the critical angle, θ_c , at which the reflectivity of a DBR strongly decreases. The structure of a DBR and the critical angle are illustrated in **Figure 12.11** (a) and (b), respectively. Note that the outside medium is a semiconductor with refractive index \bar{n}_0 . For normal incidence ($\theta = 0^\circ$), the Bragg condition is fulfilled at the Bragg wavelength, λ_{Bragg} , which is located at the center of the high-reflectivity stop band. The Bragg wavelength shifts with angle of incidence, as shown in **Figure 12.11** (c). However, the width of the stop band does not depend on the angle of incidence as long as it is small. We can thus write the following condition for the critical angle θ_c

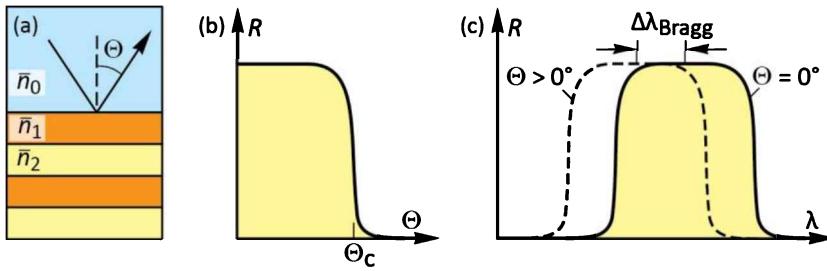


Fig. 12.11: (a) DBR structure used in calculation. (b) Reflectivity versus angle of incidence and critical angle at which reflectivity decreases. (c) DBR reflectivity versus wavelength for two angles of incidence.

$$\Delta\lambda_{Bragg} = \lambda_{Bragg}(\theta = 0^\circ) - \lambda_{Bragg}(\theta_c) = \frac{1}{2} \Delta\lambda_{stop\ band}. \quad (12.26)$$

Using the expression for the angle-dependent Bragg wavelength and the expression for the width of the stop band given earlier in this chapter, one can write

$$\lambda_{Bragg}(\theta = 0^\circ) \left[1 - \cos \left(\frac{\bar{n}_0}{\bar{n}_1} \theta_c \right) \right] = 2\lambda_{Bragg}(\theta = 0^\circ) \frac{\Delta\bar{n}}{\bar{n}_1 + \bar{n}_2} \frac{1}{\bar{n}_0}. \quad (12.27)$$

Dividing both sides of the equation by $\lambda_{Bragg}(\theta = 0^\circ)$ and solving the equation for θ_c yields

$$\theta_c = \frac{\bar{n}_1}{\bar{n}_0} \arccos \left[1 - \left(\frac{2 \Delta\bar{n}}{\bar{n}_1 + \bar{n}_2} \frac{1}{\bar{n}_0} \right) \right]. \quad (12.28)$$

Using the approximation $\cos x \approx 1 - (1/2)x^2$ (valid near $x = 0$) and $\arccos x \approx [2(1-x)]^{1/2}$ (valid near $x = 1$), the following approximation is obtained

$$\theta_c \approx \frac{\bar{n}_1}{\bar{n}_0} \sqrt{\frac{2}{\bar{n}_0} \frac{2 \Delta \bar{n}}{\bar{n}_1 + \bar{n}_2}}. \quad (12.29)$$

The equation shows that the critical angle strongly depends on the refractive index of the outside medium, \bar{n}_0 , with $\theta_c \propto (\bar{n}_0)^{-3/2}$. Thus, the critical angle is small for outside media with a high refractive index. For this reason, it is difficult if not impossible to attain omnidirectional reflection characteristics with a DBR if the outside medium is a high-index semiconductor.

As a numerical example, we consider an AlAs / GaAs DBR ($\bar{n}_{\text{AlAs}} = 3.0$; $\bar{n}_{\text{GaAs}} = 3.5$) with GaP as the outside medium ($\bar{n}_{\text{GaP}} = 3.1$). Insertion of these values in the equation for the critical angle yields $\theta_c = 20.5^\circ$, which is relatively close to normal incidence. Even high-contrast DBRs, such as SiO₂ / Si DBRs, do not have omnidirectional characteristics if the outside medium is a high-index semiconductor.

Different strategies have been employed to optimize DBRs. Chiou *et al.* (2000) used a composite DBR in an AlGaInP LED, that is, two types of DBRs stacked on top of each other, namely a non-absorbing (Al_{0.4}Ga_{0.6})_{0.5}In_{0.5}P / Al_{0.5}In_{0.5}P DBR resonant at the peak emission wavelength of 590 nm and an additional high-contrast absorbing AlAs / GaAs DBR. This DBR was located below the phosphide DBR. The arsenide DBR was resonant at a wavelength about 10% longer than the peak emission wavelength in order to reflect light incident at off-normal angles. The authors found a substantial improvement in light output with the composite DBR.

Non-periodic DBRs have a wider stop band and thereby a high reflectivity over a wider range of angles. To find optimum non-periodic DBR structures, numerical simulations and optimization procedures have been conducted (see, for example, Li *et al.*, 1999).

As shown by the calculation above, an increase in index contrast will provide a wider range of angles within which a high reflectivity is maintained. The use of such high-contrast DBRs in LEDs has been proposed by Chiou *et al.* (2003) who disclosed the use of an AlGaAs / Al_xO_y DBR in an LED. Al₂O₃ has a refractive index of about 1.75 whereas Al-rich AlGaAs has a refractive index of about 3.25 thus providing a large index contrast of $\Delta \bar{n} = 1.5$. The Al_xO_y layers of the DBR were fabricated from epitaxially grown AlAs by using an oxidation process that is performed in water vapor at temperatures of 400~450°C. The Al_xO_y layers of the DBR are not conductive thus necessitating non-oxidized AlAs openings in the Al_xO_y layers, which provide a current path between substrate and active layers.

DBRs can have a high electrical resistance for current transport perpendicular to the layers. The resistance can pose a substantial problem in LED and laser structures, which manifests itself in a high forward voltage. Early experiments on VCSELs or vertical-cavity surface-emitting lasers (Jewell *et al.*, 1989; Koyama *et al.*, 1989) revealed forward voltages as high as 30.0 V (Jewell, 1992) and prevented these first devices from lasing in continuous-wave mode. The high resistance is caused by abrupt heterojunctions which pose barriers for carrier transport. Fortunately, heterojunction barriers can be completely eliminated by parabolic compositional grading (Schubert *et al.*, 1992a, 1992b). Such compositional grading is now routinely used in DBRs so that the resistance issue of DBR heterojunctions no longer is a concern.

12.4 Omnidirectional reflectors

Electrically conductive high-reflectivity omnidirectional reflectors are highly desirable. With air as the outside medium, omnidirectional reflection characteristics can be demonstrated by using high-contrast DBRs. The high-contrast materials Si ($\bar{n} \approx 3.5$ at $\lambda = 1 \mu\text{m}$) and SiO₂ ($\bar{n} \approx 1.46$) are natural candidates for such ODRs. The optical properties of Si/SiO₂ distributed Bragg omnidirectional reflectors (DB-ODRs) and other material systems have been investigated (Chen *et al.*, 1999; Bruyant *et al.*, 2003).

Whereas omnidirectional TE reflectivity is readily obtained in DB-ODRs, the Brewster angle, at which the TM reflectivity decreases to zero, is an impediment in achieving omnidirectional characteristics for TM waves. An outside angle range of $0^\circ \leq \theta \leq 90^\circ$ (in a low-index outside material such as air) results in an angle range $0^\circ \leq \theta \ll 90^\circ$ inside the material (i.e. inside the DBR). The inside range may not include the Brewster angle under which circumstances a DBR becomes omnidirectionally reflective.¹⁰

Highly omnidirectional reflection characteristics were also obtained with DB-ODRs using polystyrene and Te layers (Fink *et al.*, 1998). Owing to the large difference in refractive indices, $\bar{n}_{\text{polystyrene}} = 1.8$ and $\bar{n}_{\text{Te}} = 5.0$, the interfacial Brewster angle θ_B is not accessible for light incident from air, resulting in a complete photonic bandgap in the wavelength range 10 to 15 μm .

Another intriguing approach uses birefringent polymers with two different refractive indices parallel and vertical to the layer planes (Weber *et al.*, 2000). By adjusting the differences between

¹⁰ θ is used as angle in the outside medium (i.e. within air), whereas θ is used as angle in the inside medium (i.e. within the semiconductor).

the vertical and in-plane indices, the value of the Brewster angle can be controlled. Brewster angles up to 90° (grazing incidence) and even imaginary values are possible, resulting in a high reflectivity for TM-polarized light at virtually all angles of incidence.

Unfortunately, the applicability of the above-mentioned DB-ODRs to LEDs is limited due to the insulating electrical characteristics of the constituent materials.

Metallic layers are capable of reflecting light over a wide range of wavelengths and incident angles with the high-reflectivity band being limited to frequencies below the plasma frequency of the free-electron gas (for historical reference articles, see Drude, 1904; Lorentz 1909). However, electron oscillations induced by the incident light waves not only result in reflection but also in absorption caused by electron–phonon scattering. Thus, pure metal reflectors have significant reflection losses, particularly when used on high-index materials.

An electrically conductive reflector that has high-reflectivity and omnidirectional characteristics is shown in *Figure 12.12* (Schubert, 2001). It consists of three layers, a semiconductor, a dielectric layer, and a metal layer (triple-layer ODR). The dielectric layer is perforated by an array of microcontacts that provide electrical conductivity. The dielectric layer has a refractive index as low as possible to provide a high index contrast to the semiconductor and the metal. The metal has a complex refractive index with a large extinction coefficient.

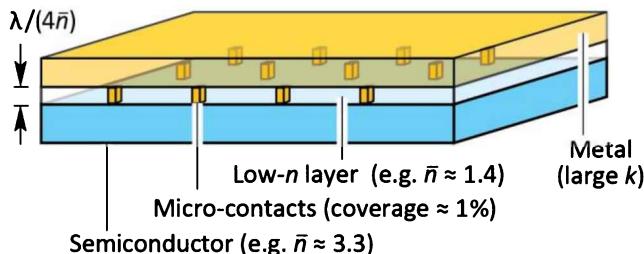


Fig. 12.12: Structure of omnidirectional reflector consisting of semiconductor, low-refractive index dielectric layer, and metal layer. The dielectric is perforated by an array of microcontacts providing electrical conductivity (after Gessmann *et al.*, 2003).

The reflectivity of the triple-layer ODR calculated by the matrix method is shown in *Figure 12.13* versus wavelength and versus angle along with the reflectivity of two DBRs (Gessmann *et al.*, 2003). The triple-layer ODR provides a broad reflectivity and omnidirectional characteristics. A small dip in the reflectivity is found for the TM wave at an incidence angle of about 30° due to the reduced reflectivity of the semiconductor / dielectric interface at the Brewster angle. The angle-integrated reflectivity of the triple-layer ODR is very high and can exceed 99%.

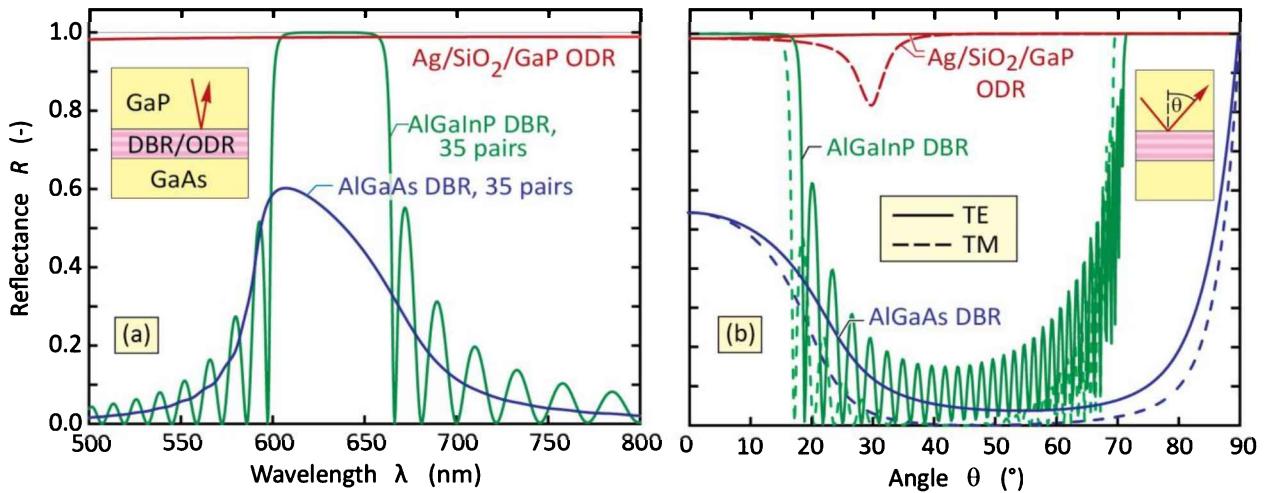


Fig. 12.13: (a) Calculated reflectance at normal-incidence versus wavelength and (b) reflectance versus angle of incidence for an omnidirectional reflector (ODR), a transparent AlGaNp/AlInP DBR, and an absorbing AlGaAs/GaAs DBR (after Gessmann *et al.*, 2003).

At normal incidence ($\theta = \theta = 0^\circ$)¹¹, the reflectivity of the triple-layer ODR can be calculated analytically and is given by

$$R_{ODR} = \frac{[(\bar{n}_s - \bar{n}_{li})(\bar{n}_{li} + \bar{n}_m) + (\bar{n}_s + \bar{n}_{li})k_m]^2 + [(\bar{n}_s - \bar{n}_{li})k_m + (\bar{n}_s + \bar{n}_{li})(\bar{n}_{li} - \bar{n}_m)]^2}{[(\bar{n}_s + \bar{n}_{li})(\bar{n}_{li} + \bar{n}_m) + (\bar{n}_s - \bar{n}_{li})k_m]^2 + [(\bar{n}_s + \bar{n}_{li})k_m + (\bar{n}_s - \bar{n}_{li})(\bar{n}_{li} - \bar{n}_m)]^2} \quad (12.30)$$

where \bar{n}_{li} and \bar{n}_s are the refractive indices of the dielectric and semiconductor, respectively, and $N_m = \bar{n}_m + ik_m$ is the complex refractive index of the metal. The equation applies to a thickness of $\lambda_0/(4\bar{n}_{li})$ for the low-index dielectric layer, i.e. a quarter-wave layer. For an AlGaNp / SiO₂ / Ag structure emitting at 630 nm, the equation yields a normal-incidence reflectivity $R_{ODR}(\theta = \theta = 0^\circ)$ of 98.8% compared to a value of 96.1% for a structure without a dielectric layer.

650 nm AlGaNp LEDs using a triple-layer ODR have been demonstrated. The surface coverage of the microcontact array was 1%. The light-output power versus current characteristics of the ODR-LED and several reference devices are shown in *Figure 12.14* (Gessmann *et al.*, 2003). Comparison of the ODR-LED output characteristics with those of DBR-LEDs yields that the ODR-LED provides higher output powers.

¹¹ θ is used as angle in the outside medium (i.e. within air), whereas θ is used as angle in the inside medium (i.e. within the semiconductor).

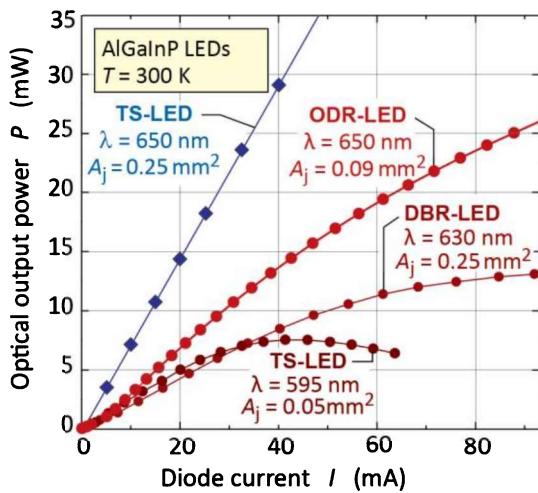


Fig. 12.14: Light-output power versus injection current of different types of LEDs. The ODR device has a higher output power than the DBR device (after Gessmann *et al.*, 2003).

ODR-LEDs have also been demonstrated in the GaInN material system (Kim *et al.*, 2004). The ODR consisted of a RuO₂ ohmic contact to p-type GaN, a quarter-wave-thick SiO₂ low-index layer perforated by an array of microcontacts, and an Ag layer. Calculations predict a 98% angle-averaged reflectivity at $\lambda = 450\text{ nm}$ for an GaN/SiO₂/Ag ODR, much higher than that for an Al_{0.25}Ga_{0.75}N/GaN distributed Bragg reflector (49%) and Ag (94%). It was shown that the RuO₂/SiO₂/Ag ODR has a higher reflectivity than Ni/Au and even Ag reflectors, leading to a higher light-extraction efficiency of the GaInN ODR-LED. The electrical properties of the ODR-LED were found to be comparable to those LEDs using conventional Ni/Au contacts. A comparison of the electrical and optical characteristics of the GaInN ODR device is shown in **Figure 12.15**.

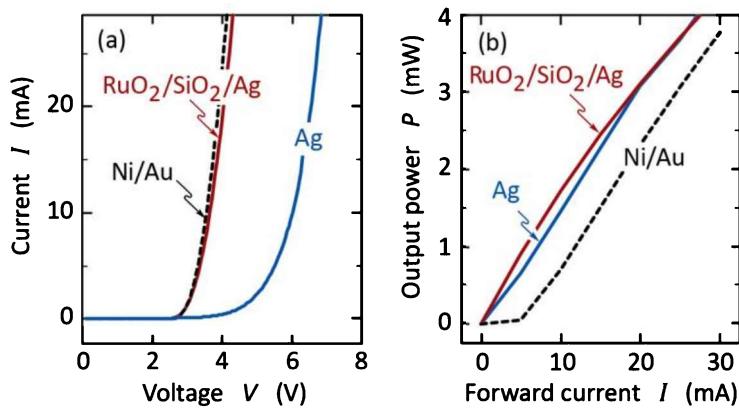


Fig. 12.15: Current-versus-voltage and light-output-versus-current characteristics of a GaInN LED with a GaInN/RuO₂/SiO₂/Ag omnidirectional reflector (after Kim *et al.*, 2004).

Experimental setup:
LED
PIN photo detector

12.5 Specular, diffuse, and mixed reflectors

It is useful to distinguish between three types of reflectors, namely (i) specular reflectors, (ii) lambertian reflectors, and (iii) mixed reflectors. Their characteristics can be described as follows:

Specular reflectors are shown in **Figure 12.16 (a)**. The angle of the reflected light is equal to the angle of the incident light. Specular reflectors are deterministic, i.e. the angle of reflection is pre-determined by the angle of the incident light ray. Specular reflectors have a smooth surface, that is, in the ideal case perfectly flat.

Diffuse reflectors, also called **lambertian reflectors** or **matte reflectors**, are shown in **Figure 12.16 (b)**. The reflected intensity of diffuse reflectors is distributed over a wide range of angles (180°), independent of the light ray's angle of incidence. One can show that for diffuse reflectors, any angle of reflection is equally likely. That is, diffuse reflectors are non-deterministic and the reflection angle of a particular incident photon is not pre-determined. Diffuse reflectors have a randomly roughened surface texture.

Mixed reflectors, also called **spread reflectors**, are shown in **Figure 12.16 (c)**. Such reflectors have both a specular and a diffuse component. The relative strength of these two components can vary. Mixed reflectors that have both a partially roughened and partially smooth surface texture.

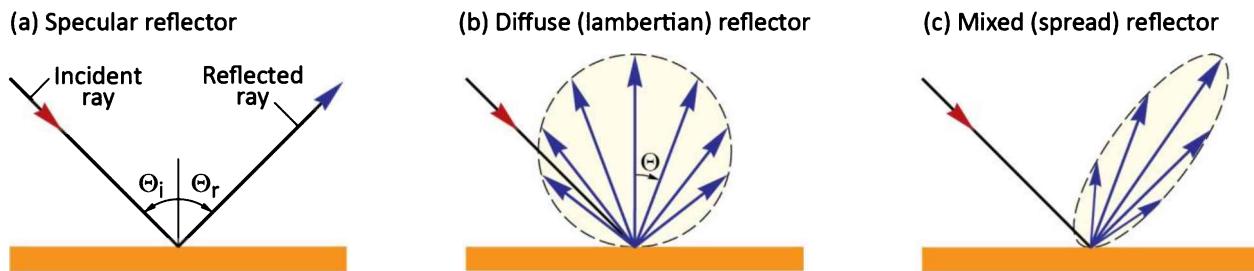


Fig. 12.16: (a) Specular reflector, (b) diffuse (lambertian) reflector, and (c) mixed (spread) reflector. The reflected power of a lambertian reflector follows a $\cos \Theta$ dependence.

Next, we will discuss the properties of lambertian surface sources and lambertian surface reflectors and subsequently discuss implications for LEDs. Lambertian surface sources are motivated by the frequently observed *experimental fact that the source radiance* (i.e. the optical power emitted per steradian per unit surface area of the source) *is a constant*, independent of the viewing angle. That is, the radiance and luminance of a lambertian source are independent of the viewing angle. The sun is the prime example of a lambertian source. As shown in **Figure 12.17 (a)**, the sun's surface has the same brightness (luminance), irrespective of the viewing angle. That is, the brightness (luminance) is the same for, e.g., the normal-incidence viewing angle and an oblique-incidence viewing angle with respect to the sun's surface. For the same reason, the moon, shown in **Figure 12.17 (b)**, can be considered a good example of a lambertian reflector. Light-diffusing reflectors randomize the propagation direction of incoming photons and are called **diffuse** or **lambertian reflectors**.

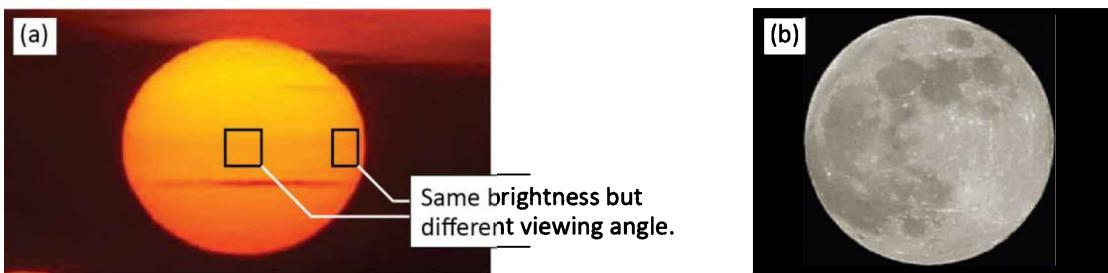


Fig. 12.17: (a) The sun's surface brightness is independent of viewing angle with respect to the sun's surface. It is a good example of a lambertian source. (b) The moon is a good example of a lambertian reflector.

We now assume that a lambertian surface source has an intensity (i.e. optical power emitted per steradian) along the direction given by the angle θ is given by

$$I = I_n \cos \theta \quad (12.31)$$

where I_n is the intensity emitted normal to the reflector surface. The angular $\cos \theta$ dependence of the equation is known as **Lambert's cosine law**.

Next, we will show that Lambert's cosine law results in a source luminance (or radiance, in radiometric units) that is *independent* of the viewing angle with respect to the source surface. Assume a lambertian surface source with area A . The projected area visible to an observer positioned at angle θ is given by $A \cos \theta$. Thus, the luminance found by the observer is given by

$$\text{Luminance} = \frac{I_n A \cos \theta}{A \cos \theta} = I_n \quad (12.32)$$

where $A \cos \theta$ is the surface area seen by the observer. Thus, the luminance is constant, independent of the viewing angle. This fact is corroborated by the photographs shown in **Figure 12.17**.

Light propagating in a layer clad by specular reflectors will be waveguided within the layer. However, light propagating in a layer clad by a lambertian reflector can be out-coupled into free space as shown in **Figure 12.18**. As a lambertian reflector follows the $\cos \theta$ dependence, the probability of reflected light falling into the escape cone defined by the angle θ_c is given by

$$p = \frac{\int_0^{\theta_c} I_n \cos \theta \ 2\pi \sin \theta \ d\theta}{\int_0^{90^\circ} I_n \cos \theta \ 2\pi \sin \theta \ d\theta} = \frac{\int_0^{\theta_c} \sin(2\theta) \ d\theta}{\int_0^{90^\circ} \sin(2\theta) \ d\theta} = \frac{1 - \cos(2\theta_c)}{2} \quad (12.33)$$

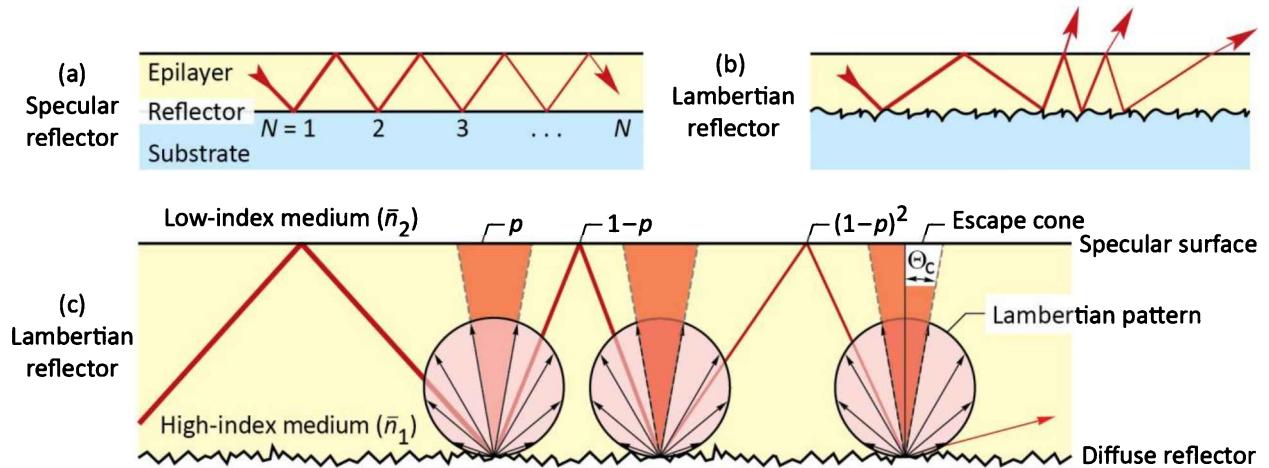


Fig. 12.18: (a) Optical ray, propagating in epilayer, guided by specular epilayer/substrate interface and epilayer/air interface. (b) and (c) Optical ray guided by lambertian reflector at epilayer/substrate interface and specular epilayer/air interface.

Using Snell's law (that is, $\bar{n}_1 \sin \theta_c = \bar{n}_2$, where \bar{n}_1 is the refractive index of the waveguide and $\bar{n}_1 > \bar{n}_2$), one obtains

$$p = \frac{1 - \cos [2 \arcsin (\bar{n}_2 / \bar{n}_1)]}{2} = \left(\frac{\bar{n}_2}{\bar{n}_1} \right)^2. \quad (12.34)$$

Assuming a lambertian reflector with unit reflectivity ($R = 1.0$), the light intensity inside the semiconductor decreases according to a geometric series. After N reflection events, the intensity of the light ray will have fallen to $(1 - p)^N$. Defining N as the number of reflection events after which the light intensity has decreased to $1/e$, we can write the equation

$$(1 - p)^N = 1/e. \quad (12.35)$$

Solving the equation for N allows one to calculate the average number of reflection events before the light ray will escape into the free space surrounding the semiconductor. One obtains

$$N = - \left[\ln \left(1 - \frac{\bar{n}_2^2}{\bar{n}_1^2} \right) \right]^{-1} \quad (12.36)$$

As an example, we consider $\bar{n}_1 = 2.5$ (GaN) and $\bar{n}_2 = 1.0$ (air) and obtain $N = 5.7$. That is, light escapes from the waveguide after about six diffuse reflection events. Thus, the introduction of diffuse reflectors into LED structures is a fruitful strategy in attaining laterally scalable emitters

that do not suffer from an efficiency penalty usually associated with the up-scaling of the LED chip dimension (Kim *et al.*, 2006).

Mechanical roughening of reflective surfaces generally results in a change from specular to diffuse reflection characteristics. Such reflective surfaces can be, e.g., metallic surfaces or dielectric surfaces. In addition, porous silica (optionally coated with a metal) is known to have diffuse reflection characteristics as multiple refraction, reflection, and scattering events at pore / silica interfaces randomize the photon propagation direction.

The fabrication of *ideal* lambertian reflectors requires substantial roughness (greater than the wavelength of light, λ). Ideal diffuse reflectors have a reflection characteristic that is independent of the angle of incidence. *Real* surface-textured reflectors usually have a mixed specular-diffuse reflection characteristic with a *preferential* direction along the specular-reflection direction. **Figure 12.19** shows schematic reflection-contour plots of (i) a diffuse, (ii) a mixed specular-diffuse, and (iii) a specular reflector for a 45° angle of incidence.

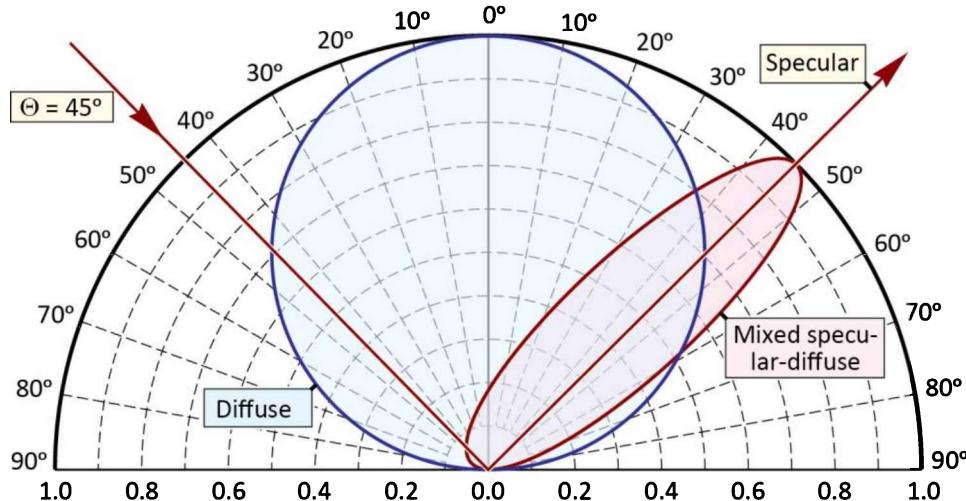


Fig. 12.19: Schematic reflection contour plots of a specular reflector, diffuse reflector, and a mixed specular-diffuse reflector for an angle of incidence of 45° .

Figure 12.20 shows the measured reflected intensity versus reflection angle for a planar smooth Ag reflector and a surface-textured Ag reflector. The surface roughening was achieved by lithography using 700 nm diameter polystyrene spheres, followed by ion-beam etching and Ag deposition. Inspection of the figure reveals a strong diffuse background for the textured reflector that is about 100-times higher than for the smooth reflector. However, the textured reflector still exhibits a specular component. The relative strength of the diffuse and specular components of mixed specular-diffuse reflectors can be assessed using a model developed by Xi *et al.* (2006).

Using this model, it is found that the mixed specular-diffuse reflector, shown in *Figure 12.20*, has a diffuse power ratio of $P_{\text{diffuse}} / (P_{\text{specular}} + P_{\text{diffuse}}) = 42.8\%$. The root-mean-square (rms) roughness of the partially diffuse reflector shown in *Figure 12.20* is 21.2 nm. By further increasing the rms roughness, diffusive power ratios of 100% can be attained.

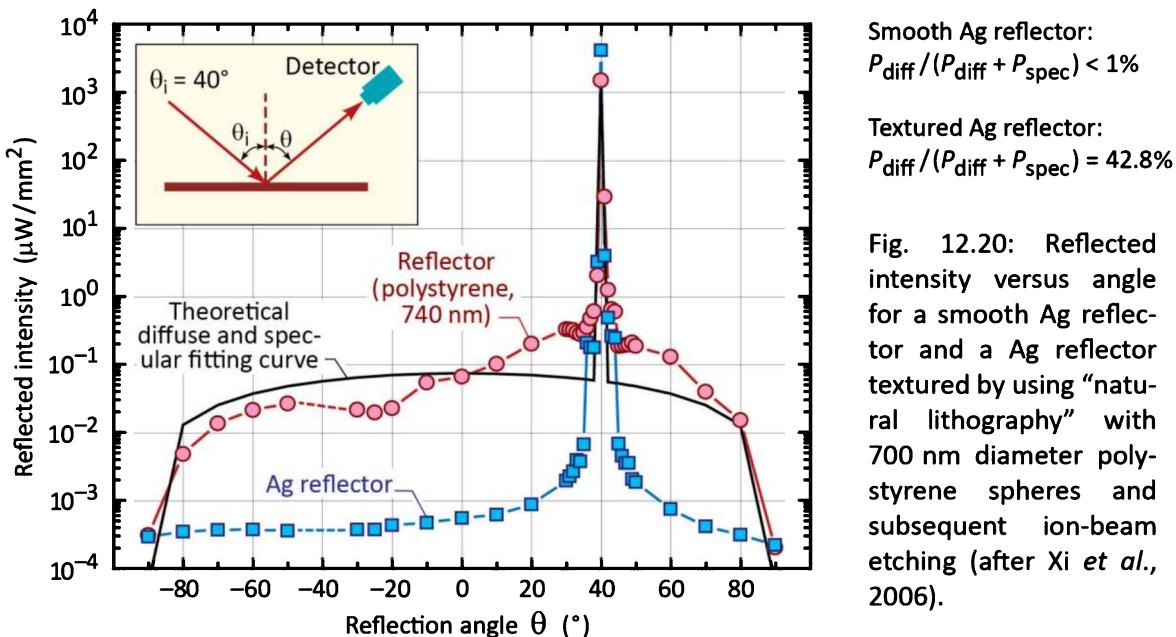


Fig. 12.20: Reflected intensity versus angle for a smooth Ag reflector and a Ag reflector textured by using “natural lithography” with 700 nm diameter polystyrene spheres and subsequent ion-beam etching (after Xi *et al.*, 2006).

Exercise: Lambertian reflectors in LEDs. Assume a lambertian reflector with reflectivity 1.0 that is incorporated in a lossless GaAs LED structure with refractive index of 3.5. Assume that the outside medium is air. Calculate the critical angle of the escape cone, the probability that a reflected light ray falls within the escape cone and the average number of reflection events before a photon escapes from the high-index GaAs layer.

Solution: Critical angle $\Theta_c = 16.6^\circ$; Probability of escape $p = 8.2\%$; Average number of reflection events before a photon escapes $N = 11.7$.

Would a hypothetical planar reflector that reflected light coming from any incoming direction towards the surface normal be useful? Is there a physical principle that prevents a reflector from reflecting light in such a way?

Solution: Although such a reflector would be very useful, such a reflector would unfortunately violate the **conservation of radiance theorem** (previously called the conservation of brightness theorem), which states that it is impossible to increase the radiance of light by a passive optical system beyond a value of L/n^2 where L is the radiance in vacuum and n is the refractive index of the medium in which the light propagates.

12.6 References

- Adachi S. in *Properties of Gallium Arsenide* EMIS Datareview Ser. **2**, 513 INSPEC (IEE, New York, 1990)
- Adachi S., Kato H., Moki A., and Ohtsuka K. "Refractive index of $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$ quaternary alloys" *Journal of Applied Physics* **75**, 478 (1994)
- Babic D. I., Piprek L., and Bowers J. E. in *Vertical-Cavity Surface-Emitting Lasers* edited by C. W. Wilmsen, H. Temkin, and L. A. Coldren (Cambridge University Press, Cambridge, 1999)
- Bell L. *The Telescope* this book discusses metal reflectors in a historical context; see pp. 220–227 (McGraw Hill, New York, 1922)
- Björk G., Yamamoto Y., and Heitmann H. "Spontaneous emission control in semiconductor microcavities" in *Confined Electrons and Photons* edited by E. Burstein and C. Weisbuch (Plenum Press, New York, 1995)
- Born M. and Wolf E. *Principles of Optics* 6th edition (Pergamon Press, New York, 1989)
- Bruyant A., Lérondel G., Reece P. J., and Gal M. "All-silicon omnidirectional mirrors based on one-dimensional photonic crystals" *Applied Physics Letters* **82**, 3227 (2003)
- Chen K. M., Sparks A. W., Luan H.-C., Lim D. R., Wada K., and Kimerling L. C. "SiO₂ / TiO₂ omnidirectional reflector and microcavity resonator via the sol-gel method" *Applied Physics Letters* **75**, 3805 (1999)
- Chiou S.-W., Lee C. P., Huang C. K., and Chen C. W. "Wide angle distributed Bragg reflectors for 590 nm amber AlGaInP light-emitting diodes" *Journal of Applied Physics* **87**, 2052 (2000)
- Chiou S.-W., Chang H., Chen T.-P., and Chang C.-S. "Light emitting diodes and fabrication method thereof" US Patent 6,552,369 (2003)
- Coldren L. A. and Corzine S. W. *Diode Lasers and Photonics Integrated Circuits* (John Wiley and Sons, New York, 1995)
- Drude P. "Optische Eigenschaften und Elektronen Theorie I" ("Optical properties and electron theory I") *Annalen der Physik* **14**, 677 (1904); see also "Optische Eigenschaften und Elektronen Theorie II" ("Optical properties and electron theory II") *Annalen der Physik* **14**, 936 (1904)
- Gessmann Th., Schubert E. F., Graff J. W., Streubel K., and Karnutsch C. "Omni-directionally reflective contacts for light-emitting diodes" *IEEE Electron Devices Letters* **24**, 683 (2003)
- Fink Y., Winn J. N., Fan S., Chen C., Michel J., Joannopoulos J. D., Thomas E. L. "A dielectric omnidirectional reflector" *Science* **282**, 1679 (1998)
- Hecht J. *Understanding Fiber Optics* (Pearson Education, Upper Saddle River NJ, 2001)
- Horng R. H., Wuu D. S., Wei S. C., Huang M. F., Chang K. H., Liu P. H., and Lin K. C. "AlGaInP / AuBe / glass light-emitting diodes fabricated by wafer bonding technology" *Applied Physics Letters* **75**, 154 (1999a)
- Horng R. H., Wuu D. S., Wei S. C., Tseng C. Y., Huang M. F., Chang K. H., Liu P. H., and Lin K. C. "AlGaInP light-emitting diodes with mirror substrates fabricated by wafer bonding" *Applied Physics Letters* **75**, 3054 (1999b)
- Jewell J. L., Huang K. F., Tai K., Lee Y. H., Fischer R. J., McCall S. L., and Cho A. Y. "Vertical cavity single quantum well laser" *Applied Physics Letters* **55**, 424 (1989)
- Jewell J. L. personal communication (1992)
- Karim A. M. "Wafer bonded 1.55 μm vertical cavity laser arrays for wavelength division multiplexing" Dissertation, University of California Santa Barbara (2001)
- Kato T., Susawa H., Hirotani M., Saka T., Ohashi Y., Shichi E., and Shibata S. "GaAs / GaAlAs surface emitting IR LED with Bragg reflector grown by MOCVD" *Journal of Crystal Growth* **107**, 832 (1991)
- Kepler J. *Dioptre* (1611)
- Kim J. K., Gessmann T., Luo H., and Schubert E. F. "GaN light-emitting diodes with RuO₂ / SiO₂ / Ag omnidirectional reflector" *Applied Physics Letters* **84**, 4508 (2004)
- Kim J. K., Luo H., Xi Y., Shah J. M., Gessmann T., and Schubert E. F. "Light extraction in GaN light-emitting diodes using diffuse omnidirectional reflectors" *Journal of the Electrochemical Society* **153**, G105 (2006)
- Kish F. A. and Fletcher R. M. "AlGaInP light-emitting diodes" in *High Brightness Light-Emitting Diodes* edited by G. B. Stringfellow and M. G. Crawford, Semiconductors and Semimetals **48** (Academic, San

- Diego, 1997)
- Koyama F., Kinoshita S., and Iga K. "Room temperature continuous wave lasing characteristics of a GaAs vertical-cavity surface-emitting laser" *Applied Physics Letters* **55**, 221 (1989)
- Li H., Gu G., Chen H., and Zhu S. "Disordered dielectric high reflectors with broadband from visible to infrared" *Appl. Phys. Lett.* **74**, 3260 (1999)
- Lorentz H. A. *The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat* (Teubner, Leipzig, Germany, 1909)
- Margalith T., Buchinsky O., Cohen D. A., Abare A. C., Hansen M., DenBaars S. P., and Coldren L. A. "Indium tin oxide contacts to gallium nitride optoelectronic devices" *Applied Physics Letters* **74**, 3930 (1999)
- McCall S. L., Levi A. F. J., Slusher R. E., Pearton S. J., and Logan R. A. "Whispering-gallery mode microdisk lasers" *Applied Physics Letters* **60**, 289 (1992)
- Mergel D., Stass W., Ehl G., and Barthel D. "Oxygen incorporation in thin films of $\text{In}_2\text{O}_3:\text{Sn}$ prepared by radio frequency sputtering" *Journal of Applied Physics* **88**, 2437 (2000)
- Mochizuki M. "Film mirror for solar heat power generation, method of manufacturing film mirror for solar heat generation, and reflection device for solar heat power generation" US Patent Application Publication 2013 / 0011666 (2013)
- Palik E. D. *Handbook of Optical Constants of Solids* (Academic Press, San Diego, 1998)
- Ray S., Banerjee R., Basu N., Batabyal A. K., and Barua A. K. "Properties of tin doped indium oxide thin films prepared by magnetron sputtering" *Journal of Applied Physics* **54**, 3497 (1983)
- Saleh B. E. A. and Teich M. C. *Fundamentals of Photonics*, 2nd edition (Wiley-Interscience, New York, 2007)
- Schubert E. F., Tu L.-W., Zydzik G. J., Kopf R. F., Benvenuti A., and Pinto M. R. "Elimination of heterojunction band discontinuities by modulation doping" *Applied Physics Letters* **60**, 466 (1992a)
- Schubert E. F., Tu L.-W., and Zydzik G. J. "Elimination of heterojunction band discontinuities" US Patent No. 5,170,407 (1992b)
- Schubert E. F. "Light-emitting diode with omni-directional reflector" US Patent application 60 / 339,335 (2001)
- Schubert E. F. "Light-emitting diode with omni-directional reflector" US Patent 6,784,462; issued Aug. 31 (2004)
- Sheu J. K., Su Y. K., Chi G. C., Jou M. J., and Chang C. M. "Effects of thermal annealing on the indium tin oxide Schottky contacts of n-GaN" *Applied Physics Letters* **72**, 3317 (1998)
- Shin J. H., Shin S. H., Park J. I., and Kim H. H. "Properties of dc magnetron sputtered indium tin oxide films on polymeric substrates at room temperature" *Journal of Applied Physics* **89**, 5199 (2001)
- Smith G. M., Forbes D. V., Coleman J. J., and Verheyen J. T. "Optical properties of reactive ion etched corner reflector strained-layer InGaAs–GaAs–AlGaAs quantum-well lasers" *IEEE Photonics Technol. Lett.* **5**, 873 (1993)
- Spisz E. W., Weigand A. J., Bowman R. L., and Jack J. R. "Solar absorptances and spectral reflectances of 12 metals for temperatures ranging from 300 K to 500 K" NASA Technical Note TN D-5353 (August 1969)
- Streubel K., personal communication (2000)
- Tu L. W., Schubert E. F., Kopf R. F., Zydzik G. J., Hong M., Chu S. N. G., and Mannaerts J. P. "Vertical cavity surface emitting lasers with semitransparent metallic mirrors and high quantum efficiencies" *Applied Physics Letters* **57**, 2045 (1990)
- Weber M. F., Stover C. A., Gilbert L. R., Nevitt T. J., and Onderkirk A. J. "Giant birefringent optics in multilayer polymer mirrors" *Science* **287**, 2451 (2000)
- Xi Y., Li X., Kim J. K., Mont F., Gessmann Th., Luo H., and Schubert E. F. "Quantitative assessment of diffusivity and specularity of surface-textured reflectors for light extraction in light-emitting diodes" *Journal of Vacuum Science and Technology A* **24**, 1627 (2006)
- Yariv A. *Quantum Electronics* 3rd edition (John Wiley and Sons, New York, 1989)
- Yariv A. and Nakamura M. "Periodic structures for integrated optics" *IEEE Journal of Quantum Electronics* **QE-13**, 233 (1977)

13

13 – Definition of efficiencies

The efficiency of a light-emitting device may be limited by a variety of independent factors. These factors include (i) the material quality of the active region (associated with the radiative efficiency and the internal quantum efficiency), (ii) the ability to extract light from the LED chip (associated with the light-extraction efficiency), the (iii) magnitude of the contact resistance (associated with the forward-voltage efficiency), and (iv) the emission spectrum (associated with the luminous efficacy of radiation). The present chapter will define different efficiencies that are relevant for light-emitting devices.

The efficiency loss mechanisms of a semiconductor LED chip, shown in *Figure 13.1*, include: The forward-voltage efficiency (VfE), the injection efficiency (IE), the radiative efficiency of the active region (RE), the light-extraction efficiency (LEE), and the power-conversion efficiency (PCE). These efficiencies will be discussed below.

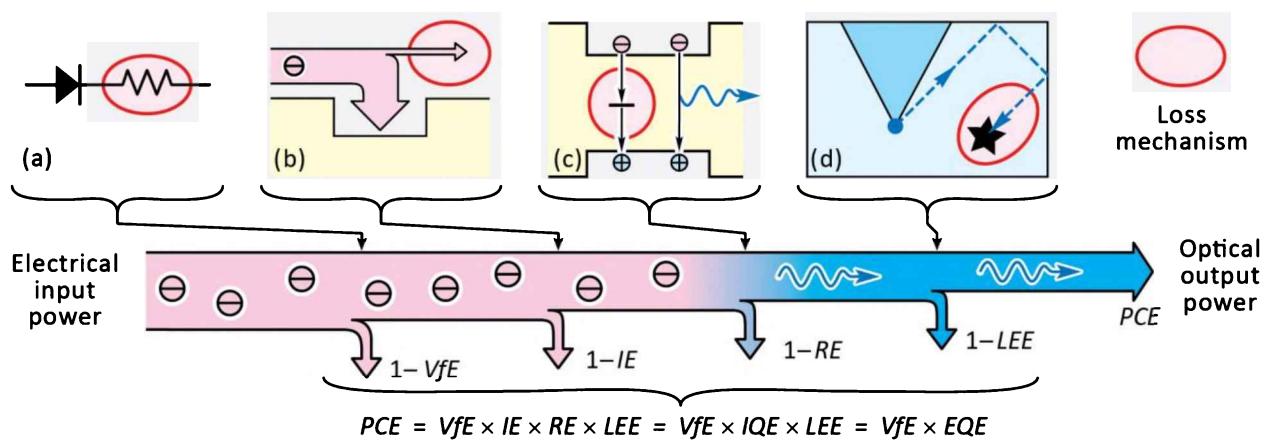


Fig. 13.1: Limitations to the efficiency of a semiconductor LED chip: (a) Forward-voltage efficiency, VfE ; (b) Injection efficiency, IE ; (c) Radiative efficiency, RE ; (d) Light-extraction efficiency, LEE . Also given are the internal quantum efficiency (IQE), external quantum efficiency (EQE), and power conversion efficiency (PCE).

With respect to semantics, we note that an **efficiency** is commonly measured in percent (%). An **efficacy** is commonly measured in units other than percent (%), e.g. in lm/W. However, this rule is not followed consistently by the technical community, in part due to the fact that many

languages other than English do not have distinct words for “efficiency” and “efficacy”. Finally, we note that the term ***quantum efficiency*** is commonly used when counting quanta, for example, when comparing the light quanta (photons) emitted by an LED with the charge quanta (electrons) injected into an LED.

13.1 Radiative efficiency (RE) of active region

Electrons and holes residing in the active region can recombine either radiatively (that is, accompanied by the emission of a photon) or non-radiatively (that is, without emission of a photon). Each of these processes has an associated lifetime; the reciprocal of the lifetime is a recombination probability per unit time or simply recombination rate. The radiative efficiency (RE) of the active region, is given by

$$\begin{aligned} RE &= \frac{\text{number of photons emitted from active region per second}}{\text{number of electrons injected into active region per second}} \\ &= \frac{\tau_{\text{radiative}}^{-1}}{\tau_{\text{radiative}}^{-1} + \tau_{\text{non-radiative}}^{-1}} \end{aligned} \quad (13.1)$$

A radiative and a non-radiative recombination process and the associated lifetimes are illustrated in **Figure 13.2**.

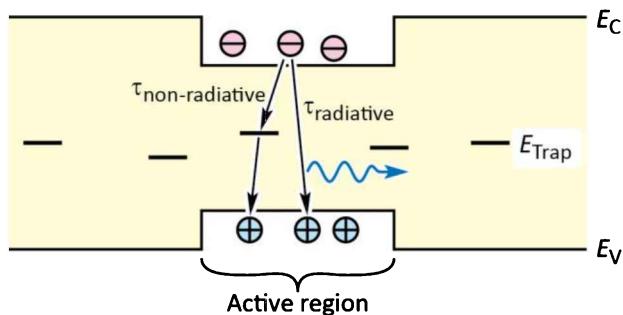


Fig. 13.2: Radiative and trap-mediated non-radiative recombination in the active region of an LED. Each recombination process has an associated lifetime.

Measurement of the radiative efficiency (RE)

The radiative efficiency (RE) can be determined experimentally by the following methods:

In a ***first method*** to measure the *RE*, an optical source is used to selectively excite the semiconductor active region in a broad range of temperatures including a cryogenic temperature, e.g. 4 K or 77 K (see, for example, Watanabe *et al.*, 2003; Fuhrmann *et al.*, 2006). The selective optical excitation is accomplished by using an excitation source having a photon energy greater

than the bandgap of the active region but smaller than the bandgap of the confinement layers. Then the photoluminescence intensity of the sample is measured as a function of temperature including the cryogenic temperature. It is assumed that the radiative efficiency is 100% at the lowest cryogenic temperature, consistent with the temperature dependence of Shockley–Read–Hall (SRH) recombination. The relative decrease in luminescence intensity then is directly representative of the decrease in radiative efficiency. An example of a radiative-efficiency measurement is shown in **Figure 13.3**.

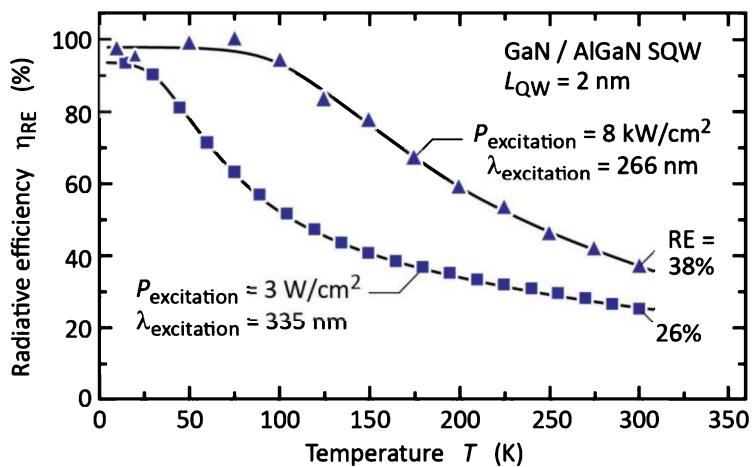


Fig. 13.3: Radiative efficiency (RE) versus temperature for a GaN/AlGaN single QW (SQW) at low and high excitation densities. The RE is assumed to be 100% as the temperature approaches 0 K (after Fuhrmann *et al.*, 2006).

The applicability of the method requires that a “plateau” is found at low temperatures, that is, that there is a range of excitation powers where the efficiency saturates. Such plateau is indeed displayed in **Figure 13.3** at $T < 100$ K for an excitation power density of 8 kW/cm^2 . Furthermore, the measurement should be carried out for QWs surrounded by an n-n or a p-p junction, but not by a p-n junction, so that carrier escape from the active region is minimized (Lin *et al.*, 2012).

Assuming that $RE \rightarrow 100\%$ for $T \rightarrow 0$ K, the temperature dependence of the radiative efficiency was described by the equation (Bimberg *et al.*, 1971; Wang *et al.*, 2007)

$$RE = \frac{1}{1 + \sum_i N_i \exp(-E_i/kT)} \quad (13.2)$$

where E_i are the activation energies of the corresponding nonradiative recombination centers, and N_i are constants related to the concentration of these centers.

A **second method** to determine the RE is based on the selective excitation of the active region and the dependence of photoluminescence intensity on excitation intensity (Ryu *et al.*, 2006; Ding *et al.*, 2007; Dai *et al.*, 2009; Ban *et al.*, 2011). When radiative recombination is dominant,

photoluminescence depends linearly on excitation. However, when non-radiative recombination is dominant, photoluminescence depends quadratically on excitation as inferred from rate equations that include an SRH term (“ A_n ”) and a radiative term (“ Bn^2 ”). Thus, the degree of non-linearity of photoluminescence-intensity-versus-excitation-intensity dependence, allows one to deduce the radiative efficiency (Dai *et al.*, 2009; Ban *et al.*, 2011).

A **third method** to determine the *RE* is based on the width of the *EQE*-versus-carrier-concentration curve that can be deduced from the *EQE*-versus-injection-current curve (Lin *et al.* 2012). The wider this curve, the higher is the radiative efficiency. If the curve develops a plateau, such plateau is an indication of a near-100% radiative efficiency (Lin *et al.* 2012).

For the purpose of the *RE* measurements, it is frequently assumed in the technical literature that the injection efficiency is equal to 100%. In this case the $IQE = RE$, as will become apparent later in this chapter. For this reason, in the context of *RE* measurements, the technical literature frequently uses *IQE* and *RE* synonymously.

13.2 Injection efficiency (*IE*) of carriers into active region

Most but not all carriers are injected into the light-emitting active region of an LED. That is, some carriers may leak out of the active region or transverse the active region without being captured by the active region. The processes of (*i*) injection into the active region and (*ii*) the transport over the active region are illustrated in **Figure 13.4**.

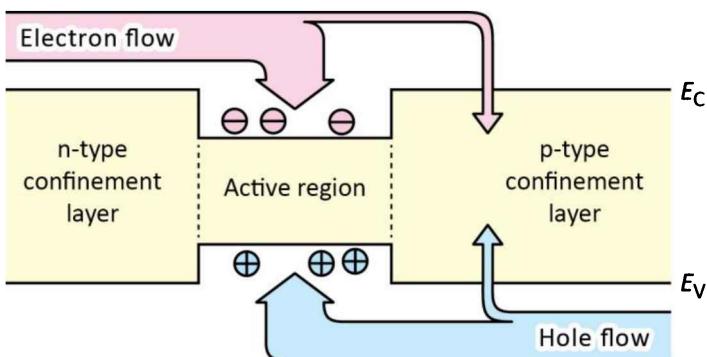


Fig. 13.4: Injection of carriers into the light-emitting active region of an LED with a fraction of electrons leaking out of the active region. They can be assumed to recombine non-radiatively in the p-type confinement region.

As a matter of principle, both electrons and holes can leak out of the active region. However, in practice, it is mostly the electrons that leak out (due to their usually lighter mass and higher mobility). The carrier ***injection efficiency*** (*IE*) is defined as the probability that electrons are injected into the active region, that is

$$IE = \frac{\text{number of electrons injected into active region per second}}{\text{number of electrons injected into LED per second}}. \quad (13.3)$$

In the steady state, the number of electrons injected into the active region is equal to the number of holes injected into the active region. Therefore, the injection efficiency of holes always is equal to the injection efficiency of electrons.

Lack of efficient injection can be caused, for example, by weak carrier confinement, i.e. low heterostructure barriers that confine carriers to the active region. Furthermore, the large polarization fields in GaN heterostructures can promote carrier escape. What happens to electrons *not* injected into the active region? It can be generally assumed that these electrons recombine non-radiatively in either the confinement regions or at the contacts.

13.3 Internal quantum efficiency (IQE)

The **internal quantum efficiency** (IQE) is defined as

$$\begin{aligned} IQE &= \frac{\text{number of photons emitted from active region per second}}{\text{number of electrons injected into LED per second}} \\ &= \frac{P_{\text{internal}} / (hv)}{I / e} \end{aligned} \quad (13.4)$$

where I is the injection current and P_{internal} is the optical power emitted from the active region (internal emission).

The active region of an *ideal* LED emits *one photon* for *every electron* injected. Each charge quantum-particle (electron) produces one light quantum-particle (photon). Thus, the active region of an *ideal* LED has an IQE of unity (100%).

It can be verified that the internal quantum efficiency is the product of radiative efficiency and injection efficiency, that is

$$IQE = RE \times IE \quad (13.5)$$

The separation of internal quantum efficiency into radiative efficiency and injection efficiency is important to identify limitations to the IQE.

Measurement of the internal quantum efficiency (IQE)

A method to determine the IQE is based on the measurement of the external quantum efficiency

(*EQE*) and the inference of the light-extraction efficiency (*LEE*) from ray-tracing simulations or other considerations (Krames *et al.*, 2007). Since the external quantum efficiency is the product of internal quantum efficiency and light-extraction efficiency ($EQE = IQE \times LEE$), the internal quantum efficiency can be determined.

13.4 Light-extraction efficiency (*LEE*)

Photons emitted by the active region should escape from the LED semiconductor chip. In an ideal LED, all photons emitted by the active region are also emitted into free space. Such an LED has unity light-extraction efficiency. However, in a real LED, not all the optical power emitted from the active region is emitted into free space. Some photons may never leave the semiconductor chip. This is due to several possible loss mechanisms. For example, light emitted by the active region can be reabsorbed in the substrate of the LED (assuming that the substrate is absorbing at the emission wavelength). Light may be incident on a metallic contact surface and some light may be absorbed by the metal. In addition, the phenomenon of *total internal reflection*, also referred to as the *trapped-light phenomenon*, reduces the ability of light to escape from the semiconductor. The ***light-extraction efficiency*** (*LEE*) is defined as

$$\begin{aligned} LEE &= \frac{\text{number of photons emitted into free space per second}}{\text{number of photons emitted from active region per second}} \\ &= \frac{P_{LED} / (hv)}{P_{internal} / (hv)} \end{aligned} \tag{13.6}$$

where P_{LED} is the optical power emitted by the LED into free space.

The light-extraction efficiency can be a severe limitation for high-performance LEDs. It is quite difficult to increase the extraction efficiency beyond 50% without resorting to sophisticated device processing. The light-extraction efficiency cannot be measured directly. It is frequently inferred from ray-tracing simulations.

13.5 External quantum efficiency (*EQE*)

The external quantum efficiency gives the ratio of the number of useable light particles to the number of injected charge particles. The ***external quantum efficiency*** (*EQE*) is defined as

$$EQE = \frac{\text{number of photons emitted into free space per second}}{\text{number of electrons injected into LED per second}}$$

$$= \frac{P_{LED} / (hv)}{I / e} . \quad (13.7)$$

It can be verified that the external quantum efficiency is the product of internal quantum efficiency and light-extraction efficiency, that is

$$EQE = IQE \times LEE . \quad (13.8)$$

The external quantum efficiency of an LED can be determined by measuring the optical power emitted by the LED and the injection current of the LED.

13.6 Forward-voltage efficiency (VfE)

The *forward-voltage efficiency* is also referred to as the *electrical efficiency* of an LED. If an LED has a parasitic series resistance (e.g. and excessive contact resistance), the forward voltage will be larger than the ideal value, which is $V_f = hv/e$, where hv is the photon energy. Any ohmic resistance will convert input power to thermal power. The thermal power generated in a parasitic series resistance is given by

$$P_R = I \left(V_f - \frac{hv}{e} \right) . \quad (13.9)$$

Thus, the fraction of the input electrical power available to generate light, called the *forward voltage efficiency* (VfE), is given by

$$VfE = \frac{IV_f - P_R}{IV_f} = \frac{IV_f - I(V_f - hv/e)}{IV_f} = \frac{hv}{eV_f} . \quad (13.10)$$

The forward-voltage efficiency can be determined by using the above equation and by inserting the measured emission energy and forward voltage.

A more detailed analysis of the forward-voltage efficiency can be based on the Shockley diode equation. This equation reveals that a current can flow even if the applied voltage is less than hv/e , i.e. less than E_g/e , so that the forward-voltage efficiency exceeds 100% (!). The operation of LEDs at forward-voltage efficiencies exceeding 100% is undoubtedly of great interest and has been discussed in detail by Santhanam *et al.* (2012).

In theory, operating an LED having an IQE of 100% in the range where $VfE > 100\%$, would yield a power-conversion efficiency greater than 100%. As a result, the LED will cool down and

could be used as a micro-refrigerator (Santhanam *et al.*, 2012). The cooling can be understood by considering that, for $VfE > 100\%$, only the high energy electrons (having energy of several kT) will be injected into the active region. The selective removal of high-energy electrons from the electron gas will lower the temperature of the LED below its ambient temperature. Although this goal has not been reached in practice, it remains an interesting aspect of LED technology.

The operating regime where $VfE > 100\%$, is limited to relatively small currents. At such small currents, Shockley-Read-Hall (SRH) recombination severely limits the IQE , so that the power-conversion efficiency of the LED does not exceed 100%. Accordingly, a significant cooling of the LED lattice temperature has not been demonstrated.

13.7 Power-conversion efficiency (PCE)

The **power-conversion efficiency** (PCE) of a light-emitting device is defined as the optical output power emitted by the LED, $P_{optical}$, divided by the electrical input power, that is

$$PCE = \frac{P_{optical}}{P_{electrical}} = \frac{P_{optical}}{IV} \quad (13.11)$$

where IV is the electrical power provided to the LED (the product of current I and voltage V). Informally, the power-conversion efficiency is also called the **wall-plug efficiency**.

It can be verified that the power-conversion efficiency is the product of external quantum efficiency and forward voltage efficiency, that is

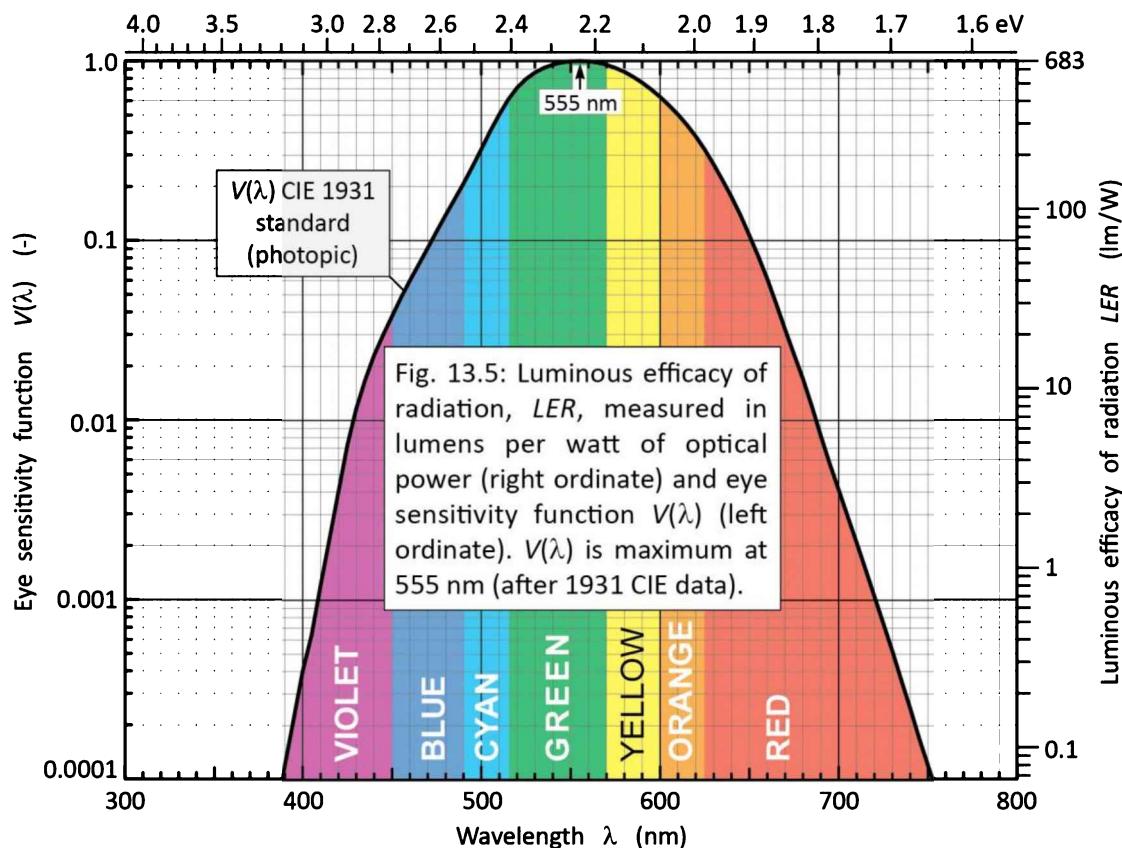
$$PCE = EQE \times VfE . \quad (13.12)$$

The power-conversion efficiency of a light-emitting device can be assessed by measuring the electrical input power and the optical output power.

13.8 Luminous efficacy of radiation (LER)

For many visible-spectrum sources, such as general lighting sources and indicator-light sources, the intended recipient of the light is the human eye. By definition, the human eye is sensitive in the visible spectral range but has no sensitivity in the ultraviolet (UV) and infrared (IR) range. Therefore, when converting the power of a light source to the power of a light source as perceived by the human eye, the spectral dependence of the human eye response must be taken into account.

The human eye sensitivity curve is shown in **Figure 13.5** (CIE, 1931, 1932). The left-hand ordinate shows the relative eye sensitivity curve $V(\lambda)$. The curve has a maximum at $\lambda = 555$ nm where its value is normalized to the unit value, that is, $V(555 \text{ nm}) = 1.0$. The right-hand ordinate shows the conversion between radiometric units (watts) to photometric units (lumens). The illustration shows that 1 W of optical power at 555 nm corresponds to 683 lm. 1 W of optical power at e.g. 700 nm corresponds to a much smaller value, namely 2.7 lm, illustrating the strong wavelength dependence of human eye sensitivity.



The power of a light source as perceived by the human eye is generally referred to as the luminous flux of the source. The luminous flux is given by

$$\Phi_{lum} = 683 \frac{\text{lm}}{\text{W}} \int_{\lambda} V(\lambda) P(\lambda) d\lambda \quad (13.13)$$

where $P(\lambda)$ is the power spectral density, i.e. the light power emitted per unit wavelength, and the prefactor 683 lm / W is a normalization factor. The optical power emitted by the source is

$$P_{optical} = \int_{\lambda} P(\lambda) d\lambda . \quad (13.14)$$

The ***luminous efficacy of radiation*** (*LER*) is defined as

$$LER = \frac{\Phi_{lum}}{P_{optical}} . \quad (13.15)$$

This quantity depends only on the spectrum of the light and is independent of any electrical properties of the light-emitting device.

In the green spectral range at 555 nm, $LER = 683 \text{ lm/W}$. Monochromatic light at the very fringes of the visible spectrum (violet and red) have a much lower luminous efficacy of radiation, e.g. a few lm/W.

For white light, the planckian black-body spectrum is a common model. The Sun and incandescent lamps have, with good approximation, a planckian spectrum. Planckian white light have a maximum luminous efficacy of radiation of 93 lm/W at a color temperature of 6300 K (Tsao *et al.*, 2010). The relatively low efficacy of planckian sources is due to the fact that these sources have significant spectral components in the UV and IR.

Tri-chromatic white light, which is composed of emission lines in the red, green, and blue, generally have good color rendering capabilities; trichromatic light can have a luminous efficacy of radiation of 300~330 lm/W depending on the spectral width of the emission lines (Chhajed *et al.*, 2005). Thus, a luminous efficacy of about $LER = 330 \text{ lm/W}$ can be considered as an approximate upper limit for white light with good color rendering capabilities (Chhajed *et al.*, 2005).

Di-chromatic white light consists of two complementary emission lines: blue and yellow. A luminous efficacy of radiation of about $LER = 450 \text{ lm/W}$ can be considered to be an approximate upper limit for di-chromatic white light (MacAdam, 1950; Li *et al.*, 2003). The color rendering capabilities of di-chromatic white light are low and insufficient for most illumination applications.

13.9 Luminous efficacy of a source (*LES*)

The luminous efficacy of a light source is a key figure of merit for white-light sources. The ***luminous efficacy of a source*** (*LES*) is defined as the ratio of the luminous flux emitted by a source to the electrical input power consumed by the source, that is

$$LES = \frac{\Phi_{lum}}{P_{electrical}} = \frac{\Phi_{lum}}{I V} . \quad (13.16)$$

The luminous efficacy of a source depends on electrical-to-optical conversion efficiency of the source as well as the efficacy of the radiation emitted by the source. For a source with perfect electrical-to-optical conversion efficiency, it is $LES = LER$. Thus, based on the preceding section, white LED sources with good color-rendering capabilities have an upper limit for the luminous source efficacy of about $LES = LER \approx 330 \text{ lm/W}$. If the color-rendering capability is not considered, the upper limit for the luminous source efficacy of about $LES = LER \approx 450 \text{ lm/W}$.

The luminous efficacy of a white light source not only determines the energy consumption of the source but also other parameters such as the heat generated during operation, the thermal heat-sinking requirements, reliability, and operating costs. For this reason, the luminous source efficacy is well publicized figure of merit for white LEDs.

Demonstrated luminous source efficacies of white sources with good color rendering properties range from 12 lm/W for incandescent sources to values greater than 200 lm/W for LED-based sources.

13.10 Efficiency of wavelength-converter materials (phosphors)

Wavelength-converters (λ -converters) is synonymous with **phosphors**. They are materials that convert photons of a shorter wavelength to photons of a longer wavelength. The mechanisms determining the efficiency of a phosphor, shown in **Figure 13.6**, include: (i) the phosphor quantum efficiency, i.e. the quantum efficiency of the λ -converter material and (ii) the inherent and unavoidable quantum-mechanical-energy loss and its associated efficiency incurred during wavelength conversion.

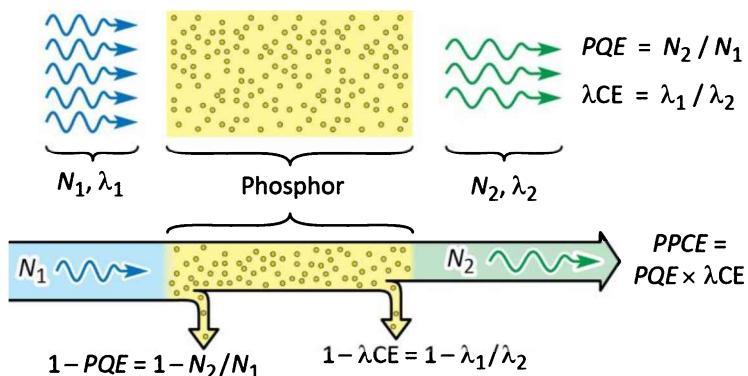


Fig. 13.6: N_1 short-wavelength photons (wavelength λ_1) absorbed by the phosphor, and N_2 long-wavelength converted photons (λ_2) emanating from the phosphor. PQE = phosphor quantum efficiency. λCE = wavelength conversion efficiency. $PPCE$ = phosphor power-conversion efficiency.

The **phosphor quantum efficiency** (PQE) of the converter material, is given by

$$PQE = \frac{\text{number of photons the phosphor emits into free space per second}}{\text{number of photons the phosphor absorbs per second}}. \quad (13.17)$$

For LED applications, solid inorganic phosphors, such as YAG:Ce, are most common. An inorganic phosphor generally is more stable than an organic phosphor.

The inherent **wavelength-conversion loss** (sometimes called **quantum deficit** or **Stokes shift**) incurred when converting a photon with wavelength λ_1 to a photon with wavelength λ_2 ($\lambda_1 < \lambda_2$) is given by

$$\Delta E = h\nu_1 - h\nu_2 = \frac{hc}{\lambda_1} - \frac{hc}{\lambda_2}. \quad (13.18)$$

Thus, the wavelength-conversion efficiency (λCE) is given by

$$\lambda CE = \frac{h\nu_2}{h\nu_1} = \frac{\lambda_1}{\lambda_2} \quad (13.19)$$

where λ_1 is the wavelength of the photon absorbed by the phosphor and λ_2 is the wavelength of the photon emitted by the phosphor. Note that wavelength-conversion loss is fundamental in nature. The loss cannot be overcome with conventional phosphor materials.

As an example, consider the wavelength-conversion loss of a YAG:Ce phosphor converting 460 nm light into 550 nm light; the λ -conversion efficiency is 84%. The wavelength-conversion loss is highest for wavelength conversion from the UV to the red. For example, the conversion from UV (405 nm) to red (625 nm) has a λ -conversion efficiency of 65%. The low λ -conversion efficiency represents a driving force to employ multiple LEDs (rather than phosphors) in highly efficient lighting systems.

The phosphor power-conversion efficiency (PPCE) is the product of Eqns. (13.17) and (13.19), that is

$$PPCE = \lambda CE \times QE. \quad (13.20)$$

The inherent wavelength-conversion loss is the reason that the theoretical efficiency limit of λ -converter-based white LEDs (including phosphor-based white LEDs) is fundamentally lower than for white-light sources based on multiple LEDs (e.g. RGB LEDs).

Exercise: LED efficiency. Consider an LED with a threshold voltage of $V_{th} = E_g/e = 2.0$ V with a differential resistance of $R_s = 20 \Omega$, so that the I - V characteristic in the forward direction is given by $V = V_{th} + IR_s$. When the device is operated at 20 mA it emits a light power of 4.0 mW of energy $h\nu = E_g$. Determine the (a) external quantum efficiency, (b) internal quantum efficiency, and (c) power conversion efficiency, assuming that the light-extraction efficiency is 50%.

Solution: (a) External quantum efficiency: Number of emitted photons per second is given by $4.0 \text{ mW}/2.0 \text{ eV} = 4.0 \times 10^{-3} \text{ CV s}^{-1}/(2.0 \times 1.602 \times 10^{-19} \text{ CV}) = 1.25 \times 10^{16} \text{ s}^{-1}$. Number of injected electrons per second is given by $20 \text{ mA}/e = 20 \times 10^{-3} \text{ C s}^{-1} / (1.602 \times 10^{-19} \text{ C}) = 1.25 \times 10^{17} \text{ s}^{-1}$. Thus, the external quantum efficiency is $1.25 \times 10^{16} \text{ s}^{-1} / 1.25 \times 10^{17} \text{ s}^{-1} = 10\%$. (b) Using $EQE = IQE \times LEE$ and using the light-extraction efficiency value of 50%, the internal quantum efficiency is 20%. (c) The power conversion efficiency is given by $P_{optical}/(IV) = 4 \text{ mW} / [20 \text{ mA} \times (2.0 \text{ V} + 20 \text{ mA} \times 20 \Omega)] = 4 \text{ mW}/48 \text{ mW} = 8.33\%$.

13.11 References

- Ban K., Yamamoto J.-I., Takeda K., Ide K., Iwaya M., Takeuchi T., Kamiyama S., Akasaki I., and Amano H. “Internal quantum efficiency of whole-composition-range AlGaN multiquantum wells” *Applied Physics Express* **4**, 052101 (2011)
- Bimberg D., Sondergeld M., and Grobe E. “Thermal Dissociation of Excitons Bounds to Neutral Acceptors in High-Purity GaAs” *Physical Review B* **4**, 3451 (1971)
- CIE “Commission Internationale de l’Eclairage Proceedings 1931” (Cambridge University Press, Cambridge UK, 1932); see also *Wikipedia* entry titled “CIE 1931 color space” <http://en.wikipedia.org/wiki/CIE_1931_color_space> (accessed 2010)
- Chhajed S., Xi Y., Li Y.-L., Gessmann T., and Schubert E. F. “Influence of junction temperature on chromaticity and color-rendering properties of trichromatic white-light sources based on light-emitting diodes” *Journal of Applied Physics* **97**, 054506 (2005)
- Dai Q., Schubert M. F., Kim M. H., Kim J. K., Schubert E. F., Koleske D. D., Crawford M. H., Lee S. R., Fischer A. J., Thaler G., and Banas M. A. “Internal quantum efficiency and nonradiative recombination coefficient of GaInN / GaN multiple quantum wells with different dislocation densities” *Applied Physics Letters* **94**, 111109 (2009)
- Ding D., Johnson S. R., Wang J. B., Yu S. Q., and Zhang Y. H. “Determination of spontaneous emission quantum efficiency in InGaAs / GaAs quantum well structures” *Proceedings SPIE* **6841**, 68410D (2007)
- Fuhrmann D., Retzlaff T., Rossow U., Bremers H., Hangleiter A., Ade G., and Hinze P. “Large internal quantum efficiency of In-free UV-emitting GaN / AlGaN quantum-well structures” *Applied Physics Letters* **88**, 191108 (2006)
- Krames M. R., Shchekin O. B., Mueller-Mach R., Mueller G. O., Zhou L., Harbers G., and Crawford M. G. “Status and future of high-power light-emitting diodes for solid-state lighting” *IEEE Journal of Display Technology* **3**, 160 (2007)
- Li Y.-L., Gessmann Th., Schubert E. F., and Sheu J. K. “Carrier dynamics in nitride-based light-emitting p-n junction diodes with two active regions emitting at different wavelengths” *Journal of Applied Physics* **94**, 2167 (2003)
- Lin G.-B., Shan Q., Birkel A. J., Cho J., Schubert E. F., Crawford M. H., Westlake K. R., and Koleske D. D. “Method for determining the radiative efficiency of GaInN quantum wells based on the width of efficiency-versus-carrier-concentration curve” *Applied Physics Letters* **101**, 241104 (2012)
- MacAdam D. L. “Maximum attainable luminous efficiency of various chromaticities” *Journal of the Optical Society of America (Letters to the Editor)* **40**, 120 (1950)
- Ryu H. Y., Ha K. H., Chae J. H., Kim K. S., J. Son K., Nam O. H., Park Y. J., and Shim J. I. “Evaluation of radiative

- efficiency in InGaN blue-violet laser-diode structures using electroluminescence characteristics” *Applied Physics Letters* **89**, 171106 (2006)
- Santhanam P., Gray, Jr. D. J., and Ram R. J. “Thermoelectrically pumped light-emitting diodes operating above unity efficiency” *Physical Review Letters* **108**, 097403 (2012)
- Tsao J. Y., Coltrin M. E., Crawford M. H., and Simmons J. A. “Solid-state lighting: An integrated human factors, technology, and economic perspective” *Proceedings of the IEEE* **98**, 1162 (2010)
- Wang Y., Pei X. J., Xing Z. G., Guo L. W., Jia H. Q., Chen H., and Zhou J. M. “Effects of barrier growth temperature ramp-up time on the photoluminescence of InGaN/GaN quantum wells” *Journal of Applied Physics* **101**, 033509 (2007)
- Watanabe S., Yamada N., Nagashima M., Ueki Y., Sasaki C., Yamada Y., Taguchi T., Tadatomo K. , Okagawa H., and Kudo H. “Internal quantum efficiency of highly-efficient InGaN-based near-ultraviolet light-emitting diodes” *Applied Physics Letters* **83**, 4906 (2003)

14

14 – Designs for high internal quantum efficiency

There are two general possibilities for attaining high internal quantum efficiency. The first possibility is to enhance the radiative recombination probability, and the second possibility is to decrease the non-radiative recombination probability. This can be accomplished in different ways which will be discussed in this chapter.

14.1 Double heterostructures

The lifetimes derived from the bimolecular rate equation show that the radiative rate increases with the free-carrier concentration for both the low excitation limit as well as the high excitation limit. It is therefore important that the region in which recombination occurs has a high carrier concentration. Double heterostructures are an excellent way to achieve such high carrier concentrations. A double heterostructure (DH) consists of the active region in which recombination occurs and two confinement layers cladding the active region. A double heterostructure LED structure is shown schematically in *Figure 14.1*.

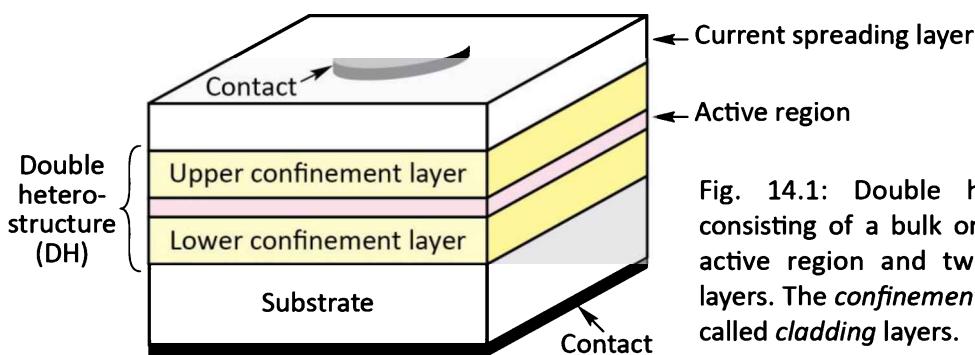


Fig. 14.1: Double heterostructure consisting of a bulk or quantum well active region and two confinement layers. The **confinement layers** are also called **cladding layers**.

The two **cladding** or **confinement layers** have a larger bandgap than the active region. If the bandgap difference between the active and the confinement regions is ΔE_g , then the band discontinuities occurring in the conduction and valence bands follow the relation

$$E_g|_{cladding} - E_g|_{active} = \Delta E_g = \Delta E_C + \Delta E_V . \quad (14.1)$$

Both band discontinuities, ΔE_C and ΔE_V , should be much larger than kT in order to avoid carrier escape from the active region into the confinement regions.

The effect of a double heterostructure on the carrier concentration is illustrated in **Figure 14.2**. Under forward bias, carriers diffuse across the p-n junction. In the case of a homostructure, minority carriers are distributed over the electron and hole diffusion lengths as illustrated in **Figure 14.2 (b)**. In III–V semiconductors, diffusion lengths can be 10 μm or longer.

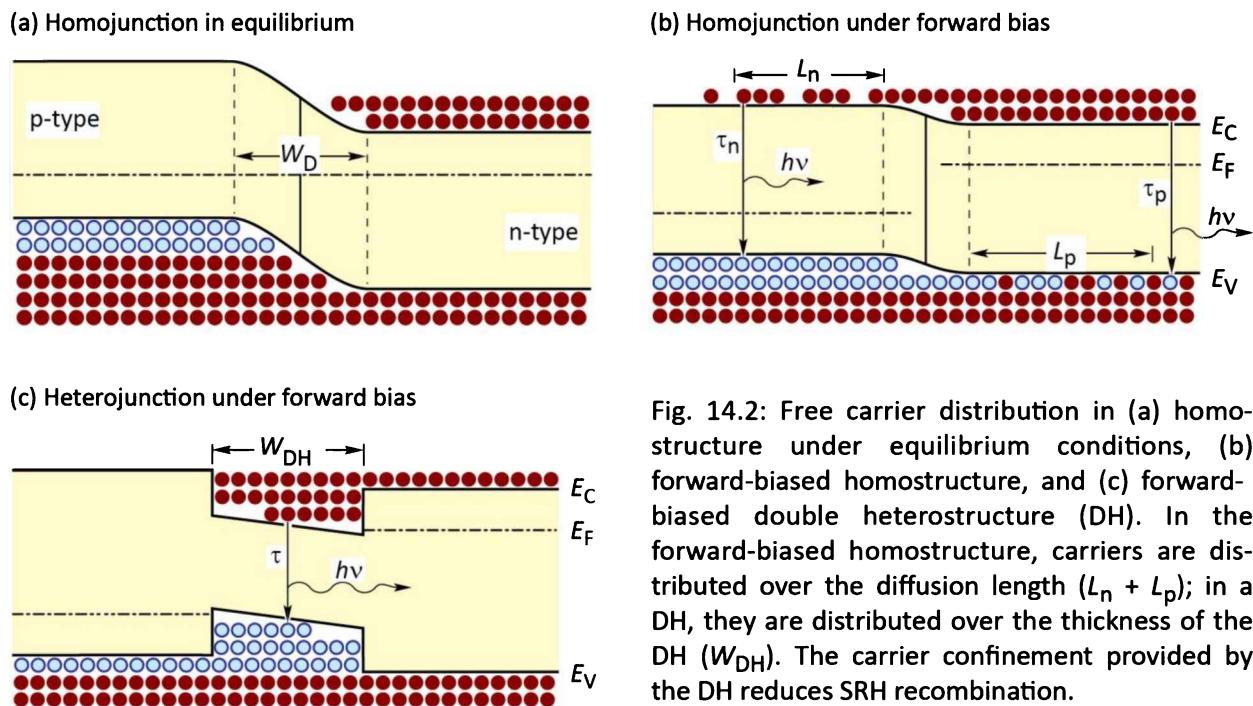


Fig. 14.2: Free carrier distribution in (a) homostructure under equilibrium conditions, (b) forward-biased homostructure, and (c) forward-biased double heterostructure (DH). In the forward-biased homostructure, carriers are distributed over the diffusion length ($L_n + L_p$); in a DH, they are distributed over the thickness of the DH (W_{DH}). The carrier confinement provided by the DH reduces SRH recombination.

The wide distribution of carriers and the correspondingly low carrier concentration (particularly towards the end of the diffusion tail) can be avoided by the employment of double heterostructures. Carriers are confined to the active region of width W_{DH} , as shown in **Figure 14.2 (c)**. To attain good carrier confinement, the barrier heights should be much greater than the thermal energy kT .

Confining carriers to a thin active region will result in high carrier concentrations. Note that the radiative efficiency (RE), given by $RE = Bn^2 / (A_{SRH}n + Bn^2)$, in the limit of high carrier concentrations, approaches 100%. That is, the limit of RE for $n \rightarrow \infty$ is 100% (where we

neglected other high-carrier-concentration effects, such as carrier leakage and Auger recombination). Accordingly, the confinement of carriers to a thin active region is beneficial for attaining a high radiative efficiency.

The advantage of the DH design is further illustrated in *Figure 14.3*. The DH structure has a much smaller number of defects in the recombination region so that non-radiative Shockley-Read-Hall recombination via deep levels is much less significant in a DH (and in quantum-well active regions). At the present time, virtually all high-efficiency LEDs use double heterostructure designs.

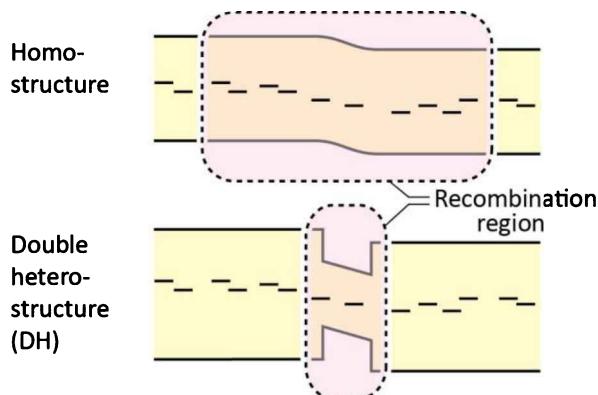


Fig. 14.3: In a double heterostructure (or QW structure), the recombination region, and thus the number of defects contained in this region, is much smaller than in a homostructure. Therefore, non-radiative SRH recombination will play a much smaller role in DH active regions (as well as QW active regions).

The term “double heterostructures” is frequently used for active layers with thickness of 100 to 500 nm. The term “quantum well” active region is used for active layers with thickness of 3 to 100 nm. Single quantum well (SQW) and multi-quantum well (MQW) active regions provide additional carrier confinement, which can further improve the internal quantum efficiency. If a MQW active region is used, the barriers between the wells may impede the flow of carriers between adjacent wells. Thus, the barriers in a MQW active region need to be sufficiently “transparent” (low and / or thin barriers) in order to allow for efficient carrier transport between the wells and to avoid the inhomogeneous distribution of carriers within the active region.

The thickness of the active region in a DH has a strong influence on the internal quantum efficiency of an LED. Typical active region thicknesses are a few tenths of a micrometer and even thinner for quantum well active regions. The dependence of device efficiency on the active region thickness is shown in *Figure 14.4* (Sugawara *et al.*, 1992). Inspection of the figure shows that the optimum thickness for an AlGaInP active region is between 0.15 and 0.75 μm . At the present time MQW active regions are more common than DH active regions.

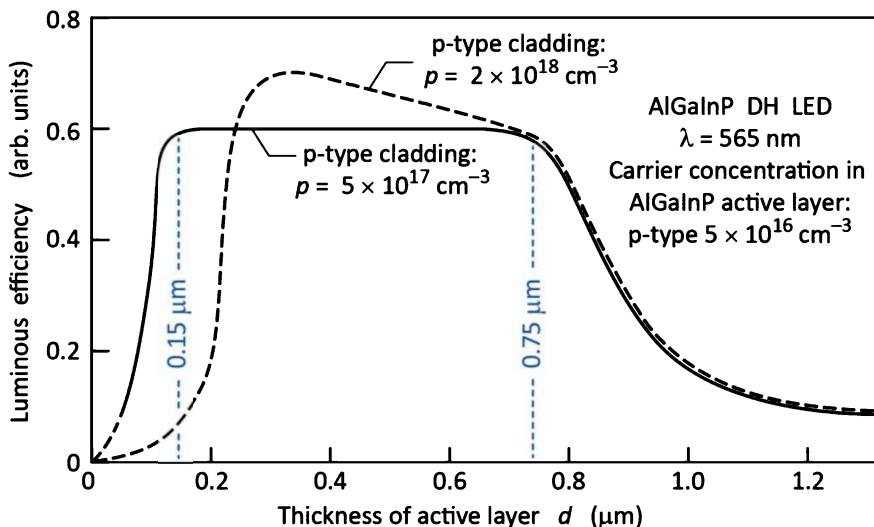


Fig. 14.4: Dependence of the luminous efficiency of an AlGaNp double heterostructure LED emitting at 565 nm on the active layer thickness. The optimum active region thickness is 0.15 to $0.75 \mu\text{m}$ (after Sugawara *et al.*, 1992).

If the active region in a double heterostructure becomes too thick, e.g. larger than the diffusion length of carriers, the advantage of the double heterostructure is lost and carriers are distributed as they are in homojunctions. On the other hand, if the active region of a double heterostructure becomes too thin, the active region tends to overflow at high injection current levels.

14.2 Doping of active region

Doping of the active region and confinement layers plays a crucial role in the efficiency of double heterostructure (DH) LEDs. The influence of the doping on the internal efficiency is multifaceted. Firstly, we consider the doping of the active region.

The active region of III–V arsenide and phosphide DH LEDs must not be heavily doped. Heavy doping with either p-type or n-type dopants would place the p-n junction effectively at the edge of the DH well region, i.e. at the active/confinement interface, thereby promoting carrier spill-over into one of the confinement regions. Diffusion of carriers into the cladding region decreases the radiative efficiency. Thus, heavy doping of the active region is rarely done in III–V arsenide and phosphide DH LEDs.

Therefore, it is required that the active region is *either* doped at a level *lower* than the doping concentration in the confinement regions *or* left undoped. Typically, the doping concentration in intentionally doped active regions is either p-type or n-type in the 10^{16} to low 10^{17} cm^{-3} range. Frequently the active layer is left undoped.

P-type doping of the active region is more common than n-type doping of the active region, due to the generally longer electron-minority-carrier diffusion length compared with the hole-

minority-carrier diffusion length. (Note that electrons generally have a higher mobility than holes in III–V semiconductors.) Thus, p-type doping of the active region ensures a more uniform carrier distribution throughout the active region.

The dependence of device efficiency on the doping concentration of the active region in an AlGaInP DH LED is shown in **Figure 14.5** (Sugawara *et al.*, 1992). Inspection of the figure shows that a high quantum efficiency is obtained only for n-type doping concentrations below $5 \times 10^{16} \text{ cm}^{-3}$, and below $1 \times 10^{17} \text{ cm}^{-3}$ for p-type active layers.

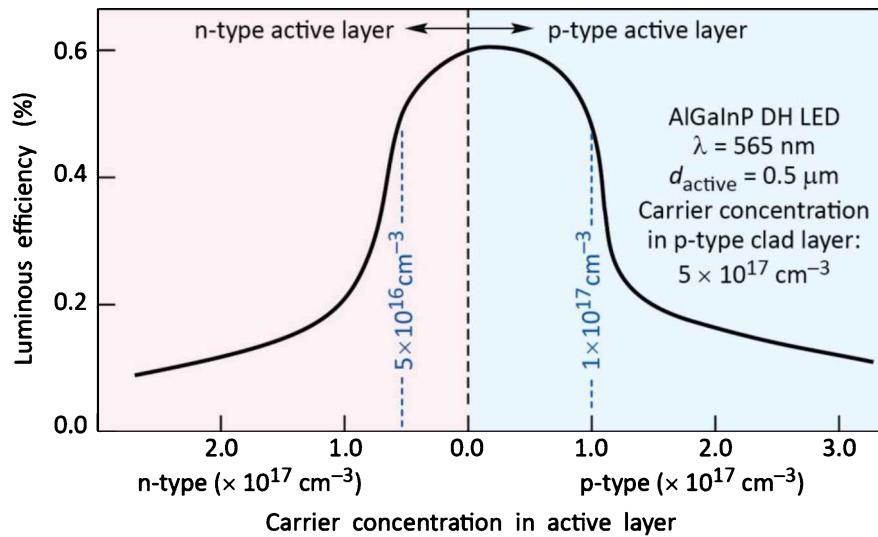


Fig. 14.5: Dependence of the luminous efficiency of an AlGaInP double heterostructure LED emitting at 565 nm on the active layer doping concentration. A low doping concentration in the active region is desirable (after Sugawara *et al.*, 1992).

Figure 14.5 also reveals that light p-type doping of the active region is preferable over light n-type doping. Therefore, most active layers of LEDs and lasers are lightly doped with acceptors. Electrons with their larger diffusion length are also more likely to diffuse into the p-type confinement layer if the active region is doped n-type.

Intentional doping of the active region can have advantages as well as disadvantages. The carrier lifetime depends on the concentration of majority carriers. In the low-excitation regime, the radiative carrier lifetime decreases with increasing free carrier concentration, i.e. doping concentration. As a result, the radiative efficiency increases. An example of a material whose luminescence efficiency increases with doping concentration is Be-doped GaAs. It is well known that the radiative efficiency of Be-doped GaAs increases with the Be doping level in the moderate doping concentration range.

On the other hand, dopants may, especially at high concentrations, introduce defects that act as recombination centers. High concentrations of intentional dopants lead to an increased concentration of native defects due to the dependence of the native and non-native defect concentrations on the Fermi level (see, for example, Longini and Greene, 1956; Baraff and Schluter, 1985; Walukiewicz, 1988, 1989, 1994; Neugebauer and Van de Walle, 1999).

The epitaxial growth process may also depend on doping. Doping atoms can act as *surfactants*, i.e. surface-active reagents. For example, a surfactant can increase the surface diffusion coefficient, thereby improving crystal quality. There are many other ways by which surfactants influence the growth process. These processes are generally not understood in great detail. However, the improvement of crystal quality has been found in a number of cases. During the growth of InGaN, for example, a marked improvement of crystal quality has been found upon doping with silicon (Nakamura *et al.*, 1996, 1998). The quantum barriers in III–V nitride multiple-quantum well structures are in fact often doped with silicon at a high level, e.g. $2 \times 10^{18} \text{ cm}^{-3}$. The increase in device efficiency found for high doping may be due to the screening of internal polarization fields which reduces the overall potential drop within the active region.

14.3 p-n junction displacement

The displacement of the p-n junction from its intended location into the cladding layer can be a significant problem in DH LED structures. Usually, the lower confinement layer is n-type, the upper confinement layer is p-type, and the active region is undoped or lightly doped with n- or p-type dopants. However, if dopant redistribution occurs, the p-n junction can be displaced into one of the confinement layers. The diffusion of dopants can occur during growth and be caused by high growth temperature, a long growth time, or a strongly diffusing dopant. Dopants can redistribute due to diffusion, segregation, and drift.

Frequently, acceptors from the top confinement layer diffuse into the active region and also into the lower confinement layer. Impurities such as Zn (zinc) and Be (beryllium) are small atoms that can easily diffuse through the crystal lattice. In addition, Zn and Be are known to have a strongly concentration-dependent diffusion coefficient. If a certain critical concentration is exceeded, Zn and Be acceptors diffuse very rapidly. As a result, the device will not work well and will not emit light at the intended emission wavelength.

An example of a Zn acceptor profile in a GaInAsP/InP double heterostructure measured by secondary ion mass spectrometry (SIMS) is shown in **Figure 14.6**. In **Figure 14.6 (a)**, the top confinement layer has a moderate intended doping concentration of $2 \times 10^{17} \text{ cm}^{-3}$. The Zn profile indicates that the Zn is mostly confined to the upper confinement layer, even though some Zn has evidently diffused into the active region. In **Figure 14.6 (b)**, however, the top confinement layer has a high intended doping concentration of $2 \times 10^{19} \text{ cm}^{-3}$. The profile indicates that Zn has strongly diffused into the active region. The p-n junction is displaced to the edge of the active region. As a result, the device shown in **Figure 14.6 (b)** has a much lower quantum efficiency than the device shown in **Figure 14.6 (a)**.

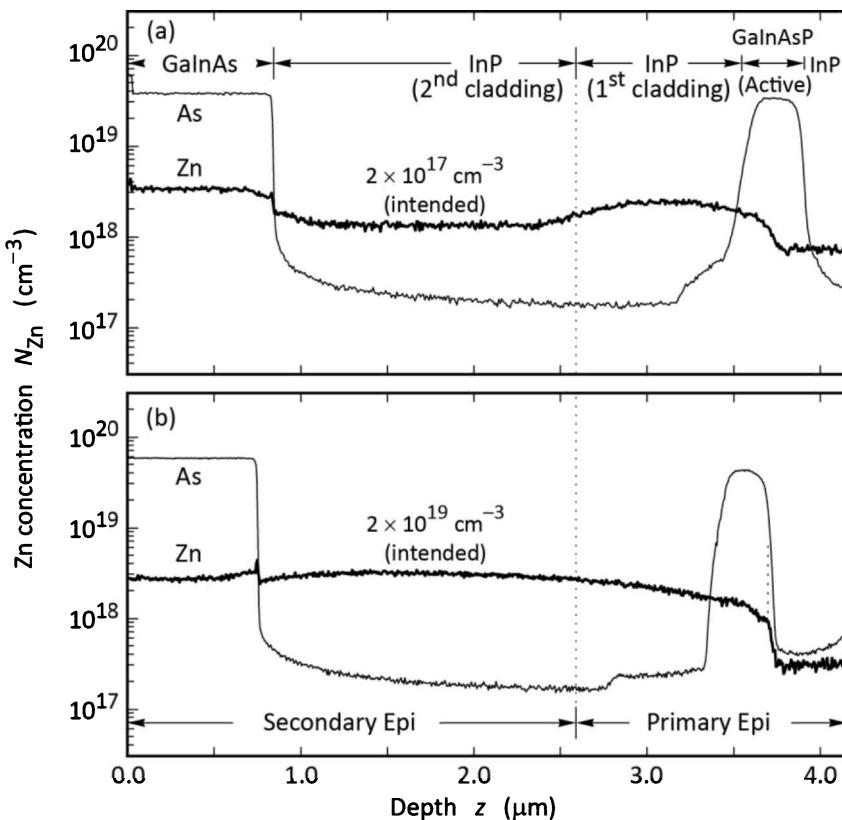


Fig. 14.6: Secondary ion mass spectrometry (SIMS) profile of Zn in a GaInAsP/InP double heterostructure. The structure uses Zn as a p-type dopant. Part (a) shows no p-n junction displacement. Part (b) shows p-n junction displacement caused by high Zn doping of the upper cladding region (after Schubert *et al.*, 1995).

A model explaining the p-n junction displacement in the GaInAsP/InP DH structure is illustrated in **Figure 14.7** (Schubert *et al.*, 1995). In this model, it is assumed that the Zn diffusion coefficient increases rapidly above a *critical* concentration N_{critical} . If this concentration is exceeded during growth, Zn will redistribute until the concentration falls below the critical concentration. As a result, Zn can diffuse into and through the active region of the double heterostructure. It is

remarkable that p-n junction displacement can occur, even if the intended Zn concentration in the *confinement region in the vicinity of the active region* is quite low.

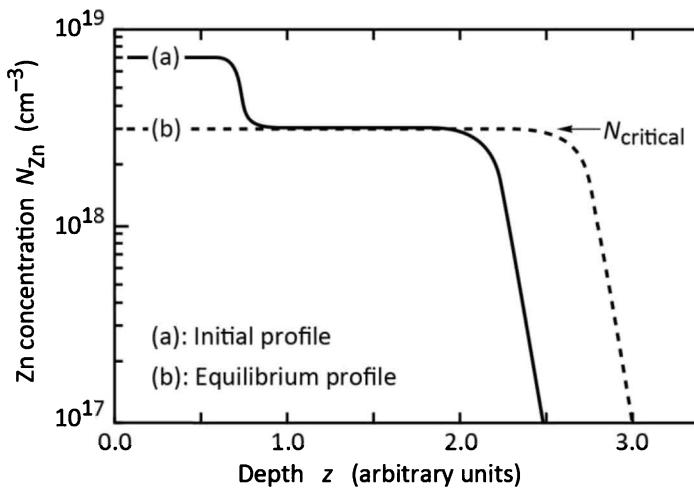


Fig. 14.7: P-n junction displacement process caused by excessive doping of the cladding region. If the acceptor dopant has a highly concentration-dependent diffusion constant and the diffusion constant increases strongly above a critical concentration, N_{critical} , p-n junction displacement occurs in the active region (after Schubert *et al.*, 1995).

14.4 Doping of the confinement regions

Doping of the confinement regions has a strong influence on the efficiency of double heterostructure LEDs. The resistivity of the confinement regions is one factor in determining the doping concentration in the confinement layers. The resistivity should be low to avoid resistive heating of the confinement regions.

Another factor is the residual doping concentration in the active region. The active region has a residual doping concentration, even if not intentionally doped. Typical doping concentrations in the active region are in the $10^{15}\sim 10^{16} \text{ cm}^{-3}$ range. The doping concentration in the confinement layers must be higher than the doping concentration of the active region to define the location of the p-n junction.

The influence of the confinement-layer doping concentration on the internal quantum efficiency was analyzed by Sugawara *et al.* (1992). The results for n-type and p-type doping are shown in *Figures* 14.8 and 14.9, respectively. The figures reveal that there is an optimum doping range for the confinement regions. For the n-type confinement regions, the optimum doping concentration ranges from 1×10^{16} to $2 \times 10^{17} \text{ cm}^{-3}$, as inferred from *Figures* 14.8.

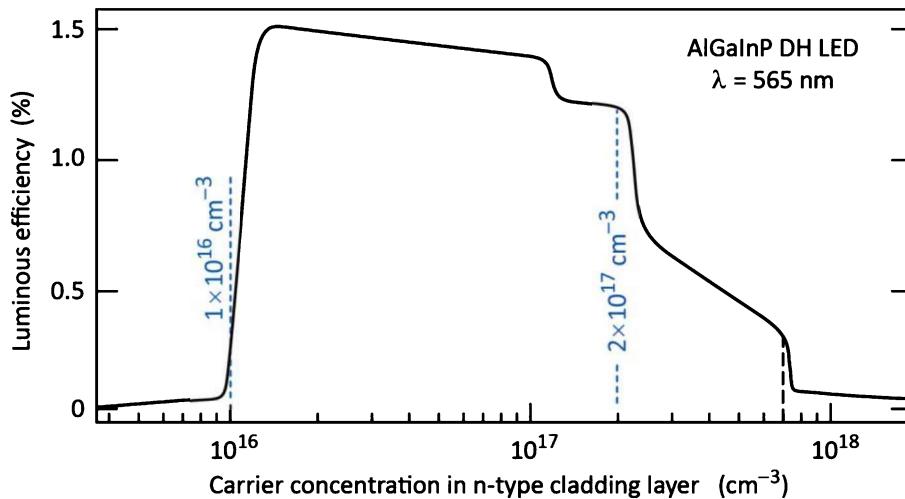


Fig. 14.8: Dependence of the luminous efficiency of an AlGaN_P double heterostructure (DH) LED emitting at 565 nm on n-type confinement layer doping concentration (after Sugawara *et al.*, 1992).

For p-type confinement regions, the optimum doping concentration ranges from 5×10^{17} to $2 \times 10^{18} \text{ cm}^{-3}$ (clearly higher than in the n-type cladding region), as inferred from *Figures 14.9*. The reason for this marked difference could again lie in the larger diffusion length of electrons than that of holes. A high p-type concentration in the cladding region keeps electrons in the active region and prevents them from diffusing deep into the confinement region.

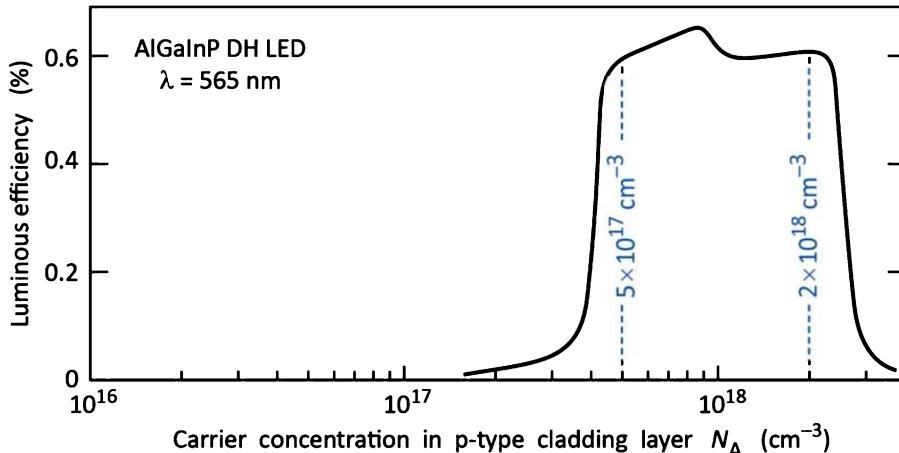


Fig. 14.9: Dependence of the luminous efficiency of an AlGaN_P double heterostructure (DH) LED emitting at 565 nm on the p-type confinement layer doping concentration (after Sugawara *et al.*, 1992).

The carrier leakage out of the active region into the p-type cladding layer in double heterostructure lasers was investigated by Kazarinov and Pinto (1994). It was shown that the electron leakage out of the active region is more severe than the hole leakage. The difference is due to the generally higher diffusion constant of electrons compared with holes.

The band diagram of a double heterostructure under forward bias conditions is shown in *Figure 14.10*. The figure illustrates that barriers are formed by the depletion layers at the two confinement–active-region interfaces. Compositional grading of these interfaces can be used to reduce these barriers.

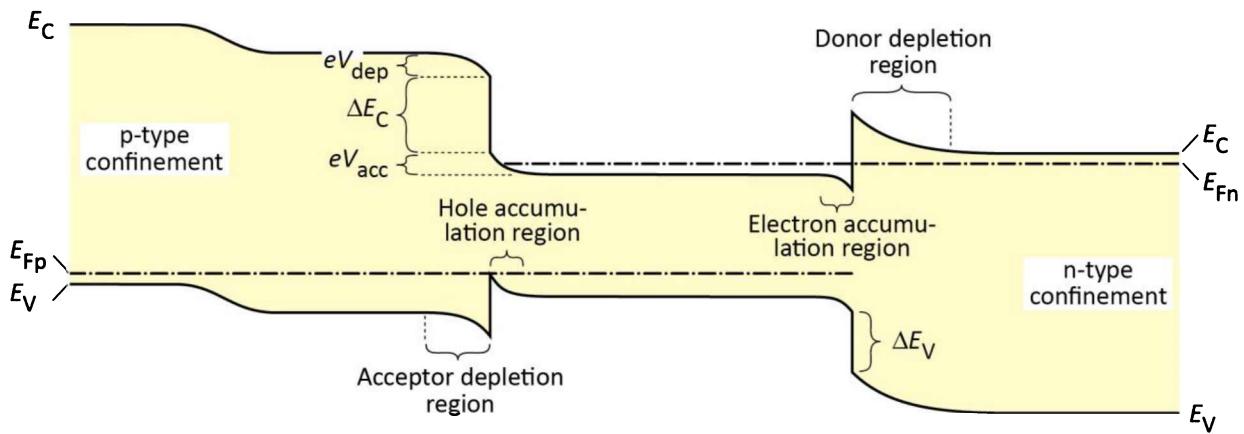


Fig. 14.10: Band diagram of a forward-biased double heterostructure. The p-type confinement layer consists of a lightly doped layer close to the active region and a higher doped layer further away from the active layer (adapted from Kazarinov and Pinto, 1994).

The influence of the cladding layer doping concentration on the radiative efficiency of a double heterostructure laser at threshold is shown in *Figure 14.11* (Kazarinov and Pinto, 1994). Inspection of the figure reveals that the confinement layer doping has a severe influence on the luminous efficiency. Low doping concentrations in the p-type confinement layers facilitate electron escape from the active region, thereby lowering the internal quantum efficiency.

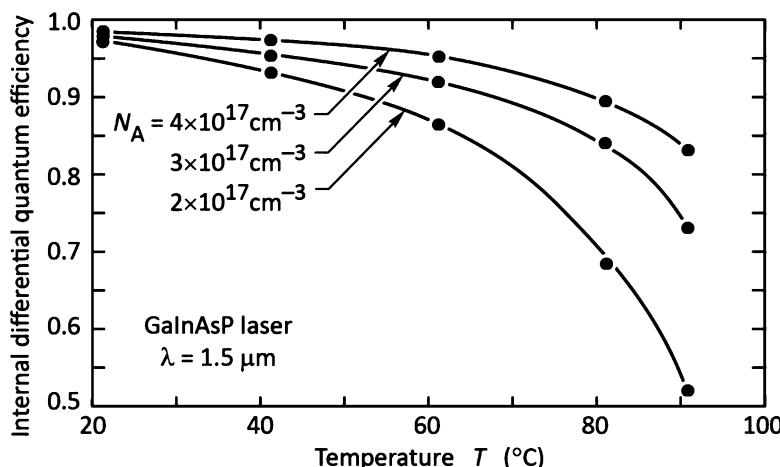


Fig. 14.11: Dependence of the internal differential quantum efficiency (emitted photons per injected electron) on temperature for different p-type doping levels in the cladding layer (after Kazarinov and Pinto, 1994).

14.5 Non-radiative recombination

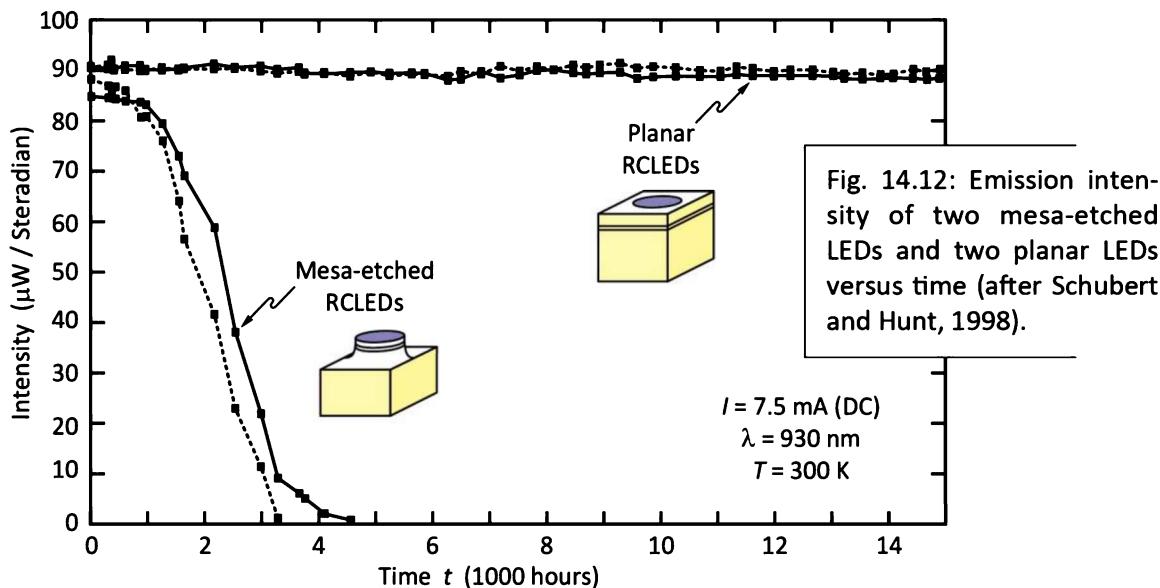
It is imperative that the material comprising the active region of light-emitting devices is of very high crystal quality. **Deep levels** caused by point defects, unwanted impurities, dislocations, and other defects must have a very low concentration.

Similarly, **surface recombination** must be kept at the lowest possible levels. This can be accomplished by keeping free surfaces several diffusion lengths away from regions in which both

electrons and holes reside, i.e. any surfaces must be “out of reach” of the active region.

Mesa-etched LEDs and lasers, in which the mesa etch exposes the active region to air, generally have low internal efficiencies due to surface recombination. Surface recombination also leads to a reduction of the lifetime of LEDs. Surface recombination generates heat at the semiconductor surface, which can lead to structural defects such as dark-line defects that further reduce the efficiency of LEDs.

Figure 14.12 shows the light intensity of two mesa-etched and two planar LEDs versus time. Inspection of **Figure 14.12** reveals that (i) the mesa-etched LED’s light intensity at $t = 0$ is slightly lower than that of the planar structure and (ii) the mesa-etched LED’s lifetime is much shorter. In the planar device, electron–hole recombination occurs below the top metal contact far away from the sidewall surfaces of the device. Thus, no intensity reduction or degradation mediated by surface recombination is expected for the planar device.



Note that the presence of surfaces does *not* reduce the radiative efficiency if only *one type of carrier* is present, e.g. near the top contact of the device. Surfaces in such unipolar regions do not have any deleterious effects.

14.6 Lattice matching

In double heterostructures, the active region material is different from the confinement layer material. However, both materials should have the same crystal structure and lattice constant. If the two semiconductors do not have the same lattice constant, then defects will occur at or near

the interface between the two semiconductors. **Figure 14.13** illustrates that **dangling bonds** can result as a consequence of mismatched semiconductors.

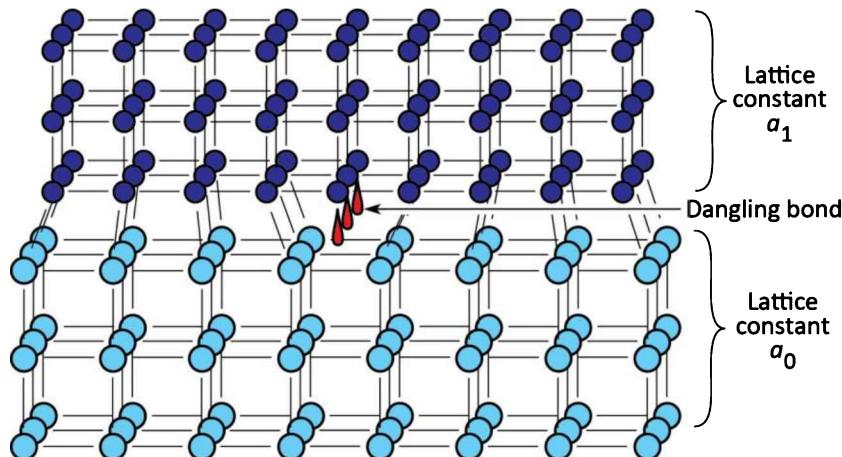


Fig. 14.13: Illustration of two crystals with mismatched lattice constant resulting in dislocations at or near the interface between the two semiconductors.

Inspection of **Figure 14.13** reveals that a *row* of dangling bonds can occur at the interface of two mismatched materials. In III–V arsenides (e.g. GaInAs / GaAs) and phosphides (e.g. GaP / GaAsP), such **misfit dislocation lines** are straight lines parallel to the substrate surface. In III–V nitrides (e.g. GaInN / GaN), such misfit dislocation lines are approximately straight lines perpendicular to the substrate surface. Misfit dislocation lines, which are extended defects, can be made visible by cathodo-luminescence. Usually a **cross-hatched** pattern is observed when performing cathodo-luminescence measurements on mismatched structures.

The effect of misfit dislocation lines on radiative recombination can be assessed by cathodo-luminescence measurements, as shown in **Figure 14.14** for a GaInAs / GaAs structure. The cathodo-luminescence micrograph exhibits a cross-hatch pattern of dark lines. The lines appear darker than the surrounding areas since carriers recombine non-radiatively at these dislocation lines.



Fig. 14.14: Cathodo-luminescence image of a 0.35 μm thick $\text{Ga}_{0.95}\text{In}_{0.05}\text{As}$ layer grown on a GaAs substrate. The dark lines forming a cross-hatch pattern are due to misfit dislocations (after Fitzgerald *et al.*, 1989).

Misfit dislocations may not occur directly *at* the interface between the mismatched materials but may start to form *near* the interface. This is because the mismatched crystal grown on top of the semiconductor will initially be *elastically* strained, and thus will assume the same in-plane lattice constant as the underlying substrate. This situation is shown in **Figure 14.15**, where a thin layer is strained so that it has the same in-plane lattice constant as the underlying material.

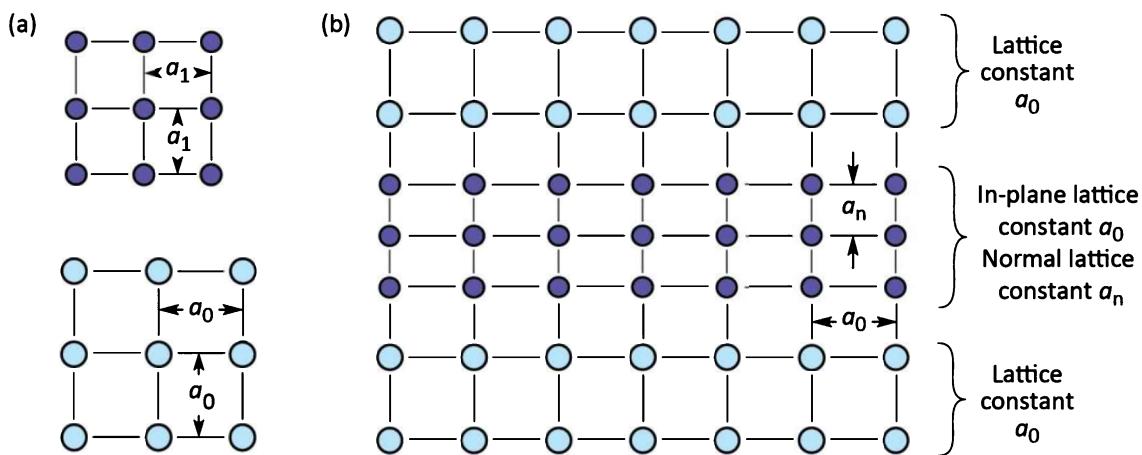


Fig. 14.15: (a) Cubic-symmetry crystals with equilibrium lattice constant a_1 and a_0 . (b) Thin, coherently strained layer with equilibrium lattice constant a_1 sandwiched between two semiconductors with equilibrium lattice constant a_0 . The coherently strained layer assumes an in-plane lattice constant a_0 and a normal lattice constant a_n .

However, once the energy needed to strain the lattice exceeds the energy required to form misfit dislocations, the thin film relaxes to its equilibrium lattice constant by forming misfit dislocations. The layer thickness at which misfit dislocations are formed is called the ***critical thickness***. It has been calculated by Matthews and Blakeslee (1976). If the layer is thinner than the critical thickness given by the Matthews–Blakeslee law, a thin dislocation-free layer can be grown even if the layers have different lattice constants.

The density of misfit dislocation lines per unit length is proportional to the lattice mismatch. Consequently, the efficiency of LEDs is expected to drop as the lattice mismatch is increased. **Figure 14.16** shows the decrease of the optical intensity of an AlGaInP LED grown on a GaAs substrate. The AlGaInP material used for high-brightness red LEDs is matched to the GaAs substrate. **Figure 14.16** reveals that the light output drops strongly as the mismatch, defined as $\Delta a/a$, exceeds 3×10^{-3} .

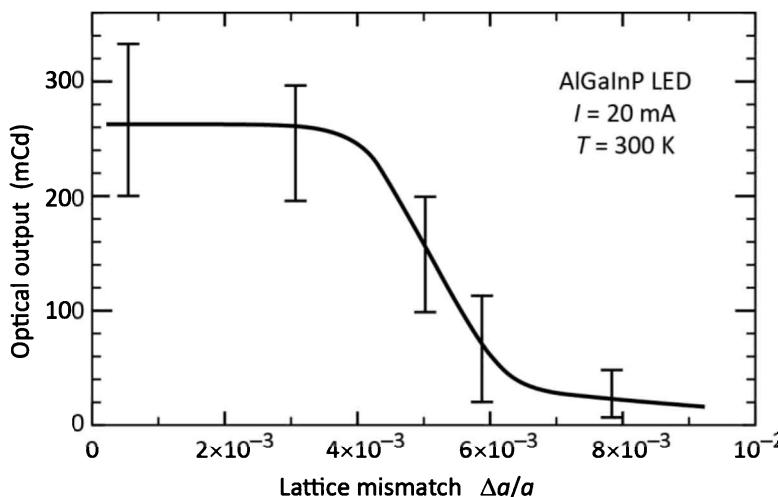


Fig. 14.16: Optical output intensity of an AlGaNp LED driven with an injection current of 20 mA versus the lattice mismatch between the AlInGaP active region and the GaAs substrate (after Watanabe and Usui, 1987).

Red GaAsP LEDs are grown on GaAs substrates and the active layers are mismatched to the substrate. As a result, red GaAsP LEDs are low cost but also low efficiency. The homojunction GaAsP LEDs grown on GaAs substrates are the lowest-cost red devices available.

Whereas the III–V arsenide and III–V phosphide material family is strongly affected by surface recombination and lattice mismatch, the III–V nitride material family is less so. One of the reasons for the insensitivity of the GaN material family to dislocation defects is the lower electrical activity that dislocations in these materials appear to exhibit. Another reason is the smaller diffusion length of carriers in the GaN material family. If the mean distance between dislocations is larger than the diffusion length, in particular the hole diffusion length, non-radiative recombination at these dislocations will not be severe. Another model explaining the high efficiency of GaInN is based on compositional fluctuations of the ternary alloy, which localize carriers preventing them from diffusing towards dislocation lines.

14.7 References

- Baraff G. A. and Schluter M. "Electronic structure, total energies, and abundances of the elementary point defects in GaAs" *Physical Review Letters* **55**, 1327 (1985)
- Fitzgerald E. A., Watson G. P., Proano R. E., Ast D. G., Kirchner P. D., Pettit G. D., and Woodall J. M. "Nucleation mechanism and the elimination of misfit dislocations at mismatched interfaces by reduction of growth area" *Journal of Applied Physics* **65**, 2220 (1989)
- Kazarinov R. F. and Pinto M. R. "Carrier transport in laser heterostructures" *IEEE Journal of Quantum Electronics* **30**, 49 (1994)
- Longini R. L. and Greene R. F. "Ionization interaction between impurities in semiconductors and insulators" *Physical Review* **102**, 992 (1956)
- Matthews J. W. and Blakeslee A. E. "Defects in epitaxial multilayers: III. Preparation of almost perfect multilayers" *Journal of Crystal Growth* **32**, 265 (1976)
- Nakamura S., Mukai T., and Iwasa N, "Light-emitting GaN-based compound semiconductor device" US

- Patent 5,578,839 (1996)
Nakamura S., Mukai T., and Iwasa N, "Light-emitting GaN-based compound semiconductor device" US Patent 5,747,832 (1998)
Neugebauer J. and Van de Walle C. G. "Chemical trends for acceptor impurities in GaN" *Journal of Applied Physics* **85**, 3003 (1999)
Schubert E. F., Downey S. W., Pinzone C., and Emerson A. B. "Evidence of very strong inter-epitaxial-layer diffusion in Zn doped GaInPAs/InP structures" *Applied Physics A* **60**, 525 (1995)
Schubert E. F. and Hunt N. E. J. "15,000 hours stable operation of resonant-cavity light-emitting diodes" *Applied Physics A* **66**, 319 (1998)
Sugawara H., Ishikawa M., Kokubun Y., Nishikawa Y., Naritsuka S., Itaya K., Hatakoshi G., Suzuki M., "Semiconductor light emitting device" US Patent 5,153,889, issued Oct. 6 (1992)
Walukiewicz W. "Fermi level dependent native defect formation: consequences for metal–semiconductor and semiconductor–semiconductor interfaces" *Journal of Vacuum Science Technology B*, **6**, 1257 (1988)
Walukiewicz W. "Amphoteric native defects in semiconductors" *Applied Physics Letters* **54**, 2094 (1989)
Walukiewicz W. "Defect formation and diffusion in heavily doped semiconductors" *Physical Review B* **50**, 5221 (1994)
Watanabe H. and Usui A. "Light emitting diode" US Patent 4,680,602, issued July 14 (1987)

15

15 – Designs for high light-extraction efficiency

Owing to the high refractive index of semiconductors, light incident on a planar semiconductor–air interface is totally internally reflected, if the angle of incidence is sufficiently large. Snell’s law gives the critical angle of total internal reflection. As a result of total internal reflection, light can be “trapped” inside the semiconductor. Light trapped in the semiconductor will eventually be reabsorbed, e.g. by the substrate, active region, cladding layer, or by a metallic contact.

If the light is absorbed by the substrate, the electron–hole pair will most likely recombine non-radiatively due to the inherently low efficiency of substrates. If the light is absorbed by the active region, the electron–hole pair may re-emit a photon or recombine non-radiatively. For active regions with internal quantum efficiencies of less than 100%, a reabsorption event by the active region reduces the efficiency of the LED.

The *external quantum efficiency (EQE)* of an LED is the product of the *internal quantum efficiency (IQE)* and the *light-extraction efficiency (LEE)* i.e.

$$EQE = IQE \times LEE. \quad (15.1)$$

The light-extraction efficiency thus plays an important role in increasing the power-conversion efficiency of LEDs.

15.1 Absorption of below-bandgap light in semiconductors

To obtain high light-extraction efficiency and avoid absorption of light, all semiconductor layers other than the active region should have a bandgap energy larger than the photon energy. This can be done in different ways, for example by using double heterostructures, window layers, and other structures that will be discussed below. In this section, we discuss the absorption of light if the energy of the light is *below* the energy gap of the semiconductor.

Naively, one would assume that a semiconductor can absorb light only if the photon energy is higher than the bandgap energy and that the semiconductor is transparent for photon energies below the bandgap. However, semiconductors do absorb below-bandgap light, although with a

much lower absorption coefficient.

The absorption coefficient versus energy of an idealized semiconductor and a real semiconductor is shown schematically in **Figure 15.1**. In the idealized semiconductor at low temperatures, the absorption coefficient versus energy is given by (Pankove, 1971)

$$\alpha \propto (E - E_g)^{1/2} \quad (\text{direct gap}) \quad (15.2\text{a})$$

$$\alpha \propto (E - E_g)^2 \quad (\text{indirect gap}). \quad (15.2\text{b})$$

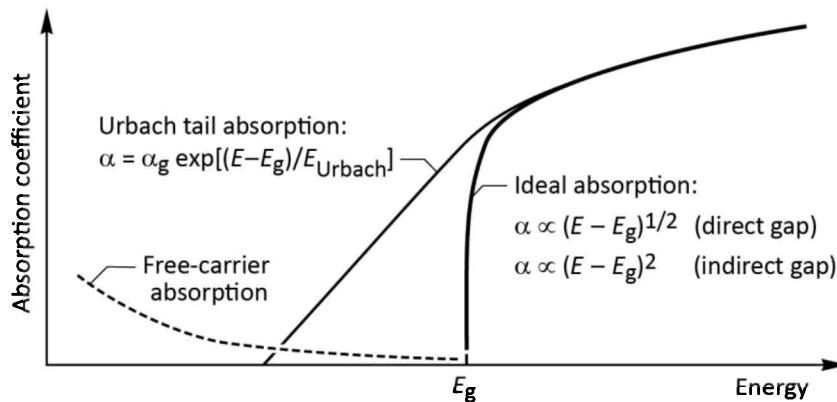


Fig. 15.1: Absorption coefficient of a semiconductor with bandgap E_g versus energy. The “Urbach tail” dominates absorption near yet below the bandgap. Absorption further below the bandgap is dominated by free-carrier absorption.

The idealized semiconductor has a zero band-to-band absorption coefficient at the bandgap energy ($E = E_g$). The absorption strength in a *real* semiconductor, for below-bandgap light, can be expressed in terms of an exponentially decaying absorption strength. In this absorption tail, called the **Urbach tail**, the absorption coefficient versus energy is given by

$$\alpha = \alpha_g \exp \left[(E - E_g) / E_{\text{Urbach}} \right] \quad (15.3)$$

where α_g is the experimentally determined absorption coefficient at the bandgap energy and E_{Urbach} is the characteristic energy (here called the **Urbach energy**), which determines how rapidly the absorption coefficient decreases for below-bandgap energies.

Urbach (1953) measured the absorption tail for different temperatures and showed that the **Urbach energy** is approximately kT , the thermal energy. The temperature dependence of the Urbach tail led Knox (1963) to the conclusion that the below-bandgap transitions are phonon-assisted transitions. Thus, the Urbach energy is given by

$$E_{\text{Urbach}} = kT. \quad (15.4)$$

The Urbach tail can be caused by mechanisms other than phonon-assisted absorption. Any mechanism introducing a potential fluctuation will lead to local variations of the semiconductor band edges. As a result of these fluctuations, the bandgap energy fluctuates as well, and below-bandgap transitions can occur.

The most common potential fluctuations are fluctuations caused by random dopant distribution and local variations of the chemical composition of a ternary or quaternary alloy semiconductor.

Potential fluctuations caused by random dopant distribution can be calculated using Poisson statistics. The magnitude of these fluctuations is given by (see, for example, Schubert *et al.*, 1997)

$$\Delta E_{Urbach} = \frac{2e^2}{3\epsilon} \sqrt{(N_D^+ + N_A^-) \frac{r_s}{3\pi}} e^{-3/4} \quad (15.5)$$

where r_s is the screening radius.

Potential fluctuations caused by random compositional fluctuations can be calculated using binomial statistics. The magnitude of these fluctuations is given by (see, for example, Schubert *et al.*, 1984)

$$\Delta E_{alloy} = \frac{dE_g}{dx} \sqrt{\frac{x(1-x)}{4a_0^{-3}V_{exc}}} \quad (15.6)$$

where x is the alloy composition of the ternary semiconductor alloy, a_0 is the semiconductor lattice constant, and V_{exc} is the excitonic volume of the electron–hole pair.

It depends on the specific case at hand, which of the different physical effects dominates in the formation of the Urbach tail. Generally, binary semiconductors such as GaP or GaAs have a smaller tail than alloys such as AlGaAs or GaAsP. Furthermore, lightly doped semiconductors have a smaller Urbach tail than heavily doped semiconductors.

For energies sufficiently lower than the bandgap energy, absorption due to the Urbach tail is negligibly small and ***free-carrier absorption*** becomes the dominant absorption mechanism. As the name suggests, a free carrier is excited to a higher energy by absorption of a photon. The absorption transition must conserve momentum. Whereas photons have very small momenta, electrons must undergo a momentum change when excited higher within a parabolic band. This momentum change is provided by acoustic phonons, optical phonons, or by impurity scattering.

Free-carrier absorption is proportional to the free-carrier concentration, since a free carrier

is required for an absorption event. Theoretical considerations in terms of the classical Drude free-electron model further show that free-carrier absorption increases as the square of the incident wavelength (Pankove, 1971). Thus, the proportionalities

$$\alpha_{fc} \propto n\lambda^2 \quad \text{and} \quad \alpha_{fc} \propto p\lambda^2 \quad (15.7)$$

are valid for n-type and p-type semiconductors, respectively. Theoretical considerations in terms of a quantum mechanical treatment show that the absorption coefficient is proportional to $\lambda^{3/2}$, $\lambda^{5/2}$, and $\lambda^{7/2}$ depending upon whether acoustic phonon scattering, optical phonon scattering, or ionized impurity scattering is involved in the momentum conservation process, respectively (Swaminathan and Macrander, 1991).

In n-type and p-type GaAs, the room-temperature free-carrier absorption coefficient near the bandgap energy ($\lambda \approx 950$ nm) can be expressed as (Casey and Panish, 1978)

$$\alpha_{fc} = 3 \text{ cm}^{-1} \frac{n}{10^{18} \text{ cm}^{-3}} + 7 \text{ cm}^{-1} \frac{p}{10^{18} \text{ cm}^{-3}}. \quad (15.8)$$

Inspection of the equation indicates that the free-carrier absorption coefficient can be of the order of 10 cm^{-1} at high carrier concentrations. Approximate values for the free-carrier absorption coefficient in several compound semiconductors are given in **Table 15.1**.

Table 15.1: Free-carrier absorption coefficient (α_{fc}) of n-type semiconductors:
 (a) after Ioffe (2002); (b) after Wiley and DiDomenico (1970); (c) after Casey and Panish (1978); (d) after Kim and Bonner (1983) and Walukiewicz *et al.* (1980);
 (e) data are extrapolated using the proportionality $\alpha_{fc} \propto n\lambda^2$.

Material	Wavelength	Electron concentration	α_{fc}
GaN	1.0 μm	$1 \times 10^{18} \text{ cm}^{-3}$	40 cm^{-1} (a, e)
GaP	1.0 μm	$1 \times 10^{18} \text{ cm}^{-3}$	22 cm^{-1} (b, e)
GaAs	1.0 μm	$1 \times 10^{18} \text{ cm}^{-3}$	3.0 cm^{-1} (c)
InP	1.0 μm	$1.1 \times 10^{18} \text{ cm}^{-3}$	2.5 cm^{-1} (d)

In LEDs, free-carrier absorption can affect the intensity of waveguided modes radiating out of the side of the chip. Free-carrier absorption also plays a role in LEDs with transparent semiconductor substrates. Such transparent substrates have a typical thickness greater than 100 μm . If the doping concentration of a transparent substrate is high, free-carrier absorption will reduce the

light-output power. If it is low, the substrate becomes resistive. Thus, a compromise needs to be made between the different doping requirements of a transparent substrate. For thin layers, such as confinement layers, free-carrier absorption effects are negligibly small if the optical path length within the layer is short.

15.2 Double heterostructures

Virtually all LED structures employ double heterostructures. They consist of two *confinement* layers and the *active* region. The band diagram of a double heterostructure is shown in **Figure 15.2**. The active region has a smaller bandgap energy than the two confinement regions. As a result, the confinement regions are *transparent* to the light emitted by the active region. Since the confinement regions are relatively thin, they can be considered, for all practical purposes, to be totally transparent.

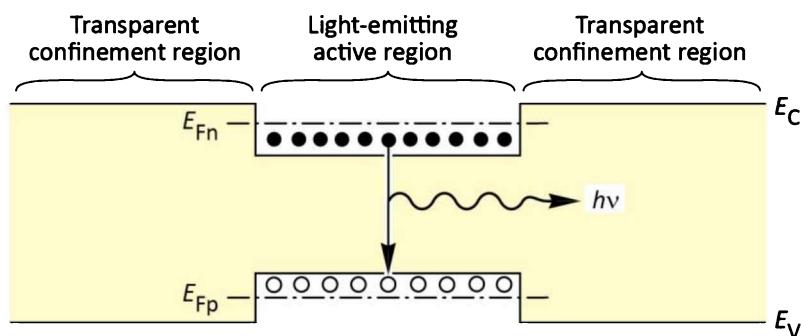


Fig. 15.2: Double heterostructure with optically transparent confinement regions. Reabsorption in the active region is unlikely due to the high carrier concentration in the active region and the resulting Burstein–Moss shift of the absorption edge.

Reabsorption of light by the active region in the current-injected area below the top contact can also be neglected. The active region is, under normal injection conditions, injected with high current densities so that the electron and hole quasi-Fermi levels rise into the bands, as illustrated in **Figure 15.2**. As a result, the active region is practically transparent for near-bandgap emission under high injection conditions.

Note, however, that the active region is in equilibrium sufficiently far away from the current-injected region. These regions are not current injected and thus will absorb near-bandgap light emitted by the active region. To reduce optical losses by absorption, the active region should have a high internal quantum efficiency to make re-emission of absorbed photons likely.

15.3 Shaping of LED dies

One of the most important problems facing high-efficiency LEDs is the occurrence of **trapped light** within a high-index semiconductor. The occurrence of trapped light is illustrated in **Figure 15.3** (Wikipedia, 2013). A light ray emitted by the active region will be subject to total internal reflection, as predicted by Snell's law. In the high-index approximation, the angle of total internal reflection is given by

$$\alpha_{crit} = \bar{n}_s^{-1} \quad (15.9)$$

where \bar{n}_s is the semiconductor refractive index and the critical angle α_{crit} is given in radians. For high-index semiconductors, the critical angle is quite small. For example, for a refractive index of 3.3, the critical angle for total internal reflection is only 17° . Thus, most of the light emitted by the active region is trapped inside the semiconductor. The trapped light is most likely to be absorbed by the thick substrate. Once absorbed, the electron–hole pair is likely to recombine non-radiatively due to the comparatively low quality and efficiency of the substrate.

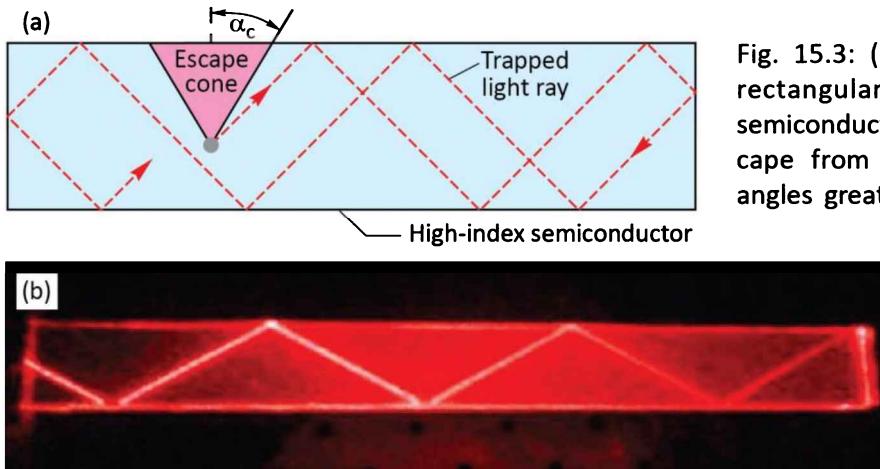


Fig. 15.3: (a) “Trapped light” in a rectangular-parallelepiped-shaped semiconductor. Light is unable to escape from the semiconductor for angles greater than α_c due to total internal reflection (TIR). (b) Red light subject to TIR propagating in a slab of material (after Wikipedia, 2013).

The light-escape problem has been known since the infancy of LED technology in the 1960s. It has also been known that the geometrical shape of the LED die plays a critical role. The optimum LED would be spherical in shape with a point-like light-emitting region in the center of the LED. Such a spherical LED is shown in **Figure 15.4 (a)**. Light emanating from the point-like active region is incident at a normal angle at the semiconductor–air interface. As a result, total internal reflection does not occur in such LEDs. Note, however, that the light is still subject to Fresnel reflection at the interface unless the sphere is coated with an anti-reflection coating.

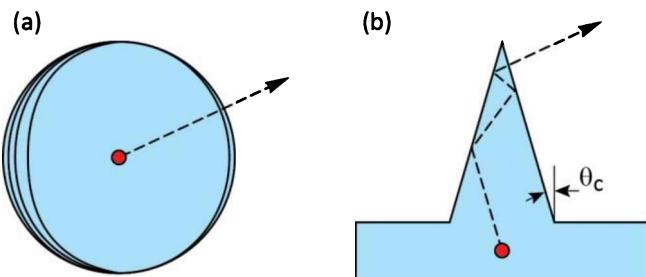


Fig. 15.4: Schematic illustration of different geometric shapes for LEDs with perfect extraction efficiency. (a) Spherical LED with a point-like light-emitting region at the center of the sphere. (b) A cone-shaped LED.

LEDs with a hemispherical dome-like structure (Carr and Pittman, 1963) as well as other shapes, e.g. inverted, truncated cones (Franklin and Newman, 1964; Loebner, 1973) have been demonstrated to improve extraction efficiency over conventional designs, i.e. rectangular parallelepiped chips. However, the practical utility of such devices has not been realized primarily due to the high cost associated with shaping of individual LED dies.

Unfortunately, spherical LEDs with a point-like light source in the center of the LED are somewhat impractical devices. Semiconductor fabrication technology is, in view of the flat substrates used in epitaxial growth, a *planar* technology. Thus, spherical LEDs are difficult to fabricate using conventional planar technology.

Another interesting LED structure is a cone-shaped structure, as shown schematically in **Figure 15.4 (b)**. Light rays are emanating from the active region at or below the base of the cone. The light rays incident at the cone–air boundary are either transmitted through the semiconductor–air interface or guided by the cone. The guided rays undergo multiple reflections. As light rays undergo multiple reflections, they will form a progressively increasing angle of incidence at the semiconductor–air interface. As a result, light guided by the cone will eventually have near-normal incidence and escape from the cone. Although an interesting concept, cone-shaped LEDs are difficult to fabricate and manufacture.

A very common LED structure has the shape of a rectangular parallelepiped as shown in **Figure 15.5 (a)**. Such LED dies are fabricated by cleaving the wafer, e.g. a GaAs wafer, along its natural cleavage planes. The LEDs have a total of six escape cones, two of them perpendicular to the wafer surface, and four of them parallel to the wafer surface. The bottom escape cone will be absorbed by the substrate if the substrate has a lower bandgap than the active region. The four in-plane escape cones will be at least partially absorbed by the substrate. Light in the top escape cone will be obstructed by the top contact, unless a thick current-spreading layer is employed. Thus, a rectangular parallelepiped LED clearly is a structure with low light-extraction efficiency. However, a substantial advantage of such LEDs is the low manufacturing cost.

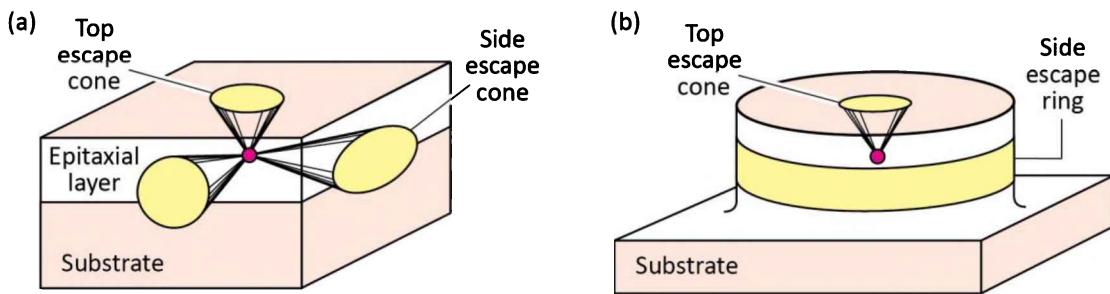


Fig. 15.5: (a) Rectangular parallelepiped LED chip having a total of six escape cones. (b) Cylindrical LED chip having a top escape cone and a side escape ring.

An LED with a cylindrical shape is shown in **Figure 15.5 (b)**. The cylindrical LED has the advantage of higher extraction efficiency compared with a cube-shaped LED. An escape ring, as shown in **Figure 15.5 (b)**, replaces the four in-plane escape cones of the rectangular LED, which results in a substantial improvement of the extraction efficiency. Cylindrical-shaped LEDs require one more processing step (etching step) compared with rectangularly shaped LEDs.

Commercially available die-shaped devices include **pedestal-shaped** GaInN/SiC LEDs with trade name “Aton” (Osram, 2001) and **truncated inverted pyramid** (TIP) AlGaInP/GaP LEDs (Krames *et al.*, 1999). Photographs and schematic structures of the two devices are shown in **Figure 15.6**. Ray traces indicated in the figure show that light rays entering the base of the pyramids escape from the semiconductor after undergoing one or multiple internal reflections. The pedestal and TIP geometries reduce the mean photon path length within the die, and thus reduce internal absorption losses. The increase in efficiency of pedestal-shaped devices is about a factor of 2 compared to rectangular-parallelepiped-shaped devices (Osram, 2001).

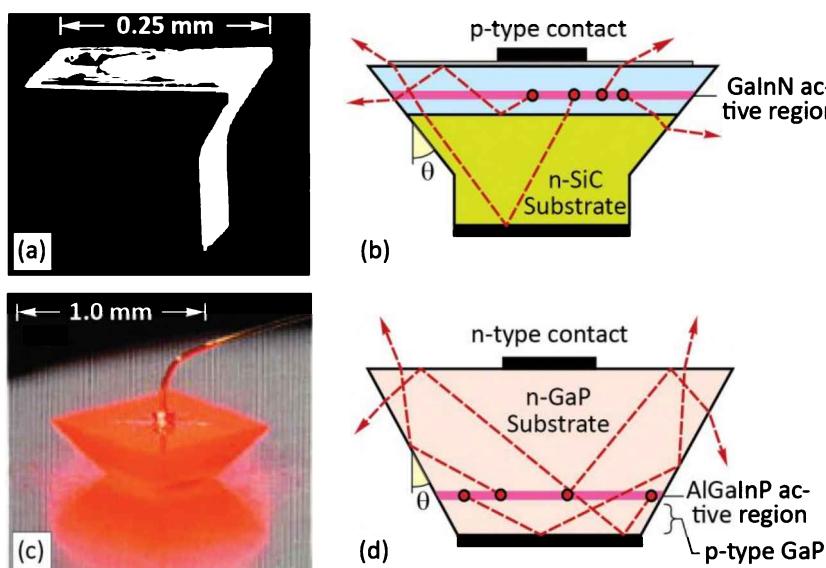


Fig. 15.6: Die-shaped devices: (a) Blue GaInN emitter on SiC substrate with trade name “Aton”. (b) Schematic ray traces illustrating enhanced light extraction. (c) Micrograph of truncated inverted pyramid (TIP) AlGaInP/GaP LED. (d) Schematic diagram illustrating enhanced extraction (after Osram, 2001; Krames *et al.*, 1999).

The geometric shape of the LED, in particular the side-wall angle, is chosen in such a way that trapping of light is minimized. Ray tracing computer models are employed to maximize the escape probability from the semiconductor. The optimum angle of the TIP sidewall tilt is 35° (Krames *et al.*, 1999). The TIP LED is a high-power LED with a large p-n junction area of 500 μm × 500 μm. The luminous source efficiency of TIP LEDs exceeds 100 lm/W and is one of the highest ever achieved with LEDs.

The TIP LED performance versus injection current is shown in **Figure 15.7** (Krames *et al.*, 1999). A peak luminous efficiency of 102 lm/W was measured for orange-spectrum ($\lambda \approx 610$ nm) devices at an injection current of 100 mA. This luminous efficiency exceeds that of most fluorescent (50~100 lm/W) and all metal-halide (68~95 lm/W) lamps. In the amber color regime, the TIP LED provides a photometric efficiency of 68 lm/W ($\lambda \approx 598$ nm). This efficiency is comparable to the source efficiency of 50 W high-pressure sodium discharge lamps. A peak external quantum efficiency of 55% was measured for red-emitting ($\lambda \approx 650$ nm) TIP LEDs. Under pulsed operation (1% duty cycle), an efficiency of 60.9% was achieved (data not shown), which sets a lower bound on the extraction efficiency of these devices.

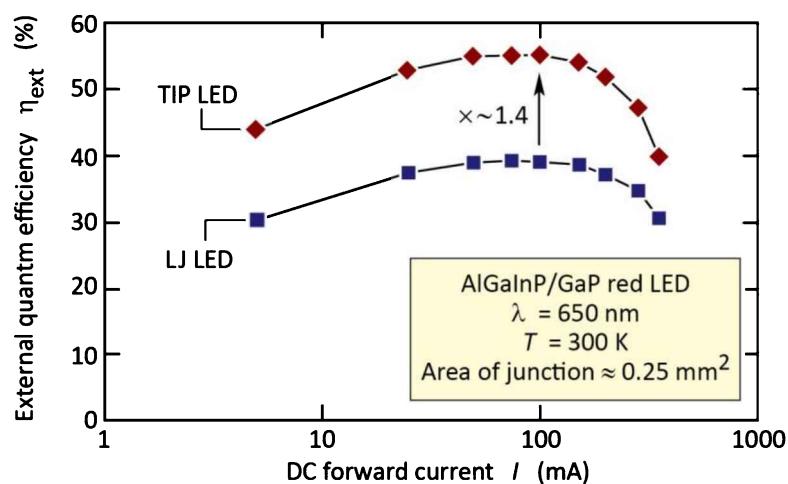


Fig. 15.7: External efficiency versus forward current for red-emitting (650 nm) truncated inverted pyramid (TIP) LEDs and large junction (LJ) LEDs mounted in power-lamp packages. The TIP LED exhibits a 1.4 times improvement in extraction efficiency compared with the LJ device, resulting in a peak external quantum efficiency of 55 % at 100 mA (after Krames *et al.*, 1999).

The rectangular parallelepiped and cylindrical LED structures can be fabricated with wafer-scale processing steps. Advanced structures, such as the TIP LED, are also fabricated by wafer-scale processes, namely a sawing process employing a beveled dicing blade (Haitz, 1992). The manufacturing cost of LEDs requiring die-level processing steps is much higher compared with LEDs manufactured with wafer-scale processes.

The use of tapered output couplers in LEDs has been reported by Schmid *et al.* (2000, 2001, 2002). For GaAs-based infrared devices, external quantum efficiencies near 50% have been demonstrated with devices having a tapered output coupler. However, due to the area requirements of the output coupler, the current-injected region of the device is much smaller than the overall device area and thus the total emission power is quite small.

15.4 Textured semiconductor surfaces

A popular method to increase the light-extraction efficiency is the use of ***roughened*** or ***textured semiconductor surfaces***, first introduced by Bergh and Saul (1973) for GaP-based devices. The authors showed that to a first approximation, no more light will escape from a rough surface plane at the first incidence than would escape from a smooth surface plane. However, since the reflections are at random angles, a similar proportion of light will escape the second and subsequent times the light ray is incident on the roughened surface. This is in contrast to a smooth surface plane which preserves reflection angles and thus tends to trap the reflected light.

Figure 15.8 shows the enhanced out-coupling of a light ray through a roughened surface.

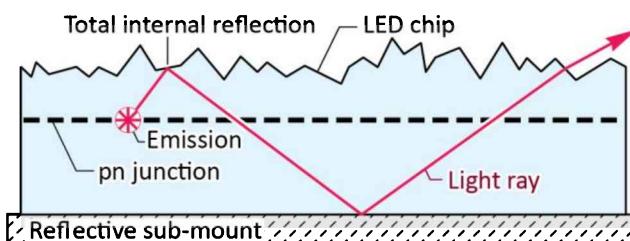


Fig. 15.8: LED chip with roughened surface enabling out-coupling of light (after Bergh and Saul, 1973).

The characteristic length-scale of the surface texture can be either much larger than the wavelength of light (Bergh and Saul, 1973) or similar to the wavelength of light (see, for example, Schnitzer *et al.*, 1993; Windisch *et al.*, 1999, 2000, 2001, 2002). External quantum efficiencies near 50% have been reported for GaAs-based surface-textured devices. A detailed discussion of properties and fabrication of microstructured surfaces was given by Sinzinger and Jahns (1999).

For many years, it had been assumed that GaN cannot be etched by any wet chemical etch. However, Stocker *et al.* (1998a) showed that, although the c-plane of GaN (Ga-face) lacks etchability, non-c-plane surfaces (such as the a-plane, m-plane, and -c-planes) *can* be etched by wet chemical etching solutions including KOH and H₃PO₄. Stocker *et al.* (1998a) also showed that

these etches are crystallographic in nature, which allows one to form smooth crystal surfaces that are suited for laser facets (Stocker *et al.* 1998b, 2000). Textured surfaces created by crystallographic etching are very well suited to enhance light extraction. Several publications reported an increase in light extraction in GaN-based LEDs with surface texturing achieved by crystallographic wet chemical etching (Gao *et al.*, 2004; Fujii *et al.*, 2004; Jung *et al.*, 2010; Horng *et al.*, 2011). The textured GaN surfaces shown in **Figure 15.9** reveal the exposure of a variety of crystal planes that are obtained as a result of crystallographic etching (Haerle, 2004a; Jung *et al.*, 2010).

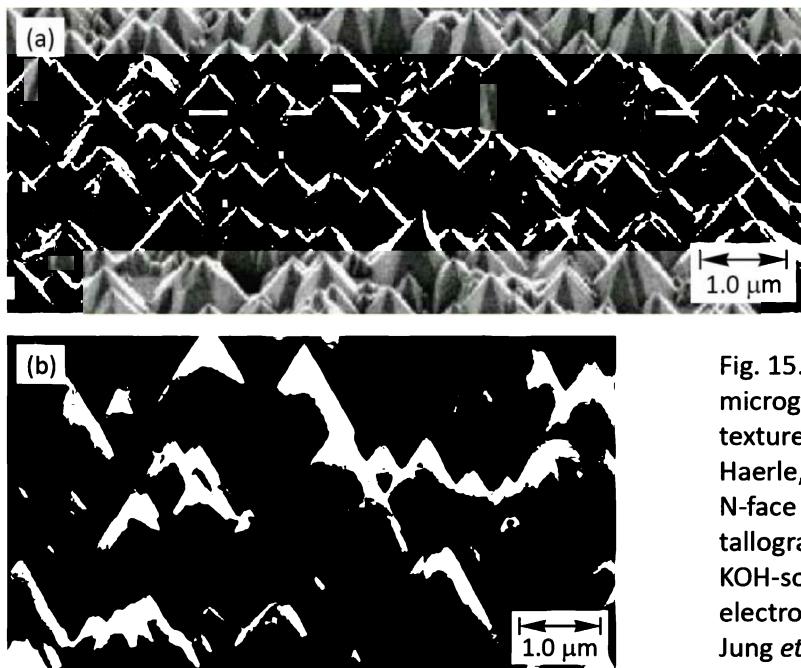


Fig. 15.9: (a) Scanning electron micrograph (SEM) of a strongly textured GaN surface (after Haerle, 2004). (b) SEM of N-face GaN surface after crystallographic etching using a KOH-solution-based photo-electrochemical process (after Jung *et al.*, 2010).

The transition from a smooth surface to a strongly textured surface and the implication on ray propagation is shown in **Figure 15.10**. A perfectly smooth surface, i.e. a **specular** surface, results in waveguided modes that cannot escape, as shown in **Figure 15.10 (a)**. For a strongly **scattering** surface, i.e. a diffuse, lambertian surface, only one or a few scattering events are sufficient to out-couple the light as shown in **Figure 15.10 (c)**. The intermediate case is shown in **Figure 15.10 (b)**. Because the light-out-coupling efficiency depends on the degree of diffusivity, it is desirable to quantify the degree of surface diffusivity (for a discussion of the quantification of the diffuseness of a surface, see the chapter entitled “Reflectors” in this book).

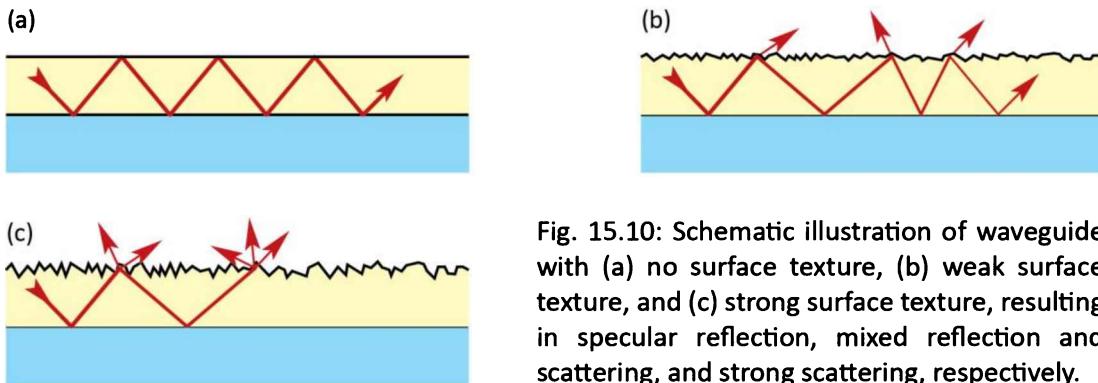


Fig. 15.10: Schematic illustration of waveguide with (a) no surface texture, (b) weak surface texture, and (c) strong surface texture, resulting in specular reflection, mixed reflection and scattering, and strong scattering, respectively.

15.5 GaN thin-film LEDs

GaN **thin-film LEDs** are a class of LEDs allowing for very high light-extraction efficiencies. Due to their unique role, they are referred to by several names including ***thin-film flip chip (TFFC) LEDs***, ***ThinGaN LEDs***, or ***vertical LEDs***. In GaN thin-film LEDs, the epitaxial film is delaminated from its original substrate. By the delamination, a greater control can be exerted on the surfaces of the epitaxial layer system. As a result, thin-film LEDs offer the highest performance among the different types of GaN LEDs. The schematic fabrication process for a GaN thin-film LED is shown in **Figure 15.11**.

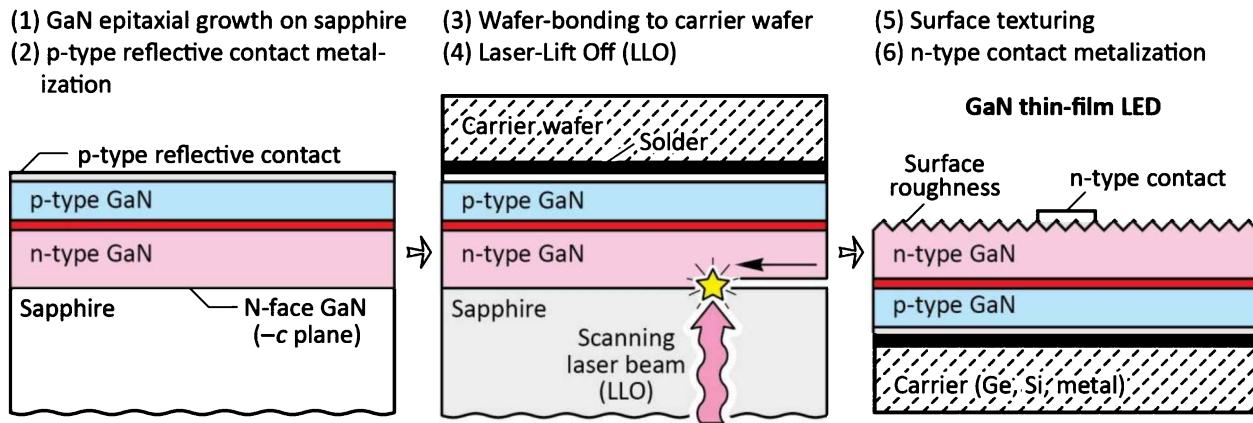


Fig. 15.11: Schematic fabrication process of GaN-based thin-film LEDs. Bonding the epitaxial wafer to a carrier wafer, removing the sapphire substrate, and mounting the device n-side up (flip-chip), allows for roughening of the N-face GaN surface ($-c$ plane).

The typical fabrication process for GaN thin-film LEDs includes the following steps:

- **Epitaxial growth** of the GaN LED on a substrate, commonly sapphire, using the layer sequence: n-type cladding layer, active region, and p-type cladding layer.

- The **p-type reflective contact** has two functionalities: First it serves as a broad-area ohmic contact. Second, it serves as a reflector and comprises a highly reflective metal such as Ag or Al. Two-layer metallizations are common such as NiZn / Ag where very thin NiZn (e.g. 2.5 nm) serves as ohmic contact and thick Ag (e.g. 200 nm) serves as optical reflector (Song *et al.*, 2003).
- In a **wafer-bonding process**, the p-type reflective contact is bonded to a carrier wafer by means of a solder. A Pd-In alloy solder was reported in the literature (Wong *et al.*, 2000). The carrier wafer can be selected from a wide variety of materials including Ge, Si, or a metal.
- **Substrate delamination** can be accomplished through a laser-lift-off (LLO) process. In this process, short-wavelength high-power pulsed laser light (248 nm) is absorbed by GaN and induces rapid thermal decomposition of an interfacial GaN layer. A scanning motion of the laser allows one to delaminate the epitaxial layer of an entire wafer (Wong *et al.*, 1998, 1999, 2000; Kelly *et al.*, 1999).
- **Surface texturing** is accomplished by crystallographic wet chemical etching (either with or without photo-enhancement). Crystallographic wet chemical etches can be used on the N-face GaN surface ($-c$ plane) but not on the Ga-face GaN surface (c plane) (Stocker *et al.*, 1998a). The microscopic optical inspection of textured chip surfaces of thin-film GaInN LEDs reveals that the chip surface appears *white*, which is an indication of a strongly textured or diffuse surface (Cree, 2011).
- The **n-type contact** is deposited on the textured surface. The Ti / Al / Ti / Au and Ti / Al / Ni / Au are common n-type metallization schemes (Luther *et al.*, 1997).

There are several variations of the thin-film GaN LED process. In one variation, both contacts, n-type and well as p-type, are on the lower side of the device. This configuration has the advantage that there are no metal electrodes on the top surface of the device that could hinder the extraction of light. Indeed light extraction for thin-film GaN-based LEDs can be very high. Light-extraction efficiencies of 75% were reported in 2004 (Haerle *et al.*, 2004b) and values of 80~95% may be possible with thin-film LEDs having both ohmic contacts on the lower side of the chip.

The increase in light extraction from a surface-textured GaInN device is shown in **Figure 15.12** (Haerle, 2004a). The author estimated the increase in output power to be 40~50%. Similar results were reported independently by Gao *et al.* (2004) and Fujii *et al.* (2004). Note that the interference fringes, observed for the GaN device having the smooth surface, completely vanish for the surface-textured film. The interference fringes are caused by the Fabry–Perot cavity formed by the GaN–air interface and the sapphire–GaN interface which form the two reflectors

of the cavity (Billeb *et al.*, 1997).

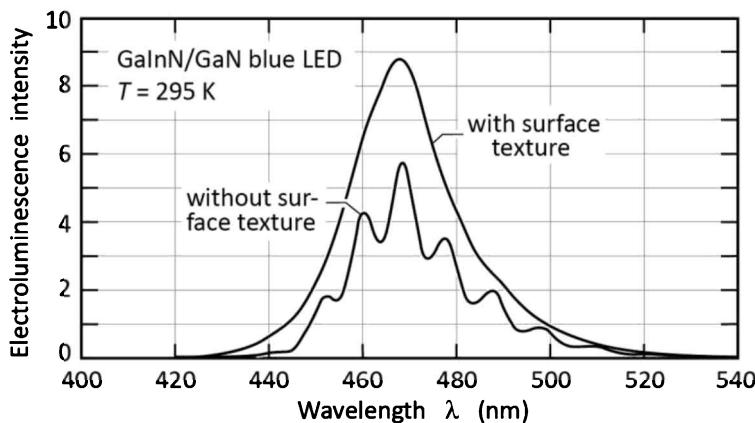


Fig. 15.12: Emission spectrum of GaInN blue LED with and without surface texture. The spectrum exhibits Fabry–Perot interference fringes for the device with a smooth surface (after Haerle, 2004).

For any LEDs with two top contacts, such as GaInN/GaN LEDs grown on sapphire substrates, both regular packaging (epi-side up) and ***flip-chip*** packaging (epi-side down) is used. Flip-chip packaging can be accomplished using ***solder-bump bonding*** without the wire bonds that are used in conventional packaging for LEDs with one or two top contact pads. An advantage of flip-chip packaging of GaInN/GaN LEDs is that the metal pads are not hindering the extraction of light radiating from the active region.

In packaged epi-side up GaInN LEDs, a large-area p-type top contact that covers the entire p-type mesa results in a uniform current distribution in the active region. However, such a large top contact can hinder the extraction of light. This fundamental trade-off can be avoided by flip-chip packaging, which is frequently used for high-power devices.

15.6 Cross-shaped contacts and other contact geometries

Different requirements need to be satisfied by the top contact. In regular LEDs, the top contact provides a pad for the bonding wire. The pad is usually circular with a typical diameter of 100 µm. The top contact pad also provides a low-resistance ohmic contact to the current-spreading layer.

Typical top contact geometries are shown in **Figure 15.13**. The simplest geometry is just a circular contact pad, as shown in **Figure 15.13 (a)**. A cross-shaped contact, as shown in **Figure 15.13 (b)**, provides a more uniform current distribution over the entire area of the active region. Note, however, that no or little current should flow to the *edges* of the LED die to avoid surface recombination.

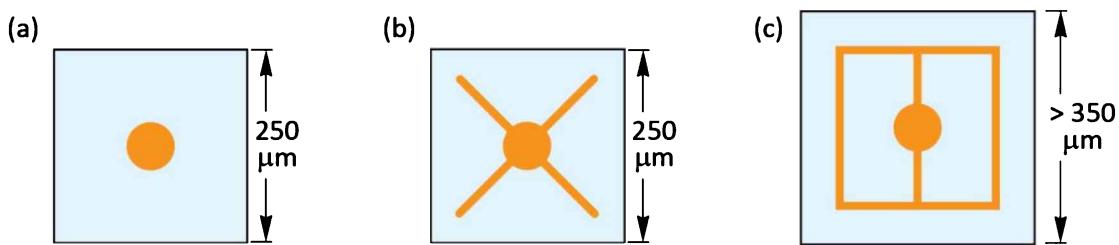


Fig. 15.13: Top view of an LED die with (a) a circular contact also serving as a bond pad and (b) a cross-shaped contact with a circular bond pad. (c) Typical contact geometry used for larger LED dies.

For large-area LEDs, a simple circular pad or a cross-shaped top contact are insufficient for uniform current distribution. In such large-area devices, patterns that include a ring, as shown in **Figure 15.13 (c)**, are better suited to providing uniform current distribution.

The area of the top contact is kept small so that the light emanating from the active region is not hindered by the opaque contact. However, the contact resistance scales with the contact area so that the top contact area cannot be scaled down arbitrarily.

15.7 Transparent substrate technology

Visible-spectrum $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$ LEDs with typical operating wavelengths of 560~660 nm, are grown lattice matched on GaAs substrates. Since the energy gap of GaAs is $E_g = 1.424$ eV ($\lambda_g = 870$ nm) at room temperature, GaAs substrates are absorbing at these emission wavelengths. As a result, the light emitted towards the substrate will be absorbed by the thick GaAs substrate. Thus, the extraction efficiency of AlGaNp/GaAs LEDs grown on GaAs substrates is low.

The extraction efficiency of AlGaNp LEDs can be increased substantially by removal of the GaAs substrate and bonding of the epitaxial layer to a GaP substrate (Kish *et al.*, 1994). GaP is an indirect-gap semiconductor with $E_g = 2.24$ eV ($\lambda_g = 553$ nm). Thus, GaP does not absorb light with $\lambda > 553$ nm emitted by the AlGaNp active region.

The fabrication process of the AlGaNp LEDs wafer-bonded to a GaP substrate is illustrated schematically in **Figure 15.14**. The AlGaNp double heterostructure (DH) is initially grown by OMVPE on a GaAs substrate. Subsequently, a thick GaP window layer (~50 μm) is grown on the DH by chloride VPE, a growth technique that allows for the low-cost growth of thick epitaxial layers. The GaAs substrate is then removed using a wet chemical selective etching process (Adachi and Oe, 1983; Kish *et al.*, 1994). During the removal of the GaAs substrate, the thick GaP window layer serves as a temporary mechanical support for the DH. The DH with the GaP window is then wafer-bonded to a GaP substrate.

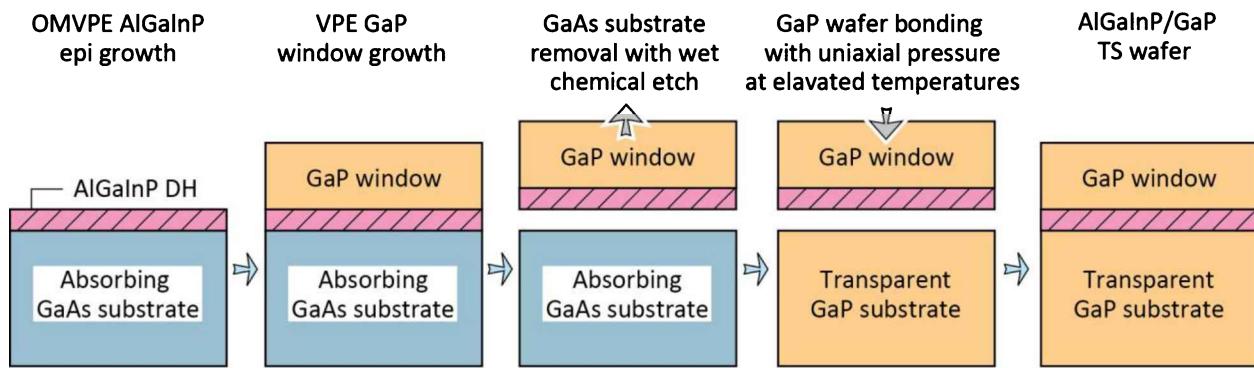


Fig. 15.14: Fabrication process for wafer-bonded transparent substrate (TS) AlGaNp/GaP LEDs. After selective removal of the GaAs growth substrate, elevated temperature and uniaxial pressure are applied to form a TS LED wafer (after Kish *et al.*, 1994).

Wafer-bonding processes require a high degree of cleanliness, the lack of any particles between the wafers, and the absence of native surface oxides on both wafers. Frequently, the gap between two wafers is filled with a contact liquid. Rotating the wafers at a high rate spins out the contact liquid. Kish *et al.* (1995) and Hoefler *et al.* (1996) reported an AlGaNp-to-GaP wafer-bonding process, suitable for 50 mm (2 inch) GaP substrates. Uniaxial pressure and elevated temperatures (750~1 000°C) are used in this process (Hoefler *et al.*, 1996). Kish *et al.* (1995) showed that the achievement of low-resistance ohmic conduction across wafer-bonded interfaces is critically dependent upon the crystallographic alignment of the bonded wafer surfaces, irrespective of the lattice mismatch between the surfaces. Furthermore, Kish *et al.* (1995) showed that the crystallographic surface orientation of the bonded surfaces must be nominally matched while simultaneously maintaining rotational alignment of the wafers. Low diode forward voltages of 2.2 V for AlGaNp/GaP LEDs are routinely achieved under high-volume manufacturing conditions with the process. The reliability of wafer-bonded LEDs is comparable to monolithic AlGaNp/GaAs LEDs. Usually the technical details of wafer-bonding processes are proprietary and not known to the general public.

The forward voltage is a critical figure of merit for wafer-bonded p-n junction devices. A low voltage indicates true semiconductor-to-semiconductor chemical bonding and an absence of interfacial oxide layers. The forward current–voltage characteristic of commercial absorbing-substrate (AS) and transparent-substrate (TS) AlGaNp LEDs is shown in *Figure 15.15*. Inspection of the figure reveals that TS LEDs have a higher forward voltage and series resistance compared with AS devices.

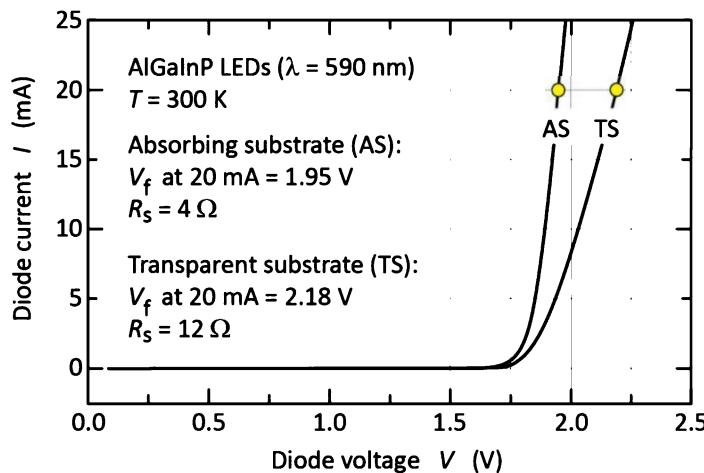
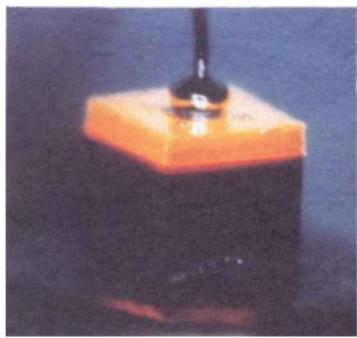


Fig. 15.15: Current–voltage characteristic, forward voltage, and series resistance of absorbing-substrate (GaAs) and transparent-substrate (GaP) LEDs with AlGaNp active regions.

The higher forward voltage in TS devices is probably due to either ohmic losses occurring at the wafer-bonded interface or in the GaP substrate. A detailed analysis by O’Shea *et al.* (2001) confirmed that a forward-voltage penalty results from the wafer-bonded interface, particularly if this interface is either contaminated with carbon or crystallographically misaligned. The n-type doping concentration in the GaP substrate is kept moderately low, to minimize free-carrier absorption.

Figure 15.16 compares AlGaNp/GaAs AS LEDs with AlGaNp/GaP TS LEDs emitting in the amber wavelength range (Kish and Fletcher, 1997). The substrate appears dark for the AS LED, in contrast to the TS LED. Transparent substrate AlGaNp/GaP LEDs have a factor of 1.5~3.0 higher external efficiency compared with AS AlGaNp/GaAs LEDs.

(a) Absorbing Substrate (AS) LED



(b) Transparent Substrate (TS) LED

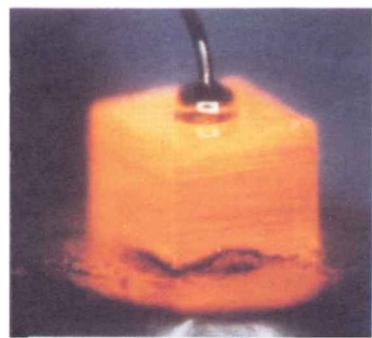


Fig. 15.16: (a) Amber AlGaNp LED with a GaP window layer and absorbing GaAs substrate (AS). (b) Amber AlGaNp LED with a GaP window layer and a transparent GaP substrate (TS) fabricated by wafer bonding. Conductive Ag-loaded die-attach epoxy can be seen at bottom (after Kish and Fletcher, 1997).

15.8 Anti-reflection optical coatings

Anti-reflection (AR) coatings are frequently used in communication LEDs to reduce the Fresnel reflection at the semiconductor–air interface. For normal incidence, the intensity reflection coefficient for normal incidence is given by

$$R = \frac{(\bar{n}_{sc} - \bar{n}_{air})^2}{(\bar{n}_{sc} + \bar{n}_{air})^2} \quad (15.10)$$

where \bar{n}_{sc} and \bar{n}_{air} are the refractive indices of the semiconductor and of air, respectively.

For normal incidence, Fresnel reflection at the semiconductor-air interface can be reduced to zero, if the AR coating has the following optimum parameters:

$$\begin{aligned} \text{Thickness} &= \lambda/4 = \lambda_0 / (4 \bar{n}_{AR}) \\ \text{Refractive index} &= \bar{n}_{AR} = \sqrt{\bar{n}_{sc} \bar{n}_{air}} . \end{aligned} \quad (15.11)$$

An AR coating with optimum thickness and refractive index is shown in **Figure 15.17**. The refractive indices and the transparency ranges of different dielectric materials suitable for AR coatings are given in **Table 15.2**.

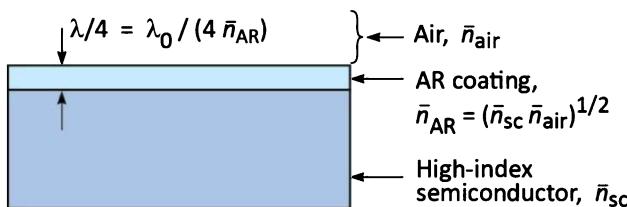


Fig. 15.17: Illustration of optimum thickness and refractive index of an anti-reflection (AR) coating.

Table 15.2: Refractive index and transparency range of common dielectrics suitable as anti-reflection coatings (after Palik, 1998).

Dielectric material	Refractive index	Transparency range
SiO ₂ (silica)	1.45	> 0.15 μm
Al ₂ O ₃ (alumina)	1.76	> 0.15 μm
TiO ₂ (titania)	2.50	> 0.35 μm
Si ₃ N ₄ (silicon nitride)	2.00	> 0.25 μm
ZnS (zinc sulfide)	2.29	> 0.34 μm
CaF ₂ (calcium fluoride)	1.43	> 0.12 μm

15.9 References

- Adachi S. and Oe K. "Chemical etching characteristics of (001) GaAs" *Journal of the Electrochemical Society* **130**, 2427 (1983)
- Bergh A. A. and Saul R. H. "Surface roughening of electroluminescent diodes" US Patent 3,739,217 issued on June 12 (1973)
- Billeb A., Grieshaber W., Stocker D., Schubert E. F., and Karlicek R. F. Jr. "Microcavity effects in GaN

- epitaxial layers" *Applied Physics Letters* **70**, 2790 (1997)
- Carr W. N. and Pittman G. E. "One-Watt GaAs p-n junction infrared source" *Applied Physics Letters* **3**, 173 (1963)
- Casey Jr. H. C. and Panish M. B. *Heterostructure Lasers, Part A: Fundamental Principles* pp. 46, 47, and 175 (Academic Press, San Diego, 1978)
- Cree Corporation. Visual inspection of the surface of a GaN LED chip manufactured by Cree Corporation reveals that the surface is white, indicating that the chip surface is diffusive (2011)
- Franklin A. R. and Newman R. "Shaped electroluminescent GaAs diodes" *Journal of Applied Physics* **35**, 1153 (1964)
- Fujii T., Gao Y., Sharma R., Hu E. L., DenBaars S. P., and Nakamura S. "Increase in the extraction efficiency of GaN-based light-emitting diodes via surface roughening" *Applied Physics Letters* **84**, 855 (2004)
- Gao Y., Fujii T., Sharma R., Fujito K., DenBaars S. P., Nakamura S., and Hu E. L. "Roughening hexagonal surface morphology on laser lift-off (LLO) N-face GaN with simple photo-enhanced chemical wet etching" *Japanese Journal of Applied Physics* **43**, L637 (2004)
- Haerle V. "Naturally textured GaN surface" *China Hi-Tech Fair (CHTF) Shenzhen, China, October 12~17 (2004a)*
- Haerle V., Hahn B., Kaiser S., Weimar A., Bader S., Eberhard F., Plössl A., and Eisert D. "High brightness LEDs for general lighting applications using the new ThinGaN-Technology" *Physica Status Solidi (a)* **201**, 2736 (2004)
- Haitz R. "Light-emitting diode with diagonal faces" US Patent 5,087,949 issued on February 11 (1992)
- Hoefer G. E., Vanderwater D. A., DeFevere D. C., Kish F. A., Camras M. D., Steranka F. M., and Tan I.-H. "Wafer bonding of 50-mm diameter GaP to AlGaN_xP light-emitting diode wafers" *Applied Physics Letters* **69**, 803 (1996)
- Horng R.-H., Lu Y.-A., and Wuu D.-S. "Light extraction investigation for thin-film GaN light-emitting diodes with imbedded electrodes" *IEEE Photonics Technology Letters* **23**, 54 (2011)
- Ioffe Institute (Saint Petersburg, Russia) database on compound semiconductors available at www.ioffe.rssi.ru/SVA/NSM/Semicond/ (2002)
- Jung Y., Baik K. H., Ren F., Pearton S. J., and Kima J. "Effects of photo-electrochemical etching of N-polar and Ga-polar gallium nitride on sapphire substrates" *Journal of the Electrochemical Society* **157**, H676 (2010)
- Kelly M. K., Vaudo R. P., Phanse V. M. Görgens L., Ambacher O., and Stutzmann M. "Large free-standing GaN substrates by hydride vapor phase epitaxy and laser-induces liftoff" *Japanese Journal of Applied Physics* **38**, L217 (1999)
- Kim O. K. and Bonner W. A. "Infrared reflectance and absorption of n-type InP" *Journal of Electronic Materials* **12**, 827 (1983)
- Kish F. A. and Fletcher R. M. "AlGaN_xP light-emitting diodes" in *High Brightness Light-Emitting Diodes* edited by G. B. Stringfellow and M. G. Crawford, Semiconductors and Semimetals **48** (Academic, San Diego, 1997)
- Kish F. A., Steranka F. M., DeFevere D. C., Vanderwater D. A., Park K. G., Kuo C. P., Osentowski T. D., Peanasky M. J., Yu J. G., Fletcher R. M., Steigerwald D. A., Crawford M. G., and Robbins V. M. "Very high-efficiency semiconductor wafer-bonded transparent-substrate (Al_xGa_{1-x})_{0.5}In_{0.5}P/GaP light-emitting diodes" *Applied Physics Letters* **64**, 2839 (1994)
- Kish F. A., Vanderwater D. A., Peanasky M. J., Ludowise M. J., Hummel S. G., and Rosner S. J. "Low-resistance ohmic conduction across compound semiconductor wafer-bonded interfaces" *Applied Physics Letters* **67**, 2060 (1995)
- Knox R. S. *Theory of Excitons* (Academic Press, New York, 1963)
- Krames M. R., Ochiai-Holcomb M., Höfler G. E., Carter-Coman C., Chen E. I., Tan I.-H., Grillot P., Gardner N. F., Chui H. C., Huang J.-W., Stockman S. A., Kish F. A., Crawford M. G., Tan T. S., Kocot C. P., Hueschen M., Posselt J., Loh B., Sasser G., and Collins D. "High-power truncated-inverted-pyramid (Al_xGa_{1-x})_{0.5}In_{0.5}P/GaP light-emitting diodes exhibiting > 50% external quantum efficiency" *Applied Physics Letters* **75**, 2365 (1999)

- Loebner E. E. "The future of electroluminescent solids in display applications" *Proceedings of the IEEE* **61**, 837 (1973)
- Luther B. P., Mohney S. E., Jackson T. N., Khan M. A., Chen Q., and Yang J. W. "Investigation of the mechanism for ohmic contact formation in Al and Ti/Al contacts to n-type GaN" *Applied Physics Letters* **70**, 57 (1997)
- O'Shea J. J., Camras M. D., Wynne D., and Hoefler G. E. "Evidence for voltage drops at misaligned wafer-bonded interfaces of AlGaN/P light-emitting diodes by electrostatic force microscopy" *Journal of Applied Physics* **90**, 4791 (2001)
- Osram Opto Semiconductors Corporation, Regensburg, Germany "Osram Opto enhances brightness of blue InGaN-LEDs" Press Release (January 2001)
- Palik E. D. *Handbook of Optical Constants of Solids* (Academic Press, San Diego, 1998)
- Pankove J. I. *Optical Processes in Semiconductors*" p. 75 and section on Urbach tail (Dover, New York, 1971)
- Schmid W., Eberhard F., Jager R., King R., Joos J., and Ebeling K. "45% quantum-efficiency light-emitting diodes with radial outcoupling taper" *Proceedings of the SPIE* **3938**, 90 (2000)
- Schmid W., Scherer M., Jager R., Strauss P., Streubel K., and Ebeling K. "Efficient light-emitting diodes with radial outcoupling taper at 980 and 630 nm emission wavelength" *Proceedings of the SPIE* **4278**, 109 (2001)
- Schmid W., Scherer M., Karnutsch C., Plobl A., Wegleiter W., Schad S., Neubert B., and Streubel K. "High-efficiency red and infrared light-emitting diodes using radial outcoupling taper" *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 256 (2002)
- Schnitzer I., Yablonovitch E., Caneau C., Gmitter T. J., and Scherer A. "30% external quantum efficiency from surface-textured, thin-film light-emitting diodes" *Applied Physics Letters* **63**, 2174 (1993)
- Schubert E. F., Goebel E. O., Horikoshi Y., Ploog K., and Queisser H. J. "Alloy broadening in photoluminescence spectra of AlGaAs" *Physical Review B* **30**, 813 (1984)
- Schubert E. F., Goepfert I. D., Grieshaber W., and Redwing J. M. "Optical properties of Si-doped GaN" *Applied Physics Letters* **71**, 921 (1997)
- Sinzinger S. and Jahns J. *Microoptics* (Wiley-VCH, New York, 1999)
- Song J.-O., Leem D.-S., Kwak J. S., Nam O. H., Park Y., and Seong T.-Y. "Low-resistance and highly-reflective Zn-Ni solid solution / Ag ohmic contacts for flip-chip light-emitting diodes" *Applied Physics Letters* **83**, 4990 (2003)
- Stocker D. A., Schubert E. F., and Redwing J. M. "Crystallographic wet chemical etching of GaN" *Applied Physics Letters* **73**, 2654 (1998a)
- Stocker D. A., Schubert E. F., Grieshaber W., Boutros K. S., and Redwing J. M. "Facet roughness analysis for InGaN/GaN lasers with cleaved facets" *Applied Physics Letters* **73**, 1925 (1998b)
- Stocker D. A., Schubert E. F., and Redwing J. M. "Optically pumped InGaN/GaN lasers with wet-etched facets" *Applied Physics Letters* **77**, 4253 (2000)
- Swaminathan V. and Macrander A. T. *Materials Aspects of GaAs and InP Based Structures*" (Prentice Hall, Englewood Cliffs, 1991)
- Urbach F. "The long-wavelength edge of photographic sensitivity of the electronic absorption of solids" *Physical Review* **92**, 1324 (1953)
- Walukiewicz W., Lagowski J., Jastrzebski L., Rava P., Lichtensteiger M., Gatos C. H., and Gatos H. C. "Electron mobility and free-carrier absorption in InP; determination of the compensation ratio" *Journal of Applied Physics* **51**, 2659 (1980)
- Wikipedia, "Total internal reflection" <http://en.wikipedia.org/wiki/Total_internal_reflection> (accessed November 5, 2013)
- Wiley J. D. and DiDomenico Jr. M. "Free-carrier absorption in n-type GaP" *Phys. Rev. B* **1**, 1655 (1970)
- Windisch R., Schoberth S., Meinlschmidt S., Kiesel P., Knobloch A., Heremans P., Dutta B., Borghs G., and Doehler G. H. "Light propagation through textured surfaces" *Journal of Optics A: Pure and Applied Optics* **1**, 512 (1999)
- Windisch R., Dutta B., Kuijk M., Knobloch A., Meinlschmidt S., Schoberth S., Kiesel P., Borghs G., Doehler

- G. H., and Heremans P. "40% efficient thin-film surface textured light-emitting diodes by optimization of natural lithography" *IEEE Transactions on Electron Devices* **47**, 1492 (2000)
- Windisch R., Rooman C., Kuijk M., Borghs G., and Heremans P. "Impact of texture-enhanced transmission on high-efficiency surface-textured light-emitting diodes" *Applied Physics Letters* **79**, 2315 (2001)
- Windisch R., Rooman C., Dutta B., Knobloch A., Borghs G., Doepler G. H., and Heremans P. "Light-extraction mechanisms in high-efficiency surface-textured light-emitting diodes" *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 248 (2002)
- Wong W. S. and Sands T., and Cheung N. W. "Damage-free separation of GaN thin films from sapphire substrates" *Applied Physics Letters* **72**, 599 (1998)
- Wong W. S., Sands T., Cheung N. W., Kneissl M., Bour D. P., Mei P., Romano L. T., and Johnson N. M. "Fabrication of thin-film InGaN light-emitting diode membranes by laser lift-off" *Applied Physics Letters* **75**, 1360 (1999)
- Wong W. S., Sands T., Cheung N. W., Kneissl M., Bour D. P., Mei P., Romano L. T., and Johnson N. M. " $\text{In}_x\text{Ga}_{1-x}\text{N}$ light emitting diodes on Si substrates fabricated by Pd–In metal bonding and laser lift-off" *Applied Physics Letters* **77**, 2822 (2000)

16

16 – Spontaneous emission from resonant cavities

16.1 Modification of spontaneous emission

Radiative transitions, i.e. transitions of electrons from an initial quantum state to a final state and the simultaneous emission of a light quantum, are one of the most fundamental processes in optoelectronic devices. There are two distinct ways by which the emission of a photon can occur, namely by *spontaneous* and *stimulated* emission. These two processes were first postulated by Einstein (1917).

Stimulated emission is employed in semiconductor lasers and superluminescent LEDs. It was realized in the 1960s that the stimulated emission mode can be used in semiconductors to drastically change the radiative emission characteristics. The efforts to harness stimulated emission resulted in the first room-temperature operation of semiconductor lasers (Hayashi *et al.*, 1970) and the first demonstration of a superluminescent LED (Hall *et al.*, 1962).

Spontaneous emission implies the notion that the recombination process occurs *spontaneously*, that is without a means to influence this process. In fact, spontaneous emission has long been believed to be uncontrollable. However, research in microscopic optical resonators, where spatial dimensions are of the order of the wavelength of light, showed the possibility of controlling the spontaneous emission properties of a light-emitting medium. The changes of the emission properties include the spontaneous emission rate, spectral purity, and emission pattern. These changes can be employed to make more efficient, faster, and brighter semiconductor devices. The changes in spontaneous emission characteristics in resonant cavity (RC) and photonic crystal (PC) structures were reviewed by Joannopoulos (1995).

Microcavity structures have been demonstrated with different active media and different microcavity structures. The first microcavity structure was proposed by Purcell (1946) for emission frequencies in the radio frequency (RF) regime. Small metallic spheres were proposed as the resonator medium. However, no experimental reports followed Purcell's theoretical publication. In the 1980s and 1990s, several microcavity structures were realized with different

types of optically active media. The emission media included organic dyes (De Martini *et al.*, 1987; Suzuki *et al.*, 1991), semiconductors (Yablonovitch *et al.*, 1988; Yokoyama *et al.*, 1990), rare-earth-doped silica (Schubert *et al.*, 1992b; Hunt *et al.*, 1995b), and organic polymers (Nakayama *et al.*, 1993; Dodabalapur *et al.*, 1994). In these publications, clear changes in spontaneous emission were demonstrated including changes in spectral, spatial, and temporal emission characteristics.

At the beginning of the 1990s, current-injection ***resonant-cavity light-emitting diodes*** (RCLEDs) were first demonstrated in the GaAs material system (Schubert *et al.*, 1992a) and subsequently in organic light-emitting materials (Nakayama *et al.*, 1993). Both publications reported an emission line narrowing due to the resonant cavities. RCLEDs have many advantageous properties when compared with conventional LEDs, including higher brightness, increased spectral purity, and higher efficiency. For example, the spectral power density in RCLEDs was shown to be enhanced by more than one order of magnitude (Hunt *et al.*, 1992, 1995a).

The change in optical gain in VCSELs due to the enhancement in spontaneous emission was analyzed by Deppe and Lei (1992). The comparison of a *macrocavity*, in which the cavity is much *longer* than the emission wavelength ($\lambda \ll L_{cav}$), with a *microcavity* ($\lambda \approx L_{cav}$) revealed that the gain can be enhanced by factors of 2~4 for typical GaAs emission linewidths at room temperature (50 nm). Thus, laser threshold currents can be lower in microcavity structures due to the higher gain.

It is important to distinguish between emission *inside* the cavity and emission *out of* the cavity. The enhancement of the spontaneous emission *inside* the cavity and emission through one of the mirrors *out of* the cavity can be very different. At moderate values of the cavity finesse, the spontaneous emission inside and out of the cavity is enhanced. However, for very high finesse cavities (see, for example, Jewell *et al.*, 1988), the overall emission out of the cavity *decreases* (Schubert *et al.*, 1996). In the limit of very high reflectivity reflectors ($R_1 = R_2 \rightarrow 100\%$), the emission out of the cavity becomes zero. This effect will be discussed in detail below.

A device in which all the spontaneous emission occurs into a single optical mode has been proposed by Kobayashi *et al.* (1982, 1985). This device has been termed a ***zero-threshold laser*** (Yokoyama, 1992) and a ***single-mode LED*** (Yablonovitch, 1994). In a conventional laser, only a small portion of the spontaneous emission couples into a single state of the electromagnetic field

controlled by the laser cavity. The rest is lost to free-space modes that radiate out of the side of the laser. The idea of a thresholdless laser is simple. It assumes a wavelength-size cavity in which only one optical mode exists. Thus, spontaneous as well as stimulated emission couples to this optical mode. The thresholdless laser should lack a threshold, i.e. the clear distinction between the spontaneous and the lasing regime which is observed in the light-output versus current characteristic of conventional lasers. Clearly, the prospects of such a device are intriguing. Even though several attempts to demonstrate a thresholdless laser have been reported (Yokoyama *et al.*, 1990; Yokoyama 1992; Numai *et al.*, 1993), a thresholdless laser has not yet been demonstrated.

16.2 Fabry–Perot resonators

The simplest form of optical cavity consists of two coplanar mirrors separated by a distance L_{cav} . About one century ago, Fabry and Perot (1899) were the first to build and analyze optical cavities with coplanar reflectors. These cavities had a large separation between the two reflectors, i.e. $L_{cav} \gg \lambda$. However, if the distance between the two reflectors is of the order of the wavelength of light, new physical phenomena occur, including the enhancement of the optical emission from an active material inside the cavity. Very small cavities, with typical dimensions of $L_{cav} \approx \lambda$, will be denoted as *microcavities*.

Coplanar microcavities are the simplest form of optical microcavities and their properties are summarized below. For a detailed discussion of the optical properties of Fabry–Perot cavities, the reader is referred to the literature (Coldren and Corzine, 1995; Saleh and Teich, 1991). Fabry–Perot cavities with two reflectors of reflectivity R_1 and R_2 are shown in **Figures 16.1(a)** and **(b)**. Plane waves propagating inside the cavity can interfere constructively and destructively resulting in stable (allowed) optical modes and attenuated (disallowed) optical modes, respectively. (Note that the photon length is much longer than the microcavity length.) For lossless (non-absorbing) reflectors, the transmittance through the two reflectors is given by $T_1 = 1 - R_1$ and $T_2 = 1 - R_2$. Taking into account multiple reflections inside the cavity, the transmittance through a Fabry–Perot cavity can be expressed in terms of a geometric series. The transmitted light intensity (transmittance) is then given by

$$T = \frac{T_1 T_2}{1 + R_1 R_2 - 2\sqrt{R_1 R_2} \cos(2\phi)} \quad (16.1)$$

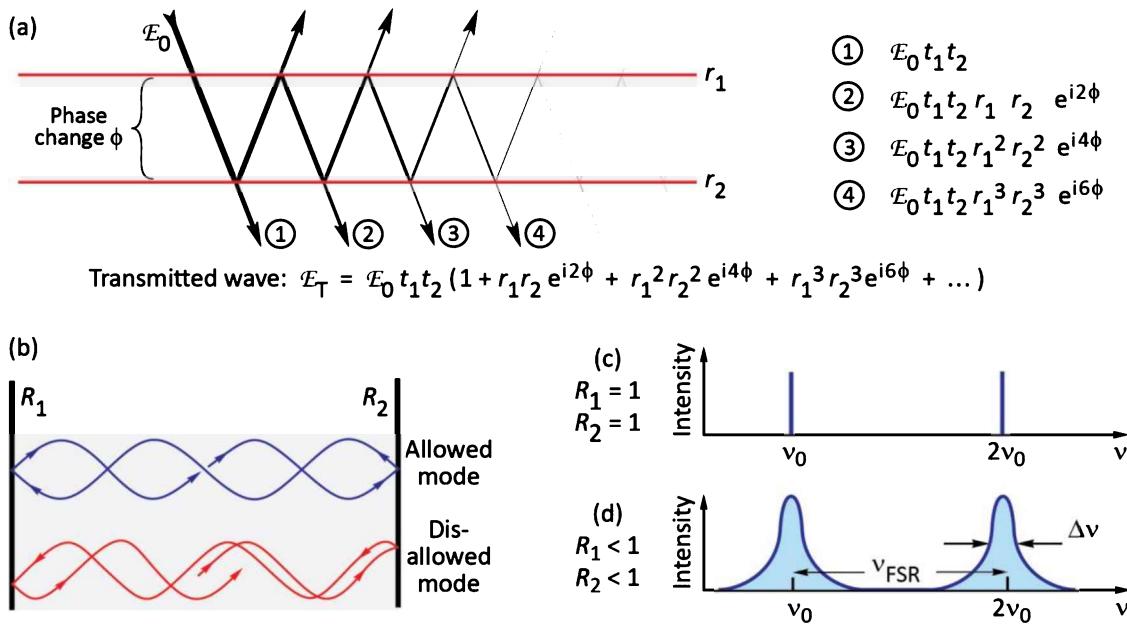


Fig. 16.1: (a) Transmission of a light wave with electric field amplitude E_0 through a Fabry–Perot resonator. (b) Schematic illustration of allowed and disallowed optical modes in a Fabry–Perot cavity consisting of two coplanar reflectors. Optical mode density for a resonator with (c) no mirror losses ($R_1 = R_2 = 100\%$) and (d) mirror losses.

where ϕ is the phase change of the optical wave for a single pass between the two reflectors. Phase changes at the reflectors are neglected. The maxima of the transmittance occur if the condition of constructive interference is fulfilled, i.e. if $2\phi = 0, 2\pi, 4\pi \dots$. Insertion of these values into Eqn. (16.1) yields the transmittance maxima as

$$T_{max} = \frac{T_1 T_2}{(1 - \sqrt{R_1 R_2})^2} . \quad (16.2)$$

For asymmetric cavities ($R_1 \neq R_2$), it is $T_{max} < 1$. For symmetric cavities ($R_1 = R_2$), the transmittance maxima are unity, i.e. $T_{max} = 1$.

Exercise: Transmission through a Fabry–Perot cavity. Derive Eqn. (16.1) by calculating the transmitted wave intensity in terms of a geometric series as illustrated in **Figure 16.1 (a)**.

Solution: As illustrated in **Figure 16.1 (a)**, the amplitude of the electric field of the transmitted wave, E_T , is obtained by the following sum

$$\begin{aligned} E_T &= E_0 t_1 t_2 + E_0 t_1 t_2 r_1 r_2 e^{i2\phi} + E_0 t_1 t_2 r_1^2 r_2^2 e^{i4\phi} + E_0 t_1 t_2 r_1^3 r_2^3 e^{i6\phi} + \dots \\ &= E_0 t_1 t_2 (1 + r_1 r_2 e^{i2\phi} + r_1^2 r_2^2 e^{i4\phi} + r_1^3 r_2^3 e^{i6\phi} + \dots) \end{aligned}$$

where t and r are the electric field Fresnel transmittance and reflectance coefficients,

respectively, and $\phi = 2\pi (\bar{n} d/\lambda)$ is the phase change incurred by the wave when traveling the distance between the two reflectors. Using the formula for the geometric series, i.e. $1 + x + x^2 + x^3 + \dots = 1/(1-x)$, we obtain

$$t = \frac{\mathcal{E}_T}{\mathcal{E}_0} = \frac{t_1 t_2}{1 - r_1 r_2 e^{i2\phi}} = \frac{t_1 t_2}{1 - r_1 r_2 \cos(2\phi) - i r_1 r_2 \sin(2\phi)}$$

Making the transition from electric-field amplitude to electric-field intensity yields

$$\begin{aligned} T = |t|^2 &= \left| \frac{A}{B - iC} \right|^2 = \frac{A^2}{B^2 + C^2} = \frac{t_1^2 t_2^2}{1 - 2r_1 r_2 \cos 2\phi + r_1^2 r_2^2 (\cos^2 2\phi + \sin^2 2\phi)} \\ &= \frac{T_1 T_2}{1 + R_1 R_2 - 2\sqrt{R_1 R_2} \cos 2\phi} \end{aligned}$$

what was to be shown.

Near $\phi = 0, 2\pi, 4\pi \dots$, the cosine term in Eqn. (16.1) can be expanded into a power series ($\cos 2\phi \approx 1 - 2\phi^2$). One obtains

$$T = \frac{T_1 T_2}{(1 - \sqrt{R_1 R_2})^2 + \sqrt{R_1 R_2} 4\phi^2}. \quad (16.3)$$

Equation (16.3) indicates that near the maxima, the transmittance can be approximated by a lorentzian function. The transmittance T in Eqn. (16.3) has a maximum at $\phi = 0$. The transmittance decreases to half of the maximum value at $\phi_{1/2} = [1 - (R_1 R_2)^{1/2}] / [4(R_1 R_2)^{1/2}]^{1/2}$. For high values of R_1 and R_2 , i.e. $R_1 \approx 1$ and $R_2 \approx 1$, it is $\phi_{1/2} = (1/2)[1 - (R_1 R_2)^{1/2}]$.

The cavity finesse, F , is defined as the ratio of the transmittance peak separation to the transmittance full-width at half-maximum, i.e.

$$F = \frac{\text{peak separation}}{\text{peak width}} = \frac{\pi}{2\phi_{1/2}} = \frac{\pi \sqrt[4]{R_1 R_2}}{1 - \sqrt{R_1 R_2}} \approx \frac{\pi}{1 - \sqrt{R_1 R_2}}. \quad (16.4)$$

Inspection of Eqn. (16.4) shows that the finesse becomes very large for high values of R_1 and R_2 .

The wavelength and frequency of light are practically more accessible than the phase. Equations (16.1) ~ (16.4) can be converted to wavelength and frequency using

$$\phi = 2\pi \frac{\bar{n} L_{cav}}{\lambda} = 2\pi \frac{\bar{n} L_{cav} \nu}{c} \quad (16.5)$$

where L_{cav} is the length of the cavity, λ is the wavelength of light in vacuum, ν is the frequency

of light, and \bar{n} is the refractive index inside the cavity. **Figures** 16.1 (c) and (d) show the transmittance through a cavity with $R = 1$ and $R < 1$, respectively. In the frequency domain, the transmittance peak separation is called the *free spectral range* ν_{FSR} , as shown in **Figure** 16.1 (d). The finesse of the cavity in the frequency domain is then given by $F = \nu_{FSR}/\Delta\nu$.

Frequently the *cavity quality factor* Q rather than the finesse is used. The cavity Q is defined as the ratio of the transmittance peak frequency to the peak width. Using this definition and Eqn. (16.4), one obtains

$$Q = \frac{\text{peak frequency}}{\text{peak width}} = \frac{2\bar{n} L_{cav}}{\lambda} \frac{\pi \sqrt[4]{R_1 R_2}}{1 - \sqrt{R_1 R_2}} \approx \frac{2\bar{n} L_{cav}}{\lambda} \frac{\pi}{1 - \sqrt{R_1 R_2}} \quad (16.6)$$

where the peak width is measured in units of frequency.

Figure 16.2 shows an example of a reflectance spectrum of a microcavity consisting of a four-pair Si / SiO₂ distributed Bragg reflector (DBR) deposited on a Si substrate, a SiO₂ center region, a 2.5-pair Si / SiO₂ DBR top reflector. The resonance wavelength of the cavity is approximately 1.0 μm. The reflectance of the cavity does not approach zero at the resonance wavelength due to the unequalness of the reflectivities of the two reflectors.

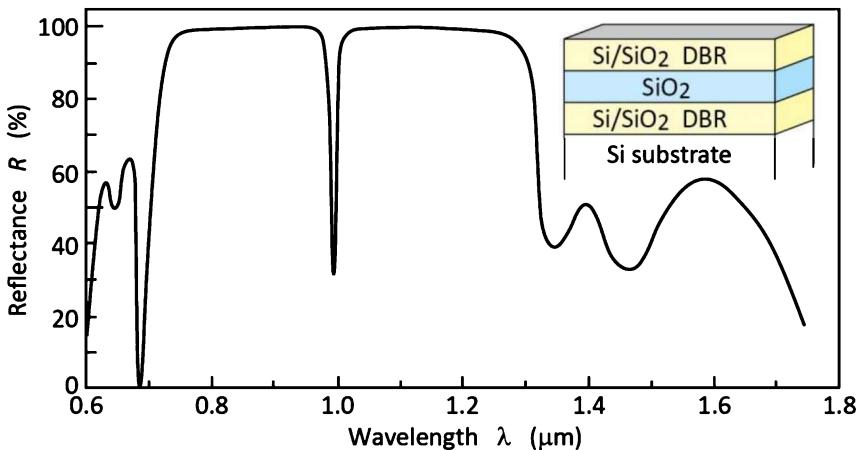


Fig. 16.2: Reflectance of a Fabry–Perot cavity consisting of two Si/SiO₂ reflectors and a SiO₂ center region. At the resonance wavelength ($\lambda \approx 1000$ nm), the reflectivity has a sharp dip.

16.3 Optical mode density in a one-dimensional resonator

In this section, the enhancement of spontaneous emission will be calculated based on the changes of the *optical mode density* in a one-dimensional (1D) resonator, i.e. a coplanar Fabry–Perot microcavity. We first discuss the basic physics causing the changes of the spontaneous emission from an optically active medium located inside a microcavity and then derive analytical formulas for the spectral and integrated emission enhancement. The spontaneous radiative

transition rate in an optically active, homogeneous medium is given by (see, for example, Yariv, 1982)

$$W_{spont} = \tau_{spont}^{-1} = \int_0^\infty \underbrace{W_{spont}^{(\ell)}}_{\substack{\text{transition} \\ \text{rate of} \\ \text{material}}} \underbrace{\rho(v_\ell)}_{\substack{\text{optical} \\ \text{mode} \\ \text{density}}} dv_\ell \quad (16.7)$$

where $W_{spont}^{(\ell)}$ is the spontaneous transition rate into the optical mode ℓ (this transition rate is a property of only the emitting material) and $\rho(v_\ell)$ is the optical mode density. Assuming that the optical medium is homogeneous, the spontaneous emission lifetime, τ_{spont} , is the inverse of the spontaneous emission rate. However, if the optical mode density in the device depends on the spatial direction, as in the case of a cavity structure, then the emission rate given in Eqn. (16.7) depends on the direction. Equation (16.7) can be applied to some small range of solid angle along a certain direction, for example the direction perpendicular to the reflectors of a Fabry–Perot cavity. Thus, Eqn. (16.7) can be used to calculate the emission rate along a specific direction, in particular the optical axis of a cavity.

The spontaneous emission rate into the optical mode ℓ , $W_{spont}^{(\ell)}$, contains the dipole matrix element of the two electronic states involved in the transition (Yariv, 1982). Thus, $W_{spont}^{(\ell)}$ will not be changed by placing the optically active medium inside an optical cavity. However, the optical mode density, $\rho(v_\ell)$, is strongly modified by the cavity. Next, the changes in optical mode density will be used to calculate the changes in spontaneous emission rate.

We first compare the optical mode density in free space with the optical mode density in a microcavity. For simplicity, we restrict our considerations to the one-dimensional case, i.e. to the case of a coplanar Fabry–Perot microcavity. Furthermore, we restrict our considerations to the emission along the optical axis of the cavity.

In a one-dimensional homogeneous medium, the density of optical modes per unit length per unit frequency is given by

$$\rho^{1D}(v) = \frac{2 \bar{n}}{c} \quad (16.8)$$

where \bar{n} is the refractive index of the medium. Equation (16.8) can be derived using a similar formalism commonly used for the derivation of the mode density in free space. The constant optical mode density given by Eqn. (16.8) is shown in **Figure 16.3 (a)**.

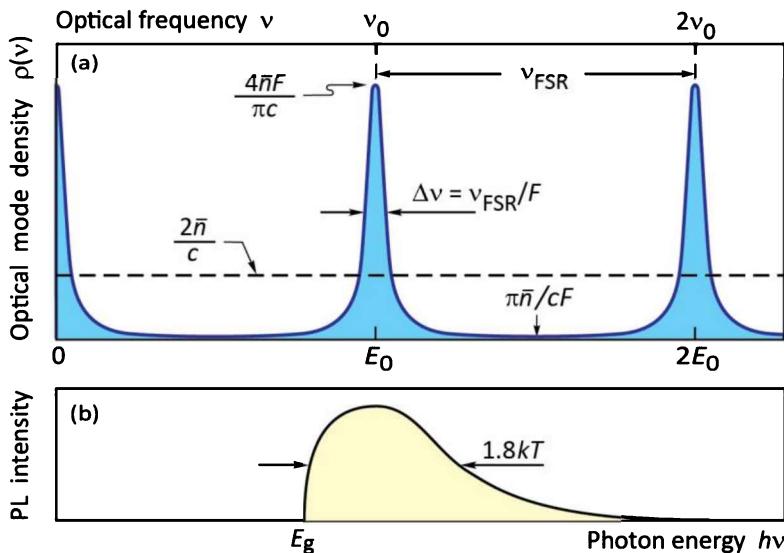


Fig. 16.3: (a) Optical mode density of a one-dimensional planar microcavity (solid line) and of homogeneous one-dimensional space. (b) Theoretical shape of the luminescence spectrum of bulk semiconductors.

Exercise: Optical mode density. Derive Eqn. (16.8), i.e. optical mode density in a 1D space. Also derive the 3D and 2D optical mode density.

Solution: To derive the 3D optical mode density, we consider a cubic volume with length L , and volume $V = L^3$, as shown in **Figure 16.4**. The possible k vectors of an optical wave inside the volume are given by

$$k = \frac{2\pi}{\lambda_0} m = \frac{2\pi}{2L} m = \frac{\pi}{L} m \quad \text{for } m = 1, 2, 3 \dots$$

where m is the mode index and λ_0 is the wavelength of the fundamental optical mode. Note that the modes are equidistant in k space. Thus, the “volume element” of one mode in k space is given by

$$dk_x dk_y dk_z = (\pi/L)^3 .$$

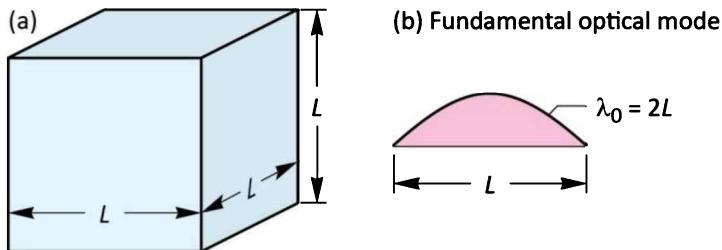


Fig. 16.4: (a) Volume element used to derive the optical mode density. (b) Fundamental optical mode.

Furthermore, the spherical volume in k -space defined by the wave vector k is given by $V = (4/3)\pi k^3$. Since k is restricted to *positive* values, the volume is reduced to only the positive quadrant and we use only one eighth of the space, i.e.

$$V = \frac{1}{8} \frac{4}{3} \pi k^3 .$$

Thus, the number of modes is obtained by dividing the volume in k space used by one optical mode through the volume element in k space, i.e.

$$N_k = \frac{\frac{1}{8} \frac{4}{3} \pi k^3}{(\pi/L)^3} \times 2 = \frac{1}{3} \frac{1}{\pi^2} k^3 V .$$

where the factor of 2 (“ $\times 2$ ”) is due to the two possible polarizations of the optical mode. Using the free-space dispersion relationship $k = \omega (\epsilon\mu)^{1/2} = 2\pi\nu \bar{n}/c$, one obtains

$$N_\nu = \frac{1}{3} \frac{1}{\pi^2} \left(\frac{2\pi\nu \bar{n}}{c} \right)^3 V .$$

Thus, the density of optical modes per unit volume per unit frequency is given by

$$\rho(\nu) = \frac{1}{V} \frac{dN_\nu}{d\nu} = 8\pi \nu^2 \frac{\bar{n}^3}{c^3} \quad \text{3D optical mode density}$$

For two degrees of freedom (2D case), we derive the *2D optical mode density* as follows: The area in k space of one optical mode is given by

$$dk_x dk_y = (\pi/L)^2 .$$

The area in k -space defined by the wave vector k is given by $\frac{1}{4}\pi k^2$. Thus

$$N_k = \frac{\frac{1}{4}\pi k^2}{(\pi/L)^2} \times 2 = \frac{1}{2\pi} k^2 A \quad \text{and} \quad N_\nu = \frac{1}{2\pi} \left(\frac{2\pi\nu \bar{n}}{c} \right)^2 A = 2\pi \left(\frac{\nu \bar{n}}{c} \right)^2 A$$

where the area $A = L^2$. We thus obtain the density of optical modes per unit area per unit frequency

$$\rho(\nu) = \frac{1}{A} \frac{dN_\nu}{d\nu} = 4\pi \nu \frac{\bar{n}^2}{c^2} \quad \text{2D optical mode density}$$

For one degree of freedom (1D case), we derive the *1D optical mode density* as follows: The length in k space of one optical mode is given by

$$dk_x = \pi/L .$$

The length in k -space defined by the wave vector k is given by $\frac{1}{2}k$. Thus

$$N_k = \frac{\frac{1}{2}k}{\pi/L} \times 2 = \frac{1}{\pi} k L \quad \text{and} \quad N_\nu = \frac{1}{\pi} \frac{2\pi\nu \bar{n}}{c} L = \frac{2\nu \bar{n}}{c} L .$$

We thus obtain the density of optical modes per unit length per unit frequency

$$\rho(\nu) = \frac{1}{L} \frac{dN_\nu}{d\nu} = \frac{2\bar{n}}{c} \quad \text{1D optical mode density}$$

what was to be shown.

In planar microcavities, the optical modes are discrete and the frequencies of these modes are integer multiples of the fundamental mode frequency. The optical mode density of a planar microcavity is shown schematically in *Figure 16.3 (a)*. The fundamental and first excited mode

occur at frequencies of ν_0 and $2\nu_0$, respectively. For a cavity with two metallic reflectors (no distributed Bragg reflectors) and a π phase shift of the optical wave upon reflection, the fundamental frequency is given by $\nu_0 = c/(2\bar{n} L_{cav})$, where c is the speed of light in vacuum and L_{cav} is the length of the cavity. In a *resonant microcavity*, the emission frequency of an optically active medium located inside the cavity equals the frequency of one of the cavity modes.

The optical mode density along the cavity axis can be derived using the relation between the mode density in the cavity and the optical transmittance through the cavity, $T(\nu)$,

$$\rho(\nu) = K T(\nu) \quad (16.9)$$

where K is a constant. The value of this constant will be determined by a normalization condition, i.e. by considering a single optical mode. Using Eqn. (16.1), the transmission through a Fabry–Perot cavity can be written as

$$T(\nu) = \frac{T_1 T_2}{1 + R_1 R_2 - 2\sqrt{R_1 R_2} \cos(4\pi \bar{n} L_{cav} \nu/c)} \quad (16.10)$$

The transmittance has maxima at $\nu = 0, \nu_0, 2\nu_0 \dots$, and minima at $\nu = \nu_0/2, 3\nu_0/2, 5\nu_0/2 \dots$. The lorentzian approximation of a transmittance maximum at $\nu = 0$ is obtained by expanding the cosine term in Eqn. (16.10) into a power series using $\cos x \approx 1 - (x^2/2)$, so that

$$T(\nu) = \frac{T_1 T_2 (\sqrt{R_1 R_2})^{-1} (4\pi \bar{n} L_{cav} / c)^{-2}}{\frac{(1 - \sqrt{R_1 R_2})^2}{(\sqrt{R_1 R_2})^{-1} (4\pi \bar{n} L_{cav} / c)^2} + \nu^2} \quad (16.11)$$

Integrating $\rho(\nu)$ over all frequencies and the cavity length yields a single optical mode, i.e.

$$K \int_0^{L_{cav}} \int_{-\infty}^{\infty} \rho(\nu) d\nu dL = 1 \quad (16.12)$$

The lower and upper limit of the frequency integration can be chosen to be $\pm \infty$ since the lorentzian approximation of Eqn. (16.11) has only one maximum at $\nu = 0$. Using Eqns. (16.9), (16.11), and the integration formula $\int_{-\infty}^{\infty} (a^2 + x^2)^{-1} dx = \pi/a$ yields

$$K = \frac{(R_1 R_2)^{3/4}}{T_1 T_2} \frac{4\bar{n}}{c} (1 - \sqrt{R_1 R_2}) \quad (16.13)$$

Using Eqn. (16.9), the optical mode density of a one-dimensional cavity for emission along the cavity axis is then given by

$$\rho(\nu) = \frac{(R_1 R_2)^{3/4}}{T_1 T_2} \frac{4\bar{n}}{c} (1 - \sqrt{R_1 R_2}) T(\nu) . \quad (16.14)$$

Equation (16.14) allows one to calculate the density of optical modes at the maxima and minima. At the *maxima*, the mode density is given by

$$\rho_{max} = \frac{(R_1 R_2)^{3/4}}{1 - \sqrt{R_1 R_2}} \frac{4\bar{n}}{c} . \quad (16.15)$$

Using $(R_1 R_2)^{3/4} \approx 1$ and the expression derived for the finesse F (see Eqn. 16.4), one obtains an approximate expression for the mode density at the maxima

$$\rho_{max} \approx \frac{4}{\pi} \frac{\bar{n} F}{c} . \quad (16.16)$$

That is, the mode density at the maxima is proportional to the finesse of the cavity. At the *minima*, the mode density is given by

$$\rho_{min} = \frac{(R_1 R_2)^{3/4} (1 - \sqrt{R_1 R_2})}{(1 + \sqrt{R_1 R_2})^2} \frac{4\bar{n}}{c} . \quad (16.17)$$

Using $(R_1 R_2)^{3/4} \approx 1$ and the expression derived for the finesse F (see Eqn. 16.4), one obtains an approximate expression for the mode density at the minima

$$\rho_{min} \approx \pi \frac{\bar{n}}{cF} . \quad (16.18)$$

That is, the mode density at the minima is *inversely* proportional to the finesse of the cavity.

The comparison of the optical mode densities of a one-dimensional (1D) free space and a 1D planar cavity is shown in **Figure 16.3**. Note that the *mode density is conserved*, i.e. the areas below the 1D mode density and the 1D cavity mode density are the same.

16.4 Spectral emission enhancement

Because the emission rate at a given wavelength is directly proportional to the optical mode density (see Eqn. 16.7), the emission rate *enhancement spectrum* is given by the ratio of the 1D

cavity mode density to the 1D free space mode density. As calculated earlier, the cavity enhancement spectrum has a lorentzian lineshape. The enhancement factor at the resonance wavelength is thus given by the ratio of the optical mode densities with and without a cavity, i.e.

$$G_e = \frac{\rho_{max}}{\rho^{1D}} \approx \frac{2}{\pi} F \approx \frac{2}{\pi} \frac{\pi (R_1 R_2)^{1/4}}{1 - \sqrt{R_1 R_2}} . \quad (16.19)$$

The equation shows that a strong enhancement of the spontaneous emission rate along the cavity axis can be achieved with microcavities.

Equation (16.19) represents the *average* emission rate enhancement out of *both* reflectors of the cavity. To find the enhancement in a *single* direction, we multiply the enhancement given by Eqn. (16.19) by the fraction of the light exiting the mirror with reflectivity R_1 (i.e. $1 - R_1$) divided by the average loss of the two mirrors for one round trip in the cavity, i.e. $\frac{1}{2} [(1 - R_1) + (1 - R_2)]$. For large R_1 and R_2 , this gives for the enhancement of the emission exiting R_1

$$G_e \approx \frac{2(1 - R_1)}{2 - R_1 - R_2} \frac{2F}{\pi} \approx \frac{1 - R_1}{1 - \sqrt{R_1 R_2}} \frac{2F}{\pi} \approx \frac{2}{\pi} \frac{\pi(R_1 R_2)^{1/4} (1 - R_1)}{(1 - \sqrt{R_1 R_2})^2} \quad (16.20)$$

where we used the approximation $1 - (R_1 R_2)^{1/2} \approx \frac{1}{2}(1 - R_1 R_2) \approx \frac{1}{2}(2 - R_1 - R_2)$. Equation (16.20) represents the emission rate enhancement from a *single* reflector with reflectivity R_1 .

Next, we take into account the standing wave effect, that is, the distribution of the optically active material relative to the nodes and antinodes of the optical wave. The antinode enhancement factor ξ has a value of 2, if the active region is located exactly at an antinode of the standing wave inside the cavity. The value of ξ is unity if the active region is smeared out over many periods of the standing optical wave. Finally, $\xi = 0$ if the active material is located at a node.

The emission rate enhancement is then given by

$$G_e = \frac{\xi}{2} \frac{2}{\pi} \frac{\pi (R_1 R_2)^{1/4} (1 - R_1)}{(1 - \sqrt{R_1 R_2})^2} \frac{\tau_{cav}}{\tau} \quad (16.21)$$

where R_1 is the reflectivity of the light-exit mirror and therefore $R_1 < R_2$. Equation (16.21) also takes into account changes in the spontaneous emission lifetime in terms of τ , the lifetime without cavity, and τ_{cav} , the lifetime with cavity. The factor of τ_{cav}/τ ensures that the

enhancement decreases if the cavity lifetime is reduced as a result of the cavity. For planar microcavities, the ratio of the spontaneous lifetime with a cavity, τ_{cav} , and the lifetime without a cavity, τ , is $\tau_{cav}/\tau \geq 0.9$ (Vredenberg *et al.*, 1993). Thus, the emission lifetime is changed by only a minor amount in a planar microcavity.

16.5 Integrated emission enhancement

The total enhancement *integrated over wavelength*, rather than the enhancement at the resonance wavelength, is relevant for many practical devices. *On resonance*, the emission is enhanced along the axis of the cavity. However, sufficiently far *off resonance*, the emission is suppressed. Because the natural emission spectrum of the active medium (without a cavity) can be much broader than the cavity resonance, it is, *a priori*, not clear whether the integrated emission is enhanced at all. To calculate the wavelength-integrated enhancement, the spectral width of the cavity resonance and the spectral width of the natural emission spectrum must be determined. The resonance spectral width can be calculated from the finesse of the cavity or the cavity quality factor.

The theoretical width of the emission spectrum of bulk semiconductors is $1.8kT$ (see, for example, Schubert, 1993), where k is Boltzmann's constant and T is the absolute temperature. At room temperature, $1.8kT$ corresponds to an emission linewidth of $\Delta\lambda_n = 31$ nm for an emission wavelength of 900 nm. For a cavity resonance width of $5\sim10$ nm, one part of the spectrum is strongly enhanced, whereas the rest of the spectrum is suppressed. The integrated enhancement ratio (or suppression ratio) can be calculated analytically by assuming a gaussian natural emission spectrum. For semiconductors at 300 K, the linewidth of the natural emission is, in the case of high-finesse cavities, *larger* than the width of the cavity resonance. The gaussian emission spectrum has a width of $\Delta\lambda_n = 2\sigma(2\ln 2)^{1/2}$ and a peak value of $(\sigma(2\pi)^{1/2})^{-1}$, where σ is the standard deviation of the gaussian function. The integrated enhancement ratio (or suppression ratio) is then given by (Hunt *et al.*, 1993)

$$G_{int} = \frac{\pi}{2} G_e \Delta\lambda \frac{1}{\sigma\sqrt{2\pi}} = G_e \sqrt{\pi \ln 2} \frac{\Delta\lambda}{\Delta\lambda_n} \quad (16.22)$$

where the factor of $\pi/2$ is due to the lorentzian lineshape of the enhancement spectrum. Hence, the integrated emission enhancement depends on the natural emission linewidth of the active material. The value of G_{int} can be quite different for different types of optically active materials.

Narrow atomic emission spectra can be enhanced by several orders of magnitude (Schubert *et al.*, 1992b). On the other hand, materials having broad emission spectra such as dyes or polymers (de Martini *et al.*, 1987; Suzuki *et al.*, 1991) may not exhibit any integrated enhancement at all. Equation (16.22) also shows that the width of the resonance has a profound influence on the integrated enhancement. Narrow resonance spectral widths, i.e. high finesse values or long cavities (Hunt *et al.*, 1992), reduce the integrated enhancement.

Exercise: Spectral enhancement and integrated enhancement of a resonant-cavity structure emitting at 900 nm. As an exercise, we calculate the spectral and wavelength-integrated enhancement of a semiconductor resonant-cavity structure using Eqns. (16.21) and (16.22). Assuming reflectivities $R_1 = 90\%$ and $R_2 = 97\%$, an antinode enhancement factor of $\xi = 1.5$ and $\tau_{cav}/\tau \approx 1$, one obtains a finesse of $F = 46$, and a peak enhancement factor of $G_e = 68$ using Eqn. (16.21). Insertion of this value into Eqn. (16.22), using a cavity resonance bandwidth of $\Delta\lambda = 6.5$ nm (Schubert *et al.*, 1994), and the theoretical 300 K natural emission linewidth of $\Delta E = 1.8kT$, i.e. $\Delta\lambda = 31$ nm, one obtains a theoretical integrated enhancement factor of $G_{int} = 13$. Experimental enhancement factors of 5 have been demonstrated (Schubert *et al.*, 1994) for the reflectivity values assumed above. The lower experimental enhancement is in part due to a broader natural emission linewidth, which exceeds the theoretical value of $1.8kT$.

The spontaneous emission spectrum of a bulk semiconductor is shown schematically in **Figure 16.3 (b)**. For maximum enhancement along the cavity axis, the cavity must be in resonance with the natural emission spectrum. Note that additional broadening mechanisms (such as alloy broadening) will broaden the natural emission spectrum over and above its theoretical value, i.e. $1.8kT$. Quantum well structures have inherently narrower spectra ($0.7kT$), due to the step-function-like density of states. Low temperatures and excitonic effects can further narrow the natural emission linewidth. Thus, higher enhancements are expected for low temperatures and quantum well active regions.

16.6 Experimental emission enhancement and angular dependence

Particularly high spontaneous emission enhancements can be attained with emitters that have naturally narrow emission lines. Atomic transitions, as in rare-earth elements, have such narrow emission lines. For this reason, rare-earth doped cavities are a great system to study the emission enhancement. The emission spectrum of a high-finesse Si / SiO₂ cavity with an erbium-doped SiO₂ active layer is shown in **Figure 16.5** (Schubert *et al.*, 1992b). A distinct narrowing of the Er

emission spectrum and a giant enhancement of the emission was found. The enhancement factor was greater than 50, when compared to a non-cavity structure.

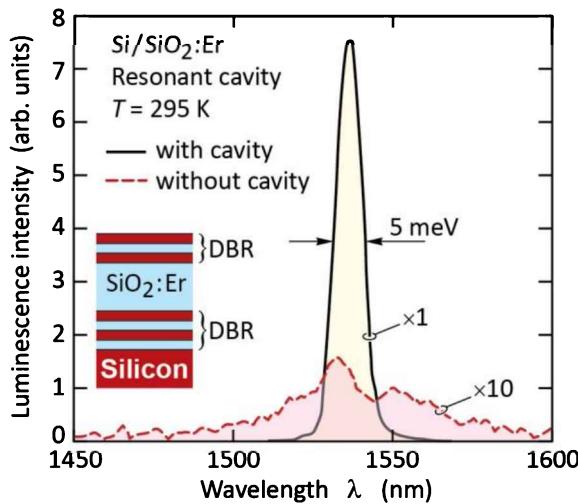


Fig. 16.5: Photoluminescence spectra of Er-doped SiO_2 . One of the spectra is for Er-doped SiO_2 located in a cavity resonant at 1540 nm. The other spectrum is without a cavity. An emission enhancement factor greater than 50 is found (after Schubert *et al.*, 1992b).

The peak emission wavelength depends on the emission angle. The angular dependence can be derived from the condition that the normal wave vector, k_{\perp} , must be constant (in resonance with the cavity), independent of the k propagation direction. This condition can be written as

$$k \cos \theta = k_{\perp} \quad (16.23)$$

or

$$\frac{2\pi}{\lambda_e(\theta)} \cos \theta = \frac{2\pi}{\lambda_{res}} \quad (16.24)$$

where λ_{res} and λ_e are the resonance and emission wavelengths, respectively, and θ is the angle inside the cavity with respect to the surface normal (polar angle), as illustrated in **Figure 16.6**.

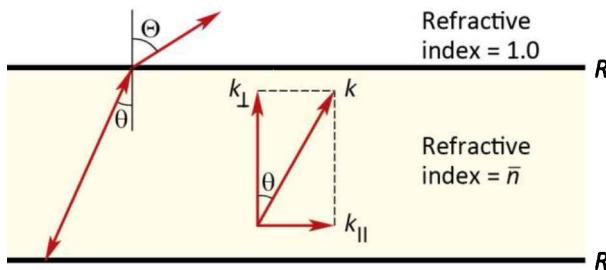


Fig. 16.6: Perpendicular and parallel component of wave vector k for light propagating in a resonant cavity.

Using Eqn. (16.24) and Snell's law ($\bar{n} \sin \theta = \sin \Theta$), the emission wavelength is given by

$$\lambda_e = \lambda_{res} \cos \left[\arcsin \left(\frac{1}{\bar{n}} \sin \theta \right) \right]. \quad (16.25)$$

For small angles, the equation can be approximated by

$$\lambda_e \approx \lambda_{res} \left(1 - \frac{\theta^2}{2 \bar{n}^2} \right). \quad (16.26)$$

The angular dependence of the peak emission wavelength is shown in **Figure 16.7**. The solid and dashed lines are the exact (Eqn. 16.25) and approximate solutions (Eqn. 16.26), respectively. A refractive index of 1.5 is used in the calculation.

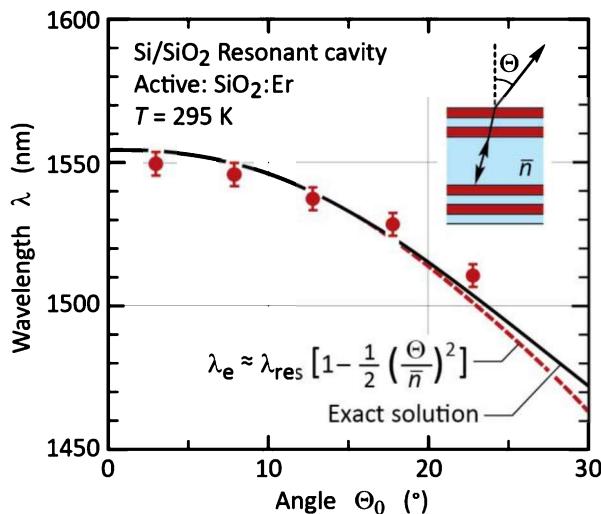


Fig. 16.7: Peak emission wavelength as a function of polar angle for a planar Si/SiO₂:Er resonant cavity (after Schubert *et al.*, 1992b).

A similar angular dependence is found for semiconductors. The emission spectra of an AlGaNp cavity are shown in **Figure 16.8** for different emission angles (Streubel *et al.*, 2002). Note that the highest intensity is found at an emission angle of 30°, indicating that the resonance wavelength of the cavity is located at the long-wavelength end of the semiconductor natural emission spectrum. This results in the highest angle-integrated emission power (Streubel *et al.*, 2002).

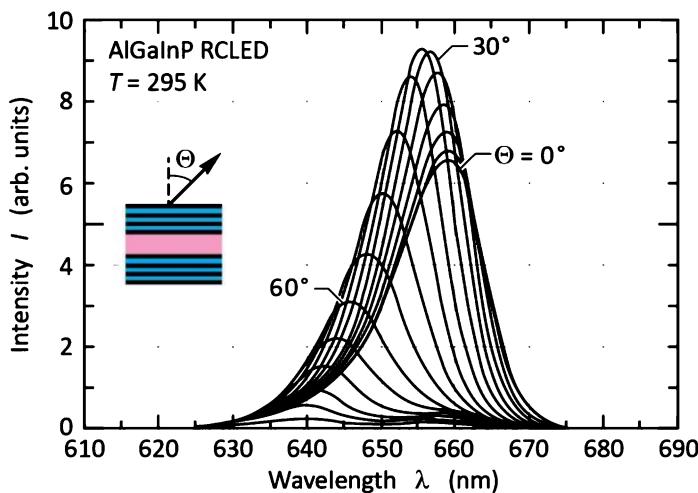


Fig. 16.8: Emission spectra of AlGaNp RCLED for different polar angles. The long-wavelength part of the QW emission is emitted in the forward direction (0°). The shorter wavelengths are emitted off-axis. When measured with an integrating sphere, an 18 nm wide spectrum (FWHM) is found (after Streubel *et al.*, 2002).

16.7 References

- Coldren L. A. and Corzine S. W. *Diode Lasers and Photonic Integrated Circuits* (John Wiley and Sons, New York, 1995)
- De Martini F., Innocenti G., Jacobovitz G. R., and Mataloni P. "Anomalous spontaneous emission time in a microscopic optical cavity" *Physical Review Letters* **59**, 2955 (1987)
- Deppe D. G. and Lei C. "Spontaneous emission and optical gain in a Fabry–Perot microcavity" *Applied Physics Letters* **60**, 527 (1992)
- Dodabalapur A., Rothberg L. J., and Miller T. M. "Color variation with electroluminescent organic semiconductors in multimode resonant cavities" *Applied Physics Letters* **65**, 2308 (1994)
- Einstein A. "On the quantum theory of radiation" (translated from German) *Zeitschrift für Physik* **18**, 121 (1917)
- Fabry G. and Perot A. "Theory and applications of a new interference method for spectroscopy" (translated from French) *Annales de Chimie et de Physique* **16**, 115 (1899)
- Hall R. N., Fenner G. E., Kingsley J. D., Soltys T. J., and Carlson R. O. "Coherent light emission from GaAs junctions" *Physical Review Letters* **9**, 366 (1962)
- Hayashi I., Panish M. B., Foy P. W., and Sumski S. "Junction lasers which operate continuously at room temperature" *Applied Physics Letters* **17**, 109 (1970)
- Hunt N. E. J., Schubert E. F., Logan R. A., and Zydzik G. J. "Enhanced spectral power density and reduced linewidth at 1.3 μm in an InGaAsP quantum well resonant cavity light emitting diode" *Applied Physics Letters* **61**, 2287 (1992)
- Hunt N. E. J., Schubert E. F., Kopf R. F., Sivco D. L., Cho A. Y., and Zydzik G. J. "Increased fiber communications bandwidth from a resonant cavity light emitting diode emitting at $\lambda = 940$ nm" *Applied Physics Letters* **63**, 2600 (1993)
- Hunt N. E. J., Schubert E. F., Sivco D. L., Cho A. Y., Kopf R. F., Logan R. A., and Zydzik G. J. "High efficiency, narrow spectrum resonant cavity light-emitting diodes" in *Confined Electrons and Photons* edited by E. Burstein and C. Weisbuch (Plenum Press, New York, 1995a)
- Hunt N. E. J., Vredenberg A. M., Schubert E. F., Becker P. C., Jacobson D. C., Poate J. M., and Zydzik G. J. "Spontaneous emission control of Er³⁺ in Si / SiO₂ microcavities" in *Confined Electrons and Photons* edited by E. Burstein and C. Weisbuch (Plenum Press, New York, 1995b)
- Jewell J. L., Lee Y. H., McCall S. L., Harbison J. P., and Florez L. T. "High-finesse AlGaAs interference filters grown by molecular beam epitaxy" *Applied Physics Letters* **53**, 640 (1988)
- Joannopoulos J. D., Meade R. D., and Winn J. N. *Photonic Crystals* (Princeton University Press, Princeton, 1995)
- Kobayashi T., Segawa T., Morimoto A., and Sueta T. *Meeting of the Japanese Society of Applied Physics Tokyo* (1982); see also Yokoyama (1992)
- Kobayashi T., Morimoto A., and Sueta T. *Meeting of the Japanese Society of Applied Physics Tokyo* (1985); see also Yokoyama (1992)
- Nakayama T., Itoh Y., and Kakuta A. "Organic photo- and electroluminescent devices with double mirrors" *Applied Physics Letters* **63**, 594 (1993)
- Numai T., Kosaka H., Ogura I., Kurihara K., Sugimoto M., and Kasahara K. "Indistinct threshold laser operation in a pnpn vertical to surface transmission electrophotonic device with a vertical cavity" *IEEE Journal of Quantum Electronics* **29**, 403 (1993)
- Purcell E. M. "Spontaneous emission probabilities at radio frequencies" *Physical Review* **69**, 681 (1946)
- Saleh B. E. A. and Teich M. C. *Fundamentals of Photonics* (John Wiley and Sons, New York, 1991)
- Schubert E. F., Wang Y. H., Cho A. Y., Tu L. W., and Zydzik G. J. "Resonant cavity light emitting diode" *Applied Physics Letters* **60**, 921 (1992a)
- Schubert E. F., Vredenberg A. M., Hunt N. E. J., Wong Y. H., Becker P. C., Poate J. M., Jacobson D. C., Feldman L. C., and Zydzik G. J. "Giant enhancement in luminescence intensity in Er-doped Si / SiO₂ resonant cavities" *Applied Physics Letters* **61**, 1381 (1992b)
- Schubert E. F. *Doping in III–V Semiconductors* p. 512 (Cambridge University Press, Cambridge UK, 1993)

- Schubert E. F., Hunt N. E. J., Micovic M., Malik R. J., Sivco D. L., Cho A. Y., and Zydzik G. J. *Science* **265**, 943 (1994)
- Schubert E. F., Hunt N. E. J., Malik R. J., Micovic M., and Miller D. L. “Temperature and modulation characteristics of resonant cavity light-emitting diodes” *IEEE Journal of Lightwave Technology* **14**, 1721 (1996)
- Streubel K., Linder N., Wirth R., and Jaeger A. “High brightness AlGaN_xP light-emitting diodes” *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 321 (2002)
- Suzuki M., Yokoyama H., Brorson S. D., and Ippen E. P. “Observation of spontaneous emission lifetime change of dye-containing Langmuir–Blodgett films in optical microcavities” *Applied Physics Letters* **58**, 998 (1991)
- Vredenberg A. M., Hunt N. E. J., Schubert E. F., Jacobson D. C., Poate J. M., and Zydzik G. J. *Physical Review Letters* **71**, 517 (1993)
- Yablonovitch E., Gmitter T. J., and Bhat R. “Inhibited and enhanced spontaneous emission from optically thin AlGaAs / GaAs double heterostructures” *Physical Review Letters* **61**, 2546 (1988)
- Yablonovitch E. personal communication (1994)
- Yariv A. *Theory and Applications of Quantum Mechanics* (John Wiley and Sons, New York, 1982) p. 143
- Yokoyama H., Nishi K., Anan T., Yamada H., Boorson S.D., and Ippen E. P. “Enhanced spontaneous emission from GaAs quantum wells in monolithic microcavities” *Applied Physics Letters* **57**, 2814 (1990)
- Yokoyama H. “Physics and device applications of optical microcavities” *Science* **256**, 66 (1992)

17

17 – Resonant-cavity light-emitting diodes

17.1 Introduction and history

The resonant-cavity light-emitting diode (RCLED) is a light-emitting diode that has a light-emitting region inside an optical cavity. The optical cavity has a thickness of typically one-half or one times the wavelength of the light emitted by the LED, i.e. a fraction of a micrometer for devices emitting in the visible or in the infrared. The resonance wavelength of the cavity coincides or is in resonance with the emission wavelength of the light-emitting active region of the LED. Thus, the cavity is a *resonant cavity*. The spontaneous emission properties from a light-emitting region located inside the resonant cavity are enhanced by the resonant-cavity effect. The RCLED is the first practical device making use of spontaneous emission enhancement occurring in microcavities.

The placement of an active region inside a resonant cavity results in multiple improvements of the device characteristics. Firstly, the light intensity emitted from the RCLED along the axis of the cavity, i.e. normal to the semiconductor surface, is higher compared with conventional LEDs. The enhancement factor is typically a factor of 2~10. Secondly, the emission spectrum of the RCLED has a *higher spectral purity* compared with conventional LEDs. In conventional LEDs, the spectral emission linewidth is determined by the thermal energy kT . However, in RCLEDs, the emission linewidth is determined by the quality factor (Q factor) of the optical cavity. As a result, the spectral emission width of the RCLED is a factor of 2~5 narrower compared with conventional LEDs. For the same reason, the wavelength shift with temperature is determined by the temperature coefficient of the optical cavity and not by the energy gap of the active material. This causes a significantly higher temperature stability of the RCLED emission wavelength compared with conventional LEDs. Thirdly, the emission far-field pattern of the RCLED is more *directed* compared with conventional LEDs. In conventional LEDs, the emission pattern is lambertian (i.e. cosine-function-like). In an RCLED, the emission pattern is directed mostly along the optical axis of the cavity.

These characteristics of RCLEDs are desirable for local-area, medium bit rate optical communication systems. LEDs play an important role in local-area (< 5 km) medium bit rate (< 1 Gbit/s) optical communication networks. In particular, plastic optical fibers are increasingly used for optical communication over short distances. The higher emission intensity and the more directed emission pattern afforded by the RCLED increase the power coupled into the optical fiber. As a result, the RCLED can transmit data over longer distances. Furthermore, the higher spectral purity of RCLEDs results in less chromatic dispersion allowing for higher bit rates.

Light-emitting diodes are the transmitter device of choice for medium bit rate optical communication over distances less than 5 km. Compared with lasers, LEDs are less expensive, more reliable, and less temperature sensitive. The RCLED has improved characteristics compared with conventional LEDs while maintaining the inherent advantages of LEDs. The reflectivity of the RCLED reflectors is lower compared with vertical-cavity surface-emitting lasers (VCSELs), thereby allowing for a lower RCLED manufacturing cost compared with VCSELs. At 650 nm, the preferred communication wavelength for plastic optical fibers, VCSELs are difficult to manufacture due to the lack of high-reflectivity reflectors.

RCLEDs are also used for high-brightness applications (Streubel *et al.*, 1998; Wirth *et al.*, 2001, 2002). In these devices, the resonance wavelength is designed to be at the long-wavelength end of the spontaneous emission spectrum of the semiconductor. This ensures that the emission intensity, integrated over all spatial directions, is maximized.

The enhanced spontaneous emission occurring in resonant-cavity structures can be beneficially employed in semiconductor and polymer LEDs. Resonant-cavity light-emitting diodes were first realized in 1992 (Schubert *et al.*, 1992a) in the GaAs material system. About a year later, RCLEDs were demonstrated in organic materials (Nakayama *et al.*, 1993).

Resonant-cavity structures with enhanced spontaneous emission also include Er-doped microcavities (Schubert *et al.*, 1992b). Owing to the inherently narrow luminescence line of intra-atomic Er radiative transitions, there is a very good overlap between the cavity optical mode and the Er luminescence line. At the present time, no Er-doped current-injection devices exist. However, the great potential of Er-doped resonant cavities makes the realization of Er-doped RCLEDs likely in the future.

17.2 RCLED design rules

The basic structure of an RCLED is shown in *Figure 17.1* and comprises two mirrors with reflectivity R_1 and R_2 . The reflectivity of the two mirrors is chosen to be unequal so that the light exits the cavity predominantly through one of the mirrors. This mirror is called the ***light-exit mirror***. Here we designate the mirror with reflectivity R_1 as the light-exit mirror. An active region is located between the mirrors, preferably at the antinode location of the standing optical wave of the cavity, as shown in *Figure 17.1*. Metal mirrors are assumed in *Figure 17.1* so that the wave amplitude is zero at the location of the mirrors.

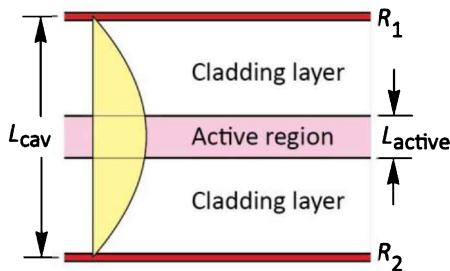


Fig. 17.1: Schematic illustration of a resonant cavity consisting of two metal mirrors with reflectivity R_1 and R_2 . The active region has a thickness L_{active} and an absorption coefficient α . Also shown is the standing optical wave. The cavity length L_{cav} is equal to $\lambda/2$.

Next, we summarize several design rules intended to maximize the spontaneous emission enhancement in resonant-cavity structures (Schubert *et al.*, 1994, 1996; Hunt *et al.*, 1995a, 1995b). These rules will provide further insight into the fundamental operating principles of RCLEDs and the differences of these devices with respect to VCSELs.

The ***first design criterion*** for RCLEDs is that the reflectivity of the light-exit reflector, R_1 , should be much lower than the reflectivity of the back reflector, i.e.

$$R_1 \ll R_2 . \quad (17.1)$$

This condition ensures that light exits the device mainly through the reflector with reflectivity R_1 . Equation (17.1) applies to the design of communication RCLEDs, where light is emitted into the small core of a multimode fiber, and display RCLEDs, where light should be emitted towards the observer.

The ***second design criterion*** calls for the shortest possible cavity length L_{cav} . In order to derive this criterion, the integrated enhancement, discussed in a preceding section, can be rewritten by using the expressions for the cavity finesse F and cavity quality factor Q . One obtains

$$G_{int} = \frac{\xi}{2} \frac{2}{\pi} \frac{1 - R_1}{1 - \sqrt{R_1 R_2}} \sqrt{\pi \ln 2} \frac{\lambda}{\Delta\lambda_n} \frac{\lambda_{cav}}{L_{cav}} \frac{\tau_{cav}}{\tau} \quad (17.2)$$

where λ and λ_{cav} are the active region emission wavelengths in vacuum and inside the cavity, respectively. Since the emission wavelength λ and the natural linewidth of the active medium, $\Delta\lambda_n$, are given quantities, Eqn. (17.2) shows that minimization of the cavity length L_{cav} maximizes the integrated intensity.

The importance of a short cavity length is elucidated by **Figure 17.2**. The optical mode densities of two different cavities, namely a short and a long cavity, are shown in **Figures 17.2 (a)** and **(b)**, respectively. Both cavities have the same mirror reflectivities and finesse. The natural emission spectrum of the active region is shown in **Figure 17.2 (c)**. The best overlap between the resonant optical mode and the active region emission spectrum is obtained for the shortest cavity.

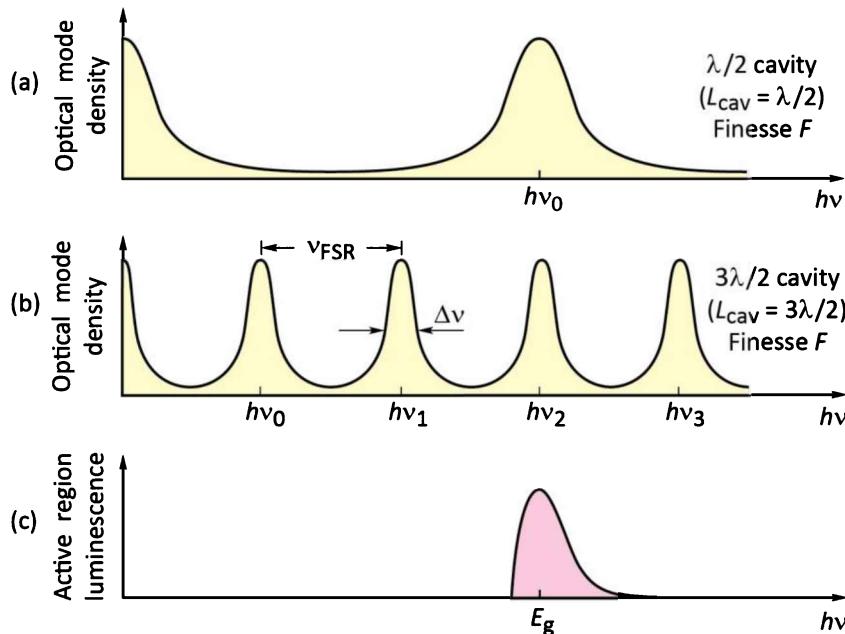


Fig. 17.2: Optical mode density for (a) a short and (b) a long cavity with the same finesse F . (c) Spontaneous free-space emission spectrum of an LED active region. The spontaneous emission spectrum has a better overlap with the short-cavity mode spectrum compared with the long-cavity mode spectrum.

The largest enhancements are achieved with the shortest cavities, which in turn are obtained if the *fundamental* cavity mode is in resonance with the emission from the active medium. The cavity length is also reduced by using a DBR with a short penetration depth, i.e. a DBR consisting of two materials with a large difference in the refractive index.

The **third design criterion** is the minimization of self-absorption in the active region. This criterion can be stated as follows: the reabsorption probability of photons emitted from the

active region into the cavity mode should be much smaller than the escape probability of photons through one of the reflectors. Assuming $R_2 \approx 1$, this criterion can be written as

$$2 \xi \alpha L_{active} < (1 - R_1) \quad (17.3)$$

where α and L_{active} are the absorption coefficient and the thickness of the active region, respectively. If the criterion of Eqn. (17.3) were not fulfilled, photons would most likely be reabsorbed by the active region. Subsequently, re-emission will, with a certain probability, occur along the lateral direction (waveguided modes), i.e. not into the cavity mode. Another possibility is that the electron–hole pairs generated by reabsorption recombine non-radiatively. In either case, reabsorption processes occurring in high-finesse cavities *reduce the cavity mode emission out of the cavity*. Thus, if the condition of Eqn. (17.3) is not fulfilled, the emission intensity of resonant cavities is lowered rather than enhanced.

Whereas the condition of Eqn. (17.3) is fulfilled in RCLEDs, it is clearly not fulfilled in VCSELs. The spontaneous emission intensities of RCLEDs and VCSELs were compared by Schubert *et al.* (1996). In this comparison, the VCSEL and the RCLED were driven by an injection current of 2 mA, which is below the threshold current of the VCSEL of $I_{th} = 7$ mA. The spontaneous emission spectra of an RCLED and a VCSEL are shown in **Figure 17.3**. The VCSEL has an AlGaAs/GaAs quantum well active region emitting at 850 nm. Both reflectors of the VCSEL are AlGaAs/AlAs DBRs. **Figure 17.3** reveals that the emission intensity of the VCSEL in the *spontaneous* regime is more than a factor of 15 lower than the emission intensity from the RCLED.

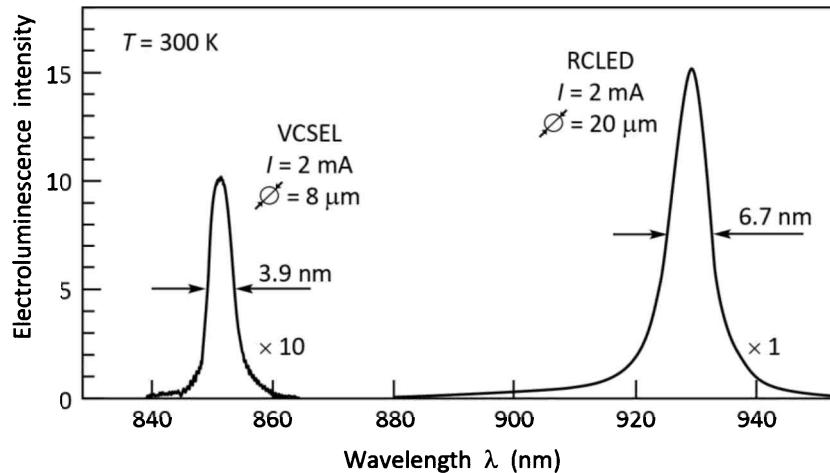


Fig. 17.3: Spontaneous electroluminescence spectrum of a vertical-cavity surface-emitting laser (VCSEL) emitting at 850 nm and of a resonant-cavity light-emitting diode (RCLED) emitting at 930 nm. The drive current for both devices is 2 mA. The VCSEL spectrum is multiplied by a factor of 10. The threshold current of the VCSEL is 7 mA (after Schubert *et al.*, 1996).

Because the magnitude of the maximum gain in semiconductors is always lower than the magnitude of the absorption coefficient in an unpumped semiconductor ($|g| < |\alpha|$), VCSELs

could not lase if the condition of Eqn. (17.3) were met. Thus, the spontaneous emission intensity in VCSELs is low and *must be low* in order to enable the device to lase. **Figure 17.3** also reveals that the emission spectral linewidth of VCSELs is narrower than that of RCLEDs. The higher spectral purity is due to the higher values of R_1 and R_2 as required for VCSELs.

The fulfillment of Eqn. (17.3) by RCLEDs also implies that these devices *cannot lase*. As stated above, it is always $|g| < |\alpha|$. Consequently, the mirror loss ($1 - R_1$) is always larger than the maximum achievable round-trip gain ($2 \xi g L_{active}$). The fundamental inability of RCLEDs to lase has been verified experimentally by pulsed injecting currents of large magnitude without finding any evidence for lasing. These considerations show that the device physics of RCLEDs and VCSELs is fundamentally different.

The arguments used above imply that the *spontaneous* emission into the fundamental cavity mode in VCSEL structures is very low due to reabsorption of photons by the active region. A reduction of the threshold current by increasing the reflectivity will be accompanied by a further decrease of the *spontaneous* emission below threshold. We therefore conclude that the so-called zero-threshold laser (Kobayashi *et al.* 1982; Yokoyama, 1992) cannot be realized by a planar resonant-cavity structure.

17.3 GaInAs/GaAs RCLEDs emitting at 930 nm

The structure of an RCLED with a GaInAs active region is shown in **Figure 17.4 (a)**. The cavity is defined by one distributed Bragg reflector (DBR) and one metallic reflector. Also included are two confinement regions and a four-quantum-well active region. The heavily doped n-type substrate is coated with a ZrO₂ anti-reflection layer (Schubert *et al.*, 1994). A picture of the first RCLED is shown in **Figure 17.4 (b)**.

The motivation for the metal reflector is twofold. Firstly, the metallic Ag reflector serves as a non-alloyed ohmic contact to the heavily doped p-type ($N_A \approx 5 \times 10^{19} \text{ cm}^{-3}$) GaAs top layer, thus effectively confining the pumped region to the area below the contact. Secondly, it was shown in the preceding section that the cavity length must be kept as short as possible for maximizing the emission enhancement. Owing to the lack of a penetration depth, metal reflectors allow for a short cavity length. Cavities with two metallic reflectors have been reported (Wilkinson *et al.*, 1995). However, optical absorption losses in the light-exit mirror can be large in a double metal mirror structure, unless very thin metallic reflectors are used (Tu *et al.*, 1990).

The lack of a p-type DBR also avoids the well-known problem of high resistance in p-type DBRs (Schubert *et al.*, 1992c; Lear and Schneider, 1996). It has been shown that parabolic grading yields the lowest ohmic resistance in DBRs. Such parabolic grading is suited to eliminating heterojunction band discontinuities (Schubert *et al.*, 1992c).

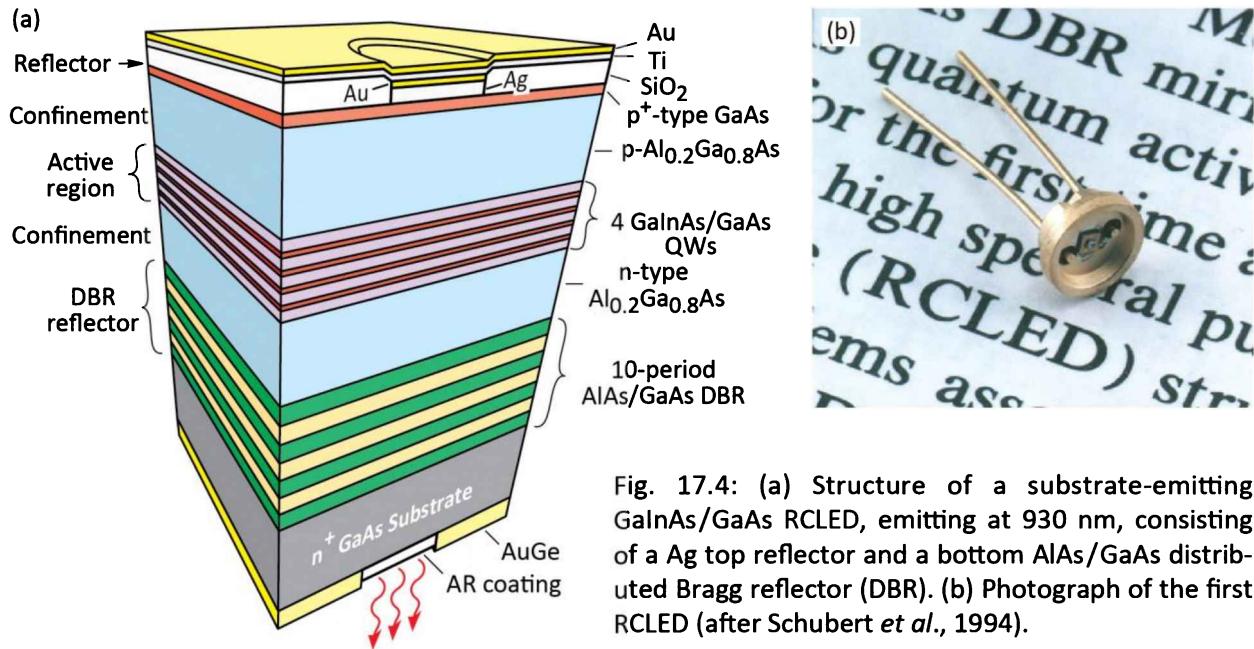


Fig. 17.4: (a) Structure of a substrate-emitting GaInAs/GaAs RCLED, emitting at 930 nm, consisting of a Ag top reflector and a bottom AlAs/GaAs distributed Bragg reflector (DBR). (b) Photograph of the first RCLED (after Schubert *et al.*, 1994).

The magnitude of the reflectivity of the DBR needs to be consistent with Eqns. (17.1) and (17.3). The Ag back mirror has a reflectivity of approximately 96%. According to Eqn. (17.1), the DBR reflectivity must be < 96%. The second criterion of Eqn. (17.3) requires that $2 \xi \alpha L_{active} < 1 - R_1$. Assuming $\xi = 1.3$, $\alpha = 10^4 \text{ cm}^{-1}$, and $L_{active} = 40 \text{ nm}$, one obtains the condition $R_1 < 90\%$. Thus, the mirror reflectivity of RCLEDs must be much lower than that of VCSELs. A high reflectivity would increase self-absorption and decrease the light output of the device as discussed earlier. De Neve *et al.* (1995) used an extensive theoretical model to calculate the mirror reflectivity. The maximum efficiency was calculated at a reflectivity of $R_1 = 50\text{--}60\%$.

The reflection and emission properties of the RCLED are shown in *Figures* 17.5 (a) and (b). The reflection spectrum of the RCLED (*Figure* 17.5 (a)) exhibits a highly reflective band for wavelengths > 900 nm and a dip in the reflectivity at the cavity resonance. The spectral width of the cavity resonance is 6.3 nm. The emission spectrum of an electrically pumped device, shown in *Figure* 17.5 (b), has nearly the same shape and width as the cavity resonance.

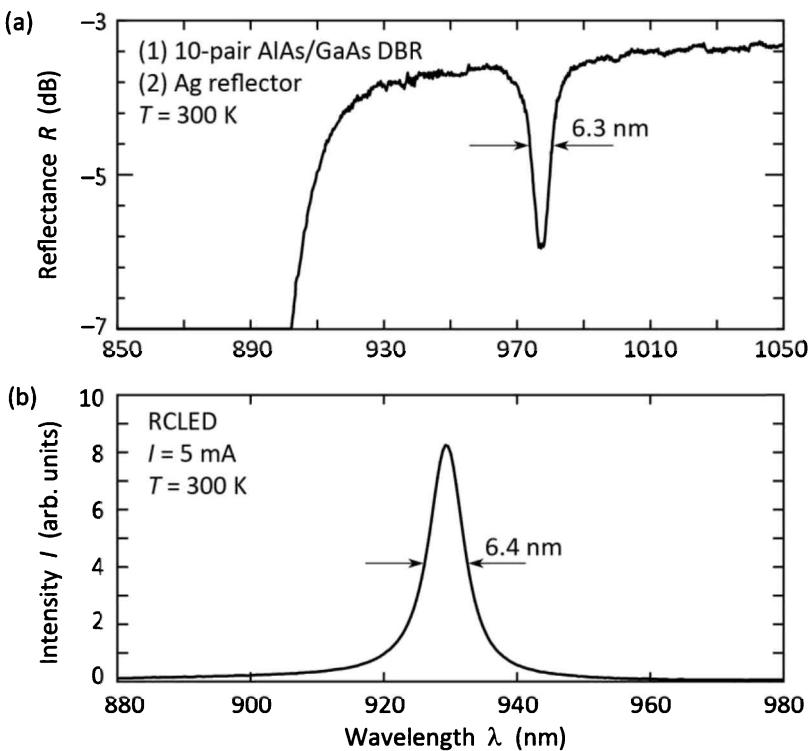


Fig. 17.5: (a) Reflectance of a resonant cavity consisting of a 10-pair AlAs/GaAs distributed Bragg reflector and an Ag reflector. (b) Emission spectrum of an RCLED consisting of a 10-pair AlAs/GaAs distributed Bragg reflector and an Ag reflector (after Schubert *et al.*, 1994).

In conventional LEDs, the spectral characteristics of the devices reflect the thermal distribution of electrons and holes in the conduction and valence bands. The spectral characteristics of light emission from microcavities are as intriguing as they are complex. However, restricting our considerations to the optical axis of the cavity simplifies the cavity physics considerably. If we assume that the cavity resonance is much narrower than the natural emission spectrum of the semiconductor, then the on-resonance luminescence is enhanced whereas the off-resonance luminescence is suppressed. The on-axis emission spectrum should therefore reflect the enhancement, that is, the resonance spectrum of the cavity. The experimental results shown in **Figure 17.5** confirm this conjecture.

Owing to the cavity, the emission spectrum of an RCLED is much narrower than the emission spectrum of regular LEDs (Schubert *et al.* 1992a; Hunt *et al.*, 1992, 1993). The spectral width of the RCLED emitted into a certain direction is given by the optical characteristics of the cavity. In contrast, the spectral width of a regular LED is about $1.8kT$, a value that is much wider than the RCLED emission spectrum. A comparison of a regular GaAs LED and a GaInAs RCLED emission spectrum is shown in **Figure 17.6**. Comparison of the spectra shows that the RCLED emission spectrum is a factor of about 10 narrower than the spectrum of the GaAs LED.

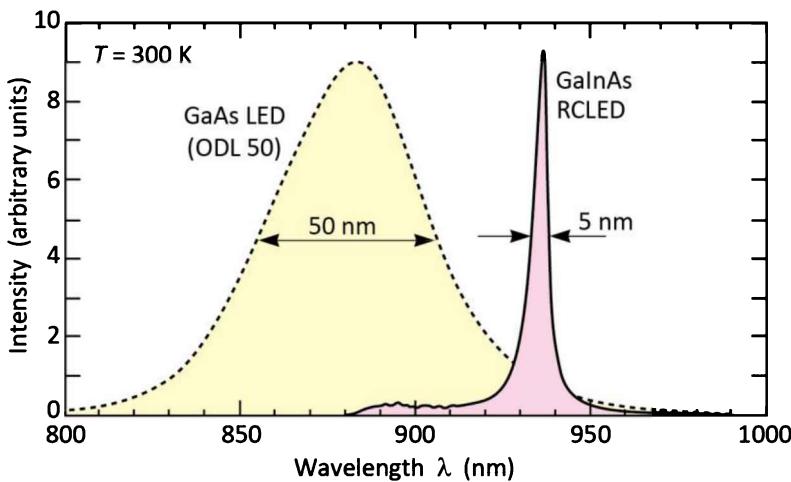


Fig. 17.6: Comparison of the emission spectra of a GaAs LED emitting at 870 nm (AT&T ODL 50 product) and a GaInAs RCLED emitting at 930 nm (after Hunt *et al.*, 1993).

A regular LED has little or no angle dependence of the emission spectrum. However, the reflective properties of DBRs and of cavities consisting of DBRs are angle-dependent. Consequently, the emission from an RCLED along a certain direction is narrower than that of a regular LED. Integrated over all directions, the RCLED has a broad emission spectrum.

A figure of merit for LEDs used in optical fiber communication systems is the photon flux density emitted from the diode at a given current, which, for a given wavelength, can be characterized in terms of the unit microwatts per unit solid angle (per steradian). The optical power coupled into a fiber is directly proportional to the photon flux density.

Figure 17.7 shows the intensity of an RCLED as a function of the injection current. For comparison, the calculated intensity of the ***ideal isotropic emitter***, which is a hypothetical device, is also shown. The ideal isotropic emitter is assumed to have an internal quantum efficiency of 100% and the device is assumed to be clad by an anti-reflection coating providing zero reflectivity ($R = 0$) for all wavelengths emitted from the active region. If the photon emission inside the semiconductor is isotropic, then the optical power per unit current per unit solid angle normal to the planar semiconductor surface is given by

$$\frac{P_{optical}}{\Omega} = \frac{1}{4\pi \bar{n}^2} \frac{hc}{e \lambda} \quad (17.4)$$

where Ω represents the unit solid angle, \bar{n} is the refractive index of the semiconductor, c is the speed of light in vacuum, e is the electronic charge, and λ is the emission wavelength in vacuum. Equation (17.4) is represented by the dashed line in **Figure 17.7**. Neither the 100% internal quantum efficiency nor the hypothetical anti-reflection coating can be reproduced in practice for fundamental reasons. Therefore, the ideal isotropic emitter represents an upper limit for the

intensity attainable with any conventional LED. Of course, even the best conventional LEDs have intensities lower than that of the ideal isotropic emitter. Also included in **Figure 17.7** is the ODL 50 GaAs LED that was frequently used for optical fiber communication in the 1990s. All devices shown in **Figure 17.7** have planar light-emitting surfaces, and no lensing is used.

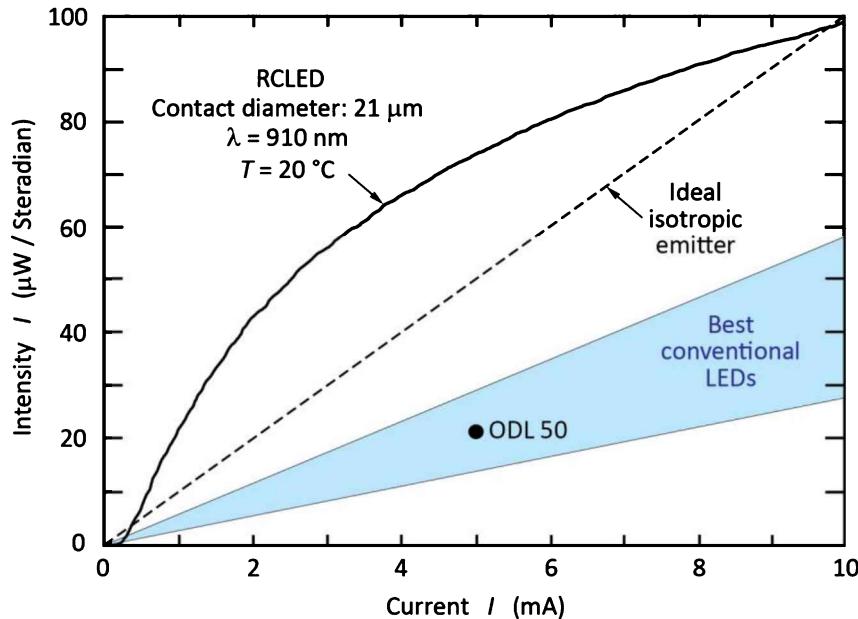


Fig. 17.7: Light-versus-current curves of a GaInAs/ GaAs RCLED and of the *ideal isotropic emitter*. The ideal isotropic emitter is a hypothetical device emitting light isotropically with a quantum efficiency of 100%. The shaded region shows the intensity of the best conventional LEDs. The ODL 50 is a commercial LED product (after Schubert *et al.*, 1994).

Figure 17.7 reveals that the RCLED provides unprecedented intensities in terms of both absolute values and slope efficiencies. The slope efficiency is 7.3 times the efficiency of the best conventional LEDs and 3.1 times the calculated efficiency of the ideal isotropic emitter. At a current of 5 mA, the intensity of the RCLED is 3.3 times that of the best conventional LEDs including the ODL 50. The high efficiencies make the RCLED well suited for optical interconnect and communication systems.

The higher spectral purity of RCLEDs reduces chromatic dispersion in optical fiber communications (Hunt *et al.*, 1993). The chromatic dispersion is directly proportional to the linewidth of the source. Since RCLEDs have linewidths 5~10 times narrower than conventional LEDs, chromatic dispersion effects, which dominate at wavelengths of 800~900 nm, are reduced as well. Hunt *et al.* (1993) showed that the bandwidth of RCLEDs is a factor of 5~10 higher than that of conventional LEDs. An RCLED-versus-LED comparison in a transmission experiment is shown in **Figure 17.8**. The results show the received signal after transmission lengths of 5 m and 3.4 km for the two devices. The fiber used is a graded-index multimode fiber with a core diameter of 62.5 μm. After a transmission length of 5 m, no marked difference is found for the two devices.

However, a substantial difference is found after a transmission length of 3.4 km. Inspection of **Figure 17.8** reveals that the RCLED exhibits much less pulse broadening as compared to the conventional LED. This difference is due to reduced material dispersion for the RCLED.

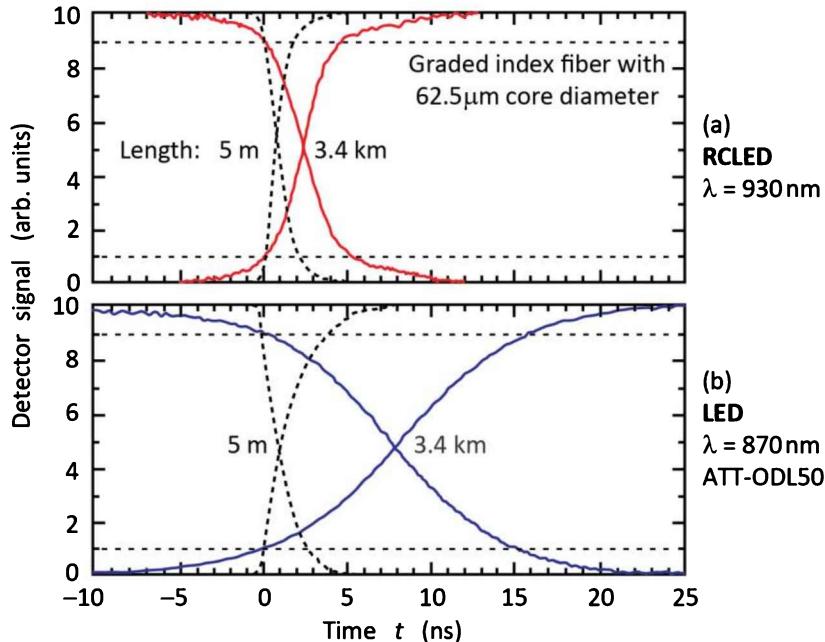


Fig. 17.8: Signal detected at the receiver end of a graded-index multimode fiber with a core diameter of 62.5 μm using (a) a GaInAs RCLED and (b) a GaAs LED source. After a short transmission distance of 5 m, no marked difference is found for the two sources. After a transmission distance of 3.4 km, the RCLED exhibits much less pulse broadening than the LED (after Hunt *et al.*, 1993).

Schubert *et al.* (1996) demonstrated the high-speed modulation capability of RCLEDs. Eye diagram measurements with a random bit pattern generator revealed wide-open eyes at 622 Mbit/s. Due to the small size of the current injected region, the parasitic capacitances of communication RCLEDs are small. It is expected that RCLEDs will be suitable for modulation frequencies exceeding 1 Gbit/s.

17.4 AlGaInP/GaAs RCLEDs emitting at 650 nm

RCLEDs have also been demonstrated in the visible wavelength range using the AlGaInP material system (Streubel *et al.*, 1998; Whitaker, 1999; Wirth *et al.*, 2001, 2002). The AlGaInP material system is commonly used for high-brightness red, orange, and yellow emitters and can be grown lattice matched on GaAs substrates. The active region of RCLEDs is an AlGaInP/GaInP multiple-quantum well structure emitting at 650 nm. The RCLEDs are suited for use in communication systems using plastic optical fibers. It is difficult to fabricate VCSELs in this wavelength range due to the unavailability of lattice-matched and transparent DBR materials with a large refractive-index contrast.

The basic structure of a top-emitting AlGaNp RCLED ($\lambda = 650$ nm) is shown in **Figure 17.9**. The device consists of an AlGaNp/GaInP MQW active region and AlGaNp cladding layers. The DBRs consist of AlAs/AlGaAs layers. The Al content of the DBR's AlGaAs layers is chosen to be sufficiently high to make the DBR transparent to the emitted light. As a result, the index contrast of the AlAs/AlGaAs DBR layers is somewhat low.

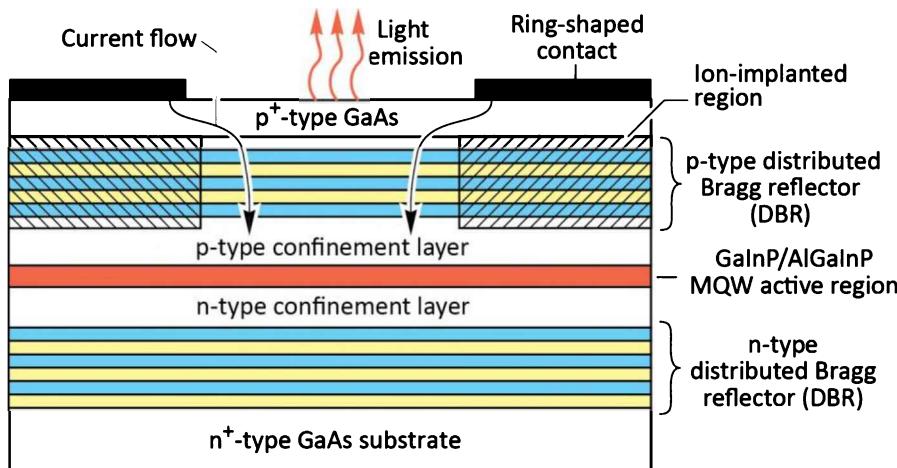


Fig. 17.9: Structure of a GaInP/AlGaNp/GaAs MQW RCLED emitting at 650 nm used for plastic optical fiber applications (after Whittaker, 1999)

The RCLED has a ring-shaped top contact configuration. The current is funneled into the center region of the ring using ion implantation to create an insulating region under the ring-shaped metallization. Hydrogen and, more frequently, oxygen implantation is used to render the semiconductor highly resistive. Note that the ion-implanted region is located in the p-type region only and does not extend into the active region, thereby avoiding the creation of defects in the active region where they would act as luminescence killers.

Packaged RCLEDs using a lensed TO package and a pig-tailed package are shown in **Figures 17.10 (a) and (b)**, respectively (Mitel Corporation, 1999). The lens is used for beam collimation, thereby enhancing the coupling efficiency to optical fibers.

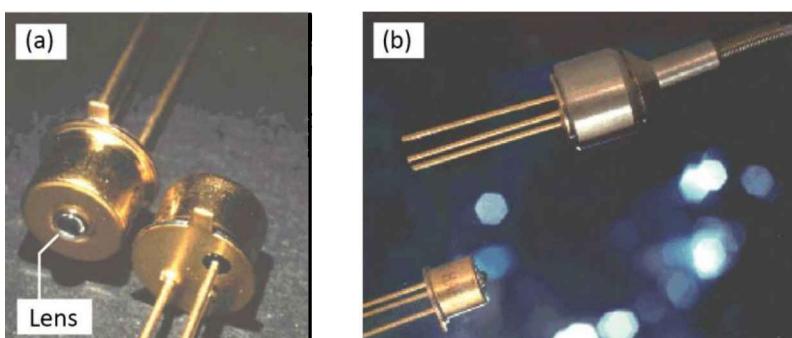


Fig. 17.10: (a) Packaged (TO package) RCLED emitting at 650 nm suited for plastic optical fiber applications. (b) Pig-tailed RCLED (courtesy of Mitel Corporation, Sweden, 1999).

Three RCLEDs under current injection conditions are shown in *Figure 17.11* (Osram Opto Semiconductors Corporation). The picture shows that the emission pattern is directed towards the surface normal of the devices. The emission wavelength is 650 nm.

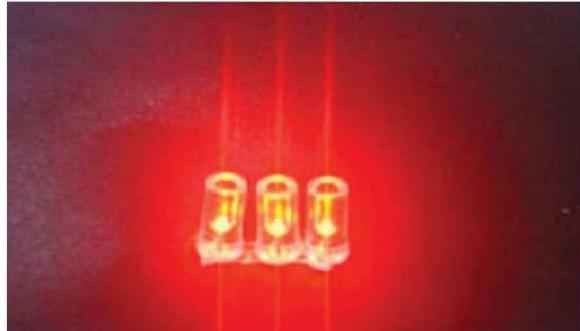


Fig. 17.11: AlGaInP/ GaAs RCLEDs emitting at 650 nm. Note the forward-directed emission pattern similar to that of a semiconductor laser (courtesy of Osram Opto Semiconductors Company, Regensburg, Germany, 1999).

Optical spectra of a 650 nm RCLED and of a conventional LED injected at different current levels are shown in *Figure 17.12*. The spectra are measured after the light is coupled into a plastic optical fiber. Thus, the magnitude of the spectra is a direct measure of the device efficiency *and* of the coupling efficiency. Inspection of the figure reveals several features. Firstly, the RCLED has a higher coupled peak power as well as integrated power than the LED. Secondly, the RCLED has a higher spectral purity than the LED.

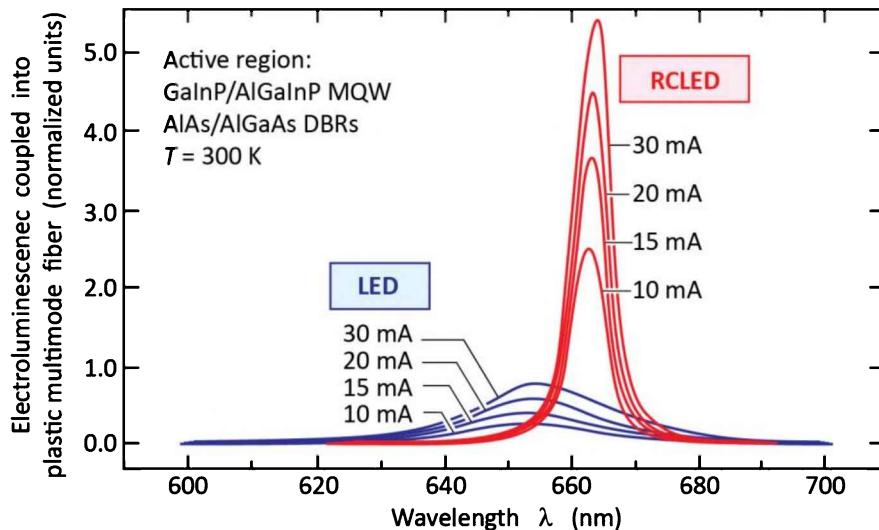


Fig. 17.12: Spectra of light coupled into a plastic optical fiber from a GaInP/AlGaInP MQW RCLED and a conventional GaInP/AlGaInP LED at different drive currents. Note the narrower spectrum and higher coupled power of the RCLED (after Streubel *et al.*, 1998).

Streubel *et al.* (1998) also reported that the emission spectrum was, at room temperature, intentionally blue-shifted with respect to the cavity resonance in order to improve the temperature stability of the RCLED output power. This cavity tuning causes a heart-shaped

(double-lobed) emission pattern at room temperature, because the resonance wavelength of the cavity decreases for off-normal emission directions. However, as the temperature increases, the natural emission spectrum from the active region shifts towards the red, so that the cavity resonance (along the normal direction) has a better overlap with the natural emission spectrum. As a result, the temperature sensitivity of the RCLED is reduced. The red-shift of the natural emission spectrum is about $0.5 \text{ nm}/^\circ\text{C}$, whereas the cavity resonance shifts at only about one-tenth of that rate.

17.5 Newer developments in RCLED devices

Resonant-cavity LEDs and generally, confined-photon emitters, increasingly involve new materials and applications. Spontaneously emitting confined-photon devices have competition from lasers, conventional LEDs, and other forms of light-emitting devices. It is only a matter of time, however, before the right materials combination at the right wavelengths, for the right application, will make commercial devices a reality. It is therefore worthwhile to mention other material systems used for confined-photon emitters.

Wilkinson *et al.* (1995) fabricated an AlGaAs/GaAs thin-film emitter with metal mirrors on a Si substrate, emitting at 880 nm. Pavesi *et al.* (1996) fabricated porous silicon RCLEDs at 750 nm wavelength. Fisher *et al.*, (1995) investigated a conjugate polymer RCLED designed for 650 nm. Hadji *et al.* (1995) have realized CdHgTe/HgTe RCLEDs operating at a wavelength of $3.2 \mu\text{m}$.

A structure of particular note is a GaAs/Al_xO_y RCLED (Huffaker *et al.*, 1995) operating at 950 nm, where the aluminum oxide in the output mirror was produced from AlAs by oxidation. This composition of output mirror allows the effective cavity length to remain small, maximizing the output enhancement in the RCLED. The back mirror of the structure was Ag.

Another particularly interesting device is the broadly tunable RCLED by Larson and Harris (1995). The top mirror is a deformable membrane which can be moved by electrostatic forces. Tunable emission was shown to range from 938 to 970 nm.

17.6 References

- De Neve H., Blondelle J., Baets R., Demeester P., Van Daele P., and Borghs G. "High efficiency planar microcavity LEDs: Comparison of design and experiment" *IEEE Photonics Technology Letters* **7**, 287 (1995)
- Fisher T. A., Lidzey D. G., Pate M. A., Weaver M. S., Whittaker D. M., Skolnick M. S., and Bradley D. D. C. "Electroluminescence from a conjugated polymer microcavity structure" *Applied Physics Letters* **67**,

- 1355 (1995)
- Hadji E., Bleuse J., Magnea N., and Pautrat J. L. "3.2 μm infrared resonant cavity light emitting diode" *Applied Physics Letters* **67**, 2591 (1995)
- Huffaker D. L., Lin C. C., Shin J., and Deppe D. G. "Resonant cavity light emitting diode with an $\text{Al}_x\text{O}_y/\text{GaAs}$ reflector" *Applied Physics Letters* **66**, 3096 (1995)
- Hunt N. E. J., Schubert E. F., Logan R. A., and Zyzdrik G. J. "Enhanced spectral power density and reduced linewidth at 1.3 μm in an InGaAsP quantum well resonant-cavity light-emitting diode" *Applied Physics Letters* **61**, 2287 (1992)
- Hunt N. E. J., Schubert E. F., Kopf R. F., Sivco D. L., Cho A. Y., and Zyzdrik G. J. "Increased fiber communications bandwidth from a resonant cavity light-emitting diode emitting at $\lambda = 940 \text{ nm}$ " *Applied Physics Letters* **63**, 2600 (1993)
- Hunt N. E. J., Schubert E. F., Sivco D. L., Cho A. Y., Kopf R. F., Logan R. A., and Zyzdrik G. J. "High efficiency, narrow spectrum resonant cavity light-emitting diodes" in *Confined Electrons and Photons* edited by E. Burstein and C. Weisbuch (Plenum Press, New York, 1995a)
- Hunt N. E. J., Vredenberg A. M., Schubert E. F., Becker P. C., Jacobson D. C., Poate J. M., and Zyzdrik G. J. "Spontaneous emission control of Er^{3+} in Si/SiO₂ microcavities" in *Confined Electrons and Photons* edited by E. Burstein and C. Weisbuch (Plenum Press, New York, 1995b)
- Kobayashi T., Segawa T., Morimoto A., and Sueta T., paper presented at the 43rd fall meeting of the Japanese Society of Applied Physics, Tokyo, Sept. (1982)
- Larson M. C. and Harris Jr. J. S. "Broadly tunable resonant-cavity light emission" *Applied Physics Letters* **67**, 590 (1995)
- Lear K. L. and Schneider Jr. R. P. "Uniparabolic mirror grading for vertical cavity surface emitting lasers" *Applied Physics Letters* **68**, 605 (1996)
- Mitel Corporation, Sweden. Photograph of RCLED is gratefully acknowledged (1999)
- Nakayama T., Itoh Y., and Kakuta A. "Organic photo- and electroluminescent devices with double mirrors" *Applied Physics Letters* **63**, 594 (1993)
- Osram Opto Semiconductors Corp., Germany. RCLED photograph is gratefully acknowledged (1999)
- Pavesi L., Guardini R., and Mazzoleni C. "Porous silicon resonant cavity light emitting diodes" *Solid State Communications* **97**, 1051 (1996)
- Schubert E. F., Wang Y.-H., Cho A. Y., Tu L.-W., and Zyzdrik G. J. "Resonant cavity light-emitting diode" *Applied Physics Letters* **60**, 921 (1992a)
- Schubert E. F., Vredenberg A. M., Hunt N. E. J., Wong Y. H., Becker P. C., Poate J. M., Jacobson D. C., Feldman L. C., and Zyzdrik G. J. "Giant enhancement of luminescence intensity in Er-doped Si/SiO₂ resonant cavities" *Applied Physics Letters* **61**, 1381 (1992b)
- Schubert E. F., Tu L. W., Zyzdrik G. J., Kopf R. F., Benvenuti A., and Pinto M. R. "Elimination of heterojunction band discontinuities by modulation doping" *Applied Physics Letters* **60**, 466 (1992c)
- Schubert E. F., Hunt N. E. J., Micovic M., Malik R. J., Sivco D. L., Cho A. Y., and Zyzdrik G. J. "Highly efficient light-emitting diodes with microcavities" *Science* **265**, 943 (1994)
- Schubert E. F., Hunt N. E. J., Malik R. J., Micovic M., and Miller D. L. "Temperature and modulation characteristics of resonant cavity light-emitting diodes" *IEEE Journal of Lightwave Technologies* **14**, 1721 (1996)
- Streubel K., Helin U., Oskarsson V., Backlin E., and Johanson A. "High-brightness visible (660 nm) resonant-cavity light-emitting diode" *IEEE Photonics Technology Letters* **10**, 1685 (1998)
- Tu L. W., Schubert E. F., Zyzdrik G. J., Kopf R. F., Hong M., Chu S. N. G., and Mannaerts J. P. "Vertical cavity surface emitting lasers with semitransparent metallic mirrors and high quantum efficiencies" *Applied Physics Letters* **57**, 2045 (1990)
- Whitaker T. "Resonant cavity LEDs" *Compound Semiconductors* **5**, 32 (1999)
- Wilkinson S. T., Jokerst N. M., and Leavitt R. P. "Resonant-cavity-enhanced thin-film AlGaAs/GaAs/AlGaAs LED's with metal mirrors" *Applied Optics* **34**, 8298 (1995)
- Wirth R., Karnutsch C., Kugler S., and Streubel K. "High-efficiency resonant-cavity LEDs emitting at 650 nm" *IEEE Photonics Technology Letters* **13**, 421 (2001)

- Wirth R., Huber W., Karnutsch C., and Streubel K. “Resonators provide LEDs with laser-like performance”
Compound Semiconductors **8**, 49 (2002)
- Yokoyama H. “Physics and device applications of optical microcavities” *Science* **256**, 66 (1992)

18

18 – Photonic crystal light-emitting diodes

Due to the phenomenon of total internal reflection, there is a fundamental difficulty in transmitting light from within thin semiconductor film with a high refractive index to a surrounding material with a low refractive index. Several strategies have been devised to overcome this difficulty. This chapter reviews several approaches to attain efficient light extraction out of a thin film, particularly the photonic crystal approach for LEDs.

18.1 The photon recycling concept for LEDs

Light-emission events occurring in a thin semiconductor film can be categorized according to two categories: (i) emission into waveguided modes (in the plane of the semiconductor film), and (ii) emission into the escape cone (perpendicular to the plane of the semiconductor). Light emitted into waveguided modes is confined to the waveguide and thus cannot be easily extracted into free space. Therefore, one possible strategy to increase light extraction would be to reduce lateral emission.

However, instead of devising ways to *reduce lateral emission*, one could devise ways to *recycle* lateral emission, thereby redirecting the emission towards the top of the device. By reabsorbing the lateral emission and creation of an electron–hole pair, one can recapture the energy into the active region, providing another chance to emit along the desired, perpendicular direction. Below, two examples of such photon-recycling devices are discussed.

An example of a photon-recycling device is the optically pumped semiconductor structure reported by Schnitzer *et al.* (1993). Consider the structure shown in **Figure 18.1 (a)**. The backside of a thin semiconductor layer is coated with a metal reflector, e.g. silver or gold. Optical excitation of the sample at low intensities, will result in a low carrier concentration, so that the active region remains mostly absorbing. Most of the light that is incident on the semiconductor–air interface will totally internally reflect, and stay within the semiconductor. Gold (Au) is a good reflector at infrared wavelengths. Silver (Ag) is a good reflector at visible and ultraviolet wavelengths. The semiconductor forming the active region is absorbing at the emission wavelength, so after some average absorption length, L_{abs} , the trapped light has a chance to re-emit.

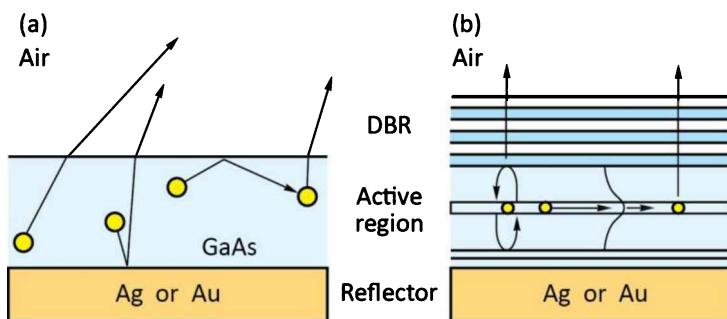


Fig. 18.1: Photon recycling LEDs. (a) Bulk epilayer on top of Ag reflector. Spontaneous emission that does not escape into air is reabsorbed and can re-emit. (b) Microcavity designed with a waveguiding active region. Waveguided light is reabsorbed after some tens of μm , and can re-emit out of the top of the device.

It was reported that such a structure can have a 72% external efficiency, which required 99.7% internal quantum efficiency (Schnitzer *et al.*, 1993). The fabrication of the structure required the active epilayer to be etched and floated off its substrate. Note, however, that the structure has no electrical contacts; the extraction efficiency of contactless structures is always higher compared with current-injection devices with contacts. An electrical device fabricated in this manner may have reliability challenges due to the fragile nature of the thin epitaxial film. A current-injection device using photon recycling was reported by Blondelle *et al.* (1995) and resulted in 16% external quantum efficiency.

The concept of photon recycling was also used by De Neve *et al.* (1995) in a GaInAs/GaAs/AlGaAs RCLED. A simplified diagram of their structure is shown in **Figure 18.1 (b)**. If the graded-composition carrier confinement region is thick enough around the quantum wells, it also acts as a graded-index waveguide. De Neve *et al.* calculated that 30% of the light emitted by the active region goes into this waveguide mode. The light is reabsorbed after some tens of μm , allowing the photons another chance to re-emit out of the top of the structure. In this way, about a 30% increase in external quantum efficiency was attained. The active region of quantum well RCLEDs is usually not thick enough to support a strongly guided optical mode. One might think that a strong waveguiding mode would take power away from the normal emission, but this does not appear to be the case. The energy tends to be at the expense of other high-angle modes instead, meaning that just modifying the waveguiding mode does little to change the external efficiency. Making use of laterally emitted light by employment of photon recycling of waveguided emission is an attractive option.

A first drawback to the photon recycling concept is that, in order to be effective, very high internal quantum efficiency is needed. That is, every photon recycling event reduces the efficiency since the internal quantum efficiency is less than 100%. A second drawback to photon recycling is that it requires a device with large enough diameter so that multiple reabsorption events can occur

within the emitting area of the device. This makes such devices less attractive for fiber-optic communications, where small diameters couple better to fibers, especially when coupling lenses are used. A third drawback is that self-absorption necessarily increases the lifetime of the spontaneous emission, thereby slowing down the maximum modulation rate of the devices.

The devices reported by De Neve *et al.* (1995) were shown to work with highest efficiency for large device diameters. At high current densities, the carrier confinement in the quantum wells is reduced, and the efficiency decreases, so larger devices tend to be more efficient. Also, their device was resonant to the long-wavelength side of the natural emission peak, rather than at the peak. This shifts the maximum intensity of the main emission lobe to an off-normal angle, rather than being on axis, but it maximizes the total amount of emission into the main lobe. The devices reported by Blondelle *et al.* (1995) and De Neve *et al.* (1995) are therefore designed for maximizing the total emission from the top of the device, rather than for the specific needs of fiber coupling. Display devices and free-space communication devices benefit from this approach. The authors achieved an external quantum efficiency of 16%, compared with the theoretical 2% for an ideal planar emitter of the same refractive index.

18.2 Thresholdless laser or single-mode LED

There have been a number of reports describing a ***thresholdless laser***, also called ***single-mode LED***, as a possible high efficiency light-emitting device (see, for example, De Martini *et al.*, 1987). A thresholdless laser is a light-emitting device that emits most or all of its light into a single spatial optical mode. Consider an optical object, such as a three-dimensional pillar, with spatial dimensions on the order of λ . The pillar would support a longitudinal optical mode but would have such a small diameter that no lateral mode could exist. When reducing the spatial dimensions of the pillar, the number of optical modes that are supported by the pillar, decreases. For sufficiently small dimensions, only a single optical mode may be supported by the pillar. As a consequence, all light will be emitted into this single mode and such a device may be called a single-mode LED.

Let us denote the probability that light is emitted into the fundamental spatial cavity mode as β . If the β of the LED were almost 1, then the light-versus-current curve of the device would be linear, and *indistinguishable* from a laser with no threshold. A comparison of the light-versus-current for a conventional laser, a high- β laser, and a truly thresholdless laser is given in **Figure 18.2**.

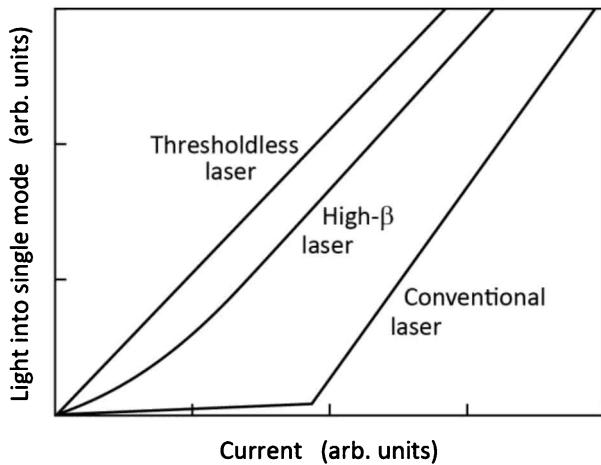


Fig. 18.2: Light-power-versus-current curves for single spatial-mode emission from (i) a conventional laser, (ii) a high β -factor laser, and (iii) a thresholdless laser. The conventional laser has a distinct current threshold. The high β -factor laser has a less distinct threshold. It would be noticeable in the spectrum and device modulation speed, however. A hypothetical thresholdless laser would have a β close to 1, and would somehow suppress all other lossy emission until the carrier density required for gain (or at least transparency) was achieved.

One may envision a thresholdless device to be fabricated from a resonant-cavity pillar. However, there could be a number of problems besides those of fabrication. Simply having a good sub-threshold intensity, giving only a small kink in the L – I curve between sub-threshold and lasing, is not good enough for a thresholdless laser. The output modulation of such a device would be much slower below threshold compared to the lasing regime. To enable lasing, there must be some gain, because the β will never be 1. For there to be gain, we must first pump the active region past transparency, since it is absorbing under no-pump conditions. This gain condition requires a certain carrier density, which, in fact, determines the threshold.

One may also envision a device such that all optical modes other than the fundamental are highly suppressed, and the internal quantum efficiency is nearly perfect. In this case, with even a small current, the carrier densities within the device would slowly build until the active region becomes transparent, and some light could get out through the fundamental optical mode. Such a device would have a very low threshold, but one would have to be careful not to ever drive the device below its threshold even momentarily, because of the long time required to build the necessary space charge inside the device for transparency. Also, if the fundamental emission is suppressed at low carrier densities, it is unlikely that the other modes would be suppressed by an even higher degree, which is what would be required not to lose the injected carriers. In fact, it is the difficulty in suppressing any side emission that makes thresholdless lasers difficult if not impossible to implement.

One could, of course, make a laser with a very small threshold by making the laser very small. The problem is that such a device would be capable of only a very small output power. For a hypothetical four-quantum-well single-mode LED, which does not rely on gain, but achieves a β

close to 1, the current densities cannot be more than about $50 \mu\text{A}/\mu\text{m}^2$ or $5 \text{kA}/\text{cm}^2$. Using the formulas for β for a dielectric pillar ($\beta > 0.5$), and this current density, one finds that less than $2 \mu\text{W}$ would be emitted into the single mode for the pillar structure. This clearly is too small for many applications, including high-speed communications in the $100\text{--}650 \text{ Mbit/s}$ regime, where $10 \mu\text{W}$ is considered a minimum for an emitter.

Another confined-photon emitter is the ***microdisk laser*** (McCall *et al.*, 1992), which is fabricated as a thin dielectric disk that couples light out the edges of the disk. Lasing modes can be described by a mode number M , where $\exp(iM\phi)$ is the form of the electric field around the cylindrical disk. Because waves can propagate both ways, M can be positive or negative. The disk can be fabricated with a thickness such that the emission perpendicular to the disk is suppressed. Small disks will only support a few modes, and therefore can have a high spontaneous emission factor β . The Q of these modes are also high enough to achieve lasing. One attractive aspect of such disks is that the lasing emission occurs in the plane of the sample, from a very small device. This could be useful for integration of many photonic devices on a single wafer. However, it is difficult to efficiently couple the output into waveguides and fibers, as it only couples evanescently. Advances have been made in improving the longevity, operating temperature range, and active-region passivation of such devices (Mohideen *et al.*, 1993). Room-temperature cw electrical pumping is still not easily attained, however.

To summarize, semiconductor RCLEDs are suited for devices with diameters larger than about $5\text{--}10 \mu\text{m}$, and should remain multimode emitters. Semiconductor vertical-cavity lasers are best fabricated at sizes that make sense technologically, without regard to the β factor of the spontaneous emission. A high- β laser may not be desirable for some traditional laser applications, since the spontaneous emission will introduce excess noise into the emission.

18.3 Photonic crystals

Photonic crystals or ***photonic bandgap structures*** are two- and three-dimensional structures that have a periodic modulation of the refractive index (John, 1987; Yablonovitch, 1987; Joannopoulos *et al.*, 1995, 1997). As a result of the modulation, the optical mode density in the structures is changed which influences a wide range of properties including optical emission, absorption, waveguiding, and scattering.

The name “photonic crystal” was chosen having in mind the analogy to “atomic crystals”. Atomic crystals strongly influence the properties of electrons propagating in a crystal. The well-known Kronig-Penney model teaches that electrons, propagating in the periodic potential of an atomic crystal, are subject to restrictions such as certain ranges of allowed energies (allowed bands such as the valence and conduction band) and certain ranges of forbidden energies (disallowed bandgaps). Electrons propagating in a periodic potential undergo Bragg reflection, if the electron’s de-Broglie wavelength is resonant with a periodicity of the crystal.

Similarly, light propagating in the periodic refractive-index profile of a photonic crystal, is subject to restrictions such as certain ranges of allowed energies (allowed photonic bands) and certain ranges of forbidden energies (disallowed photonic bandgaps). Light propagating in a periodic refractive-index profile undergo Bragg reflection, if the wavelength is resonant with a periodicity of the photonic crystal.

This physical similarity between electronic crystals and photonic crystals is caused by the mathematical similarity of the Schrödinger equation (which controls the propagation of electrons) and an electrodynamics equation derived from Maxwell’s equations (which controls the propagation of photons). Joannopoulos *et al.* (2008) listed the similarities between these two equations as shown in the **Table 18.1**.

Table 18.1: Similarity between mathematical equations governing quantum mechanics and equations governing electrodynamics (after Joannopoulos *et al.*, 2008).

	<i>Quantum mechanics</i>	<i>Electrodynamics</i>
Field	$\Psi(\vec{r}, t) = \Psi(\vec{r}) e^{-iEt/\hbar}$	$\vec{H}(\vec{r}, t) = \vec{H}(\vec{r}) e^{-i\omega t}$
Eigenvalue problem	$H \Psi = E \Psi$	$\Theta \vec{H} = (\omega/c)^2 \vec{H}$
Hermitian operator	$H = -\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r})$	$\Theta = \nabla \times \frac{1}{\epsilon(\vec{r})} \nabla \times$

Photonic crystals were shown to be of great interest for light-emitting diodes. Fan *et al.* (1997) considered a light-emitting semiconductor thin-film structure with an integrated photonic crystal. The periodicity of the crystal was chosen in such a way that the thin film does not allow for the emission of light being waveguided by the thin film. As a consequence, light can only be emitted along the film-normal direction thereby enabling high extraction efficiency. Boroditsky *et al.* (1999) proposed a thin-film LED having a current-injected central region and a photonic-crystal region

surrounding the central region. The waveguided light emitted by the central region enters the surrounding photonic-crystal region which scatters the light into free space.

Photonic crystal structures commonly consist of a series of rods or holes arranged in a regular pattern, such as a triangular pattern (i.e. a hexagonal close-packed pattern). The periodicity of the pattern can create an optical bandgap for lateral emission at certain emission energies and one or both polarizations (Burstein and Weisbuch, 1995).

Wierer *et al.* (2009) reviewed the photonic crystal concept for LEDs and the following paragraphs follow their discussion. A first approach for a photonic crystal LED is the so-called bandgap approach (Fan *et al.*, 1997); the photonic crystal penetrates through the entire device and is tuned near a photonic bandgap, inhibiting light emission into guided modes and hence increasing extraction efficiency. Although surface recombination is low in III–V nitrides compared to other III–V compounds, it is still difficult to achieve a working device in which the photonic crystal penetrates through the active region. Besides, obtaining an omni-directional bandgap is challenging due to the relatively low refractive index of GaN and the thickness of typical epitaxial layers, and so far bandgaps have only been demonstrated in membrane geometries (Wierer *et al.*, 2009).

Most scientific efforts have therefore focused on another approach: using photonic crystals as diffraction gratings. In this scheme, spontaneous emission into guided modes is allowed but these modes are subsequently out-coupled by the diffractive properties of the photonic crystal. The lattice constant of a diffractive photonic crystal is larger than in the bandgap approach – that is, the photonic crystal operates near a high Bragg order rather than at the first Bragg order. Several groups have demonstrated III–V nitride LEDs with photonic-crystal diffractive layers on the surfaces of the LEDs, showing an improvement in light extraction and in the control of the radiation pattern (Wierer *et al.*, 2009).

In the diffraction grating approach, controlling and extracting all the guided modes is a difficult task. Each guided mode propagates in the GaN at a different angle and therefore interacts with the photonic crystal differently. Having *all* the modes diffract is non-trivial and requires proper optimization for high extraction efficiency (Wierer *et al.*, 2009).

Figure 18.3 (a) shows the typical emission pattern within a flip-chip LED with GaInN / GaN multi-quantum wells (MQWs). The GaN slab is bonded to a reflector, with the MQWs placed close (approximately one wavelength) to the reflector. Inspection of the figure reveals that the emission pattern comprises two angular regions. The first region is light emitted within the light-escape cone

given by the angle

$$\phi_c = \arcsin \frac{\bar{n}_{air}}{\bar{n}_{GaN}} = 24.6^\circ. \quad (18.1)$$

where \bar{n}_{air} and \bar{n}_{GaN} are the refractive indices of air and GaN, respectively. The pattern in the light-escape cone displays broad-shaped lobes whose peaks correspond to the Fabry–Perot resonances between the GaN–reflector and GaN–air interfaces. These modes will appear as broad resonances in the far-field radiation patterns typical of resonant-cavity LEDs. Outside the extraction cone, light is emitted in a set of sharp guided modes that propagate within the GaN slab. The number of these modes increases with slab thickness and their sharpness (finesse) is a function of the material losses and reflectivity of the interfaces. For a planar LED these guided modes are likely absorbed or only collected on the edges of the LED, thereby limiting the light-extraction efficiency of LEDs that take no measures in out-coupling the waveguided modes (Wierer *et al.*, 2009).

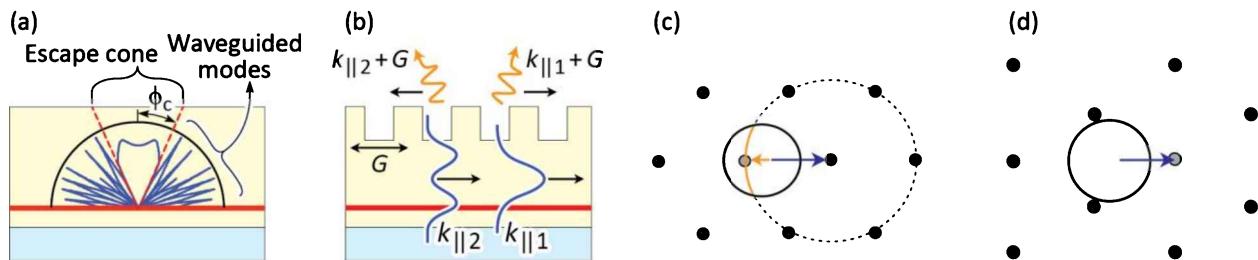


Fig. 18.3: Light emission and extraction. (a) Internal emission pattern of a GaInN quantum well in a planar GaN slab with a reflective silver mirror. The emission pattern represents the emission intensity versus polar angle. (b) Guided modes (TE1 and TE2) in the same structure with a photonic-crystal layer. Each guided mode possesses a wavevector $k_{\parallel m}$ (blue line) that can be coupled to a leaky harmonic $k_{\parallel m} + G$ (yellow line). (c, d) Ewald-shpere constructions for a mode propagating along the Γ -M (c) and Γ -K (d) directions of a photonic crystal. In (c) the arc of the circle is the locus of the diffracted harmonic when the azimuthal angle of the mode varies (after Wierer *et al.*, 2009).

Figure 18.3 (b) schematically illustrates how a photonic-crystal layer diffracts wave-guided modes into air, thereby increasing the light-extraction efficiency. Each of the guided modes can be characterized by an in-plane wavevector ($k_{\parallel m}$), related to the angle of propagation within the GaN (ϕ_m) by

$$k_{\parallel m} = \frac{2\pi}{\lambda} \bar{n}_{GaN} \sin \phi_m \approx \sqrt{\left(\frac{2\pi}{\lambda} \bar{n}_{GaN}\right)^2 - \left(\frac{m \pi}{d_{GaN}}\right)^2}. \quad (18.2)$$

where m is the mode number ($1, 2, 3 \dots$), λ is the wavelength of light, and d_{GaN} is the thickness of GaN (Wierer *et al.*, 2009).

In the presence of the photonic crystal, the guided modes become Bloch modes and the wavevectors $k_{\parallel m}$ are now coupled to other harmonics $\{k_{\parallel m} + G\}$ by the reciprocal lattice vectors G . To diffract the guided light into air, the lattice constant a of the photonic crystal needs to satisfy the diffraction condition

$$| k_{\parallel m} + p G_0 | < \frac{2\pi}{\lambda} \quad (18.3)$$

where $G_0 = 2\pi/a$ and p is an integer (determining which harmonic is responsible for diffraction to air). The Bloch mode is then referred to as a leaky mode, because its power leaks to air as it propagates. In air, the radiated harmonic behaves like a plane wave with an angle

$$\phi_{m,air} = \arcsin\left(\frac{\lambda}{2\pi} | k_{\parallel m} + p G_0 | \right) \quad (18.4)$$

The diffraction condition depends on the wavelength, lattice constant and mode propagation angle. In practice, an LED supports a few to several tens of modes (depending on the GaN thickness) and a wavelength range $\Delta\lambda \approx 20\text{--}50$ nm. For high extraction efficiency, it is necessary to extract guided light propagating in all azimuthal directions. Therefore, a 2D photonic crystal is required. **Figure 18.3 (c)** and (d) show a top view of a triangular photonic crystal in reciprocal space (Ewald construction) for the same mode propagating along two azimuthal directions. The dots represent the reciprocal lattice points of the photonic crystal with the lattice rotated 90° between **Figure 18.3 (c)** and (d), and the full circle describes the air escape cone. For a mode (blue arrow) propagating along the ΓM direction in **Figure 18.3 (c)** the diffraction condition is met: A harmonic falls in the air cone (orange arrow) and the mode becomes leaky. The mode along the ΓK direction, shown in **Figure 18.3 (d)**, remains guided because no reciprocal lattice point lies within the air cone. In general, when the azimuthal angle of the guided mode is varied, reciprocal lattice points describe a circular locus (dashed circle of **Figure 18.3 (c)**). The part of this locus that is inside the air cone can be diffracted, resulting in the arc of the circle shown in **Figure 18.3 (c)**. Thus, the type and lattice constant of the photonic crystal need to be chosen properly, ideally to ensure the extraction of all optical modes at all angles (Wierer *et al.*, 2009).

The above considerations are geometrical in nature and are intended to ensure that the diffraction condition is met for as many modes and angles as possible. Another issue worthy of consideration is the *efficiency* of the light-extraction process enabled by photonic crystals. The multimode nature of GaN layers makes this question non-trivial: the vertical structure has to be engineered to allow good coupling of leaky modes with the photonic crystal. In general, high-order modes are better diffracted than low-order modes because of a higher overlap with the photonic crystal. In addition to the vertical design, the light-extraction efficiency can also be influenced by other parameters such as the depth and size of holes forming the photonic crystal (Wierer *et al.*, 2009).

18.4 LEDs using photonic crystals

The first publication of an optically excited photonic crystal LED-like structure is due to Erchak *et al.* (2001) who used a two-dimensional photonic crystal that lied entirely inside the upper cladding layer of an asymmetric quantum well structure. The LED-like structure, shown in **Figure 18.4 (a)**, included a lower and an upper GaInP cladding layer, and a GaInAs quantum well layer. **Figure 18.4 (b)** shows the side and top view of the electric field intensity at the photonic-crystal resonance wavelength. Erchak *et al.* (2001) obtained very encouraging results for the photonic crystal LED-like structure, namely a six-fold enhancement of light extraction along the surface-normal direction.

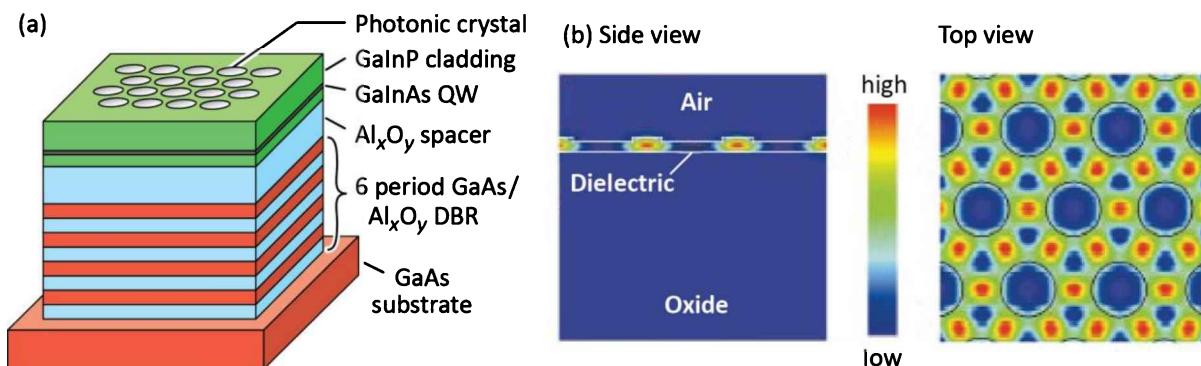


Fig. 18.4: (a) Structure of 2D photonic crystal LED (PC LED) with a triangular PC implemented in the GaInAs/GaInP material system. (b) Side view and top view of electric field intensity at resonance (790 nm) for the PC LED structure (after Erchak *et al.*, 2001).

Subsequent to the report by Erchak *et al.* (2001), the field of photonic crystal LEDs generated much interest and encouraging results were reported by numerous groups including Ichikawa and Baba

(2004), Oder *et al.* (2004), Wierer *et al.* (2004), Kim *et al.* (2005), David *et al.* (2006), Kwon *et al.* (2008), and Wierer *et al.* (2009).

A schematic of a GaN based thin-film photonic crystal LED is shown in **Figure 18.5** (Wierer *et al.*, 2009). The III–V nitride epitaxial structure comprised an InGaN / GaN MQW heterostructure grown on (0001) sapphire. It contained two GaInN QWs in which the indium (In) composition was tuned for an emitting peak wavelength of 450 nm. Silver-based metallization was used as a p-contact and reflector. The sapphire substrate was removed by a laser lift-off process and the GaN layer was thinned to a thickness of 500~1 000 nm. Planar LEDs were left smooth, and patterning and dry etching were used to form the photonic-crystal layer in the n-type GaN for the photonic crystal LEDs. Various lattice constants (200~1 000 nm), photonic crystal types (triangular, honeycomb, Archimedean A7 and A13) and depths (100~300 nm) were investigated, with a filling factor of about 0.3. The lattice constants started at the onset of diffraction to air and spanned several Bragg orders. The various lattice types were used to explore the issue of in-plane variation of the extraction and omni-directional extraction. The thickness of the film was chosen to ensure good overlap of the modes with the photonic crystal. A cross-section of the device is shown in **Figure 18.5** (a). **Figure 18.5** (b) shows a top view of a lit photonic-crystal LED. The n-type contact was patterned into a grid, and the total LED area was about $200 \times 200 \mu\text{m}^2$. Within the parameter range investigated, the best photonic-crystal LED has a very high light-extraction efficiency (73%) without using encapsulants, demonstrating a significant step forward in high-performance LEDs.

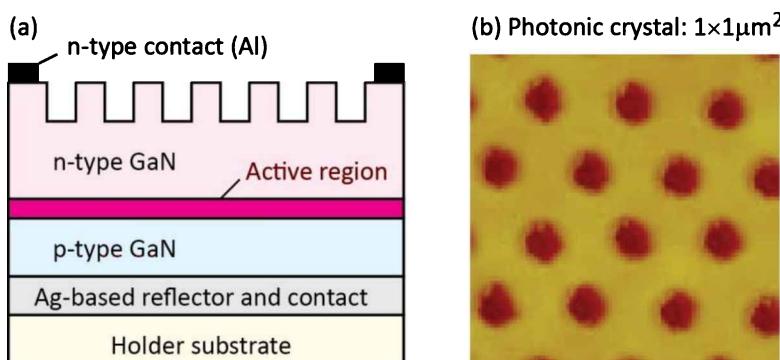


Fig. 18.5: (a) Structure of GaN-based photonic crystal LED. (b) Micrograph of photonic crystal as seen from the top view of the LED (after Wierer *et al.*, 2009).

18.5 References

Blondelle J., De Neve H., Demeester P., Van Daele P., Borghs G., and Baets R. "16% external quantum efficiency from planar microcavity LED's at 940 nm by precise matching of the cavity wavelength" *Electronics Letters* **31**, 1286 (1995)

- Boroditsky M., Krauss T. F., Cocciali R., Vrijen R., Bhat R., and Yablonovitch E. "Light extraction from optically pumped light-emitting diode by thin-slab photonic crystals" *Applied Physics Letters* **75**, 1036 (1999)
- Burstein E. and Weisbuch C. "Confined electrons and photons" (Plenum Press, New York, 1995)
- David A., Fujii T., Moran B., Nakamura S., DenBaars S. P., Weisbuch C., and Benisty H. "Photonic crystal laser lift-off GaN light-emitting diodes" *Applied Physics Letters* **88**, 133514 (2006)
- De Martini F., Innocenti G., Jacobovitz G. R., and Mataloni P. "Anomalous spontaneous emission time in a microscopic optical cavity" *Physical Review Letters* **59**, 2955 (1987)
- De Neve H., Blondelle J., Baets R., Demeester P., Van Daele P., and Borghs G. "High efficiency planar microcavity LEDs: Comparison of design and experiment" *IEEE Photonics Technology Letters* **7**, 287 (1995)
- Erchak A. A., Ripin D. J., Fan S., Rakich P., Joannopoulos J. D., Ippen E. P., Petrich G. S., and Kolodziejski L. A. "Enhanced coupling to vertical radiation using a two-dimensional photonic crystal in a semiconductor light-emitting diode" *Applied Physics Letters* **78**, 563 (2001)
- Fan S., Villeneuve P. R., Joannopoulos J. D., and Schubert E. F. "High extraction efficiency of spontaneous emission from slabs of photonic crystals" *Physical Review Letters* **78**, 3294 (1997)
- Ichikawa H. and Baba T. "Efficiency enhancement in a light-emitting diode with a two-dimensional surface grating photonic crystal" *Applied Physics Letters* **84**, 457 (2004)
- Joannopoulos J. D., Meade R. D., and Winn J. N. *Photonic Crystals* (Princeton University Press, Princeton NJ, 1995)
- Joannopoulos J. D., Villeneuve P. R. and Fan S. "Photonic crystals" *Solid State Communications* **102**, 165 (1997)
- Joannopoulos J. D., Johnson S. G., Winn J. N., and Meade R. D. *Photonic Crystals*, second edition (Princeton University Press, Princeton NJ, 2008)
- John S. "Strong localization of photons in certain disordered dielectric superlattices" *Physical Review Letters* **58**, 2486 (1987)
- Kim D.-H., Cho C.-O., Roh Y.-G., Jeon, H., Park Y. S., Cho J., Im J. S., Sone C., Park Y., Choi W. J., and Park Q.-H. "Enhanced light extraction from GaN-based light-emitting diodes with holographically generated two-dimensional photonic crystal patterns" *Applied Physics Letters* **87**, 203508 (2005)
- Kobayashi T., Segawa T., Morimoto A., and Sueta T., paper presented at the 43rd fall meeting of the Japanese Society of Applied Physics, Tokyo, Sept. (1982)
- Kwon M.-K., Kim J.-Y., Park I.-K., Kim K. S., Jung G.-Y., Park S.-J., Kim J. W., and Kim Y. C. "Enhanced emission efficiency of GaN/InGaN multiple quantum well light-emitting diode with an embedded photonic crystal" *Applied Physics Letters* **92**, 251110 (2008)
- McCall S. L., Levi A. F. J., Slusher R. E., Pearton S. J., and Logan R. A. "Whispering-gallery mode microdisk laser" *Applied Physics Letters* **60**, 289 (1992)
- Mohideen U., Hobson W. S., Pearton J., Ren F., and Slusher R. E. "GaAs / AlGaAs microdisk lasers" *Applied Physics Letters* **64**, 1911 (1993)
- Oder T. N., Kim K. H., Lin J. Y., and Jiang H. X. "III-nitride blue and ultraviolet photonic crystal light emitting diodes" *Applied Physics Letters* **84**, 466 (2004)
- Schnitzer I., Yablonovitch E., Caneau C., and Gmitter T. J. "Ultra-high spontaneous emission quantum efficiency, 99.7% internally and 72% externally, from AlGaAs / GaAs / AlGaAs double heterostructures" *Applied Physics Letters* **62**, 131 (1993)
- Wierer J. J., Krames M. R., Epler J. E., Gardner N. F., Crawford M. G., Wendt J. R., Simmons J. A., and Sigalas M. M. "InGaN/GaN quantum-well heterostructure light-emitting diodes employing photonic crystal structures" *Applied Physics Letters* **84**, 3885 (2004)
- Wierer J. J. Jr., David A. and Megens M. M. "III-nitride photonic-crystal light-emitting diodes with high extraction efficiency" *Nature Photonics* **3**, 163 (2009)
- Yablonovitch E. "Inhibited spontaneous emission in solid-state physics and electronics" *Physical Review Letters* **58**, 2509 (1987)
- Yokoyama H. "Physics and device applications of optical microcavities" *Science* **256**, 66 (1992)

19

19 – Visible-spectrum LEDs

Originally, LEDs were exclusively used for low-brightness applications such as indicator lamps. In these applications, the efficiency and the overall optical power of the LED are *not* of primary importance. However, in more recent applications, for example traffic light applications, the light emitted by LEDs must be seen even in bright sunlight and from a considerable distance. LEDs with high efficiency and brightness are required for such applications.

In this chapter, low-brightness as well as high-brightness LEDs are discussed. GaAsP and nitrogen-doped GaAsP LEDs are suitable only for low-brightness applications. AlGaAs LEDs are suitable for low- as well as high-brightness applications. AlGaNp and GaInN LEDs are used in high-brightness applications.

19.1 The GaAsP, GaP, GaAsP:N, and GaP:N material systems

The $\text{GaAs}_{1-x}\text{P}_x$ and $\text{GaAs}_{1-x}\text{P}_x:\text{N}$ material system is used for emission in the red, orange, yellow, and green wavelength range. The GaAsP system is lattice mismatched to GaAs substrates, resulting in a relatively low internal quantum efficiency. As a result these LEDs are suitable for low-brightness applications only.

$\text{GaAs}_{1-x}\text{P}_x$ was one of the first material systems used for visible-spectrum LEDs (Holonyak and Bevacqua 1962; Holonyak *et al.* 1963, 1966; Pilkuhn and Rupprecht 1965; Wolfe *et al.*, 1965; Nuese *et al.* 1966). In the early 1960s, GaAs substrates were already available. Bulk growth of GaAs substrates was initiated in the 1950s and epitaxial growth by LPE and VPE started in the 1960s. As phosphorus is added to GaAs, the ternary alloy $\text{GaAs}_{1-x}\text{P}_x$, or briefly GaAsP, is formed. The addition of phosphorus increases the bandgap of GaAs, which emits in the infrared at 870 nm. The visible wavelength range starts at about 750 nm, so that a small amount of phosphorus is sufficient to attain visible-spectrum light emitters. Note, however, that the sensitivity of the human eye is low at the edges of the visible spectrum.

A significant problem with GaAsP LEDs is the lattice mismatch between the GaAs substrate and the GaAsP epitaxial layer. A large mismatch exists between GaAs and GaP (about 3.6%) so

that many misfit dislocations occur when the critical thickness of GaAsP on GaAs is exceeded. As a result, the luminescence efficiency decreases substantially in GaAsP with increasing phosphorus content. GaAsP LEDs are therefore useful for low-brightness applications only.

It was realized early in the GaAsP work that the lattice mismatch between the GaAs substrate and the GaAsP epilayer reduces the radiative efficiency. It was also found that the radiative efficiency of the active p-n junction layer strongly depends on the growth conditions and, in particular, on the thickness of the GaAsP buffer layer (Nuese *et al.*, 1969). A thick buffer layer reduces the dislocation density by annihilation of misfit dislocations. However, the dislocation density does not approach the low dislocation density of GaAs substrates, so that even with thick GaAsP buffer layers the dislocation density is substantial.

The band structure of GaAs, GaAsP, and GaP is shown schematically in **Figure 19.1**. The figure shows that GaAsP is a direct-gap semiconductor for low phosphorus mole fractions. Beyond the direct–indirect crossover occurring at phosphorus mole fractions of about 45~50%, the semiconductor becomes indirect and the radiative efficiency drops rapidly (Holonyak *et al.*, 1963, 1966). GaP is an indirect-gap semiconductor and therefore is unsuitable as an efficient LED material.

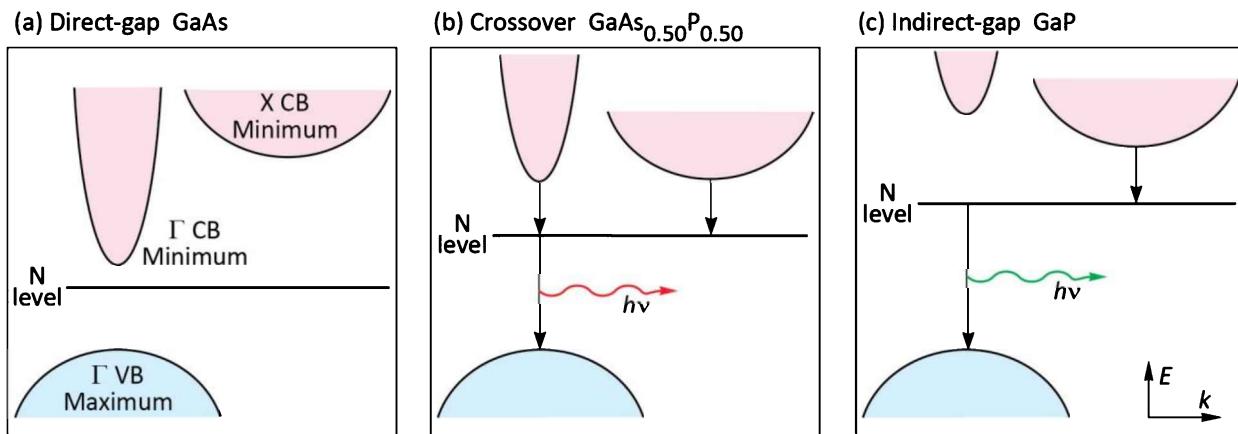


Fig. 19.1: Schematic band structure of GaAs, GaAsP, and GaP. Also shown is the nitrogen level. At a P mole fraction of 45–50%, the direct–indirect crossover occurs.

GaAsP and GaP LEDs are frequently doped with isoelectronic impurities such as nitrogen (Grimmeiss and Scholz, 1964; Logan *et al.*, 1967a, 1967b, 1971; Crawford *et al.*, 1972; Groves and Epstein 1977; Groves *et al.* 1978a, 1978b). The isoelectronic impurities form an optically active

level within the forbidden gap of the semiconductor so that carriers recombine radiatively via the nitrogen levels, as indicated in *Figure 19.1*.

Isoelectrically doped LEDs are also interesting from a fundamental point of view. They are one of the first practical applications of Heisenberg's uncertainty principle. Isoelectronic impurities have an electronic wave function that is strongly localized in position space (small Δx). Therefore, the wave function is delocalized in momentum space (large Δp). Since the level is delocalized in momentum space, two vertical transitions can occur via the isoelectronic trap, with one of them being radiative. Physically speaking, the change in momentum, occurring when an electron makes a transition from the indirect X valley of the conduction band to the central Γ valley of the valence band, is absorbed by the isoelectronic impurity atom.

The emission wavelength of undoped and nitrogen-doped GaAsP is shown in *Figure 19.2* (Craford *et al.*, 1972). The emission energy of GaAsP and GaP doped with the isoelectronic impurity nitrogen is below the bandgap of the semiconductor. *Figure 19.2* illustrates that the emission energy is about 50~150 meV below the bandgap of the semiconductor. As a result, reabsorption effects are much less likely in nitrogen-doped structures compared with LEDs based on band-edge emission. This is an advantage of LEDs doped with isoelectronic impurities.

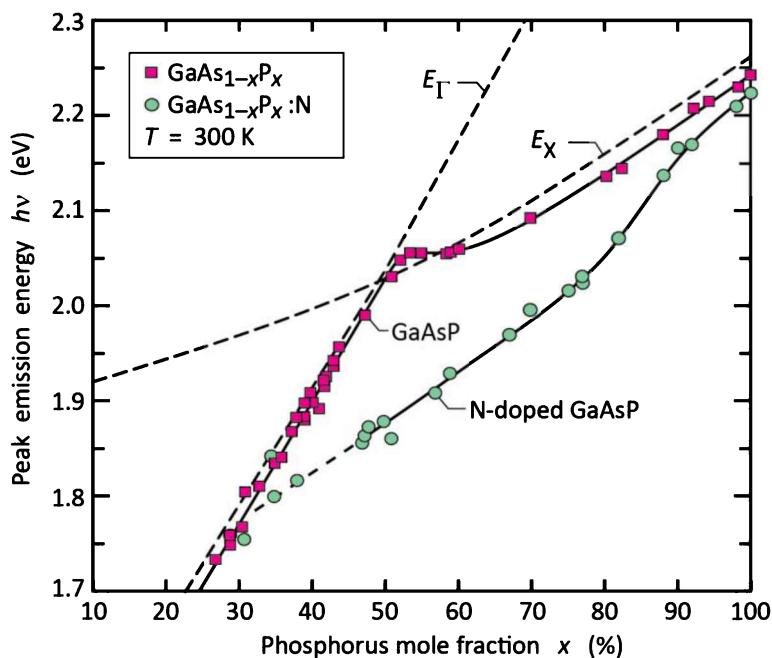


Fig. 19.2: Room-temperature peak emission energy versus alloy composition for undoped and nitrogen-doped GaAsP LEDs injected with a current density of 5 A/cm^2 . Also shown is the energy gap of the direct-to-indirect (E_Γ -to- E_X) transition. The direct–indirect cross-over occurs at $x \approx 50\%$ (after Craford *et al.*, 1972).

Groves *et al.* (1978a, 1978b) showed that this advantage is particularly pronounced if only the active region is doped with nitrogen. In this case, the region of the p-n junction plane and the

regions located within the carrier diffusion lengths from the junction plane are doped with nitrogen. Other regions, such as the confinement and window layers, are not doped with the isoelectronic impurity, so that reabsorption of light by the isoelectronic impurities is limited to the narrow active region. Quantum efficiencies of several percent can be attained with GaP:N LEDs in which the nitrogen doping is limited to the active region.

The external quantum efficiency of undoped and nitrogen-doped GaAsP is shown in **Figure 19.3** (Campbell, 1974). Only the vicinity of the active region is doped with nitrogen. The efficiency of the nitrogen-doped LEDs is strongly enhanced over the entire composition range compared with the GaAsP LEDs without nitrogen doping.

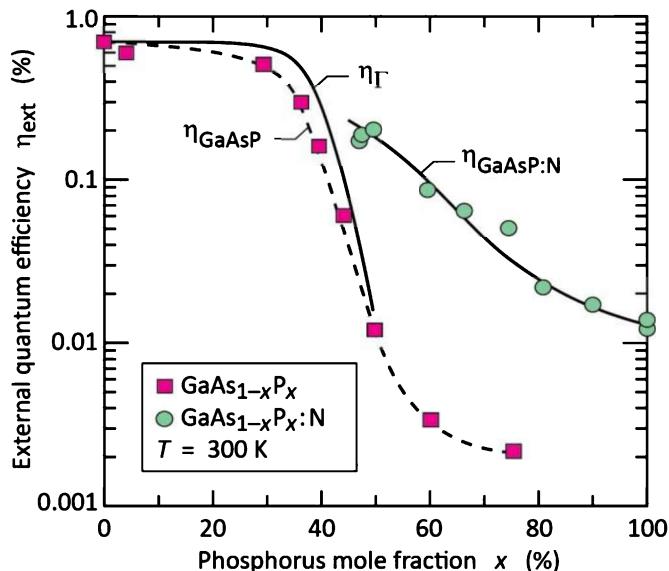


Fig. 19.3: Experimental external quantum efficiency of undoped and N-doped GaAsP versus the P mole fraction. Also shown is the calculated direct-gap (Γ) transition efficiency, η_{Γ} , and the calculated nitrogen (N) related transition efficiency, η_N (solid lines). Note that the nitrogen-related efficiency is higher than the direct-gap efficiency in the indirect bandgap ($x > 50\%$) regime (after Campbell *et al.*, 1974).

Note that the GaAsP LED efficiency decreases by more than two orders of magnitude in the composition range $x = 40\sim60\%$. This decrease is due to the direct–indirect crossover occurring in GaAsP and due to the increasing dislocation density occurring at higher phosphorus mole fractions. At a phosphorus mole fraction of 75%, the GaAsP external quantum efficiency is only 0.002%.

The external quantum efficiency of undoped and N-doped GaAsP vs. emission wavelength is shown in **Figure 19.4**. Again, only the vicinity of the active region is doped with nitrogen. The efficiency of nitrogen-doped GaAsP is higher than that of undoped GaAsP, in particular in the orange, yellow, and green wavelength range where the improvement is a factor of 2~5. In the red wavelength range, the undoped and nitrogen-doped GaAsP LEDs have similar efficiencies.

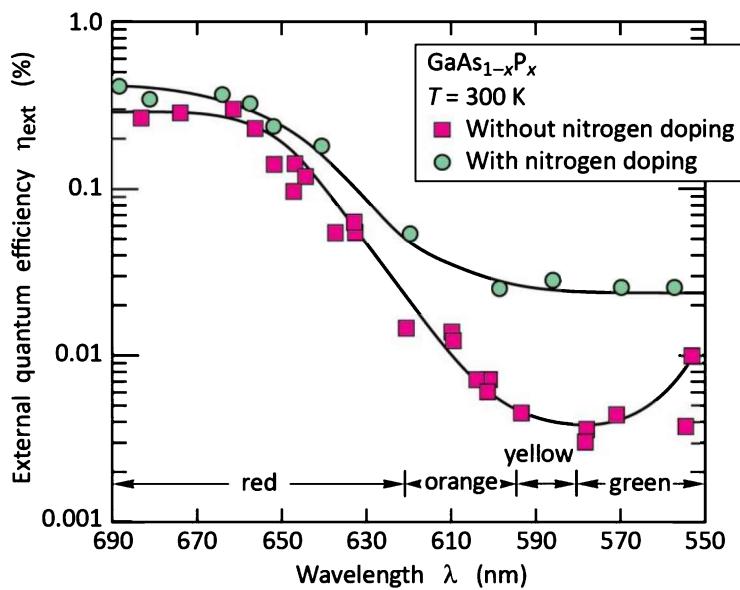


Fig. 19.4: External quantum efficiency versus emission wavelength in undoped and nitrogen-doped $\text{GaAs}_{1-x}\text{P}_x$ (after Groves *et al.*, 1978a, 1978b).

The ratio of external quantum efficiencies of undoped and nitrogen-doped GaAsP LEDs is shown in **Figure 19.5**. It is inferred from the figure that nitrogen-doped devices have a higher efficiency over the entire composition range.

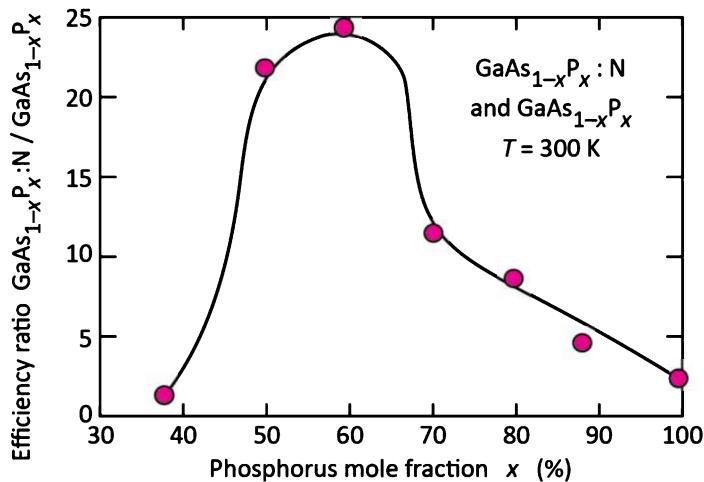


Fig. 19.5: Efficiency ratio between nitrogen-doped and undoped $\text{GaAs}_{1-x}\text{P}_x$ at 300 K (after Groves *et al.*, 1978a, 1978b).

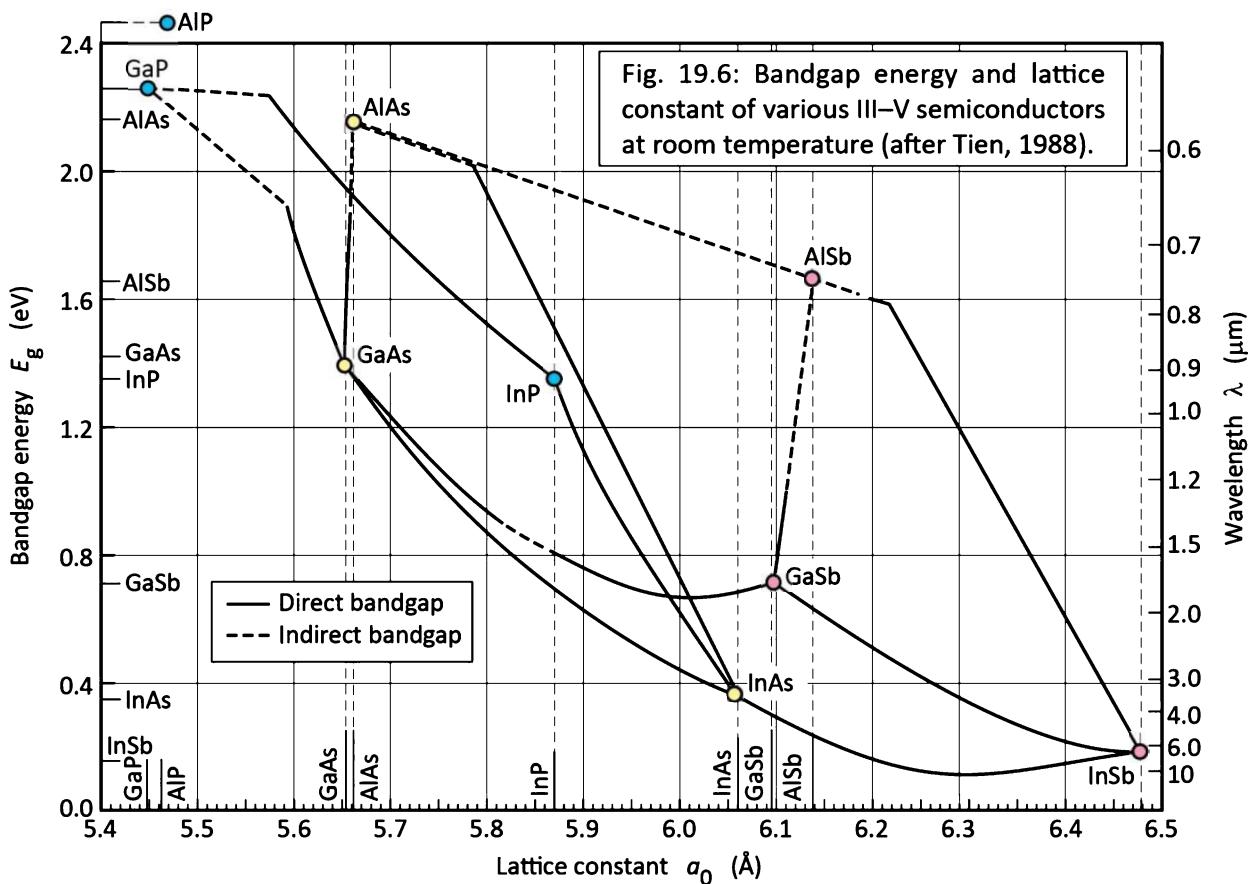
The brightness of LEDs based on isoelectronic impurity transitions is limited by the ***finite solubility*** of nitrogen. For example, nitrogen is soluble in GaP up to nitrogen concentrations of about 10^{20} cm^{-3} . Since an optical transition via a nitrogen level has a certain lifetime, the maximum nitrogen concentration limits the LED operation to a maximum current beyond which the LED efficiency decreases.

Commercial low-brightness green LEDs are based on nitrogen-doped GaP. The main application of GaP:N LEDs is indicator lamps. However, nitrogen-doped GaP LEDs are not suitable

for high-brightness applications, i.e. for applications under bright ambient light conditions such as sunlight. High-brightness green LEDs are based on GaInN.

19.2 The AlGaAs/GaAs material system

The $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ material system was developed in the 1970s and early 1980s and it was the first material system suitable for high-brightness LED applications (for a review, see Steranka, 1997). Owing to the very similar atomic radii of Al (1.82 \AA) and Ga (1.81 \AA), the material system $\text{Al}_x\text{Ga}_{1-x}\text{As}$ (or briefly AlGaAs) is lattice matched to GaAs for all Al mole fractions. The lack of dependence of the lattice constant on the Al mole fraction can be inferred from *Figure 19.6*, which shows the energy gap and lattice constant of several III–V semiconductors and of its ternary and quaternary alloys as a function of the lattice constant (adopted from Tien, 1988).



GaAs and $\text{Al}_x\text{Ga}_{1-x}\text{As}$ for Al mole fractions $x < 0.45$ are direct-gap semiconductors. The energy gap of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ versus the Al mole fraction is shown in *Figure 19.7* (Casey and Panish, 1978). For Al

mole fractions $x < 45\%$, the Γ conduction-band valley is the *lowest* minimum and the semiconductor has a *direct* gap. For $x > 45\%$, the X valleys are the *lowest* conduction-band minimum and the semiconductor becomes *indirect*.

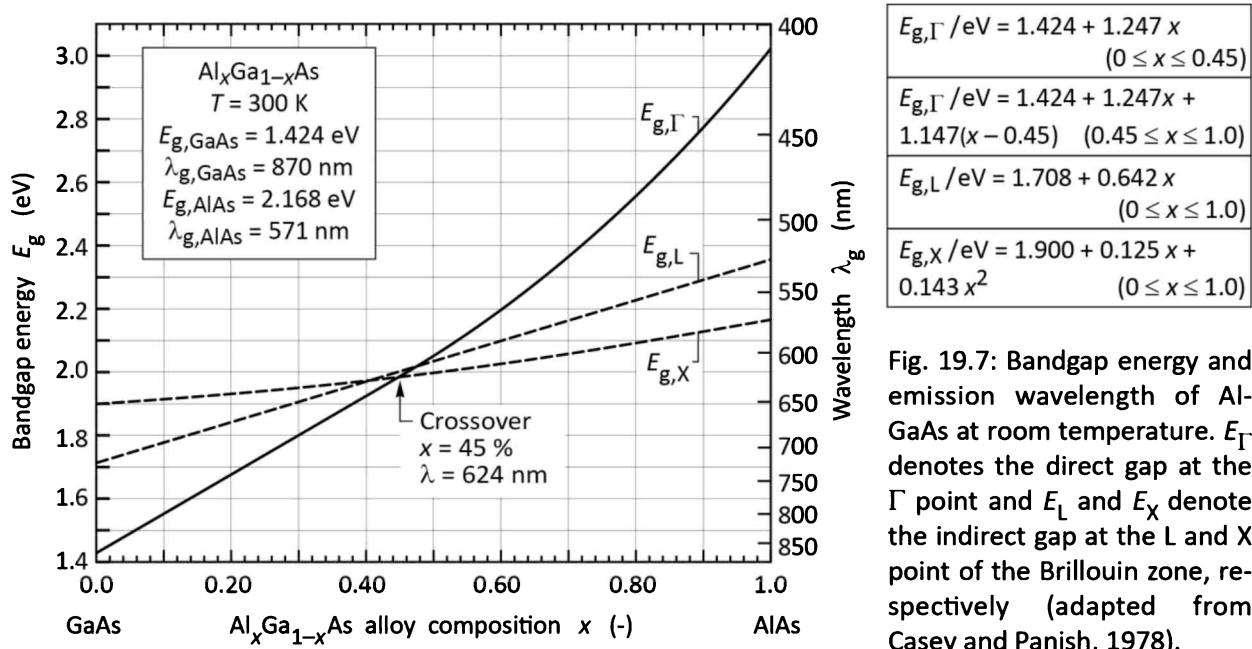


Fig. 19.7: Bandgap energy and emission wavelength of Al-GaAs at room temperature. E_Γ denotes the direct gap at the Γ point and E_L and E_X denote the indirect gap at the L and X point of the Brillouin zone, respectively (adapted from Casey and Panish, 1978).

The AlGaAs material system is suited for high-brightness visible-spectrum LEDs emitting in the red wavelength range. The direct–indirect crossover occurs at a wavelength of 621 nm. At that wavelength, the radiative efficiency of the AlGaAs system becomes quite low due to the direct–indirect transition. To maintain high efficiency, the emission energy must be several kT lower than the bandgap energy at the direct–indirect crossover point.

There are several possible strategies for AlGaAs-based red LEDs, including $\text{Al}_x\text{Ga}_{1-x}\text{As}$ bulk active regions, $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum well active regions, and $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{Al}_y\text{Ga}_{1-y}\text{As}$ ($x > y$) double heterostructure active regions. The first possibility, $\text{Al}_x\text{Ga}_{1-x}\text{As}$ bulk active regions, lacks the advantages of a heterostructure and this approach is therefore not used in high-brightness LEDs. The two other possibilities are more attractive due to the employment of heterostructures. The quantum well and double heterostructure active region is used in high-efficiency red LEDs and the two structures are shown schematically in **Figure 19.8**. In the $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum well case shown in **Figure 19.8 (a)**, size quantization is used to increase the emission energy. In the case of the $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{Al}_y\text{Ga}_{1-y}\text{As}$ double heterostructures shown in **Figure 19.8 (b)**, AlGaAs

is used for both the barrier region and the well region. A drawback of the $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum well active regions is the requirement of very thin GaAs quantum wells clad by $\text{Al}_x\text{Ga}_{1-x}\text{As}$ barriers. Vertical transport in multi-quantum well (MQW) structures can lead to non-uniform carrier distribution in the MQW active region unless the barriers are very thin. Consequently, the $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{Al}_y\text{Ga}_{1-y}\text{As}$ double heterostructure approach is usually preferred.

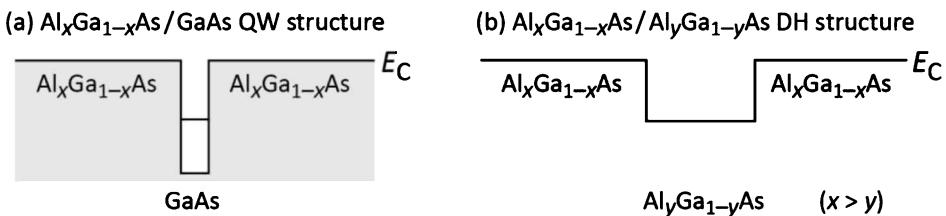


Fig. 19.8: Band diagrams of AlGaAs/GaAs structures suited for emission in the red part of the visible spectrum. (a) AlGaAs/GaAs quantum well (QW) structure with a thin GaAs well. (b) AlGaAs/AlGaAs double heterostructure (DH) with an AlGaAs active region.

AlGaAs/GaAs LEDs have been fabricated as homostructures, single heterostructures, and double heterostructures (Nishizawa *et al.*, 1983). The most efficient AlGaAs red LEDs are double-heterostructure transparent-substrate (DH-TS) devices (Ishiguro *et al.*, 1983; Steranka *et al.* 1988; Ishimatsu and Okuno, 1989). **AlGaAs DH-TS LEDs** are grown on temporary GaAs substrates and consist of a thick (e.g. 125 μm) $\text{Al}_x\text{Ga}_{1-x}\text{As}$ lower confinement layer with an Al mole fraction $x > 60\%$, an $\text{Al}_x\text{Ga}_{1-x}\text{As}$ active layer ($x = 35\%$, for red devices), and a thick (e.g. 125 μm) $\text{Al}_x\text{Ga}_{1-x}\text{As}$ upper confinement layer, also with an Al mole fraction $x > 60\%$. For devices emitting in the IR, the Al mole fractions of the active and confinement layers can be lower. After epitaxial growth, the absorbing GaAs substrate is removed by polishing and selective wet chemical etching. AlGaAs DH-TS LEDs are more than a factor of 2 brighter than double-heterostructure absorbing-substrate (DH-AS) devices (Steranka *et al.*, 1988).

In the 1980s the growth method of choice for AlGaAs DH-TS LEDs was liquid-phase epitaxy (LPE). This growth method is capable of growing, at a high growth rate, very thick high-quality AlGaAs layers with high Al content. LPE can be scaled up for high-volume production (Ishiguro *et al.*, 1983; Steranka *et al.*, 1988; Ishimatsu and Okuno, 1989). AlGaAs/GaAs DH-AS LEDs have also been grown by OMVPE (Bradley *et al.*, 1986). However, the OMVPE growth rate is lower than the LPE growth rate. OMVPE growth of thick layers, as required for DH-TS devices, is therefore

difficult. Historically AlGaAs DH-TS LEDs were the first high-brightness LEDs suitable for demanding applications such as automotive brake lights and traffic lights, which must be clearly visible under bright ambient conditions.

The reliability of AlGaAs devices is known to be lower than that of AlGaN_xP devices that do not contain any AlGaAs. High-Al-content AlGaAs layers are subject to oxidation and corrosion, thereby lowering the device lifetime. Dallesasse *et al.* (1990) reported the deterioration of AlGaAs/GaAs heterostructures by hydrolysis. Cracks, fissures, and pinholes were found after long-term exposure to room environmental conditions, especially for thick AlGaAs layers ($> 0.1 \mu\text{m}$) with a high Al content such as 85%. The authors found very thin AlGaAs layers (e.g. 20 nm) to be stable, even for Al contents of 100%. Hermetic packaging is required to avoid oxidation and hydrolysis of AlGaAs layers. Steranka *et al.* (1988) stated that some AlGaAs devices on the market have exhibited severe degradation. However, accelerated aging data taken at 55°C with an injection current of 30 mA showed no degradation at all after 1000 h of stress. Such a performance requires excellent understanding and control of the device fabrication and packaging process.

19.3 The AlGaN_xP/GaAs material system

The AlGaN_xP material system was developed in the late 1980s and early 1990s and today is the primary material system for high-brightness LEDs emitting in the long-wavelength part of the visible spectrum, i.e. in the red, orange, amber, and yellow wavelength range. The AlGaN_xP material system and AlGaN_xP LEDs have been reviewed by Stringfellow and Craford (1997), Chen *et al.* (1997), and Kish and Fletcher (1997). Further reviews and recent developments were published by Mueller (1999, 2000) and Krames *et al.* (2002).

The energy gap and the corresponding wavelength versus the lattice constant of AlGaN_xP (Chen *et al.*, 1997) is shown in **Figure 19.9**. AlGaN_xP can be lattice matched to GaAs. Replacing all As atoms in the GaAs lattice by *smaller* P atoms and some of the Ga atoms in the GaAs lattice by *larger* In atoms, forms GaInP, which at the particular composition $\text{Ga}_{0.5}\text{In}_{0.5}\text{P}$, is lattice matched to GaAs. Since Al and Ga have very similar atomic radii, the material $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$ is also lattice matched to GaAs.

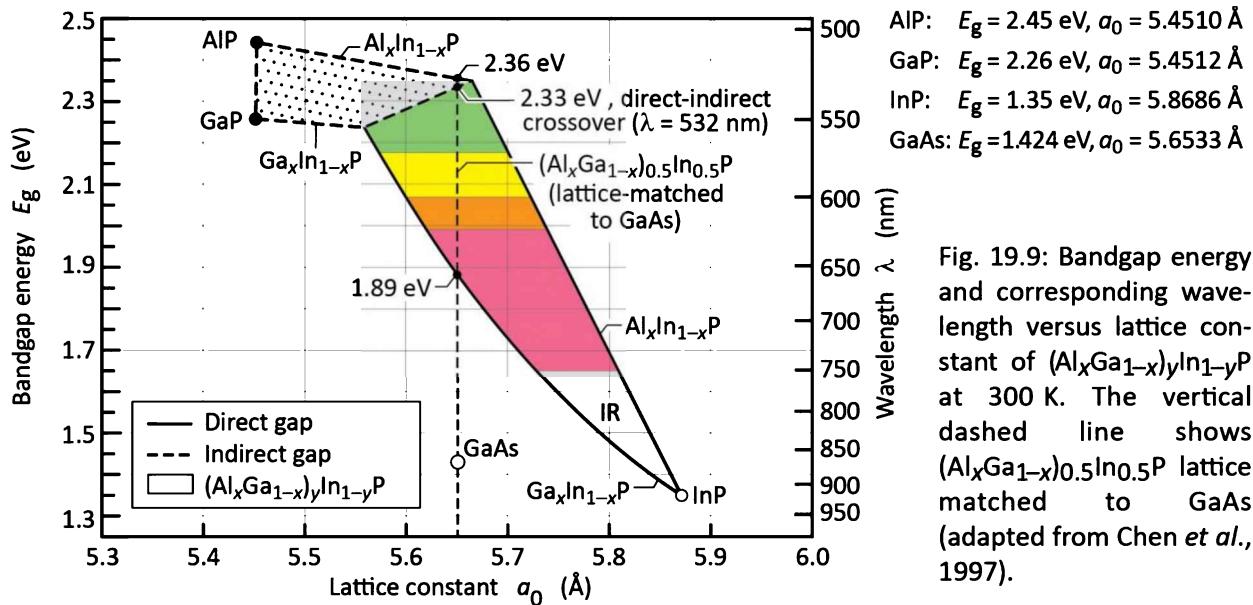


Fig. 19.9: Bandgap energy and corresponding wavelength versus lattice constant of $(Al_xGa_{1-x})_yIn_{1-y}P$ at 300 K. The vertical dashed line shows $(Al_xGa_{1-x})_0.5In_{0.5}P$ lattice matched to GaAs (adapted from Chen *et al.*, 1997).

According to Chen *et al.* (1997), $(Al_xGa_{1-x})_0.5In_{0.5}P$ has a direct bandgap for $x < 0.5$ and an indirect bandgap for $x > 0.5$. At the crossover point ($x = 0.5$), the bandgap energy is 2.33 eV, corresponding to a wavelength of 532 nm. Kish and Fletcher (1997) compiled data from Prins *et al.* (1995) and concluded that $(Al_xGa_{1-x})_0.5In_{0.5}P$ is a direct-gap semiconductor for Al mole fractions $x < 0.53$. The energy gap versus Al mole fraction is shown in **Figure 19.10** (Prins *et al.*, 1995; Kish and Fletcher, 1997). At Al mole fractions $x < 53\%$, the Γ conduction-band valley is the lowest minimum and the semiconductor has a direct gap. For $x > 53\%$, the X valleys are the lowest conduction-band minimum and the semiconductor becomes indirect. The emission wavelength at the direct-indirect crossover point is approximately 555 nm. The exact wavelength of the crossover point may depend on the degree of atomic ordering present in a particular material (Kish and Fletcher, 1997).

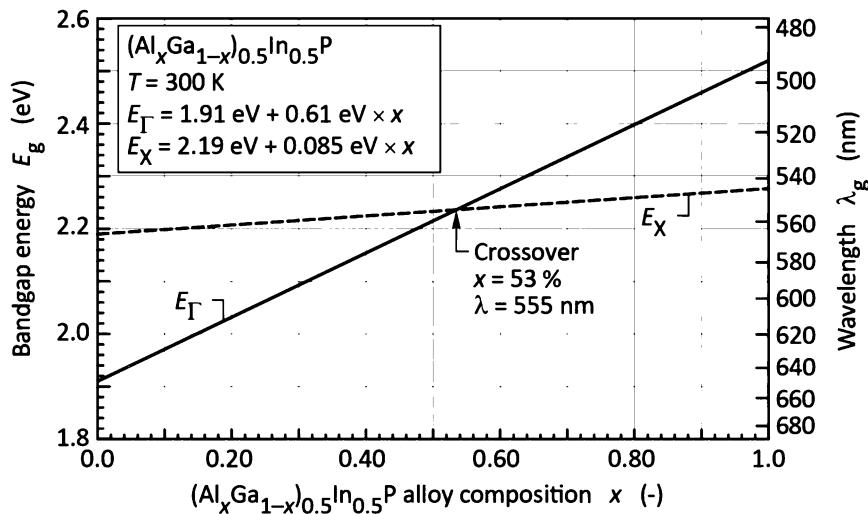


Fig. 19.10: Bandgap energy and emission wavelength of unordered AlGaN_P lattice matched to GaAs at room temperature. E_Γ denotes the direct gap at the Γ point and E_X denotes the indirect gap at the X point of the Brillouin zone (after Prins *et al.*, 1995 and Kish and Fletcher, 1997).

A contour plot of the lattice constant and the energy gap of the AlGaN_P materials system is shown in **Figure 19.11** (Chen *et al.*, 1997). The bandgap energy values and the composition of the direct–indirect crossover shown in **Figure 19.11** are slightly different from the data shown in **Figure 19.10**, which can be attributed to atomic ordering in AlGaN_P. Atomic ordering lowers the bandgap energy by values up to 190 meV (Kish and Fletcher, 1997).

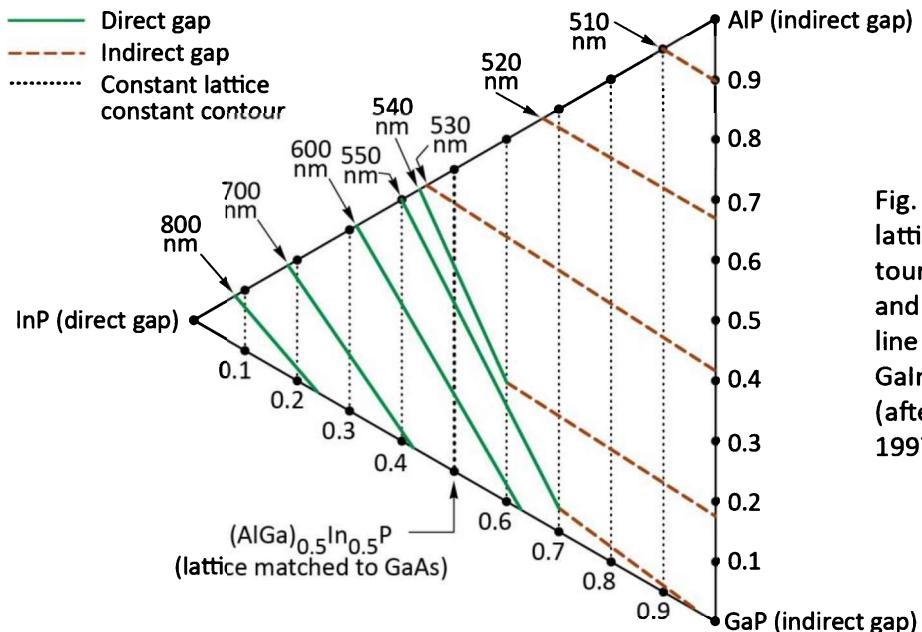


Fig. 19.11: Constant lattice constant contours (vertical lines) and constant emission line contours of the AlGaN_P materials system (after Chen *et al.*, 1997).

The AlGaN_P material system is suited for high-brightness visible-spectrum LEDs emitting in the red, orange, amber, and yellow wavelength range. At the direct–indirect crossover the radiative efficiency of the AlGaN_P system is quite low due to the direct–indirect transition. To maintain high efficiency, the emission energy should be several kT lower than the bandgap energy at the direct–indirect crossover point.

19.4 The GaInN material system

The GaInN material system was developed in the early 1990s. GaInN LEDs emitting in the blue and green wavelength range became commercially available in the late 1990s. To date GaInN is the primary material system for high-brightness blue and green LEDs. The GaInN material system and GaInN LEDs have been reviewed by Nakamura and Fasol (1997) and by Strite and Morkoc (1992).

One of the greatest surprises of the GaInN materials system is its high radiative efficiency despite the presence of a very high concentration of threading dislocations in GaInN/GaN epitaxial films. These threading dislocations are due to the lattice mismatch between the commonly used sapphire and SiC substrates and the GaN and GaInN epitaxial films. Typical densities of the threading dislocations are in the $10^7 \sim 10^9 \text{ cm}^{-2}$ range.

In the III–V arsenide and III–V phosphide material systems, misfit dislocations have disastrous consequences for the radiative efficiency. The lack of such strongly detrimental consequences in the GaInN material system is not fully understood. However, it is believed that the small diffusion length of holes and an apparently low electrical activity of dislocations in GaN and GaInN allows for high radiative efficiencies.

It has also been postulated that fluctuations of the In content in GaInN cause carriers to be localized in potential minima, thus preventing carriers from reaching dislocations. The carriers localized in potential minima will eventually recombine radiatively. Although the physical mechanisms of the high radiative efficiency of GaInN are not yet understood in detail, it is clear that the optical properties of III–V nitrides are much less affected by dislocations than those of III–V arsenides and III–V phosphides.

The bandgap energy versus the lattice constant in the nitride material family is shown in **Figure 19.12**. Inspection of the figure indicates that GaInN is, in theory, suitable for covering the entire visible spectrum. However, the growth of high-quality GaInN becomes increasingly more difficult as the In composition is increased, in part due to re-evaporation of In from the growth surface. As a result, the GaInN material system is exclusively used for ultraviolet (UV), blue, and green LEDs at the present time and rarely for longer wavelengths.

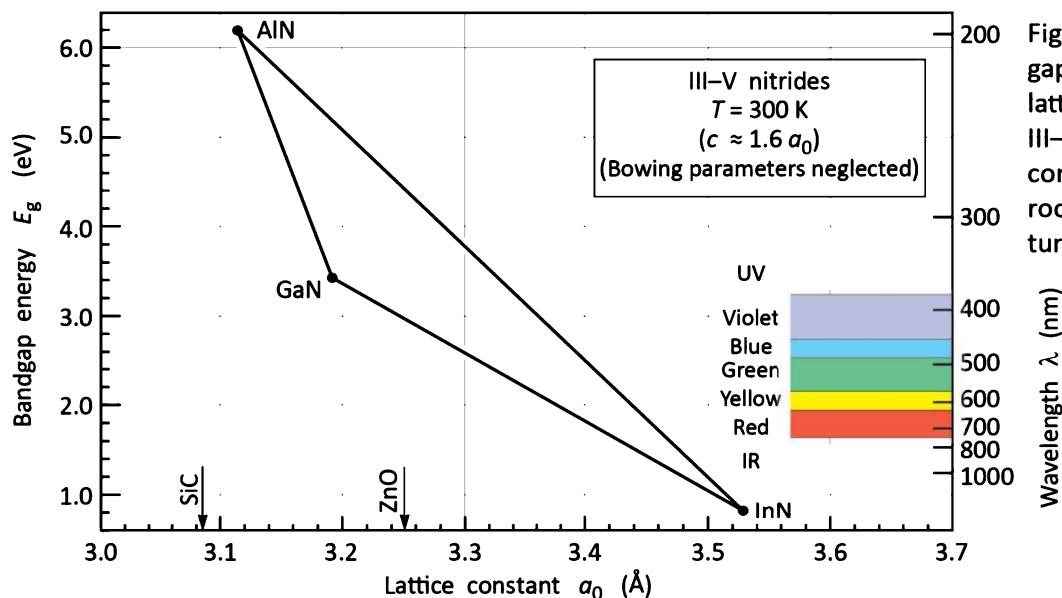


Fig. 19.12: Bandgap energy versus lattice constant of III–V nitride semiconductors at room temperature.

Prior to the year 2002, the generally accepted value for the InN bandgap energy was 1.9 eV. However, Wu *et al.* (2002a, 2002b) showed by luminescence measurements that the bandgap of InN is lower, namely between 0.7 and 0.8 eV. Luminescence measurements also indicated that the InN bandgap exhibits an unusual blue shift with increasing temperature.

19.5 Optical characteristics of high-brightness LEDs

We next review some general optical characteristics of high-brightness LEDs. The optical emission spectra of red AlGaInP and green and blue GaInN LEDs are shown in *Figure 19.13* (Toyoda Gosei, 2000). Comparison of the emission spectra reveals that the emission spectra have a broadened linewidth, particularly the green LED, which has a wider emission spectrum than either the blue or the red LED.

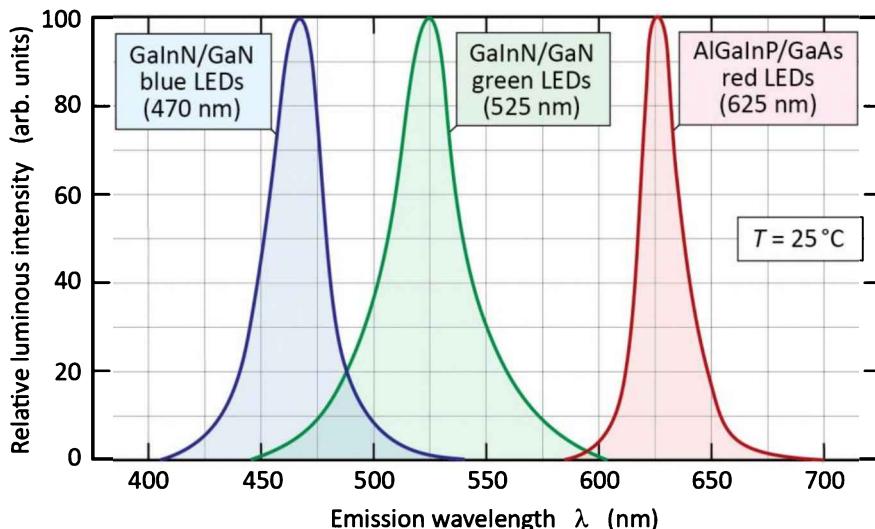


Fig. 19.13: Typical emission spectrum of GaInN/GaN blue, GaInN/GaN green, and AlGaInP/GaAs red LEDs at room temperature (after Toyoda Gosei Company, 2000).

It has been postulated that **In-rich clusters** or **quantum dots** are formed during the growth of GaInN, especially in GaInN with a high In content. It is known that the formation of such In-rich clusters depends strongly on the growth conditions.

Note that the LEDs shown in *Figure 19.13* have an active region composed of a **semiconductor alloy**. Alloy broadening, i.e. the broadening of the emission band due to random fluctuations of the chemical composition of the active material, will lead to spectral broadening that goes beyond the $1.8kT$ linewidth expected for thermally broadened emission bands.

The light power versus injection current is shown in *Figure 19.14*. A linear dependence with unit slope, i.e. unit differential quantum efficiency, is expected for light-versus-current curves in ideal LEDs. The unit-slope line is represented by the dashed line in *Figure 19.14*. The mature

AlGaN_xP LED closely follows the unit-slope line. However, the green LED has a large deviation from the unit differential quantum efficiency slope due to the lower maturity of the GaInN material system, especially with high concentrations of In.

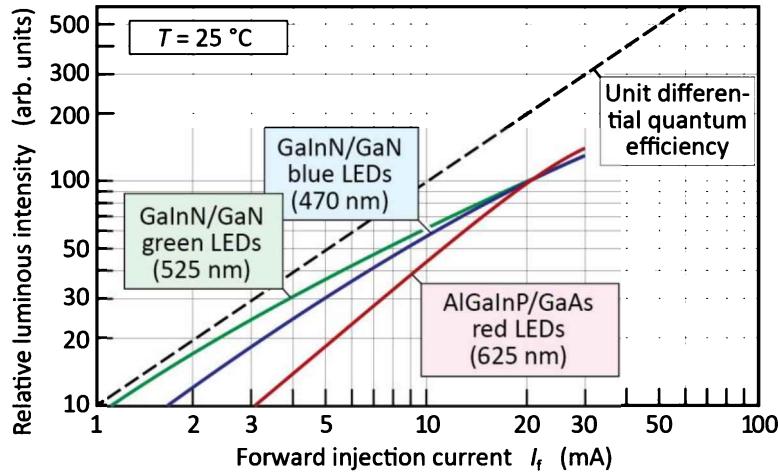


Fig. 19.14: Typical light output power versus injection current of GaInN/GaN blue, GaInN/GaN green, and AlGaN_xP/GaAs red LEDs at room temperature (adopted from Toyoda Gosei Company, 2000).

The temperature dependence of the optical emission intensity is shown in *Figure 19.15*. The figure reveals that III–V nitride diodes have a much weaker temperature dependence than the AlGaN_xP LED. Two factors contribute to the weaker temperature dependence. Firstly, the active-to-confinement barriers are higher in the wide-gap III–V nitride material system than in other III–V material systems. Consequently, carriers in the GaInN active region are very well confined. Thus, carrier leakage out of the well and carrier overflow are of little relevance in GaInN LEDs. Secondly, AlGaN_xP has a direct–indirect transition of the bandgap at about 555 nm. At elevated temperatures, the indirect valleys become increasingly populated so that the radiative efficiency decreases.

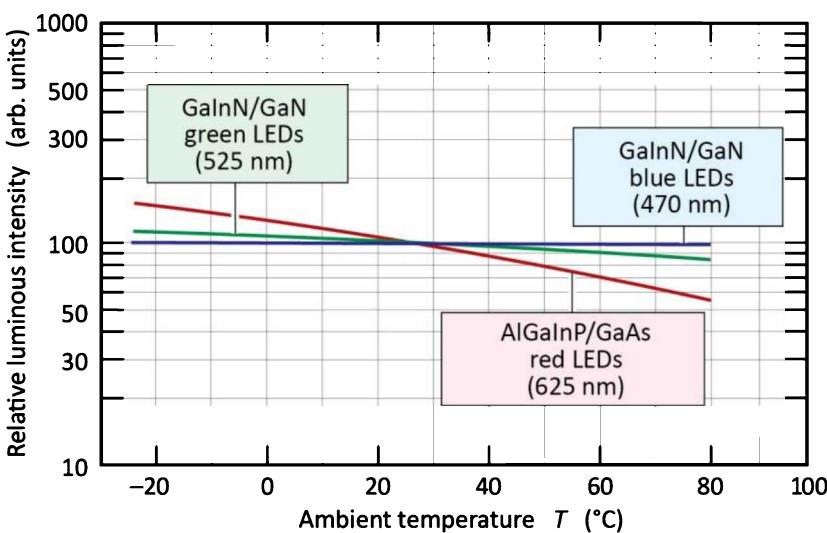


Fig. 19.15: Typical output intensity of GaInN/GaN blue, GaInN/GaN green, and AlGaN_xP/GaAs red LEDs versus ambient temperature (after Toyoda Gosei Company, 2000).

19.6 Electrical characteristics of high-brightness LEDs

The forward current–voltage (I – V) characteristics of a blue GaInN, a green GaInN, and a red AlGaNp LED are shown in **Figure 19.16**. The forward turn-on voltage scales with the emission energy, indicating a well-behaved characteristic. Closer inspection of the forward voltage (at 1 mA) of the green LED ($V_{f,green} = 2.65$ V) indicates that it is very similar to the blue LED ($V_{f,blue} = 2.75$ V) even though the emission energies of the blue and green LED are quite different ($\lambda_{blue} = 470$ nm, $h\nu_{blue} = 2.64$ eV; $\lambda_{green} = 525$ nm, $h\nu_{green} = 2.36$ eV). The small difference in forward voltage indicates that carriers probably lose energy by phonon emission when injected from the GaN barrier into the GaInN active region. They lose more energy when being injected from the GaN barrier into the In-richer active region of the green LED. The energy, dissipated by emission of phonons, is supplied by the external voltage applied to the LED.

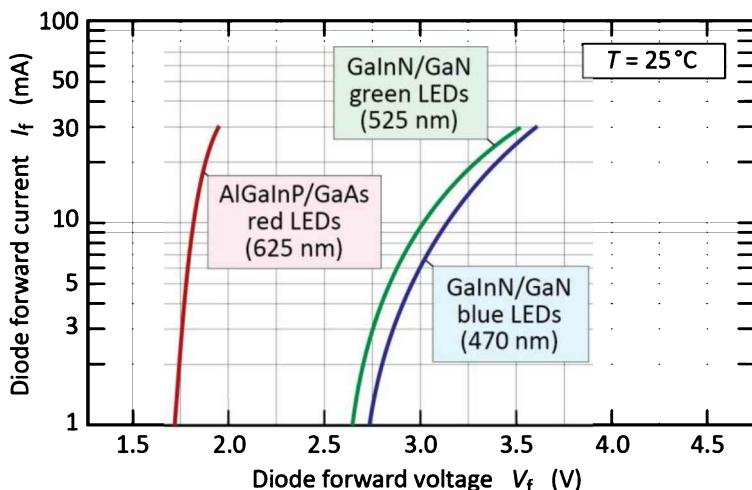


Fig. 19.16: Typical forward current–voltage (I – V) characteristic of GaInN/GaN blue, GaInN/GaN green, and AlGaNp/GaAs red LEDs at room temperature (after Toyoda Gosei Company, 2000).

The diode series resistance can be inferred from the slope of the I – V characteristics. The blue and green diodes have a higher series resistance than the red AlGaNp diode. The larger resistance in GaInN LEDs can be attributed to several factors including the “lateral” resistance in the n-type buffer layer for devices grown on sapphire substrates, strong polarization effects occurring in the nitride material family, lower p-type conductivity in the cladding layer, and higher p-type contact resistance. The lower p-type conductivity is due to the high acceptor activation energy (approximately 200 meV) in GaN and GaInN so that only a small fraction of acceptors is activated.

The temperature dependence of the forward voltage at a diode current of 30 mA is shown in **Figure 19.17**. For all diodes shown, the forward voltage decreases as the temperature is increased. The decrease in forward voltage is due to the decrease of the bandgap energy. In the blue and green GaInN diodes, the lower forward voltage is also due to the decrease in series resistance occurring at high temperatures. This resistance decrease is due to the higher acceptor activation occurring at elevated temperatures and the resulting higher conductivity of the p-type GaN and GaInN layers.

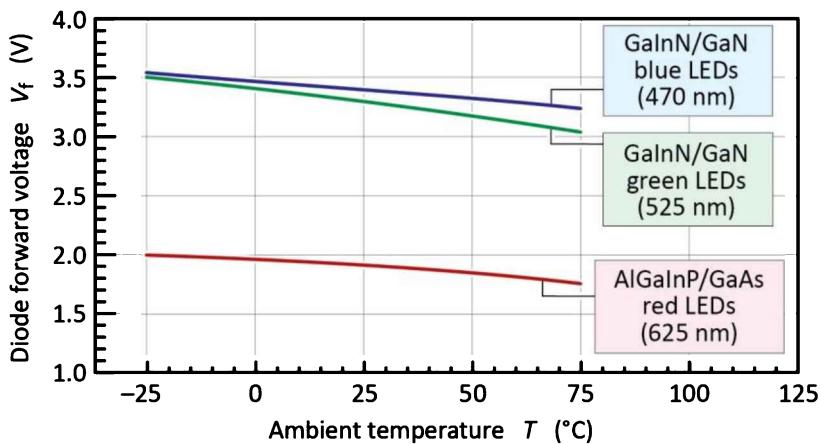


Fig. 19.17: Typical diode forward voltage at a current of 30 mA of GaInN/GaN blue, GaInN/GaN green, and AlGaN/P/GaAs red LEDs versus temperature (after Toyoda Gosei Company, 2000).

19.7 Wavelength dependence of the efficiency in the visible spectrum

The efficiency of visible-spectrum LEDs has a distinct wavelength dependence. That is, high-efficiency devices (made of AlGaNP) are available in the long-wavelength range (> 590 nm); high-efficiency devices (made of GaInN) are also available in the short-wavelength range (< 525 nm). However, in the center of the visible spectrum, particularly near 550 nm, the efficiency of the devices is significantly lower. This fact is frequently referred to as the “*green valley*” or “*green gap*” implying that high-efficiency devices (with power efficiencies $> 50\%$) are lacking in this wavelength range.

The luminous efficiency of high-brightness LEDs and of some low-cost LEDs is shown versus wavelength in **Figure 19.18** (United Epitaxy Corp., 1999). The figure indicates that yellow (590 nm) and orange (605 nm) AlGaNP and green (525 nm) GaInN LEDs are good choices for devices with high luminous efficacy.

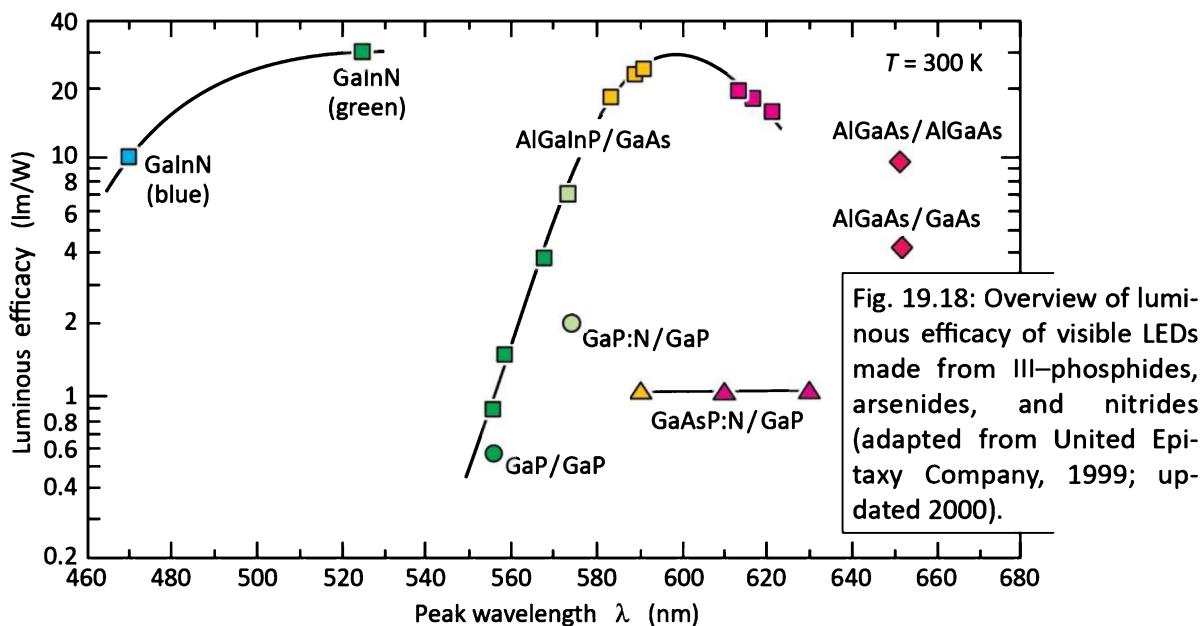


Fig. 19.18: Overview of luminous efficacy of visible LEDs made from III–phosphides, arsenides, and nitrides (adapted from United Epitaxy Company, 1999; updated 2000).

Figure 19.18 also shows that low-power and low-cost LEDs, such as GaAsP and GaP:N LEDs, have a much lower efficiency. These LEDs are not suitable for high-brightness applications due to their inherently lower quantum efficiency. GaAsP LEDs are mismatched to the GaAs substrate and therefore have a low internal efficiency. The GaP:N LEDs also have a low efficiency due to the nitrogen-impurity-assisted nature of the radiative transition.

Figure 19.19 shows the external quantum efficiency attainable with visible-spectrum LEDs, specifically AlGaN_xP long-wavelength LEDs (red, orange, yellow) and GaN short-wavelength LEDs (violet, blue, and green) (Krames *et al.*, 2007). Both, AlGaN_xP and GaN are capable of very high efficiency. However, inspection of the figure shows that while devices with high power efficiencies $> 50\%$ are available for the short and long-wavelength end of the visible spectrum (violet and red), such high efficiencies are not available near 550 nm where the efficiencies are lower than 10%. The near-550-nm-wavelength range is frequently referred to as the “green gap” to convey that this wavelength range suffers from the availability of high-efficiency LEDs.

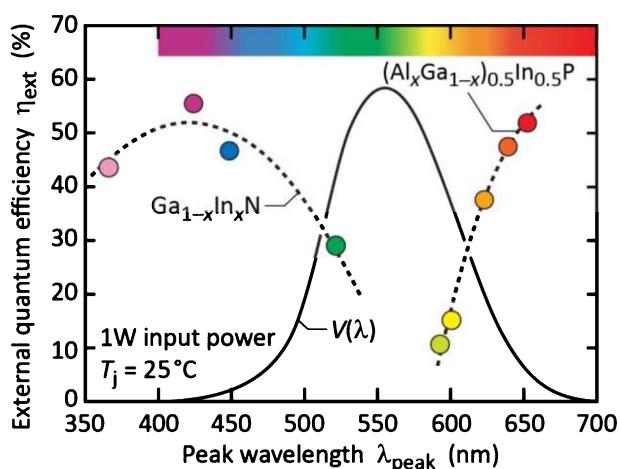


Fig. 19.19: External quantum efficiencies for high-power visible-spectrum LEDs made from the nitride and phosphide material system. $V(\lambda)$ is the luminous eye response curve of the CIE. The dashed lines are guides to the eye. (after Krames *et al.*, 2007).

19.8 References

- Bradley R. R., Ash R. M., Forbes N. W., Griffiths R. J. M., Jebb D. P., and Shepard H. E. "Metalorganic chemical vapor deposition of junction isolated GaAlAs/GaAs LED structures" *Journal of Crystal Growth* **77**, 629 (1986)
- Campbell J. C., Holonyak Jr. N., Craford M. G., and Keune D. L. "Band structure enhancement and optimization of radiative recombination in GaAsP (and InGaP:N)" *Journal of Applied Physics* **45**, 4543 (1974)
- Casey Jr. H. C. and Panish M. B. *Heterostructure Lasers, Part A* and *Heterostructure Lasers, Part B* (Academic Press, San Diego, 1978)
- Chen C. H., Stockman S. A., Peanasky M. J., and Kuo C. P. "OMVPE growth of AlGaN/P for high-efficiency visible light-emitting diodes" in *High Brightness Light Emitting Diodes* edited by G. B. Stringfellow and M. G. Craford, *Semiconductors and Semimetals* **48**, (Academic Press, San Diego, 1997)
- Craford M. G., Shaw R. W., Herzog A. H., and Groves W. O. "Radiative recombination mechanisms in GaAsP diodes with and without nitrogen doping" *Journal of Applied Physics* **43**, 4075 (1972)
- Craford M. G. "Overview of device issues in high-brightness light-emitting diodes" in *High Brightness Light Emitting Diodes* edited by G. B. Stringfellow and M. G. Craford, *Semiconductors and Semimetals* **48** (Academic Press, San Diego, 1997)
- Craford M. G. "The bright future of light-emitting diodes" Plenary talk on light emitting diodes at the MRS Fall Meeting, Boston Massachusetts December (1999)
- Dallesasse J. M., El-Zein N., Holonyak Jr. N., Hsieh K. C., Burnham R. D., and Dupuis R. D. "Environmental degradation of AlGaAs–GaAs quantum-well heterostructures" *Journal of Applied Physics* **68**, 2235 (1990)
- Grimmeiss H. G. and Scholz H. "Efficiency of recombination radiation in GaP" *Physics Letters* **8**, 233 (1964)
- Groves W. O. and Epstein A. S. "Epitaxial deposition of III–V compounds containing isoelectronic impurities" US Patent 4,001,056 (1977)
- Groves W. O., Herzog A. H., and Craford M. G. "Process for the preparation of electroluminescent III–V materials containing isoelectronic impurities" US Patent Re. 29,648 (1978a)
- Groves W. O., Herzog A. H., and Craford M. G. "GaAsP electroluminescent device doped with isoelectronic impurities" US Patent Re. 29,845 (1978b)
- Holonyak Jr. N. and Bevacqua S. F. "Coherent (visible) light emission from Ga(AsP) junctions" *Applied Physics Letters* **1**, 82 (1962)
- Holonyak Jr. N., Bevacqua S. F., Bielan C. V., and Lubowski S. J. "The "direct–indirect" transition in Ga(AsP) p-n junctions," *Applied Physics Letters* **3**, 47 (1963)
- Holonyak Jr. N., Nuese C. J., Sirkis M. D., and Stillman G. E., "Effect of donor impurities on the direct–indirect transition in Ga(AsP)" *Applied Physics Letters* **8**, 83 (1966)
- Ishiguro H., Sawa K., Nagao S., Yamanaka H., and Koike S. "High efficient GaAlAs light emitting diodes of 660 nm with double heterostructure on a GaAlAs substrate" *Applied Physics Letters* **43**, 1034 (1983)
- Ishimatsu S. and Okuno Y. "High efficiency GaAlAs LED" *Optoelectronic Devices and Technology* **4**, 21 (1989)
- Kish F. A. and Fletcher R. M. "AlGaN/P light-emitting diodes" in *High Brightness Light Emitting Diodes* edited by G. B. Stringfellow and M. G. Craford, *Semiconductors and Semimetals* **48** (Academic Press, San Diego, 1997)
- Krames M. R. et al. "High-power truncated-inverted-pyramid ($\text{Al}_x\text{Ga}_{1-x}$)_{0.5}In_{0.5}P/GaP light emitting diodes exhibiting > 50% external quantum efficiency" *Applied Physics Letters* **75**, 2365 (1999)
- Krames M. R. et al. "High-brightness AlGaN/N light emitting diodes" *Proceedings of the SPIE* **3938**, 2 (2000)
- Krames M. R., Amano H., Brown J. J., and Heremans P. L. "High-efficiency light-emitting diodes" Special Issue of *IEEE Journal of Selected Topics of Quantum Electronics* **8**, 185 (2002)
- Krames M. R., Shchekin O. B., Mueller-Mach R., Mueller G. O., Zhou L., Harbers G., and Craford M. G. "Status and future of high-power light-emitting diodes for solid-state lighting" *IEEE Journal of Display Technology* **3**, 160 (2007)

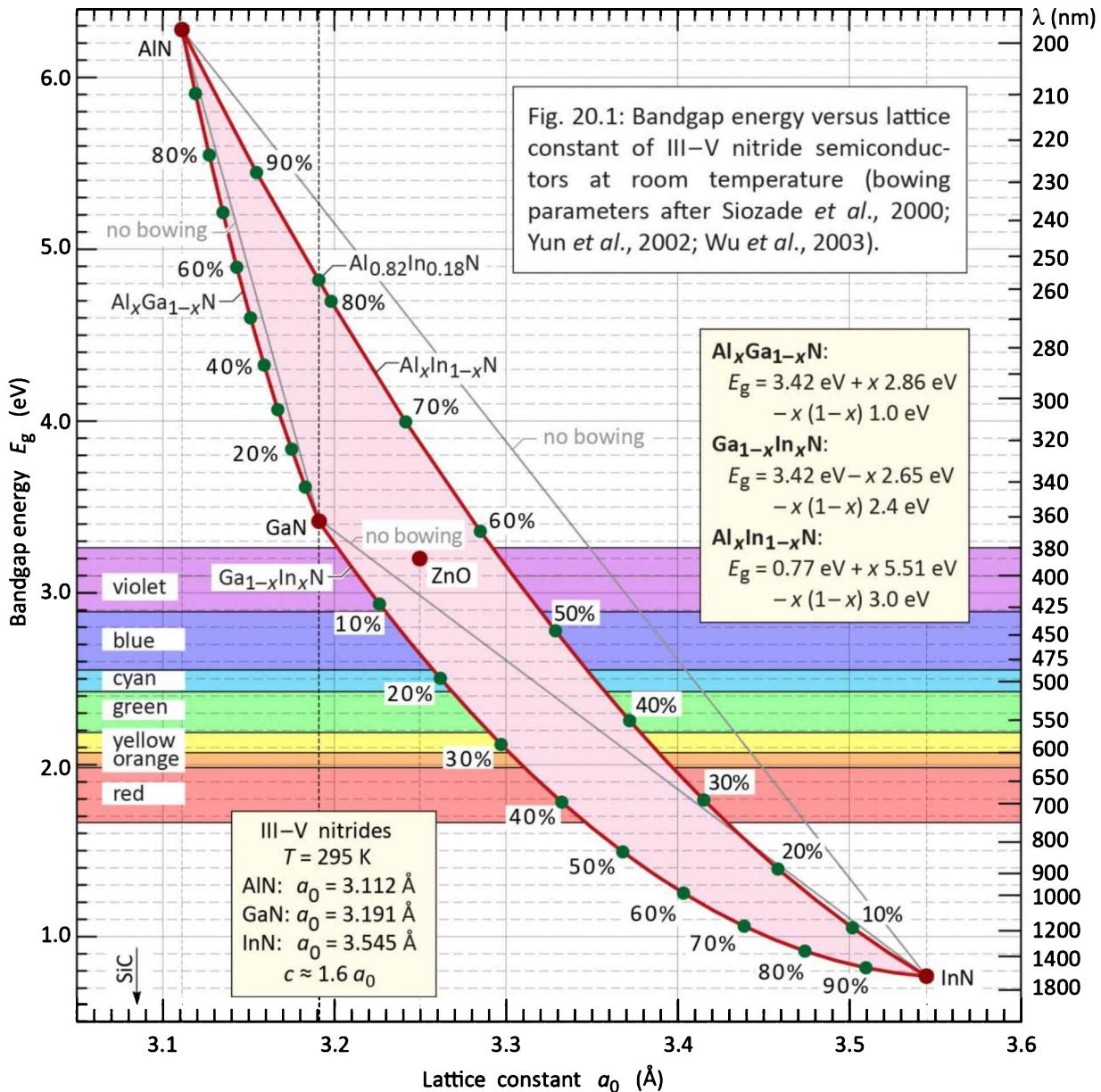
- Logan R. A., White H. G., and Trumbore F. A. "P-n junctions in compensated solution grown GaP" *Journal of Applied Physics* **38**, 2500 (1967a)
- Logan R. A., White H. G., and Trumbore F. A. "P-n junctions in GaP with external electroluminescence efficiencies ~2% at 25°C" *Applied Physics Letters* **10**, 206 (1967b)
- Logan R. A., White H. G., and Wiegmann W. "Efficient green electroluminescent junctions in GaP" *Solid State Electron.* **14**, 55 (1971)
- Mueller G. (editor) *Electroluminescence I, Semiconductors and Semimetals* **64** (Academic Press, San Diego, 1999)
- Mueller G. (editor) *Electroluminescence II, Semiconductors and Semimetals* **65** (Academic Press, San Diego, 2000)
- Nakamura S. and Fasol G. *The Blue Laser Diode* (Springer, Berlin, 1997)
- Nishizawa J., Koike M., and Jin C. C. "Efficiency of GaAlAs heterostructure red light emitting diodes" *J. Appl. Phys.* **54**, 2807 (1983)
- Nuese C. J., Stillman G. E., Sirkis M. D., and Holonyak Jr. N., "Gallium arsenide-phosphide: crystal, diffusion, and laser properties" *Solid State Electronics* **9**, 735 (1966)
- Nuese C. J., Tietjen J. J., Gannon J. J., and Gossenberger H. F. "Optimization of electroluminescent efficiencies for vapor-grown GaAsP diodes" *Journal of the Electrochemical Society: Solid State Science* **116**, 248 (1969)
- Pilkuhn M. and Rupprecht H. "Electroluminescence and lasing action in GaAsP" *Journal of Applied Physics* **36**, 684 (1965)
- Prins A. D., Sly J. L., Meney A. T., Dunstan D. J., O'Reilly E. P., Adams A. R., and Valster A. *Journal of Physics and Chemistry of Solids* **56**, 349 (1995)
- Steranka F. M., DeFevre D. C., Camras M. D., Tu C.-W., McElfresh D. K., Rudaz S. L., Cook L. W., and Snyder W. L. "Red AlGaAs light emitting diodes" *Hewlett-Packard Journal* p. 84 August (1988)
- Steranka F. M. "AlGaAs red light-emitting diodes" in *High Brightness Light Emitting Diodes* edited by G. B. Stringfellow and M. G. Crawford, *Semiconductors and Semimetals* **48** (Academic Press, San Diego, 1997)
- Stringfellow G. B. and Crawford M. G. (Editors) *High Brightness Light Emitting Diodes, Semiconductors and Semimetals* **48** (Academic Press, San Diego, 1997)
- Strite S. and Morkoc H., "GaN, AlN, and InN: A review" *Journal of Vacuum Science and Technology B* **10**, 1237 (1992)
- Tien P. K. Original version of the graph is courtesy of P. K. Tien of AT&T Bell Laboratories (1988)
- Toyoda Gosei Corporation, Japan, LED product catalog (2000)
- United Epitaxy Corporation, Taiwan, General LED and wafer product catalog (1999)
- Wolfe C. M., Nuese C. J. and Holonyak Jr. N. "Growth and dislocation structure of single-crystal Ga(AsP)" *Journal of Applied Physics* **36**, 3790 (1965)
- Wu J., Walukiewicz W., Yu K. M., Ager III J. W., Haller E. E., Lu H., Schaff W. J., Saito Y., and Nanishi Y. "Unusual properties of the fundamental bandgap of InN" *Applied Physics Letters* **80**, 3967 (2002a)
- Wu J., Walukiewicz W., Yu K. M., Ager III J. W., Haller E. E., Lu H., and Schaff W. J. "Small bandgap bowing in In_{1-x}Ga_xN alloys" *Applied Physics Letters* **80**, 4741 (2002b)

20

20 – The AlGaN material system

20.1 The AlGaN bandgap

The bandgap energy versus lattice constant of the AlGaN material system is shown in **Figure 20.1**. The AlGaN material system spans a very wide range of wavelengths covering the deep UV, near UV, visible, and even the near infrared spectral range.



Of the three binary semiconductors InN, GaN, and AlN, epitaxially grown GaN has been shown to be synthesizable with the highest quality. It has generally been difficult to synthesize Al-rich AlGaN alloys and In-rich GaInN alloys with high internal quantum efficiencies.

There has been some controversy with respect to the bandgap energy of InN, which initially had been found to be about 1.9 eV. Later, however, the bandgap energy was found to be much lower, namely 0.77 eV (Matsuoka *et al.*, 2002; Wu *et al.*, 2002).

The energy-gap bowing can be expressed in terms of a constant, a linear term ($\propto x$) and a non-linear term [$\propto x(1-x)$] according to

$$E_g^{AB} = E_g^A + (E_g^B - E_g^A)x - E_b x(1-x) \quad (20.1)$$

with E_b called the ***bowing energy*** or ***bowing parameter***. The bowing parameters used in **Figure 20.1** for AlGaN, GaInN, and AlInN are based on data published by Siozade *et al.* (2000), Yun *et al.* (2002), and Wu *et al.* (2002; 2003). Additional data on the bowing energies have become available (Walukiewicz *et al.*, 2004).

20.2 Crystal structure of III–V nitrides and origin of polarization

While group-IV semiconductors crystallize in the diamond structure, and III–V arsenides and phosphides crystallize in the zincblende structure, III–V nitrides preferentially crystallize in the wurtzite structure. All of these semiconductors have in common a tetrahedral bonding configuration, that is, atoms constituting the semiconductors have four bonds that point to the four corners of a tetrahedron.

The unit cell of the wurtzite structure is shown in **Figure 20.2** (adapted from Wikipedia, 2010). The wurtzite structure has two lattice constants, a and c , as indicated in the figure. The lattice constant c is the height of the unit cell, and a is the basal-plane lattice constant. Joining three unit cells allows one to form a hexagonal surface area, as shown in the figure. The hexagonally-shaped surface is a $+c$ -plane (Ga-face) or $-c$ -plane (N-face), as shown in the figure. Due to the hexagonal arrangement of atoms in the c -planes, the wurtzite structure is said to have hexagonal symmetry.

The wurtzite structure ***lacks inversion symmetry*** along the c -direction, that is, the physical and chemical properties of the $+c$ and $-c$ plane are different; we thus need to distinguish between these planes and the associated directions. For example, one can infer from **Figure 20.2**:

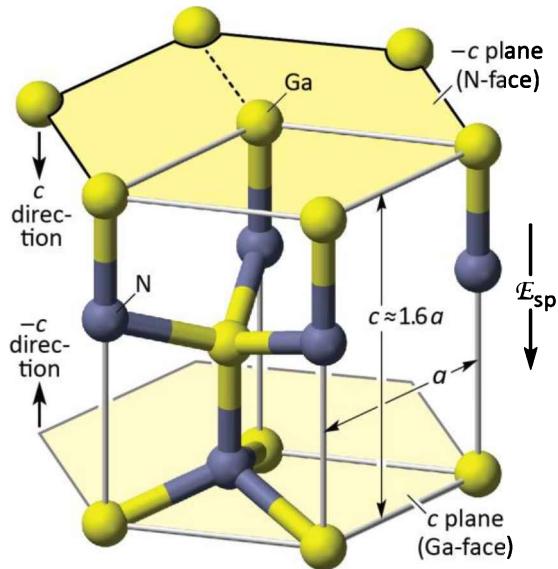


Fig. 20.2: Unit cell of GaN wurtzite crystal shown along with the hexagonal $-c$ -plane surface (top). In the wurtzite structure, Ga and N atoms are tetrahedrally bonded. An undistorted tetrahedron has an angle between the bonds of $\arccos(-1/3) = 109.47^\circ$. However, in the wurtzite structure, the tetrahedron is slightly distorted with respect to bond angle and length. Therefore, the polarizations of the three bonds pointing more horizontally no longer compensate the polarization of the vertical bond. As a result, a net polarization remains; it is denoted as the *spontaneous polarization* with an electric field E_{sp} . The polarization can increase when the crystal is strained, i.e. further distorted. The strain-induced polarization is denoted as *piezoelectric polarization* with an electric field E_{pz} (adapted from Wikipedia, 2010).

c-plane: On the $\{0001\}$ plane (c -plane or Ga-face plane), Ga atoms are bonded with three bonds (triply bonded) to the crystal. A triply bonded Ga atom is particularly stable since Ga inherently is trivalent;

$-c$ -plane: On the $\{000\bar{1}\}$ plane ($-c$ -plane or N-face plane), N atoms are bonded with three bonds (triply bonded) to the crystal. A triply bonded N atom is less stable since it has an unshared electron pair emerging from the surface plane;

Note that singly bonded atoms, such as a N atom on the c -plane (i.e. Ga face plane) or a Ga atom on the $-c$ plane (N-face plane), are inherently unstable.

These points show that the physical and chemical properties of Ga-face and N-face GaN are inherently different. A detailed discussion of these differences was given by Gatos and Lavine (1960) for III–V zincblende crystals. The high chemical stability of Ga-face GaN and the lower chemical stability of N-face GaN result from the above-listed differences in chemical bonding.

The GaN wurtzite structure is electrically polarized along the c -direction. That is, it has an electric field along the c -direction. The **polarization effect** has far-reaching consequences for devices made of this class of semiconductors. For this reason, we will explain the origin of the polarization effect below.

Consider the center Ga atom shown in **Figure 20.2**. It has four bonds, one pointing downward along the vertical direction, and three pointing along a close-to-horizontal direction. Since the Ga–N bond is located between two atoms having unequal electronegativity, the chemical bond

is naturally polarized, that is, the electron pair forming the bond tends to be closer to the more electronegative atom (i.e. the N atom). In an undistorted tetrahedrally bonded atom, the arrangement of the four bonds is symmetric so that the angle between any two bonds and is given by $\arccos(-1/3) = 109.47^\circ$. That is, in a semiconductor in which the atoms are bonded with *perfect* tetrahedral symmetry, the electric polarizations in the four bonds exactly compensate each other, thereby canceling out the polarized properties of the chemical bonds. That is, the vector sum of the four bonds inside one tetrahedron is exactly zero.

In wurtzite III–V nitride semiconductors, the tetrahedral bonding configuration is slightly distorted. For perfect tetrahedral symmetry in a wurtzite-structure semiconductor, the c/a lattice-constant ratio can be calculated to be $2\sqrt{2}/\sqrt{3} = 1.633$. In GaN, however, the tetrahedral bonding is distorted and the c/a ratio is 1.626 (Jeffrey *et al.*, 1956; Ueno *et al.*, 1994). The change in c/a ratio implies a distortion in (i) bond angle and (ii) bond length and both of these distortions have been suggested for the wurtzite lattice (Nann and Schneider, 2004; Yamanaka *et al.*, 2007; Karpov, 2010).

Let us consider a distortion in bond angle. Assume that the angle between the vertical bond and the other three bonds is less than the above-mentioned 109.47° that applies to the ideal tetrahedron (Nann and Schneider, 2004). As a result, the polarization of the three bonds pointing in the close-to-horizontal directions no longer fully compensates the polarization of the vertical bond. A similar argument can be made about the bond length: Since the vertical bond length is longer than the length of the other three bonds (Karpov, 2010), the polarization of the three bonds pointing along the close-to-horizontal direction no longer fully compensates the polarization of the vertical bond. Consequently, a net polarization remains; it is denoted as the ***spontaneous polarization***. It results in an electric field that is directed along the positive c -direction, as indicated in **Figure 20.2**.

When the wurtzite crystal is under mechanical strain, the tetrahedral bonding configuration is further distorted. The additional polarization caused by strain is denoted as the ***piezoelectric polarization***.

The natural epitaxial growth direction for wurtzite III–V nitrides is the $+c$ -direction (Ga-face). This growth direction results in high-quality crystals. III–V nitrides grown on the c -plane can be thought of having only two polarization charges (with the same magnitude of charge), one located at the $+c$ surface plane and the other one at the $-c$ surface plane. As a result of these

charges, internal electric fields occur in III–V nitrides that have a significant effect on the optical and electrical properties of the material.

The spontaneous polarization charges as well as strain-induced or piezoelectric polarization charges were calculated and are found in the literature (Bernardini *et al.*, 1997; Ambacher *et al.*, 1999; 2000; 2002). The direction of the internal electric field depends on the strain and the growth orientation (Ga-face or N-face). The polarization and electric field directions are illustrated for different cases in *Figure 20.3*.

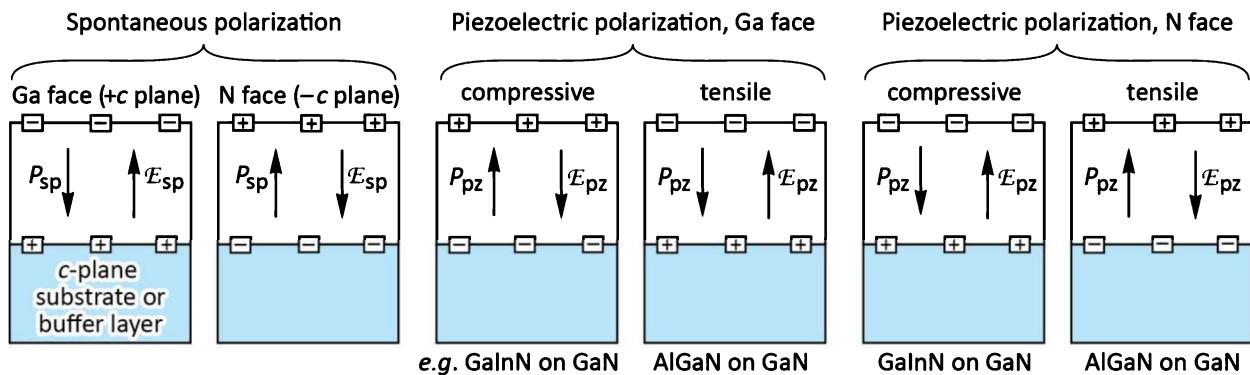


Fig. 20.3: Surface charges and direction of electric field and polarization field for spontaneous (sp) and piezoelectric (pz) polarization in III–V nitrides for Ga and N face orientation.

The **strain** in the epitaxial layer can be **compressive** or **tensile**. In the compressive-strain case, the epitaxial layer of interest is laterally compressed compared with its equilibrium configuration (“laterally” meaning “in the plane of the wafer”). For example, GaInN is compressively strained when grown on a thick relaxed GaN buffer layer (In is a larger atom than Ga). In the tensile-strain case, the epitaxial layer of interest is expanded along the lateral direction. For example, AlGaN is under tensile strain when grown on a thick relaxed GaN buffer layer (Al is a smaller atom than Ga).

The calculated magnitude of the spontaneous, piezoelectric, and total polarization charge for common III–V nitride alloys grown on relaxed GaN is shown in *Figure 20.4* (Gessmann *et al.*, 2002a and 2002b).

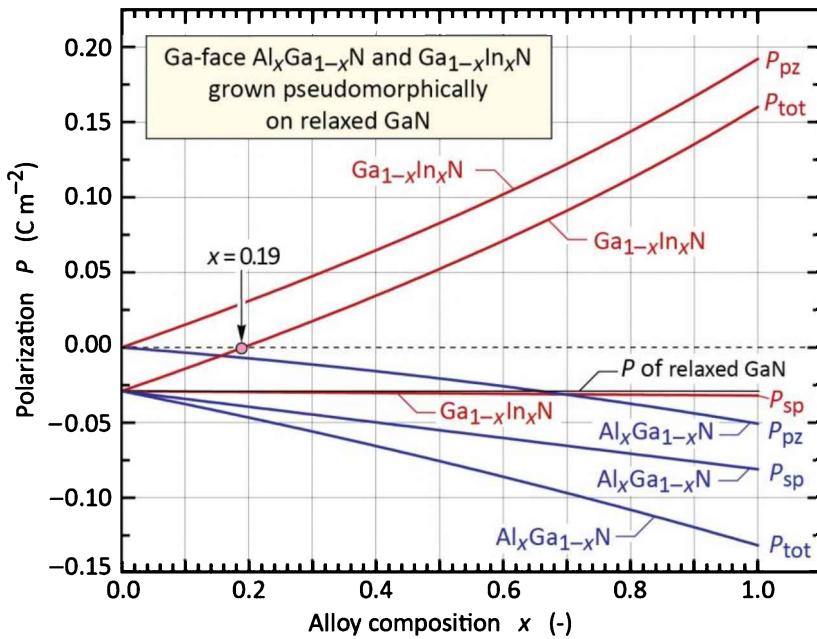


Fig. 20.4: Magnitude and direction of spontaneous and piezoelectric polarization in GaInN and AlGaN grown pseudomorphically on relaxed GaN. Relaxed GaN has a spontaneous polarization, but no piezoelectric polarization (after Gessmann *et al.*, 2002a and 2002b).

Another illustration of the GaN wurtzite crystal structure is shown in *Figure 20.5* along with significant crystal planes of the material (Ryou *et al.* 2009). These planes include the *m*-plane, *a*-plane, and *r*-plane. The *a*-plane is perpendicular to the *a*-axis, as indicated in the figure. Likewise, the *c*-plane is perpendicular to the *c*-axis.

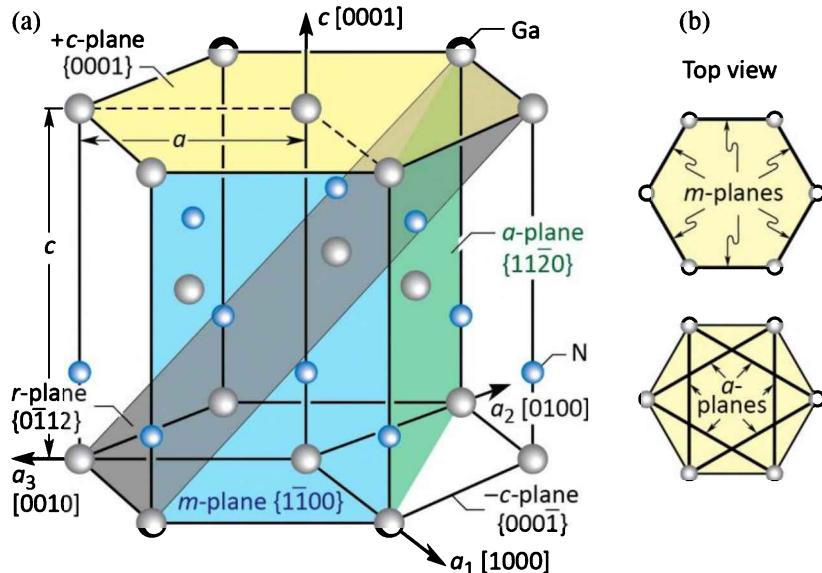


Fig. 20.5: (a) GaN lattice of the wurtzite structure including lattice constants (*a* and *c*) and relevant crystal planes. The crystal shown is Ga-face and Ga-terminated (top). The wurtzite structure lacks inversion symmetry along the *c*-direction. As a consequence, the *+c* plane and *-c* plane have different properties with respect to electrical polarization charges, crystal growth, and wet chemical etching (after Ryou *et al.*, 2009). (b) Top view showing crystal planes (atom in center of hexagon not shown).

As discussed above, epitaxial GaN has a polarization field along the *c*-growth-direction. However, there is no polarization field along the *m*-plane and *a*-plane direction, since the bond vector sum has only a *c*-component. For this reason, the *m*-plane and *a*-plane are non-polar growth

directions and the *r*-plane is a semi-polar growth direction. The epitaxial growth on these planes allows for the growth of LEDs that have no electric field along the growth direction and thus no quantum-confined Stark effect (Waltereit *et al.*, 2000). For this reason, the growth on non-polar surfaces has attracted much attention (Kim *et al.*, 2007). It has been found, however, that, overall, the properties of non-polar and semi-polar GaN are not superior to polar (*c*-plane) GaN, so that commercial GaN LEDs are based on *c*-plane GaN. Relevant crystal planes and their polarity are given in **Table 20.1**.

Table 20.1: Miller indices of some crystal planes in wurtzite GaN and their polarity.

<i>Crystal plane</i>	<i>Miller indices</i>	<i>Polarity</i>
<i>c</i> -plane	(0001)	Polar
<i>m</i> -plane	{1̄100}	Non-polar
<i>a</i> -plane	{11̄20}	Non-polar
<i>r</i> -plane	{11̄02}	Semi-polar

20.3 Consequences of polarization effects in III–V nitrides

Polarization fields affect carrier-transport and optical properties of LEDs. A consequence of polarization fields for radiative recombination in quantum well structures is shown in **Figure 20.6**. The quantum well layers have an internal electric field that spatially separates electrons and holes, thereby preventing efficient radiative recombination.

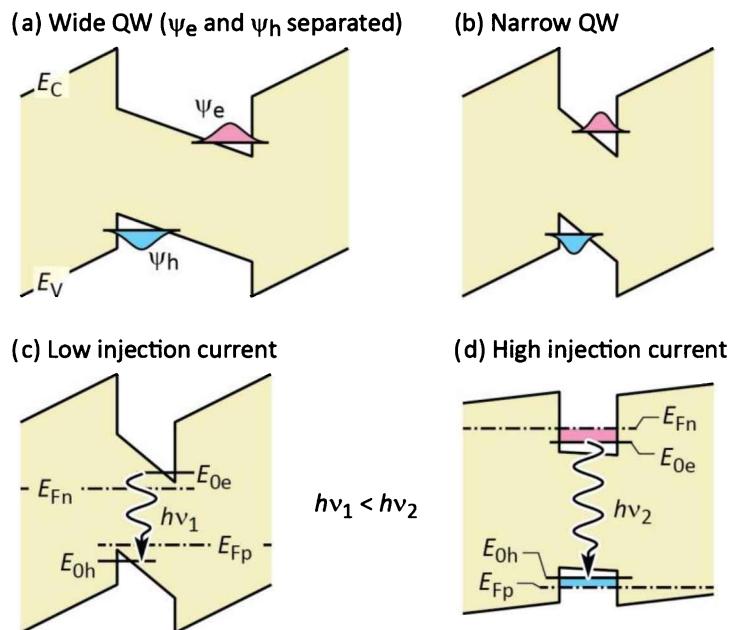


Fig. 20.6: Schematic band diagram of (a) thick and (b) thin GaInN/GaN active regions with polarization fields for Ga-face growth (substrate on left-hand side). (c) and (d): Illustration of the blue shift occurring with increasing injection current when the carrier concentration in the QW increases. High carrier concentrations screen polarization charges, flatten the bands (reduction of QCSE), and blue-shift of the emission ($h\nu_1 < h\nu_2$).

This is particularly true for thick quantum wells, e.g. > 10 nm, shown in *Figure 20.6 (a)*, but less so for narrow quantum wells, shown in *Figure 20.6 (b)*. Furthermore, the energy of the radiative transition is lowered by the electric field, as shown in *Figure 20.6 (c)*. The spectral emission band shifts to longer wavelengths as the electric field increases. The lower efficiency and lower emission energy are collectively referred to as the ***quantum-confined Stark effect*** (QCSE). The original Stark effect is named after Johannes Stark who discovered in 1913 that the spectral lines of an atom shift to lower energy when exposed to a static electric field. The QCSE refers to basically the same effect when occurring in a quantum well.

To alleviate this deleterious effect, it is imperative that the quantum well layers are kept very thin. Quantum well thicknesses of 2~3 nm (20~30 Å) are typically used to minimize such electron–hole separation effects. Nevertheless, a spectral shift occurs in GaInN/GaN QWs as the current density (or photoluminescence excitation density) increases (Takeuchi *et al.*, 1997, 1998). The shift saturates at high current densities, when the QW contains a high density of carriers that screen the electric field caused by polarization charges, as suggested in *Figure 20.6 (d)*.

The large electric fields caused by the polarization effects can be screened by a high free-carrier concentration, which can be attained through either (*i*) high doping of the active region or (*ii*) a high injection current. Screening of the internal electric field also results in a blue-shift of the emission, frequently found in GaInN LEDs as the injection current is increased (Takeuchi *et al.*, 1997). The spectral shift is illustrated in *Figure 20.7* for a GaInN/GaN LED (Lee, 2011). The spectral shift can be a problem with respect to color stability of an LED, because it is undesirable that the emission wavelength of a device changes when the current is increased.

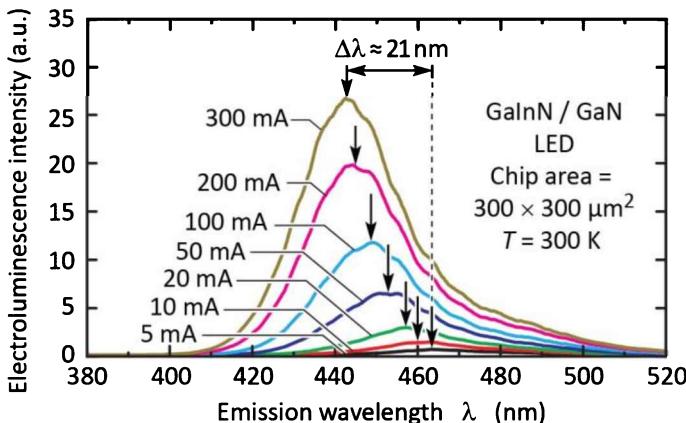


Fig. 20.7: Emission spectra of a GaInN/GaN LED for different injection currents. As the current increases, the emission shifts to shorter wavelengths (blue shift) due to screening of polarization charges (after Lee, 2011).

Although polarization fields result in the unwanted QCSE, there are also device structures that may benefit from polarization fields. For example, polarization fields can be used advantageously to reduce ohmic contact resistances in GaN devices. Polarization-enhanced contacts to p-type GaN employ a thin, compressively strained GaInN cap layer that is deposited on the p-type GaN. The electric field in the GaInN cap layer is polarized in such a way that the tunneling probability of holes through the contact barrier is enhanced (Li *et al.*, 2000; Gessmann *et al.*, 2002a and 2002b). Contact resistances as low as 1.1×10^{-6} and $2 \times 10^{-7} \Omega\text{cm}^2$ have been reported for non-annealed and annealed polarization-enhanced ohmic contacts to p-type GaN, respectively (Kumakura *et al.*, 2001; 2003). Both the concept of polarization-enhanced contacts and the very low specific contact resistances are remarkable because they represent a significant improvement over conventional contact technologies. The forward voltage of GaN-based LEDs with polarization-enhanced contacts is generally lower than the forward voltage of LEDs with conventional GaN contacts (Su, 2005).

20.4 P-type doping in III–V nitrides

While n-type doping works reasonably well in III–V nitrides, p-type doping is highly problematic in due to the following facts:

- (i) Acceptors in III–V nitrides have a **high thermal activation energy** which is $\gg kT$ at 300 K. As a result, only a small percentage of acceptors, typically less than 10%, are ionized at room temperature.
- (ii) Acceptors in III–V nitrides can be **chemically passivated by hydrogen atoms** that form a dopant–hydrogen complex. The electron that acceptors tend to accept from the valence band is provided by a hydrogen atom thereby passivating the acceptor.

These two effects will be discussed below.

In III–V arsenides and phosphides, acceptors levels are shallow and thus located close to the valence band edge. For example, C and Be acceptors in GaAs have an ionization energy of 26 and 28 meV, respectively, which is on the order of kT at room temperature ($kT = 26$ meV). As a result, most of the acceptors are ionized, as illustrated in *Figure 20.8*.

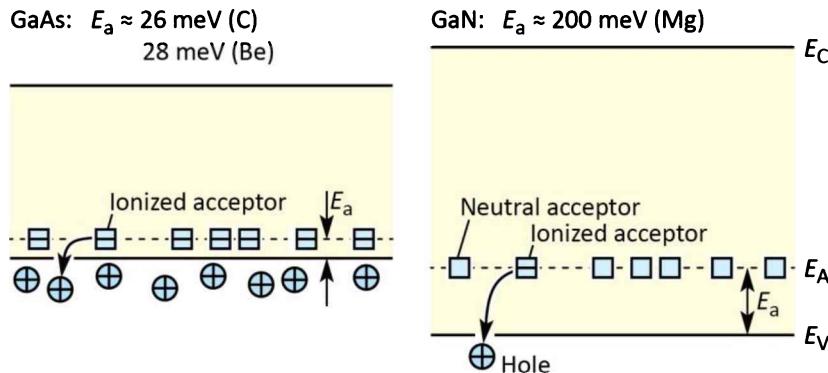


Fig. 20.8: Schematic band diagram of p-type GaAs and GaN at room temperature. In GaAs, the acceptor activation energy is low thus permitting high hole concentrations. In GaN, however, the acceptor activation energy is high which results in low hole concentrations.

The commonly used acceptor in GaN is the group-II element magnesium (Mg) which substitutes for Ga thereby forming a substitutional acceptor. In GaN, the Mg acceptor level is located about $E_a \approx 200 \text{ meV}$ above the valence band edge, as shown in **Figure 20.8**. Since $E_a \gg kT$, only a small fraction of acceptors are ionized at room temperature. The ionized hole concentration can be calculated by (Schubert *et al.* 1996)

$$p = \sqrt{(N_A/g) N_v} \exp\left(-\frac{E_a}{2kT}\right) \quad (20.2)$$

where N_A is the acceptor concentration, N_v is the effective density of states at the valence band edge, and g is the ground-state degeneracy of the acceptor. Evaluation of the equation for $T = 300 \text{ K}$ yields that the hole concentration is less than 10% of the acceptor concentration. That is, only a small fraction of the acceptors are ionized at room temperature and the free hole concentration in GaN is generally low. Typical hole concentrations in highly Mg doped GaN are in the 10^{17} cm^{-3} range.

Exercise: Activation of Mg acceptors in GaN. Mg acceptors in GaN have an activation energy of $E_a = 200 \text{ meV}$. (a) Calculate the fraction of acceptors that are ionized at 300 K for an acceptor concentration of $N_{\text{Mg}} = 10^{18} \text{ cm}^{-3}$ using the formula $p = (g^{-1} N_{\text{Mg}} N_v)^{1/2} \exp(-E_a/2kT)$ where g is the acceptor ground state degeneracy ($g = 4$) and N_v is the effective density of states at the valence band edge of GaN. (b) What would be the activation of acceptors if a hydrogen atom were bonded to each acceptor?

Solution: (a) Using the formula given above, one obtains that only about 6% of the acceptors are ionized. (b) If acceptors are passivated, p-type conductivity cannot be established.

The Mg acceptor ionization energy in GaN, GaInN, and AlGaN is shown in **Figure 20.9** (Kumakura *et al.*, 2000). Inspection of the figure reveals that the acceptor ionization energy generally

decreases in GaInN (as the In mole fraction increases) and increases in AlGaN (as the Al mole fraction increases).

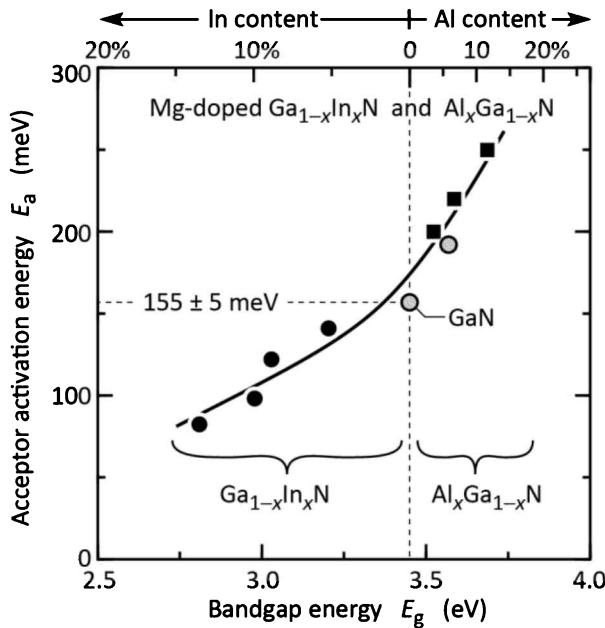


Fig. 20.9: Thermal activation energy of Mg acceptors in GaInN and AlGaN as a function of the molar In and Al content, respectively (after Kumakura *et al.*, 2000).

Mg acceptor thermal activation energy:

in GaN: $E_a = 155$ meV
(after Kumakura *et al.*, 2000)

in AlN: $E_a = 510 \sim 630$ meV
(after Nam *et al.*, 2003 and Taniyasu *et al.*, 2006)

The p-type doping problem is particularly severe in AlGaN and AlN, because the acceptor ionization energy increases with Al content. For AlN, Mg acceptor ionization energies ranging between 510 and 630 meV have been reported (Nam *et al.*, 2003 and Taniyasu *et al.*, 2006). For AlGaN with high Al content, conventional p-type doping produces hole concentrations to values $< 10^{15} \text{ cm}^{-3}$ which are too low for efficient LEDs.

In order to overcome the p-type doping limitation of Mg-doped GaN, very high Mg doping concentrations, $N_{Mg} > 10^{19} \text{ cm}^{-3}$, have been employed. However, at such high concentrations, Mg atoms form defects (such as Mg occupying an interstitial site or Mg forming a defect complex). As a consequence, many Mg atoms incorporated into GaN do not form a substitutional acceptor. In Ga_{0.96}In_{0.04}N with $N_{Mg} = 3 \times 10^{19} \text{ cm}^{-3}$, it has been estimated that only about 17% of Mg atoms form substitutional acceptors (Kumakura *et al.*, 2000). For GaN with Mg doping concentrations exceeding $6 \times 10^{19} \text{ cm}^{-3}$, it may be assumed that less than 10% of the Mg atoms form substitutional acceptors.

The doping of AlGaN/GaN and AlGaN/AlGaN superlattices (rather than GaN or AlGaN bulk material) was postulated to strongly increase the electrical activation of Mg acceptors (Schubert *et al.*, 1996). Experimentally demonstrated Mg-doped superlattices indeed increased the p-type

conductivity by a factor of 10 or more, as reported by several research groups (Goepfert *et al.*, 1999, 2000; Kozodoy *et al.* 1999a, b; Kipshidze *et al.*, 2002, 2003).

Historically, the initial demonstration of p-type conductivity in GaN has been very difficult due to ***passivation of p-type dopants*** by atomic hydrogen (H) that is generated during the epitaxial growth of GaN by vapor-phase epitaxy. Sources of H atoms and H passivation include the methyl ($-\text{CH}_3$) and ethyl groups ($-\text{C}_2\text{H}_5$) of the organo-metallics, ammonia (NH_3), as well as hydrogen from the H_2 carrier gas.

The atomic model of H passivation of a Mg acceptor in GaN is shown in *Figure 20.10*. Mg occupies a Ga site in GaN and the Mg acceptor can be passivated by a H atom that bonds to a N atom neighboring the Mg atom (see, for example, Götz *et al.*, 1995; Nickel, 1999; Van de Walle and Neugebauer, 2006). Since the Mg–N bonds no longer lack an electron, the Mg acceptor no longer strives to accept a valence electron. Thus, the Mg acceptor is passivated.

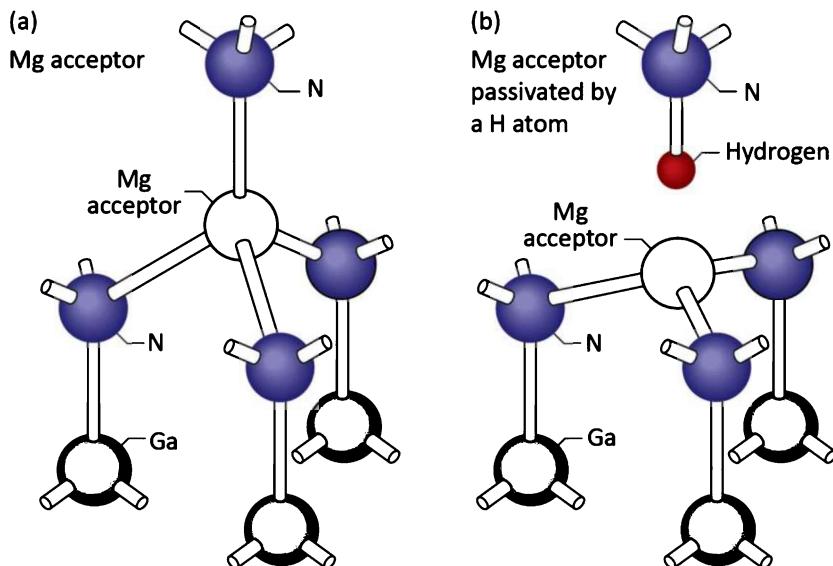


Fig. 20.10: Schematic representation of (a) a substitutional Mg acceptor on a Ga site of the GaN lattice and (b) a Mg acceptor passivated by a H atom (Mg–H complex). The Mg–H complex can be de-passivated by a high temperature anneal in a H_2 -free ambient, which causes the weak H–N bond to break and allows the H atoms to diffuse out of the GaN.

Amano *et al.* (1989) discovered that acceptor dopants can be activated (i.e. de-passivated) by low-energy electron-beam irradiation (LEEBI). This important discovery allowed for the first demonstration of p-type GaN and opened the door to the development and commercialization of GaN pn-junction technology. Subsequently, Nakamura *et al.* (1991; 1992) showed that acceptors can be activated by both, LEEBI as well as by thermal annealing. The heating caused by LEEBI is believed to have the same effect as thermal annealing. Thermal annealing is the preferred method since it is a simpler and more efficient method than LEEBI. Typical annealing conditions for MOCVD-grown p-type GaN are $675\text{--}725^\circ\text{C}$ for 5 minutes in an N_2 atmosphere;

p-type $\text{Al}_{0.30}\text{Ga}_{0.70}\text{N}$ is annealed at higher temperatures, typically at 850°C for 1~2 minutes, also in an N_2 atmosphere. Ambients other than N_2 , particularly O_2 -containing ambients, have been reported to be slightly more efficient in driving hydrogen out of GaN (Chung *et al.*, 2000; Kuo *et al.*, 2002).

During thermal annealing the relatively weak H–N bond is broken and the H atoms are driven out of the epitaxial film, as schematically illustrated in **Figure 20.11**. Hydrogen atoms are small and generally diffuse easily through the interstitial sites of a crystalline material.

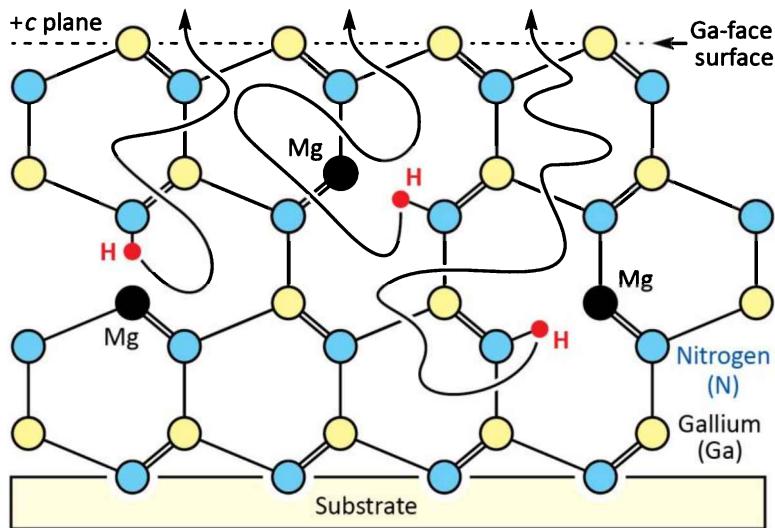


Fig. 20.11: Schematic representation of hydrogen (H) expulsion by thermal annealing of Mg-doped p-type GaN (Ga-terminated Ga-face surface for growth along c direction is shown). The H expulsion results in de-passivation of p-type GaN. The annealing is performed in an H_2 -free ambient (frequently N_2 or O_2).

20.5 Epitaxial growth of GaN on sapphire

A common substrate for GaN epitaxial growth, sapphire, is a very stable substrate in terms of its thermal, chemical, and mechanical properties. However, sapphire has the complex **corundum structure** whereas III–V nitrides crystallize in the **wurtzite structure**. Furthermore, the lattice constants of sapphire and GaN are different. As a result, GaN epitaxial films have misfit dislocations (threading and edge dislocations) that are typically on the order of $10^8\text{--}10^9 \text{ cm}^{-2}$.

Due to the mismatch between sapphire substrates and GaN epitaxial layers in terms of (i) lattice structure and (ii) lattice constant, the epitaxial growth of GaN on a sapphire substrate requires a special growth procedure that is frequently referred to as a **two-step growth process** (Amano *et al.*, 1986; 1989). In the two-step growth process, a thin buffer layer (also called nucleation layer), typically 50 nm thick, is grown on the sapphire substrate at low temperatures, e.g. 550°C . Subsequently, the wafer is heated to the typical growth temperature of GaN,

e.g. 1050°C, to initiate the GaN epitaxial growth. If the low-temperature buffer layer is *too thin*, it will not completely “wet” the substrate and the buffer layer will form islands. If the buffer layer is *too thick*, the epilayer becomes “de-coupled” from the substrate so that the GaN epitaxial layer will not grow coherently on the substrate. As a result, epitaxial growth cannot be expected to occur as a single-crystal across the entire wafer. Therefore, the thickness of the buffer layer and the temperature profile of the two-step growth process are critical parameters that need to be optimized to get high-quality GaN epitaxial growth (Amano *et al.*, 1986; 1989). **Figure 20.12** shows scanning electron micrographs of GaN grown on sapphire without a buffer layer and with an AlN buffer layer (Amano *et al.*, 1986; 1989; Amano, 2012).

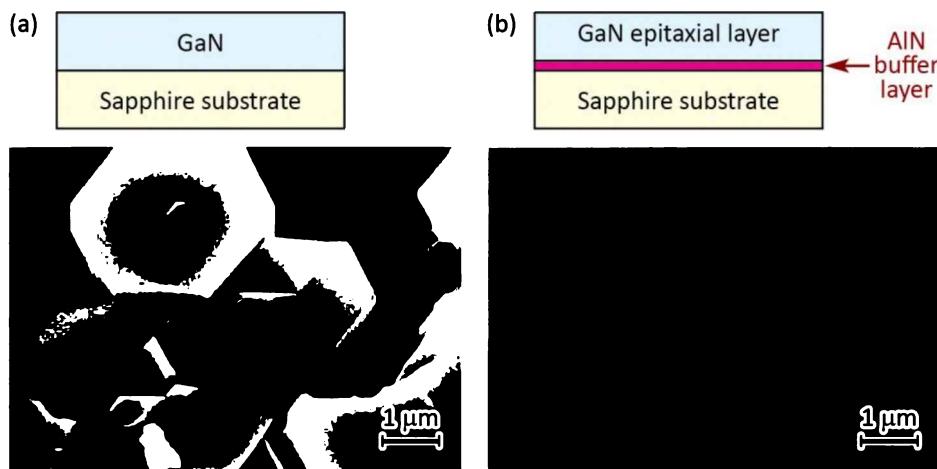


Fig. 20.12: Epitaxially grown GaN layers (a) without a buffer layer and (b) with a very thin (50 nm) AlN buffer layer grown at a low temperature, 600°C, followed by GaN grown at a higher temperature, 1040°C (Amano *et al.*, 1986, 1989; Amano, 2012).

The low-temperature buffer layer can be formed of AlN (Amano *et al.*, 1986) as well as of GaN (Nakamura, 1991). Both types of buffer layer are suited to attain high-quality epitaxial layers. For growth on sapphire, GaN buffer layers are preferred for GaN-based devices such as GaInN/GaN blue LEDs, and AlN buffer layers are preferred for AlGaN-based devices such as AlGaN/GaN UV LEDs.

The initial stages of GaN growth on sapphire are shown schematically in **Figure 20.13** (Nakamura and Fasol, 1997). The initial layer (called the faulted zone), grown at low temperatures (about 500°C) and subsequently annealed, is highly dislocated. However, dislocations undergo self-annihilation during the anneal, so that subsequent layers (called the semi-sound zone and the sound zone) have much lower dislocation densities.

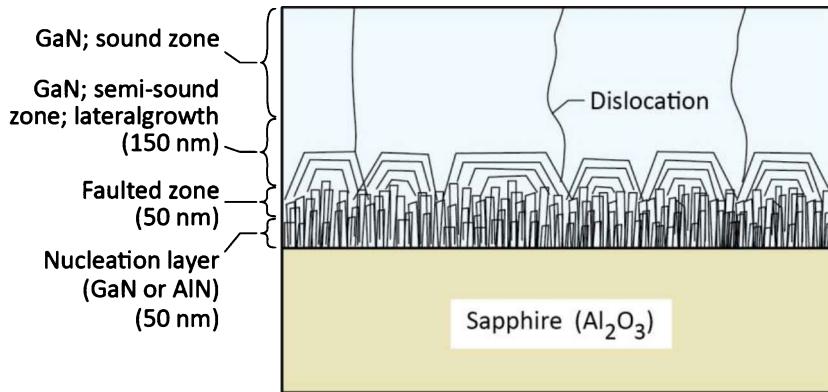


Fig. 20.13: Dislocation structure of a GaN epitaxial layer grown on sapphire by means of the GaN or AlN nucleation layer developed by Amano *et al.* (1986). The nucleation layer is normally grown at 500°C, much lower than the following GaN epitaxial layers (after Nakamura and Fasol, 1997).

A thorough review of the initial stages of GaN epitaxial growth on sapphire and an analysis of these initial stages using atomic-force microscopy and optical reflectometry was published by Koleske *et al.* (2004). The initial stages of epitaxial growth are particularly sensitive to the choice of growth parameters since the sapphire substrate and III–V nitride epitaxial layers have a different crystal structure as well as a different lattice constant. **Figure 20.14** shows the optical reflectance of the wafer during the initial stages of epitaxial growth of GaN on sapphire. The optical reflectance is measured by a light source that reflects off the wafer surface. The evolution of the reflectance includes features that can be attributed to the higher refractive index of GaN (compared with sapphire), and to a change in surface morphology (decomposition and nuclei formation) that makes the optical characteristics of the crystal surface more diffuse.

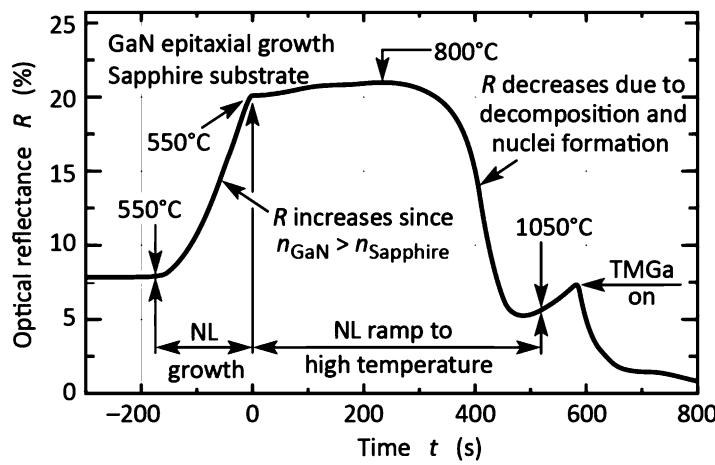


Fig. 20.14: Optical reflectance profile showing GaN nucleation layer (NL) growth, annealing, and initial high-temperature GaN growth. During the process, the reflectance initially increases and then decreases. After holding at 1050 °C for 60 s, TMGa is turned on for high-temperature growth (after Koleske *et al.*, 2005).

20.6 Dislocations in III–V nitrides

The most common substrate for GaN epitaxial growth, sapphire, is a very stable substrate in terms of its thermal, chemical, and mechanical properties. However, sapphire has the complex **corundum structure** whereas III–V nitrides crystallize in the **wurtzite structure**. Furthermore, the

lattice constants of sapphire and GaN are different. As a result, GaN epitaxial films grown on sapphire have unavoidable misfit dislocations (threading and edge dislocations) that are typically on the order of $10^8 \sim 10^9 \text{ cm}^{-2}$.

Generally, dislocation lines are electrically charged so that the region surrounding a dislocation line is either coulombically attractive or repulsive to a free carrier. The nature of the coulombic interaction (attractive or repulsive) depends on the polarity of the dislocation line and the polarity of the carrier. As an example, **Figure 20.15** shows a negatively charged dislocation line which is attractive to holes and repulsive to electrons.

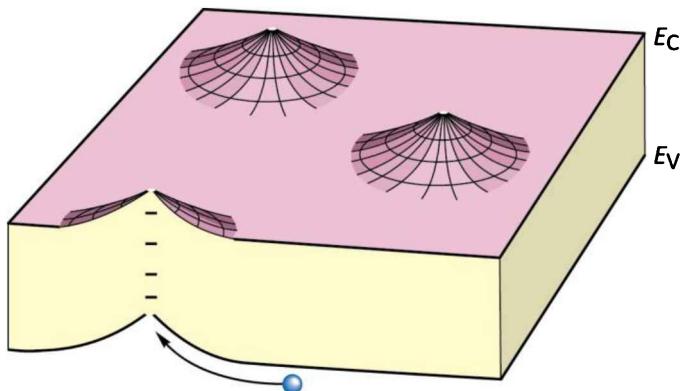


Fig. 20.15: Band diagram of semiconductor having negatively charged dislocations. Holes are attracted to dislocation lines where they must ultimately recombine with electrons.

Figure 20.16 shows the temporal development of the carrier dynamics of a positively charged dislocation line. Initially, electrons are attracted but holes are repelled due to the potential created by the dislocation. However, the continued collection of electrons will screen the dislocation potential thereby reducing the repulsive barrier for holes. As a result, electrons and holes will recombine non-radiatively via electron states of the dislocation line. In **Figure 20.16**, the electronic states of the dislocation are located within the bandgap.

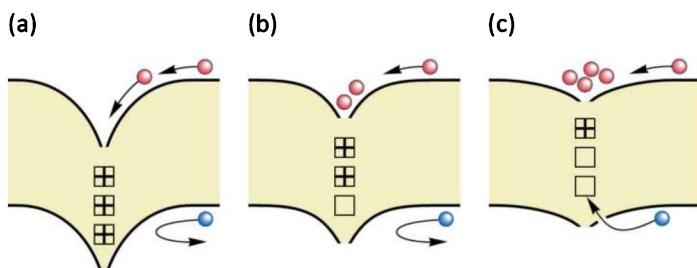


Fig. 20.16: Recombination in a positively charged dislocation. (a)–(c) Sequence shows electrons accumulating in the potential minimum thereby screening the dislocation potential and allowing holes to recombine.

A puzzling question is as to why the radiative recombination efficiency in III–V nitrides is so high despite the high density of dislocations? Several possible explanations are discussed below.

However, none of these explanations has gained general acceptance.

One possible explanation for high radiative rates in III–V nitrides is that the electronic states of the dislocation do not lie inside the forbidden gap, as shown in *Figure 20.17* (a) and (b), but lie *outside* the forbidden gap, i.e. within the allowed bands of the semiconductor, as shown in *Figure 20.17* (c) and (d). This explanation is not in conflict with dark regions surrounding dislocations, observed in cathodoluminescence experiments (Rosner *et al.*, 1997; Albrecht *et al.*, 2002). Such dark regions unequivocally show the *absence* of *radiative* recombination but do not necessarily prove the *presence* of *non-radiative* recombination. The dark regions observed in cathodoluminescence could be explained by the incomplete screening of the dislocation potential, which would result in the repulsion of either electrons or holes, and thus result in the absence of radiative recombination.

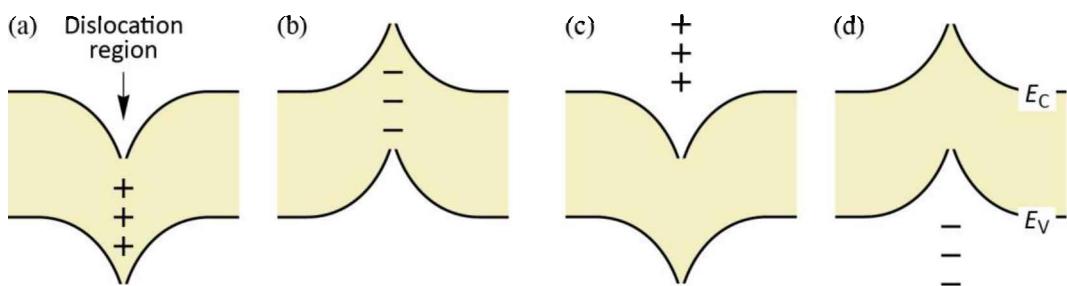


Fig. 20.17: Band diagram of dislocation: (a) donor states in gap, (b) acceptor states in gap, (c) donor states in conduction band, and (d) acceptor states in valence band. Whereas (a) and (b) lead to non-radiative recombination, (c) and (d) no not.

Another explanation for high radiative rates in III–V nitrides is compositional alloy fluctuations, alloy clustering effects (specifically In-rich GaInN clusters), and phase separation effects, all of which result in a variation of the bandgap energy and lead to local potential minima, which in turn can attract and confine carriers (Chichibu *et al.*, 1996; Nakamura and Fasol, 1997; Narukawa *et al.*, 1997a, b; Chichibu *et al.*, 2006). This explanation would be particularly suitable for ternary and quaternary alloy semiconductors such as GaInN and AlGaN. The potential minima attract and confine carriers and prevent them from diffusing towards the dislocation lines where they would recombine non-radiatively. A band diagram showing energy-gap fluctuations and ***carrier localization*** is schematically illustrated in *Figure 20.18*. However, the explanation is limited in its applicability to GaInN active regions. The model cannot explain the high radiative efficiency exhibited by GaN despite a high dislocation density (Lester *et al.*, 1995).

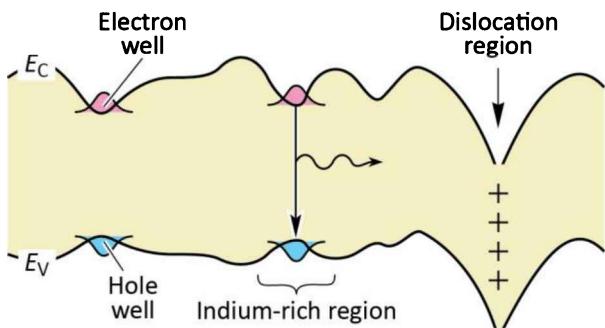


Fig. 20.18: Band diagram of GaInN having clusters of In rich regions which spatially localize carriers and prevent them from diffusing to dislocations.

Furthermore, because the direct detection of indium-composition fluctuations in GaInN by transmission-electron microscopy was shown to be difficult and even ambiguous due to high-energy electron-beam induced damage (Smeeton *et al.*, 2003), the magnitude of the indium fluctuations in GaInN still is under discussion.

Hangleiter *et al.* (2005) pointed out that the V-shaped defect that can occur at a dislocation line-defect of a GaInN/GaN quantum-well, can reduce the thickness of the quantum well. The authors proposed that the higher bandgap associated with the thinner region of the GaInN quantum well shields the dislocation line-defect from mobile carriers residing in the thicker region of the quantum well. The authors suggested that this effect explains why GaInN/GaN quantum-well structures, despite a high density of dislocation line-defects, have a high radiative efficiency. However, the suggestion that a defect (V-defect) would *enhance* radiative efficiency is counter-intuitive and contradicts conventional wisdom. In direct contradiction to Hangleiter *et al.*, Detchprohm *et al.* (2008, 2009) showed that V-defects *reduce* the efficiency of GaInN/GaN LEDs. Furthermore, the model proposed by Hangleiter *et al.* cannot explain the well-known high radiative efficiency of GaN thin films.

Although a generally accepted explanation has not yet been established, it remains a fact that the radiative efficiency in III–V nitrides, in particular GaInN/GaN blue emitters, exhibits low sensitivity to the presence of dislocations. That is, high radiative efficiencies are obtained in GaInN/GaN blue emitters despite high dislocation densities. This is illustrated in **Figure 20.19** which compares the normalized efficiency of different III–V semiconductors as a function of the dislocation density in GaAs, AlGaAs, GaP, and GaAsP. Lester *et al.* (1995) estimated a 4% radiative efficiency for GaN with a dislocation density of 10^{10} cm^{-2} . The shaded region is the estimation of the author of this book based on blue GaInN/GaN emitters. The data shown in the figure elucidate that III–V nitrides have a much higher tolerance towards dislocations as compared to

III–V arsenides and phosphides.

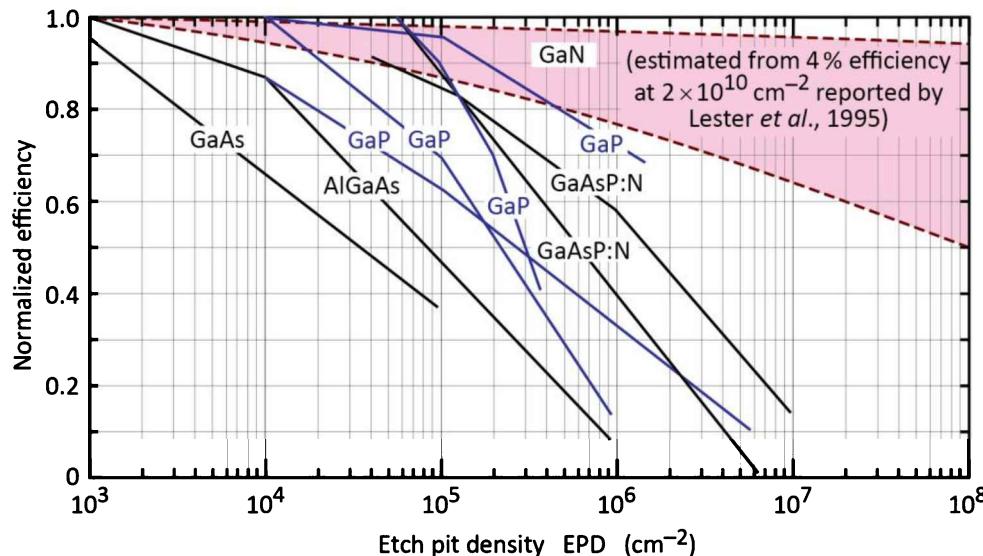


Fig. 20.19: Dependence of radiative efficiency on etch pit density (III–V arsenide and phosphide data adopted from Lester *et al.*, 1995; III–V nitride data estimated by the author of this book based on data published by Lester *et al.*, 1995).

20.7 References

- Albrecht M., Strunk H. P., Weyher J. L., Grzegory I., Porowski S., and Wosinski T. "Carrier recombination at single dislocations in GaN measured by cathodoluminescence in a transmission electron microscope" *Journal of Applied Physics* **92**, 2000 (2002)
- Amano H., Sawaki N., Akasaki I., and Toyoda Y. "Metalorganic vapor phase epitaxial growth of a high quality GaN film using an AlN buffer layer" *Applied Physics Letters* **48**, 353 (1986)
- Amano H., Kito M., Hiramatsu K., and Akasaki I. "P-type conduction in Mg-doped GaN treated with low-energy electron beam irradiation (LEEBI)" *Japanese Journal of Applied Physics* **28**, L2112 (1989)
- Amano Hiroshi, personal communication. The provision of photographs showing GaN epitaxial layers grown with and without an AlN buffer layer is gratefully acknowledged (2012)
- Ambacher O., Smart J., Shealy J. R., Weimann N. G., Chu K., Murphy M., Schaff W. J., Eastman L. F. Dimitrov R., Wittmer L., Stutzmann M., Rieger W. and Hilsenbeck J. "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization charges in N- and Ga-face AlGaN/GaN heterostructures" *Journal of Applied Physics* **85**, 3222 (1999)
- Ambacher O., Foutz B., Smart J., Shealy J. R., Weimann N. G., Chu K., Murphy M., Sierakowski A. J., Schaff W. J., Eastman L. F., Dimitrov R., Mitchell A., and Stutzmann M. "Two-dimensional electron gases induced by spontaneous and piezoelectric polarization in undoped and doped AlGaN/GaN heterostructures" *Journal of Applied Physics* **87**, 334 (2000)
- Ambacher O., Majewski J., Miskys C., Link A., Hermann M., Eickhoff M., Stutzmann M., Bernardini F., Fiorentini V., Tilak V., Schaff W. J., and Eastman L. F. "Pyroelectric properties of Al(In)GaN/GaN hetero- and quantum well structures" *Journal of Physics: Condensed Matter* **14**, 3399 (2002)
- Bernardini F., Fiorentini V., and Vanderbilt D. "Spontaneous polarization and piezoelectric constants of III–V nitrides" *Phys. Rev. B* **56**, R10 024 (1997)
- Chichibu S., Azuhata T., Sota T., and Nakamura S. "Spontaneous emission of localized excitons in InGaN single and multiquantum well structures" *Applied Physics Letters* **69**, 4188 (1996)
- Chichibu S. F., Uedono A., Onuma T., Haskell B. A., Chakraborty A., Koyama T., Fini P. T., Keller S., DenBaars S. P., Speck J. S., Mishra U. K., Nakamura S., Yamaguchi S., Kamiyama S., Amano H., Akasaki I., Han J. and Sota T. "Origin of defect-insensitive emission probability in In-containing (Al,In,Ga)N alloy semiconductors" *Nature Materials* **5**, 810 (2006)
- Chung S. H., Lachab M., Wang T., Lacroix Y., Basak D., Fareed Q., Kawakami Y., Nishino K., and Sakai S.

- "Effect of oxygen on the activation of Mg acceptor in GaN epilayers grown by metalorganic chemical vapor deposition" *Japanese Journal of Applied Physics* **39**, 4749 (2000)
- Detchprohm T., Zhu M., Xia Y., Li Y., Zhao W., Senawiratne J., and Wetzel C. "Improved performance of GaInN based deep green light emitting diodes through V-defect reduction" *Physica Status Solidi C* **5**, 2207 (2008)
- Detchprohm T., Zhu M., Zhao W., Wang Y., Li Y., Xia Y., and Wetzel C. "Enhanced device performance of GaInN-based deep green light emitting diodes with V-defect-free active region" *Physica Status Solidi C* **6**, S840 (2009)
- Gatos H. C. and Lavine M. C. "Characteristics of the {111} surfaces of the III–V intermetallic compounds" *Journal of the Electrochemical Society* **107**, 428 (1960)
- Gessmann T., Li Y.-L., Waldron E. L., Graff J. W., Schubert E. F., and Sheu J. K. "Ohmic contacts to p-type GaN mediated by polarization fields in thin $\text{In}_x\text{Ga}_{1-x}\text{N}$ capping layers" *Applied Physics Letters* **80**, 986 (2002a)
- Gessmann T., Li Y.-L., Waldron E. L., Graff J. W., Schubert E. F., and Sheu J. K. "Novel type of ohmic contacts to p-doped GaN using polarization fields in thin $\text{In}_x\text{Ga}_{1-x}\text{N}$ capping layers" *Journal of Electronic Materials* **31**, 416 (2002b)
- Goepfert I. D., Schubert E. F., Osinsky A., and Norris P. E. "Demonstration of efficient p-type doping in AlGaN/GaN superlattice structures" *Electronics Letters* **35**, 1109 (1999)
- Goepfert I. D., Schubert E. F., Osinsky A., Norris P. E., and Faleev N. N. "Experimental and theoretical study of acceptor activation and transport properties in p-type $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ superlattices" *Journal of Applied Physics* **88**, 2030 (2000)
- Götz W., Johnson N. M., Walker J., Bour, Amano D. P. H., and Akasaki I. "Hydrogen passivation of Mg acceptors in GaN grown by metalorganic chemical vapor deposition" *Applied Physics Letters* **67**, 2666 (1995)
- Hangleiter A., Hitzel F., Fuhrmann D., Rossow U., Ade G., and Hinze P. "Suppression of nonradiative recombination by V-shaped pits in GaInN/GaN quantum wells produces a large increase in the light emission efficiency" *Physical Review Letters* **95**, 127402 (2005)
- Jeffrey G. A., Parry G. S., and Mozzi R. L. "Study of the wurtzite-type binary compounds: Structures of aluminum nitride and beryllium oxide" *The Journal of Chemical Physics* **25**, 1024 (1956)
- Karpov S. Y. "Spontaneous polarization in III-nitride materials: crystallographic revision" *Physica Status Solidi C* **7**, 1841 (2010)
- Kim K.-C., Schmidt M. C., Sato H., Wu F., Fellows N., Saito M., Fujito K., Speck J. S., Nakamura S., and DenBaars S. P. "Improved electroluminescence on nonpolar *m*-plane InGaN/GaN quantum wells LEDs" *Physica Status Solidi – Rapid Research Letters* **1**, 125 (2007)
- Kipshidze G., Kuryatkov V., Borisov B., Holtz M., Nikishin S., and Temkin H. "AlGaN-based ultraviolet light-emitting diodes grown on Si ⟨111⟩" *Applied Physics Letters* **80**, 3682 (2002)
- Kipshidze G., Kuryatkov V., Zhu K., Borisov B., Holtz M., Nikishin S., and Temkin H. "AlN/AlGaN superlattice light-emitting diodes at 280 nm" *Journal of Applied Physics* **93**, 1363 (2003)
- Koleske D. D., Coltrin M. E., Cross K. C., Mitchell C. C., and Allerman A. A. "Understanding GaN nucleation layer evolution on sapphire" *Journal of Crystal Growth* **273**, 86 (2004)
- Kozodoy P., Hansen M., DenBaars S. P., and Mishra U. K. "Enhanced Mg doping efficiency in $\text{Al}_{0.2}\text{Ga}_{0.8}\text{N}/\text{GaN}$ superlattices" *Applied Physics Letters* **74**, 3681 (1999a)
- Kozodoy P., Smorchkova Y. P., Hansen M., Xing H., DenBaars S. P., Mishra U. K., Saxler A. W., Perrin R., and Mitchel W. C. "Polarization-enhanced Mg doping of AlGaN/GaN superlattices" *Applied Physics Letters* **75**, 2444 (1999b)
- Kumakura K., Makimoto T., and Kobayashi N. "Activation energy and electrical activity of Mg in Mg-doped $\text{In}_x\text{Ga}_{1-x}\text{N}$ ($x < 0.2$)" *Japanese Journal of Applied Physics* **39**, L337 (2000)
- Kumakura K., Makimoto T., and Kobayashi N. "Low resistance non-alloy ohmic contact to p-type GaN using Mg-doped InGaN contact layer" *Physica Status Solidi A* **188**, 363 (2001)
- Kumakura K., Makimoto T., and Kobayashi N. "Ohmic contact to p-GaN using a strained InGaN contact layer and its thermal stability" *Japanese Journal of Applied Physics* **42**, 2254 (2003)

- Kuo C. H., Chang S. J., Su Y. K., Chen J. F., Wu L. W., Sheu J. K., Chen C. H., and Chi G. C. "InGaN/GaN light emitting diodes activated in O₂ ambient" *IEEE Electron Device Letters* **23**, 240 (2002)
- Lee W. "Polarization engineering in GaN-based light-emitting diodes" (Doctoral Thesis, Rensselaer Polytechnic Institute, 2011)
- Lester S. D., Ponce F. A., Craford M. G., and Steigerwald D. A. "High dislocation densities in high efficiency GaN-based light-emitting diodes" *Applied Physics Letters* **66**, 1249 (1995)
- Li Y.-L., Schubert E. F., Graff J. W., Osinsky A., and Schaff W. F. "Low-resistance ohmic contacts to p-type GaN" *Applied Physics Letters* **76**, 2728 (2000)
- Matsuoka T., Okamoto H., Nakao M., Harima H., and Kurimoto E. "Optical bandgap energy of wurtzite InN" *Applied Physics Letters* **81**, 1246 (2002)
- Nakamura S. "GaN growth using GaN buffer layer" *Japanese Journal of Applied Physics* **30**, L1705 (1991)
- Nakamura S., Senoh M., and Mukai T. "Highly p-type Mg doped GaN films grown with GaN buffer layers" *Japanese Journal of Applied Physics* **30**, L1708 (1991)
- Nakamura S., Mukai T., Senoh M., and Iwasa N. "Thermal annealing effects on p-type Mg-doped GaN films" *Japanese Journal of Applied Physics* **31**, L139 (1992)
- Nakamura S., Senoh M., and Mukai T. "P-GaN/N-InGaN/N-InGaN double heterostructure blue-light-emitting diodes" *Japanese Journal of Applied Physics* **32**, L8 (1993a)
- Nakamura S., Senoh M., and Mukai T. "High-power InGaN/GaN double-heterostructure violet light emitting diodes" *Applied Physics Letters* **62**, 2390 (1993b)
- Nakamura S. "Growth of InGaN compound semiconductors and high-power InGaN/AlGaN double heterostructure violet-light-emitting diode" *Microelectronics Journal* **25**, 651 (1994)
- Nakamura S. and Fasol G. *The Blue Laser Diode* (Springer, Berlin, 1997); the authors discuss carrier localization effects on page 305 of the book. The authors discuss the initial stages of epitaxial growth on page 13 of the book.
- Nam K. B., Nakarmi M. L., Li J., Lin J. Y., and Jiang H. X. "Mg acceptor level in AlN probed by deep ultraviolet photoluminescence" *Applied Physics Letters* **83**, 878 (2003)
- Nann T. and Schneider J. "Origin of permanent electric dipole moments in wurtzite nanocrystals" *Chemical Physics Letters* **384**, 150 (2004)
- Narukawa Y., Kawakami Y., Fujita S., Fujita S., and Nakamura S. "Recombination dynamics of localized excitons in In_{0.20}Ga_{0.80}N - In_{0.05}Ga_{0.95}N multiple quantum wells" *Physical Review B* **55**, R 1938 (1997a)
- Narukawa Y., Kawakami Y., Funato M., Fujita S., Fujita S., and Nakamura S. "Role of self-formed InGaN quantum dots for exciton localization in the purple laser diode emitting at 420 nm" *Applied Physics Letters* **70**, 981 (1997b)
- Nickel N. H. (editor) *Hydrogen in Semiconductors II in Semiconductors and Semimetals*, Vol. **61** (Academic Press, Boston, 1999)
- Rosner S. J., Carr E. C., Ludowise M. J., Girolami G., and Erikson H. I. "Correlation of cathodoluminescence inhomogeneity with microstructural defects in epitaxial GaN grown by metal-organic chemical-vapor deposition" *Applied Physics Letters* **70**, 420 (1997)
- Ryou J.-H., Yoder P. D., Liu J., Lochner Z., Kim H., Choi S., Kim H. J., and Dupuis R. D. "Control of quantum-confined Stark effect in InGaN-based quantum wells" *IEEE Journal of Selected Topics in Quantum Electronics* **15**, 1080 (2009)
- Schubert E. F., Grieshaber W., and Goepfert I. D. "Enhancement of deep acceptor activation in semiconductors by superlattice doping" *Applied Physics Letters* **69**, 3737 (1996)
- Siozade L., Leymarie J., Disseix P., Vasson A., Mihailovic M., Grandjean N., Leroux M., and Massies J. "Modelling of thermally detected optical absorption and luminescence of (In,Ga)N/GaN heterostructures" *Solid State Communications* **115**, 575 (2000)
- Smeeton T. M., Kappers M. J., Barnard J. S., Vickers M. E., and Humphreys C. J. "Electron-beam-induced strain within InGaN quantum wells: False indium "cluster" detection in the transmission electron microscope" *Applied Physics Letters* **83**, 5419 (2003)
- Su Y. K., personal communication at the *Second Asia-Pacific Workshop on Widegap Semiconductors (APWS)*, Hsinchu, Taiwan, March 7~9 (2005)

- Takeuchi T., Sota S., Katsuragawa M., Komori M., Takeuchi H., Amano H., and Akasaki I. "Quantum-confined Stark effect due to piezoelectric fields in GaInN strained quantum wells" *Japanese Journal of Applied Physics* **36**, L382 (1997)
- Takeuchi T., Wetzel C., Yamaguchi S., Sakai H., Amano H., and Akasaki I. "Determination of piezoelectric fields in strained GaInN quantum wells using the quantum-confined Stark effect" *Applied Physics Letters* **73**, 1691 (1998)
- Taniyasu Y., Kasu M., and Makimoto T. "An aluminium nitride light-emitting diode with a wavelength of 210 nm" *Nature* **441**, 325 (2006)
- Ueno M., Yoshida M., Onodera A., and Shimomura O. "Stability of the wurtzite-type structure under high pressure: GaN and InN" *Physical Review B* **49**, 14 (1994)
- Van de Walle C. G. and Neugebauer J. "Hydrogen in semiconductors" *Annual Review in Materials Research* **36**, 179 (2006)
- Waltereit P., Brandt O., Trampert A., Grahn H. T., Menniger J., Ramsteiner M., Reiche M. and Ploog K. H. "Nitride semiconductors free of electrostatic fields for efficient white light-emitting diodes" *Nature* **406**, 865 (2000)
- Walukiewicz W., Li S. X., Wu J., Yu K. M., Ager III J. W., Haller E. E., Lu K., and Schaff W. J. "Optical properties and electronic structure of InN and In-rich group III-nitride alloys" *Journal of Crystal Growth* **269**, 119 (2004)
- Wikipedia < www.wikipedia.org > (accessed November 2010)
- Wu J., Walukiewicz W., Yu K. M., Ager J. W., Haller E. E., Lu H., and Schaff W. J. "Small band gap bowing in $\text{In}_{1-x}\text{Ga}_x\text{N}$ alloys" *Applied Physics Letters* **80**, 4741 (2002)
- Wu J., Walukiewicz W., Yu K. M., Ager III J. W., Li S. X., Haller E. E., Lu H., and Schaff W. J. "Universal bandgap bowing in group-III nitride alloys" *Solid State Communications* **127**, 411 (2003)
- Yamanaka T., Dutta M., and Stroscio M. A. "Spontaneous polarization effects in nanoscale wurtzite structures" *Journal of Computational Electronics* **6**, 313 (2007)
- Yun F., Reschchikov M. A., He L., King T., Morkoç H., Novak S. W., and Wei L. "Energy band bowing parameter in $\text{Al}_x\text{Ga}_{1-x}\text{N}$ alloys" *Journal of Applied Physics* **92**, 4837 (2002)

21

21 – Ultraviolet light emitters

21.1 The ultraviolet (UV) spectral range

The UV-visible boundary is at about 380 nm, where the 1978 CIE eye sensitivity curve has a value of about 0.1% of its maximum value.¹² This chapter concentrates on materials issues of III-V nitrides, on devices emitting in the ultraviolet (UV, $\lambda < 380$ nm), and on devices emitting in the violet near the UV-visible boundary (380~410 nm). For UV devices, we will differentiate between devices having a GaInN active region ($\lambda > 360$ nm) and devices having an AlGaN active region ($\lambda < 360$ nm).

There are two classifications of the UV spectrum, both of which are shown in *Figure 21.1*. The UV-A, UV-B, and UV-C classification is based on a convention established during the Second International Congress on Light in 1932 (International, 1932).

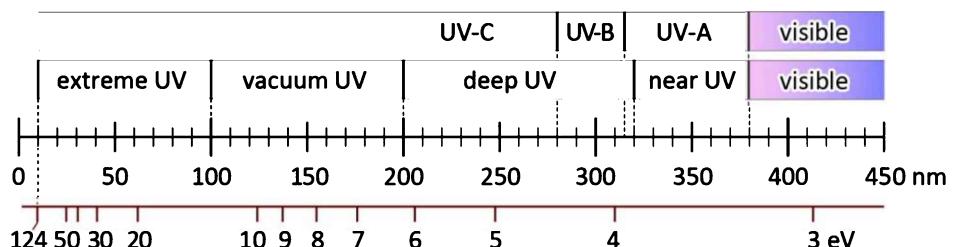


Fig. 21.1: Nomenclature of UV radiation versus wavelength (after International Congress on Light, 1932).

UV-A radiation (315~380 nm) from the sun penetrates the earth's atmosphere (including clouds) and reaches the earth's surface.

UV-B radiation (280~315 nm) from the sun is partially absorbed by the earth's ozone layer.

UV-C radiation ($< 280 \text{ nm}$) from the sun is fully absorbed by the earth's ozone layer.

Therefore, no natural light with $\lambda \leq 280$ nm exists at the earth's surface. This wavelength range is also referred to as the ***solar-blind*** range. UV-C radiation is germicidal and can be used for sterilization in health-care applications and decontamination of drinking water.

¹² The UV-visible boundary does not have an exact wavelength value. The technical literature may report wavelengths values of about 380~390 nm as the UV-visible boundary. In the 1931 CIE (x, y) chromaticity diagram, the UV-visible boundary is located at 380 nm.

Receiving normal doses of UV-A and UV-B radiation is very important because it is needed by the human body to generate Vitamin D. However, excessive doses of UV radiation result in sunburn, damage to the skin (e.g. DNA damage), and skin cancer. Furthermore, the absorption of UV radiation in the human eye, specifically the eye's lens, causes the lens to become absorptive and clouded. The clouding of the eye's lens is an illness, cataract, which can lead to total blindness.

An alternative classification of UV radiation has evolved over time. This alternative classification has the following categories:

- Extreme UV: 10 ~ 100 nm
- Vacuum UV: 100 ~ 200 nm
- Deep UV: 200 ~ 320 nm
- Near UV: 320 ~ 380 nm.

21.2 UV and violet LEDs emitting at wavelengths longer than 360 nm

UV devices emitting at wavelengths longer than 360 nm generally have GaN or GaInN active regions. GaInN LEDs with peak wavelengths ranging from 400 nm to 410 nm were reported as early as 1993 (Nakamura *et al.*, 1993a, b, 1994). The design of an early GaInN UV LED emitting at 370 nm is shown in **Figure 21.2**. The device has a GaInN/AlGaN double heterostructure active region and AlGaN electron- and hole-blocking layers (Mukai *et al.*, 1998). The active region and the carrier-blocking layers are clad by p-type and n-type GaN.

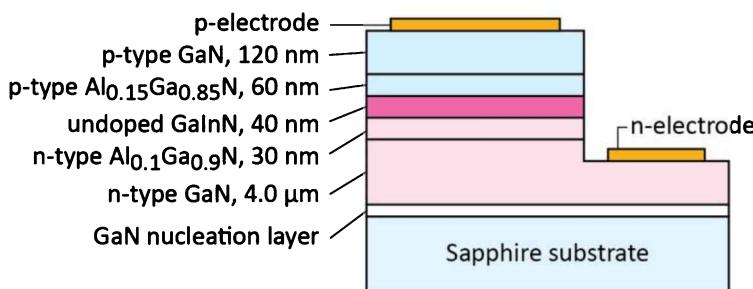


Fig. 21.2: Layer structure of GaInN UV LED grown on sapphire substrate emitting at 370 nm (after Mukai *et al.*, 1998).

Figure 21.3 shows the emission spectrum of a 375 nm UV LED (Nichia, 2006) for pulsed and continuous-wave (cw) injection. Inspection of the figure reveals a small red-shift of the peak wavelength when going from pulsed to cw injection. This shift is likely caused by junction heating, which generally leads to a lower bandgap energy and a red-shift of the emission wavelength.

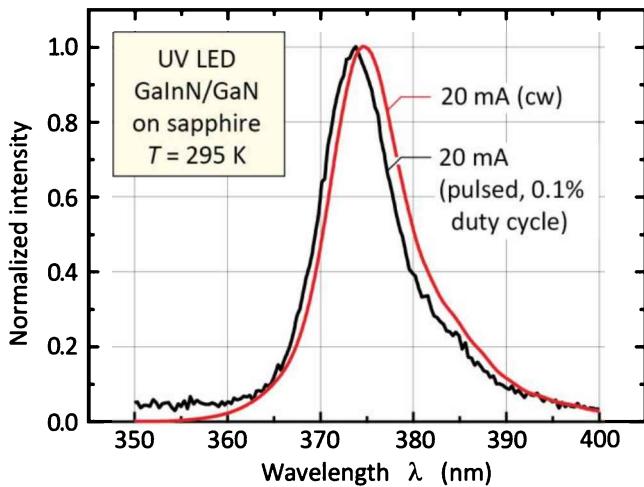


Fig. 21.3: Room temperature emission spectrum of commercial 375 nm UV LED (Nichia Company, 2006) under continuous wave (cw) and pulsed conditions.

Figure 21.4 shows the emission intensity of a 375 nm UV LED as a function of the active layer thickness. The figure reveals that optimum output power is attained at 30~50 nm active layer thickness. It can be assumed that the active region is heavily doped in order to screen the polarization fields. Quantum well active regions with quantum well thicknesses < 5 nm reduce the effect of spatial electron–hole separation and for this reason, quantum well active regions have superseded the double heterostructure designs.

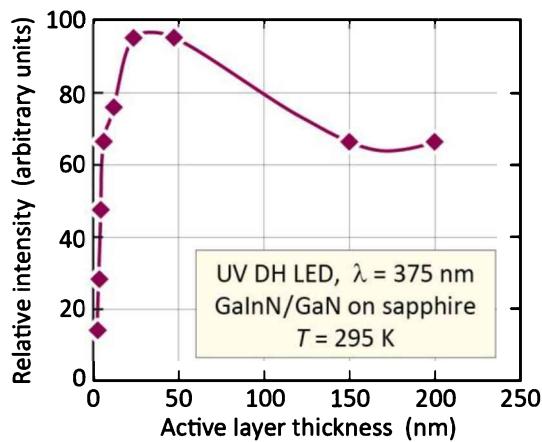


Fig. 21.4: Room temperature emission intensity as a function of GaInN active layer thickness for double heterostructure UV LEDs emitting at 375 nm (after Mukai *et al.*, 1998).

Figure 21.5 shows the dependence of the output power as a function of wavelength. As the emission wavelength decreases, a pronounced drop in the emission intensity is found. This has been attributed to the very positive influence of indium (In) incorporation on the internal quantum efficiency. The positive influence is reduced as less indium is incorporated in the active region with pure GaN active regions ($\lambda \approx 360$ nm) not benefiting at all.

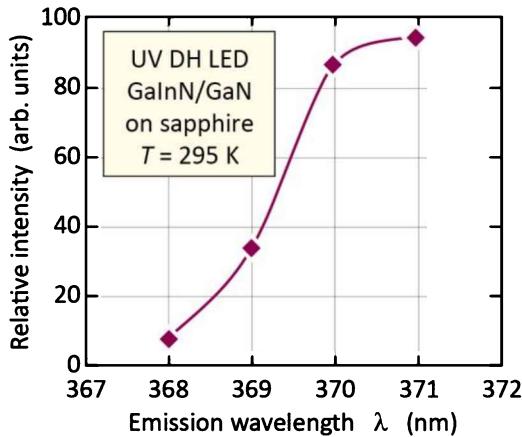


Fig. 21.5: Room temperature intensity as a function of emission wavelength for GaInN double heterostructure UV LEDs (after Mukai *et al.*, 1998).

21.3 UV devices emitting at wavelengths shorter than 360 nm

LEDs emitting at wavelengths shorter than 360 nm typically have either AlGaN active regions or Al_xGa_{1-x}N / Al_yGa_{1-y}N multiple-quantum well (MQW) active regions. The power-conversion efficiency of these devices is generally low, i.e. less than 1%, although substantial progress has been made (Zhang *et al.*, 2002a, 2003; Yasan *et al.*, 2002; Kipshidze *et al.*, 2003; Fischer *et al.*, 2004; Kim *et al.*, 2004; Oder *et al.*, 2004; Razeghi and Henini, 2004; Shakya *et al.*, 2004).

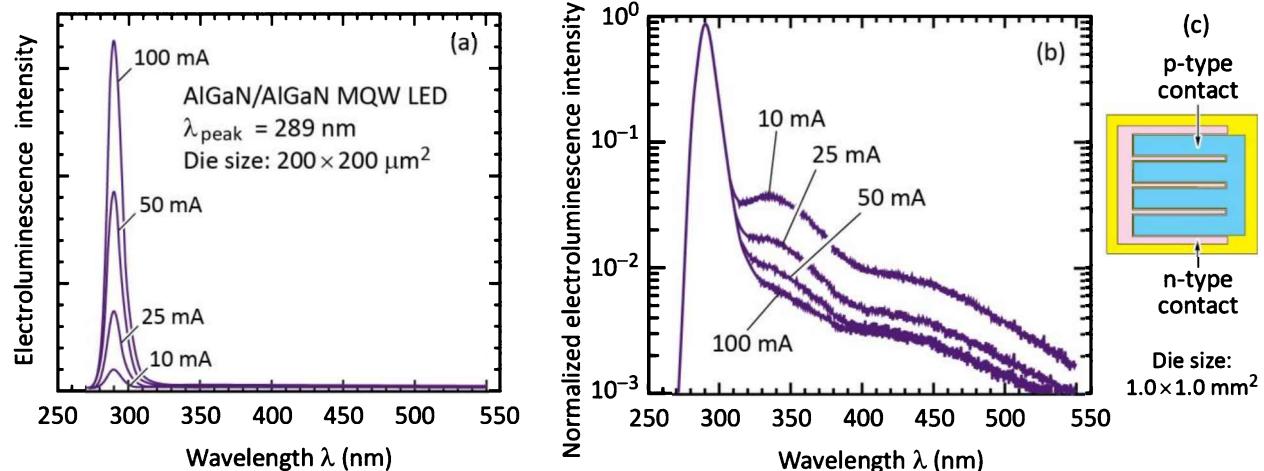


Fig. 21.6: Emission spectrum of deep-UV AlGaN/AlGaN multiple-quantum well LED for different injection currents on (a) linear and (b) logarithmic scales. An interdigitated contact geometry, as shown in (c), was used for large-area dies (after Fischer *et al.*, 2004).

Figure 21.6 shows the emission spectrum of an AlGaN/AlGaN deep-UV LED with interdigitated contact geometry (Fischer *et al.*, 2004). The active region of the device is composed of three Al_{0.36}Ga_{0.64}N quantum wells with Al_{0.48}Ga_{0.52}N barriers for emission at 290 nm. The spectrum displays one clean emission line with a peak wavelength of 289 nm. Some sub-bandgap emission

near 330 nm becomes apparent when plotting the spectrum on a logarithmic scale. The forward voltage of the $200 \times 200 \mu\text{m}^2$ and $1 \times 1 \text{ mm}^2$ devices was reported to be about 7.0 V and 6.0 V at a forward current of 20 mA, respectively.

The shortest wavelength attained with a UV LED is 210 nm (Taniyasu *et al.*, 2006). The authors reported the successful control of both n-type and p-type doping in AlN, which has a direct bandgap greater than 6.0 eV. The doping strategy allowed the authors to develop an AlN pin (p-type / intrinsic / n-type) homojunction LED, that is, the device was based on an undoped AlN active region clad by n-type AlN and p-type AlN. The emission wavelength of 210 nm is the shortest possible wavelength attainable for the $\text{Al}_x\text{Ga}_{1-x}\text{N}$ alloy system. The UV emission was attributed to electron-hole-exciton transitions.

However, the external quantum efficiency (*EQE*) of the UV LED reported by Taniyasu *et al.* (2006) was very low, an estimated $10^{-6}\%$. The low *EQE* can in part be attributed to the employment of a homojunction active region (rather than a heterojunction active region). Furthermore, due to the fact that p-type doping and n-type doping in $\text{Al}_x\text{Ga}_{1-x}\text{N}$ becomes increasingly difficult as the Al content increases, the free hole concentration and the efficiency of UV LEDs decrease as the Al content increases. This trend was comprehensively surveyed by Khan *et al.* (2008) and is shown in **Figure 21.7**. The efficiency of UV LEDs was also discussed by Hirayama *et al.* (2007) for devices emitting at wavelengths less than 280 nm, where typical *EQEs* are less than 1%.

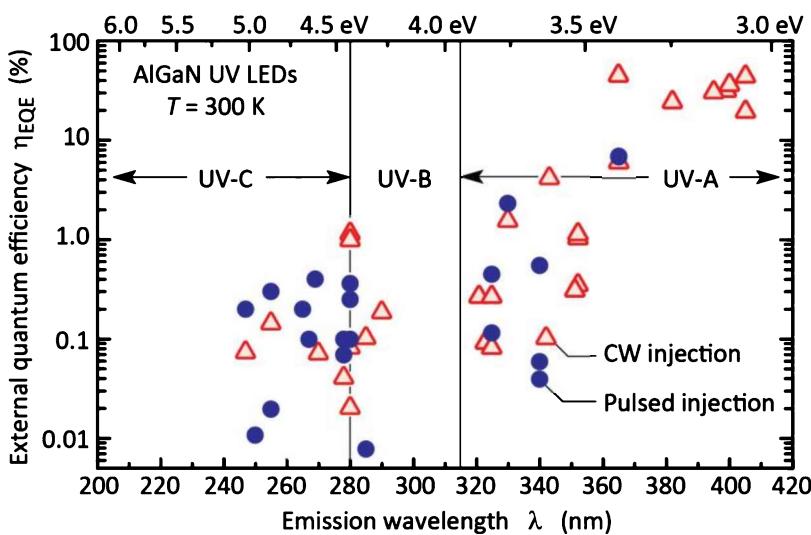


Fig. 21.7: Compilation of the external quantum efficiencies of UV LEDs operated with pulsed and continuous-wave (CW) current injection. The data points represent the results of different research groups (data compiled by Khan *et al.*, 2008).

The following issues deserve special attention in the field of AlGaN/AlGaN UV LEDs:

Affinity of aluminum to oxygen: Al has a very high affinity to O₂ making the incorporation of oxygen into AlGaN increasingly likely as the Al content increases. Oxygen forms a deep, DX-like level in Al-rich AlGaN (McCluskey *et al.*, 1998; Wetzel *et al.*, 2001).

Conductivity of AlGaN: Both the p-type and n-type conductivity of AlGaN decrease as the Al mole fraction increases, particularly for Al mole fractions exceeding 30% (Katsuragawa *et al.*, 1998; Goepfert *et al.*, 2000; Jiang and Lin, 2002). This leads to a higher resistivity in the confinement layers and higher device series resistances. A particular problem is the p-type conductivity in AlGaN. Al_xGa_{1-x}N/Al_yGa_{1-y}N superlattices have been employed to alleviate the p-type doping problem.

Lateral conductivity: In devices grown on insulating substrates, with the standard side-by-side contact configuration, the n-type AlGaN layer provides the lateral conductivity. As the resistivity of Si-doped n-type AlGaN increases with the Al content, devices generally become more resistive. To compensate for this effect, the mean distance, which the electron current flows laterally in the n-type AlGaN lateral-conduction layer, needs to be reduced. This can be accomplished by an array of micro-LEDs (Kim *et al.*, 2003; Khan, 2004) or by closely spaced fingers in interdigitated-contact geometries.

Contact resistance: Due to the high bandgap of AlGaN, contact-barrier heights are generally higher, which makes the attainment of low-resistance contacts increasingly difficult as the Al content increases.

Diffusion of acceptors: During the epitaxial growth of the top cladding layer, Mg acceptors may diffuse back into the active region, thereby lowering its radiative efficiency. Acceptor diffusion and the associated decrease in radiative efficiency may impose a limit on the maximum thickness of the p-type cladding layer.

Heterojunction barriers: Due to the larger bandgap energies, the conduction- and valence-band discontinuities of heterojunctions are generally larger than for smaller-bandgap semiconductors. Compositional grading at the heterojunction interfaces reduces the resistance of heterojunction barriers.

Light extraction: To reduce reabsorption effects, all device layers should have an Al content sufficiently high to be transparent to the emitted light.

Cracking: AlGaN films grown on relaxed GaN are, due to the smaller lattice constant of AlGaN, under tensile strain. If the films are sufficiently thick, they crack. However, cracking can be strongly reduced or even eliminated by using Al-rich strain-compensating superlattices (Hearne *et al.*, 2000; Han *et al.*, 2001; Zhang *et al.*, 2002b). Such strain-compensating superlattices reduce the lattice constant so that subsequent epitaxial layers have reduced tensile strain or are even under compressive strain. Hearne *et al.* (2000) provided a quantitative analysis of the maximum attainable thickness of a crack-free layer under tensile strain.

Optical micrographs of a 0.9 μm thick $\text{Al}_{0.15}\text{Ga}_{0.85}\text{N}$ film are shown in **Figure 21.8** (a) and (b) for growth without and with a strain-compensating superlattice, respectively. **Figure 21.8** (b) shows a virtually crack-free AlGaN layer that was grown on a strain-compensating $\text{AlN}/\text{Al}_{0.45}\text{Ga}_{0.55}\text{N}$ superlattice with 10 periods and equal well- and barrier-layer thicknesses of 100 Å.

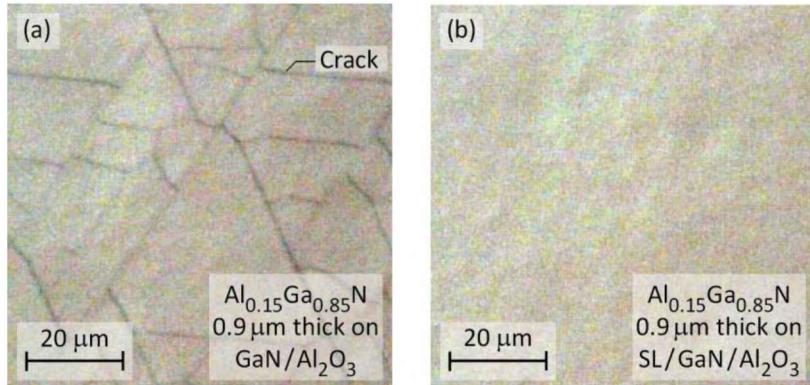


Fig. 21.8: Optical micrographs of $\text{Al}_{0.15}\text{Ga}_{0.85}\text{N}$ layer grown (a) without and (b) with a strain-compensating $\text{AlN}/\text{Al}_{0.45}\text{Ga}_{0.55}\text{N}$ superlattice (SL). The SL has 10 periods and equal well and barrier thicknesses of 10 nm. Angles between crack lines frequently are 60° and 120° due to the hexagonal symmetry of the lattice.

Exercise: Cracking. Why does cracking occur in epitaxial layers that are under biaxial tensile strain but not in epitaxial layers that are under biaxial compressive strain?

Solution: Wafer bowing and ultimately cracking of an epitaxial film that is under biaxial tensile strain releases the strain energy stored in the film. For epitaxial layers that are under compressive strain, the strain energy can be released by wafer bowing, film buckling, and film delamination. Due to the compressive strain, there is “no room” for fissures or cracks, so that cracks generally do not form in compressively strained films.

The strain energy stored in a homo-epitaxial film, that is lattice mismatched to the substrate, is proportional to the thickness of the film. As the thickness of a strained film increases, it will at some point become energetically more favorable to reduce the strain energy by creating misfit dislocations and cracks. Thus, at a certain thickness, the film will form misfit dislocations to release the strain energy. The critical thickness, at which a homo-epitaxial film starts forming misfit dislocations, is given by the **Matthews–Blakeslee law** (Matthews and Blakeslee, 1976). As the film thickness increases further, misfit dislocations do not suffice to release the strain energy, so that at some point the film will start to crack. A formula for the critical thickness at which a film under biaxial tensile strain starts to crack was given by Hearne *et al.* (2000).

AlGaN UV LEDs frequently have forward voltages $V_f \gg h\nu/e$. Depending on the device structure, the excess forward voltage may originate from the p-type contact, the p-type AlGaN confinement layer, the n-type AlGaN layer providing lateral conduction, or from unipolar heterojunctions.

For devices with low or moderate efficiency as well as for high-power devices, device packages with low thermal resistance are desirable. The heat resulting from an excess forward voltage and low quantum efficiency must be removed to avoid excessively high junction temperatures. Morita *et al.* (2004) reported a particularly well heat-sunk device, namely a structure in which the sapphire substrate was removed by laser lift-off, and, using an AuSn solder, the epilayer was directly bonded to a CuW heat sink.

21.4 Alternative approaches for UV emitters

The difficulties of the AlGaN system, particularly at high Al contents, have been reviewed in the previous section as well as in the literature (Khan *et al.*, 2008; Schubert and Cho, 2010). One of the major problems is the lack of efficient p-type doping in AlGaN. Therefore, light-emitting structures using light-emitting structures, that do not require p-type doping, are of high interest.

An electron beam allows one to generate electron-hole pairs without doping. The electron beam may impinge on the sample as shown in **Figure 21.9**. Electron-hole pairs are generated within a volume that has a spatial extent on the order of a micrometer. Since the electron beam has a high acceleration voltage, e.g. 5 kV, each electron entering the sample will have the energy of 5 keV. The generation of one electron-hole pair requires the energy E_g , that is, the bandgap energy. Therefore, one 5 keV primary electron of the electron beam will generate many secondary electron-hole pairs in the sample.

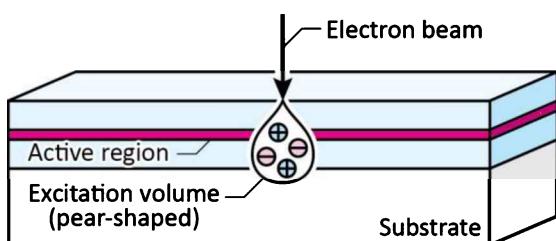


Fig. 21.9: Principle of electron-hole-pair excitation by a high-energy electron-beam impinging on a sample. The excitation volume is pear-shaped with a diameter on the order of μm .

Electron-beam excitation was demonstrated for AlGaN emitters (Ota *et al.*, 2010) as well as BN emitters (Watanabe *et al.*, 2009; Nebel, 2009) with remarkable success. For example, Ota *et al.* (2010) reported optical powers of 100 mW emitted at a wavelength of 240 nm by using a device with an $\text{Al}_x\text{Ga}_{1-x}\text{N}/\text{AlN}$ MQW active region. The power-conversion efficiency of the device was a remarkable 40%.

Another approach that would make p-type doing unnecessary is based on the large polarization fields in GaN and related materials. The large polarization fields allow for polarization-charge tunnel junctions that generate holes by tunneling. Electrons and the internally generated holes can reach the active region of the light emitter without the requirement for p-type doping and p-type contact (Schubert, 2010b). The author showed that polarization-charge tunnel junctions can carry sufficient current to be viable for ultraviolet emitters.

Sources limiting the efficiency in conventional pn-junction UV emitters include poor hole-injection efficiency and optical absorption in the p-type GaN contact layers. Polarization-charge tunnel junctions can deliver improvements in both areas, by eliminating the need for p-type contacts and by enabling the use of a p-side-down structure; such p-side-down structure has favorable transport characteristics (i.e. reduced electron leakage out of the active region) compared to a p-side-up structure (Schubert, 2010b).

21.5 References

- Fischer A. J., Allerman A. A., Crawford M. H., Bogart K. H. A., Lee S. R., Kaplar R. J., Chow W. W., Kurtz S. R., Fullmer K. W., and Figiel J. J. "Room-temperature direct current operation of 290 nm light-emitting diodes with milliwatt power levels" *Applied Physics Letters* **84**, 3394 (2004)
- Goepfert I. D., Schubert E. F., Osinsky A., Norris P. E., and Faleev N. N. "Experimental and theoretical study of acceptor activation and transport properties in p-type Al_xGa_{1-x}N/GaN superlattices" *Journal of Applied Physics* **88**, 2030 (2000)
- Han J., Waldrip K. E., Lee S. R., Figiel J. J., Hearne S. J., Petersen G. A., and Myers S. M. "Control and elimination of cracking of AlGaN using low-temperature AlGaN interlayers" *Applied Physics Letters* **78**, 67 (2001)
- Hearne S. J., Han J., Lee S. R., Floro J. A., Follstaedt D., Chason M. E., and Tsong I. S. T. "Brittle-ductile relaxation kinetics of strained AlGaN/GaN heterostructures" *Applied Physics Letters* **76**, 1534 (2000)
- Hirayamaa H., Yatabe T., Noguchi N., Ohashi T., and Kamata N. "231~261 nm AlGaN deep-ultraviolet light-emitting diodes fabricated on AlN multilayer buffers grown by ammonia pulse-flow method on sapphire" *Applied Physics Letters* **91**, 071901 (2007)
- International Congress on Light, Copenhagen, Denmark (1932)
- Jiang H. X. and Lin J. Y. "AlGaN and InAlGaN alloys – epitaxial growth, optical and electrical properties, and applications" *Opto-Electronics Review* **10**, 271 (2002)
- Katsuragawa M., Sota S., Komori M., Anbe C., Takeuchi T., Sakai H., Amano H., and Akasaki I. "Thermal ionization energy of Si and Mg in AlGaN" *Journal of Crystal Growth* **189/190**, 528 (1998)
- Khan M. A. "Deep-ultraviolet LEDs fabricated in AlInGaN using MEMOCVD" *Proceedings SPIE* **5530**, 224 (2004)
- Khan M. A., Balakrishnan K., and Katona T. "Ultraviolet light-emitting diodes based on group three nitrides" *Nature Photonics* **2**, 77 (2008)
- Kim K. H., Li J., Jin S. X., Lin J. Y., and Jiang H. X. "III-nitride ultraviolet light-emitting diodes with delta doping" *Applied Physics Letters* **83**, 566 (2003)
- Kim K. H., Fan Z. Y., Khizar M., Nakarmi M. L., Lin J. Y., and Jiang H. X. "AlGaN-based ultraviolet light-

- emitting diodes grown on AlN epilayers" *Applied Physics Letters* **85**, 4777 (2004)
- Kipshidze G., Kuryatkov V., Zhu K., Borisov B., Holtz M., Nikishin S., and Temkin H. "AlN/AlGaN superlattice light-emitting diodes at 280 nm" *Journal of Applied Physics* **93**, 1363 (2003)
- Matthews J. W. and Blakeslee A. E. "Defects in epitaxial multilayers – III; Preparation of almost perfect multilayers" *Journal of Crystal Growth* **32**, 265 (1976)
- McCluskey M. D., Johnson N. M., Van de Walle C. G., Bour D. P., Kneissl M., and Walukiewicz W. "Metastability of oxygen donors in AlGaN" *Physical Review Letters* **80**, 4008 (1998)
- Morita D., Yamamoto M., Akaishi K., Matoba K., Yasutomo K., Kasai Y., Sano M., Nagahama S.-I., and Mukai T. "Watt-class high-output-power 365 nm ultraviolet light-emitting diodes" *Japanese Journal of Applied Physics* **43**, 5945 (2004)
- Mukai T., Morita D., and Nakamura S. "High-power UV GaInN/AlGaN double-heterostructure LEDs" *Journal of Crystal Growth* **189**, 778 (1998)
- Nakamura S., Senoh M., and Mukai T. "P-GaN / N-InGaN / N-InGaN double heterostructure blue-light-emitting diodes" *Japanese Journal of Applied Physics* **32**, L8 (1993a)
- Nakamura S., Senoh M., and Mukai T. "High-power InGaN/GaN double-heterostructure violet light emitting diodes" *Applied Physics Letters* **62**, 2390 (1993b)
- Nakamura S. "Growth of InGaN compound semiconductors and high-power InGaN/AlGaN double heterostructure violet-light-emitting diode" *Microelectronics Journal* **25**, 651 (1994)
- Nebel C. E. "Tackling the deep ultraviolet" *Nature Photonics* **3**, 564 (2009)
- Nichia Company, UV LED product emitting at 375 nm (2006)
- Oder T. N., Kim K. H., Lin J. Y., and Jiang H. X. "III-nitride blue and ultraviolet photonic crystal light emitting diodes" *Applied Physics Letters* **84**, 466 (2004)
- Oto T., Banal R. G., Kataoka K., Funato M., and Kawakami Y. "100 mW deep-ultraviolet emission from AlN-based quantum wells pumped by an electron beam" *Nature Photonics* **4**, 767 (2010)
- Razeghi M. and Henini M. "Optoelectronic devices: III-nitrides" (Elsevier, Amsterdam, 2004)
- Schubert E. F. and Cho J. "UV LEDs – Electron-beam excitation" *Nature Photonics* **4**, 735 (2010)
- Schubert M. F. "Interband tunnel junctions for wurtzite III-nitride semiconductors based on heterointerface polarization charges" *Physical Review B* **81**, 035303 (2010a)
- Schubert M. F. "Polarization-charge tunnel junctions for ultraviolet light-emitters without p-type contact" *Applied Physics Letters* **96**, 031102 (2010b)
- Shakya J., Kim K. H., Lin J. Y., and Jiang H. X. "Enhanced light extraction in III-nitride ultraviolet photonic crystal light-emitting diodes" *Applied Physics Letters* **85**, 142 (2004)
- Taniyasu Y., Kasu M., and Makimoto T. "An aluminium nitride light-emitting diode with a wavelength of 210 nanometers" *Nature* **441**, 325 (2006)
- Watanabe K., Taniguchi T., Niyyama T., Miya K. and Taniguchi M. "Far-ultraviolet plane-emission handheld device based on hexagonal boron nitride" *Nature Photonics* **3**, 591 (2009)
- Wetzel C., Amano H., Akasaki I., Ager III J. W., Grzegory I., and Meyer B. K. "DX-like behavior of oxygen in GaN" *Physica B* **302~303**, 23 (2001)
- Yasan A., McClintock R., Mayes K., Darvish S. R., Kung P., and Razeghi M. "Top-emission ultraviolet light-emitting diodes with peak emission at 280 nm" *Applied Physics Letters* **81**, 801 (2002)
- Zhang J. P., Chitnis A., Adivarahan V., Wu S., Mandavilli V., Pachipulusu R., Shatalov M., Simin G., Yang J. W., and Khan M. A. "Milliwatt power deep ultraviolet light-emitting diodes over sapphire with emission at 278 nm" *Applied Physics Letters* **81**, 4910 (2002a)
- Zhang J. P., Wang H. M., Gaevski M. E., Chen C. Q., Fareed Q., Yang J. W., Simin G., and Khan M. A. "Crack-free thick AlGaN grown on sapphire using AlN/AlGaN superlattices for strain management" *Applied Physics Letters* **80**, 3542 (2002b)
- Zhang J. P., Wu S., Rai S., Mandavilli V., Adivarahan V., Chitnis A., Shatalov M., and Khan M. A. "AlGaN multiple-quantum-well-based deep ultraviolet light-emitting diodes with significantly reduced long-wave emission" *Applied Physics Letters* **83**, 3456 (2003)

22

22 – The efficiency droop

22.1 Introduction to the efficiency droop

In an *ideal* LED, the optical emission power is linearly dependent on the injection current. The quantum efficiency of an LED is proportional to the optical emission power divided by the injection current. Since the optical power depends linearly on current, the efficiency of the ideal LED is a constant, independent of current. The optical-power-versus-current and efficiency-versus-current curve of the ideal LED is shown in **Figure 22.1** by the dashed curve.

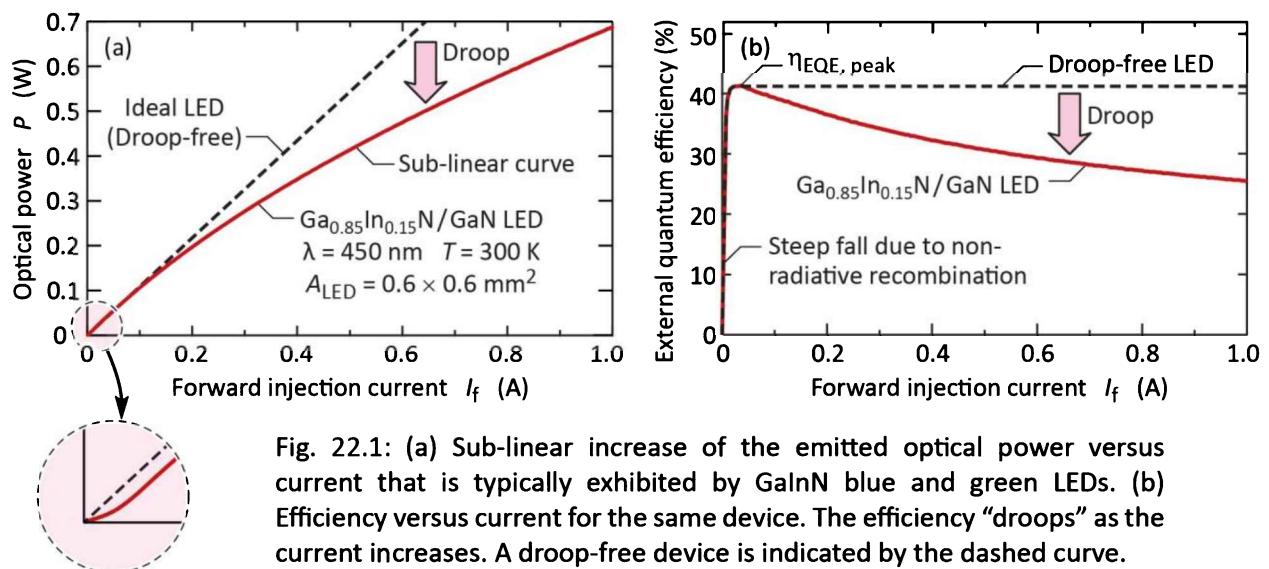


Fig. 22.1: (a) Sub-linear increase of the emitted optical power versus current that is typically exhibited by GaInN blue and green LEDs. (b) Efficiency versus current for the same device. The efficiency “droops” as the current increases. A droop-free device is indicated by the dashed curve.

However, in *real* LEDs, the optical emission power initially increases sub-linearly with the injection current. As a consequence, the efficiency has a peak value at relatively low currents and then “droops” as the current increases. The verb “to droop” in this context has the meaning “to decrease gradually”. The sub-linear increase of the optical emission power and the drooping efficiency versus current is shown in **Figure 22.1** for a GaInN/GaN LED. The efficiency droop at a certain operating current I is defined as:

$$\text{Droop} = \frac{\eta_{EQE,peak} - \eta_{EQE,I}}{\eta_{EQE,peak}}$$
(22.1)

where $\eta_{EQE,I}$ is the EQE at the current of operation, I . The magnitude of the droop can be as high as 50%. Therefore, the efficiency droop can be a very significant loss mechanism, particularly at high current values.

The efficiency has a peak value at low current values. A gradual decrease of the efficiency occurs beyond the peak value. The efficiency peak occurs at current density values, typically ranging from $0.1\sim10\text{ A/cm}^2$ (Mukai *et al.*, 1999) as illustrated in **Figure 22.2**.

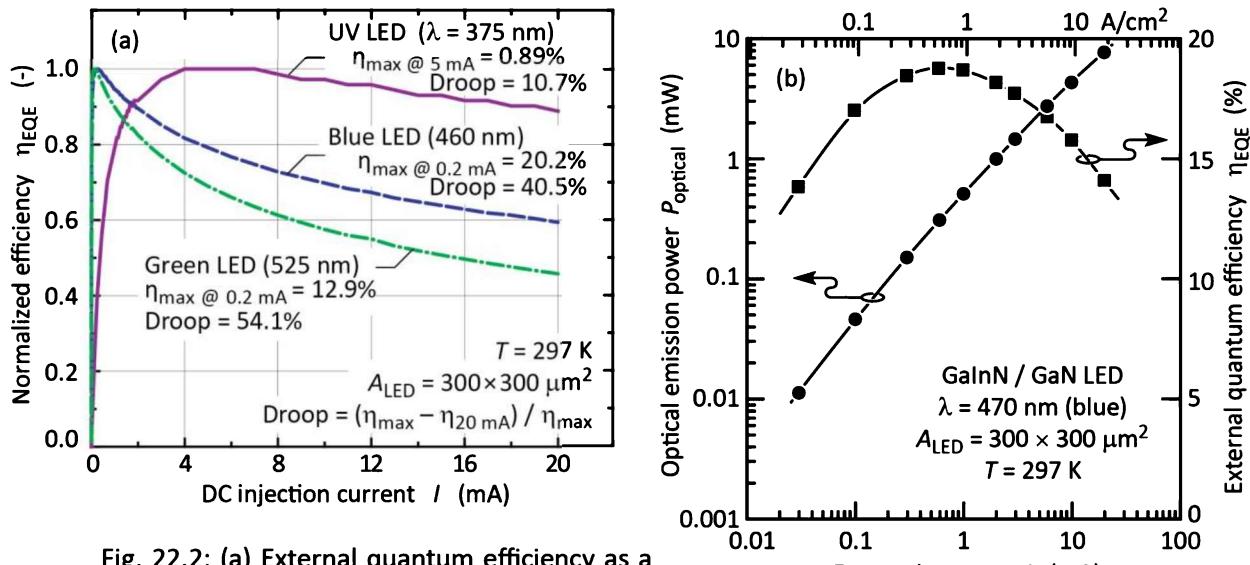


Fig. 22.2: (a) External quantum efficiency as a function of current for a GaInN UV, blue, and green LED. The maximum efficiency is reached at a low current and the efficiency “droops” as the current increases. GaInN green LEDs generally have a larger droop than GaInN blue LEDs. (b) Optical emission power and external quantum efficiency vs. current on a logarithmic scale (after Mukai *et al.*, 1999).

Among III–V nitrides devices, the efficiency droop is generally increases with wavelength (Mukai *et al.*, 1999; Baeumler *et al.*, 2007). That is, green devices can exhibit a greater droop than blue devices. This is illustrated in **Figure 22.2** which shows a violet, blue, and Green GaInN/GaN LED.

The physical origin of the efficiency droop is very different for different material systems. For devices made from III–V phosphides and arsenides, the dominant cause for the droop is heating. For devices made from III–V nitrides, multiple origins of the efficiency droop have been proposed and they will be discussed in the following sections.

22.2 The “ABC” model

The *ABC* model is a rate-equation model that is frequently employed when analyzing recombination processes in LEDs. The *ABC* model can be used for *undoped* active regions, in which the concentration of injected carriers far exceeds the equilibrium carrier concentration. In this case, the carrier concentrations in the active region under injection conditions are $n = p$. Since $n = p$, the *ABC* model can be based only on the electron concentration, n .

Electron-hole pairs inside the active region can either recombine radiatively or non-radiatively. Using the Shockley-Read-Hall process and the Auger process as the non-radiative processes, the recombination rate can be written as

$$R = A_{SRH} n + B n^2 + C_{Auger} n^3 \quad (22.2)$$

where A_{SRH} , B , and C_{Auger} represent the Shockley-Read-Hall, radiative, and Auger coefficient, respectively. Recombination described by the A_{SRH} , B , and C_{Auger} coefficients is commonly referred to as the *ABC* model. **Figure 22.3** illustrates the electron flow and the three recombination channels in the active region.

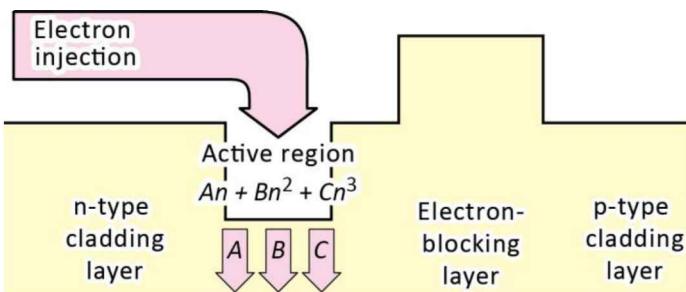


Fig. 22.3: Illustration of electrons injected into the active region and recombining in the active region, as expressed by recombination terms $A_{SRH}n + Bn^2 + C_{Auger}n^3$. Electron leakage out of the active region is not considered.

22.3 The “ABC + $f(n)$ ” model

As the carrier concentration increases, an increasing number of carriers may not be captured by the active region or may escape from the active region, i.e. leak out of the active region. The non-capture and escape of carriers from the active region is not taken into account by the *ABC* model.

A term, $f(n)$, can be added to the *ABC* model to account for the carrier loss. The additional loss term, $f(n)$, represents carriers that recombine outside the active region. Including the $f(n)$ term, we obtain

$$R = A_{SRH}n + Bn^2 + C_{Auger}n^3 + f(n) \quad (22.3)$$

where $f(n)$ stands for ***non-capture*** and for ***escape*** of carriers from the active region. That is, $f(n)$ is the recombination rate of the carriers that leak out of the active region. **Figure 22.4** illustrates the current flow in the active region including the $f(n)$ term. The figure considers only the electron flow. This is generally sufficient, since holes behave in a complementary manner as electrons; that is, for every electron that is recombining in the active region, there is a hole recombining in the active region.

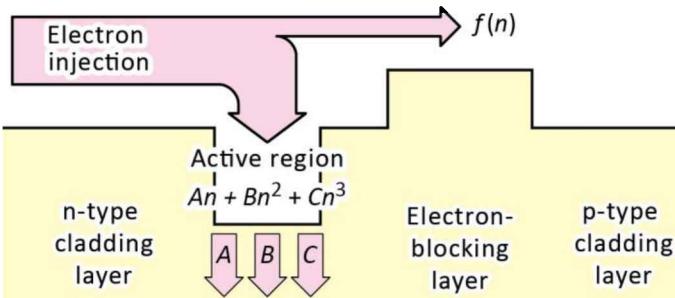


Fig. 22.4: Illustration of electrons injected into the active region and recombining in the active region, as expressed by recombination terms $A_{SRH}n + Bn^2 + C_{Auger}n^3$, and electrons leaking out of the active region, as described by the function $f(n)$.

22.4 The function $f(n)$ for drift-induced leakage

The function $f(n)$, which represents the carrier leakage out of the active region, has been derived analytically for ***drift-induced leakage***, i.e. when the leakage of carriers is mediated by electron drift in the p-type layer that occurs under high-injection conditions (Lin *et al.*, 2012). The model, illustrated in **Figure 22.5**, is based on the increasing strength of the electron drift current in the p-type cladding region as the device enters the high-injection regime and a concomitant electric field develops in the p-type cladding layer.

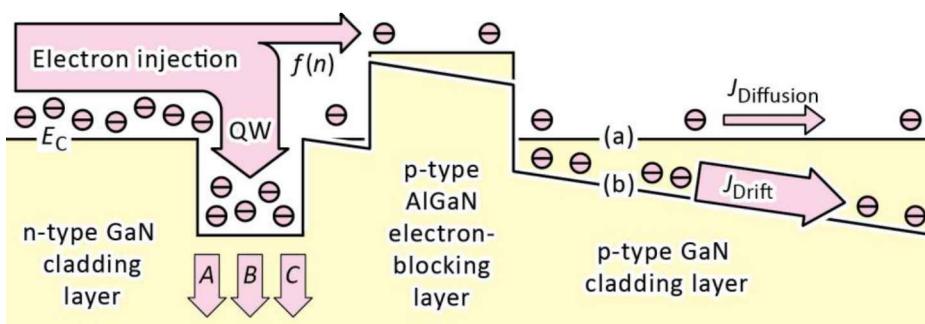


Fig. 22.5: Schematic illustration of (a) electron diffusion current occurring under low-injection conditions and (b) droop-causing electron drift current occurring under high-injection conditions.

The Shockley theory for pn junctions, which accurately predicts the exponential relationship between current and voltage, is limited to the low-level-injection regime. The low-level-injection

condition in an n⁺p junction is defined as the regime in which the injected electron concentration at the edge of the p-type neutral layer, $\Delta n_p(0)$, is much smaller than the equilibrium hole concentration, p_{p0} , that is

$$\Delta n_p(0) \ll p_{p0} . \quad (22.4)$$

This condition is meant to ensure that the conductivity of the depletion layer is much lower than the conductivity of the p-type neutral layer so that an incremental voltage applied to the pn-junction diode drops predominantly across the depletion layer, rather than the p-type neutral layer. However, if electrons and holes have strongly different mobilities, e.g. $\mu_n > 10 \mu_p$, the above condition is insufficient. In case of a large disparity between electron and hole mobility, the low injection condition needs to be generalized (Meyaard *et al.*, 2011)

$$\Delta n_p(0) \mu_n \ll p_{p0} \mu_p \quad (22.5)$$

Eqn. (22.5), which is more stringent than Eqn. (22.4), illustrates that high-level injection can be more easily reached in pn-junction diodes made of semiconductors with a strong asymmetry in electron and hole mobility ($\mu_n \gg \mu_p$). Semiconductors with such strong asymmetry, that also exhibit the efficiency droop, include GaN (exhibiting droop at 300 K) as well as AlGaN_P (exhibiting droop at cryogenic temperatures). It was shown that the onset of the efficiency droop in GaInN LEDs indeed occurs in the high-level injection regime (Meyaard *et al.*, 2011).

In a quantum-well (QW) active region, the quantity $\Delta n_p(0)$ depends on the carrier concentration in the QW, n_{QW} , according to $\Delta n_p(0) = n_{QW} \exp(-\Delta E_{Barrier}/kT) = \delta n_{QW}$, where $\Delta E_{Barrier}$ is the effective energy difference between carriers in the QW and carriers injected into the p-type layer; for good confinement, $\delta \ll 1$. Indeed, the numerical value of δ was estimated to be on the order of 0.1% (Lin *et al.*, 2012).

In pn QW and heterojunctions, the electron concentration injected into the p-type neutral region is, due to a heterojunction barrier, much smaller than in pn homojunctions. In pn heterojunctions, we denote the electron concentration injected into the p-type neutral region by $\Delta n_p(0)$. Given that, in a GaInN LED, the thickness of p-type GaN cladding layer is smaller than the electron diffusion length, we write the electron diffusion current leaking out of the active region as

$$J_{Diffusion} = \frac{e D_n \Delta n_p(0)}{L_{p-Layer}} = \frac{e D_n}{L_{p-Layer}} \delta n_{QW} \quad (22.6)$$

where $L_{p-Layer}$ is the thickness of the p-type layer, and D_n is the electron diffusion coefficient in the p-type layer. As the diode enters high-level injection, some of the applied voltage starts to drop across the low-conductivity p-type layer and a drift current arises. The drift current of electrons injected into the p-type neutral layer, at the edge of the neutral layer, is given by

$$J_{Drift} = e \mu_n \Delta n_p(0) \mathcal{E} = e \mu_n \delta n_{QW} \frac{J_{Total}}{\sigma_p} \quad (22.7)$$

where \mathcal{E} , J_{Total} , and σ_p , is the electric field in the p-type layer, the total current density of the diode, and the p-type layer conductivity ($\sigma_p = e p_{p0} \mu_p$), respectively. The drift-induced leakage current increases with the total current, and will, at a sufficiently large current, become significant. As a consequence, the injection efficiency into the active region is reduced and the device enters the droop regime.

Close to the peak-efficiency point, where radiative recombination dominates, the recombination rate can be approximated by $R \approx B n_{QW}^2$. In this region, the total current density, J_{Total} , depends on the carrier concentration in the QW according to

$$J_{Total} = e d_{active} R \approx e d_{active} B n_{QW}^2 \quad (22.8)$$

where d_{active} is the active-region thickness. Inserting Eqn. (22.8) into Eqn. (22.7), we find the following dependence of the drift-induced leakage-current density on carrier concentration

$$J_{Drift} = e \mu_n \delta n_{QW} \frac{J_{Total}}{e \mu_p p_{p0}} \approx e d_{active} \frac{\delta \mu_n}{\mu_p p_{p0}} B n_{QW}^3 = e d_{active} C_{DL} n_{QW}^3. \quad (22.9)$$

where C_{DL} is a proportionality constant associated with the lowering of the injection efficiency due to drift of electrons in the p-type layer (“drift leakage”). Since the $f(n) \propto n_{QW}^3$ dependence (drift-induced leakage, see Eqn. 22.9) is stronger than the $f(n) \propto n_{QW}$ dependence (diffusion-induced leakage, see Eqn. 22.6), the latter one may be neglected. Writing $f(n) = C_{DL} n_{QW}^3 + D_{DL} n_{QW}^4 + \dots$, allows one to identify the third-order coefficient as

$$C_{DL} = \frac{\delta \mu_n}{\mu_p p_{p0}} B \quad (22.10)$$

As a numerical example of a GaN-based LED, we choose: $p_{p0} = 5.0 \times 10^{17} \text{ cm}^{-3}$, $\mu_p = 2.5 \text{ cm}^2/(\text{V s})$, $\mu_n = 300 \text{ cm}^2/(\text{V s})$, $B = 10^{-10} \text{ cm}^3/\text{s}$, and $\delta = 0.1\%$. Using these values, we obtain $C_{DL} = 2.4 \times 10^{-29} \text{ cm}^6/\text{s}$, in agreement with experimental values (Lin *et al.*, 2012).

At even higher current densities, when the drift-induced leakage current becomes significant, the dependence of the total current density, J_{Total} , shifts from a $J_{Total} \propto n_{QW}^2$ dependence to a $J_{Total} \propto n_{QW}^3$ dependence, i.e.

$$J_{Total} = ed_{active} R \approx ed_{active} C_{DL} n_{QW}^3 \quad (22.11)$$

and consequently, the dominant term of the loss function $f(n)$ shifts from $f(n) \propto n_{QW}^3$ to $f(n) \propto n_{QW}^4$. Inserting Eqn. (22.11) into Eqn. (22.9) yields

$$J_{Drift} = e\mu_n \delta n_{QW} \frac{J_{Total}}{e\mu_p p_{p0}} \approx ed_{active} \left(\frac{\delta \mu_n}{\mu_p p_{p0}} \right)^2 B n_{QW}^4 = ed_{active} D_{DL} n_{QW}^4. \quad (22.12)$$

The equation allows one to identify the fourth-order coefficient as

$$D_{DL} = \left(\frac{\delta \mu_n}{\mu_p p_{p0}} \right)^2 B. \quad (22.13)$$

A fourth-power dependence of $f(n)$ has indeed been reported in the literature (Dai *et al.*, 2010).

The total recombination rate can now be written as

$$\begin{aligned} R &= A_{SRH} n_{QW} + B n_{QW}^2 + C_{Auger} n_{QW}^3 + f(n_{QW}) \\ &= A_{SRH} n_{QW} + B n_{QW}^2 + C_{Auger} n_{QW}^3 + C_{DL} n_{QW}^3 + D_{DL} n_{QW}^4 \end{aligned} \quad (22.14)$$

where both drift-induced reduction of the injection efficiency as well as Auger losses are included. A comparison reveals that $C_{DL} \gg C_{Auger}$, so that the C_{Auger} can be neglected. Squaring Eqn. (22.10) and dividing by Eqn. (22.13) yields

$$C_{DL}^2 / (D_{DL} B) = 1.0 \quad (22.15)$$

This relation is helpful in verifying the applicability of the drift-induced-leakage model. To extract the A , C , and D coefficients from experiments, the numerical values of the parameters d_{active} , B ,

and IQE_{peak} need to be known. However, based on Eqn. (22.15), the extracted ratio C^2/DB is a constant and thus does not depend on the numerical values of these parameters. Experimental results have confirmed that the theoretical ratio C_{DL}^2/D_{DL} is indeed close to the experimentally determined ratio C^2/D .

Finally, based on the drift-induced leakage model, we calculate the current density at which the efficiency reaches its peak value, i.e. the onset-of-droop current density. Based on the ABC model, the carrier concentration, at which the efficiency reaches its maximum, $n_{peak-IQE}$, can be expressed as (Lin *et al.*, 2012)

$$\eta_{peak-IQE} = \sqrt{A_{SRH}/C} \quad (22.16)$$

where we neglected the contribution of 4th-order terms (i.e. the D coefficient). Near the peak-efficiency point, the total recombination rate R is dominated by Bn^2 . Therefore, based on the drift-induced leakage model, the onset-of-droop current density can be expressed as

$$J_{onset-of-droop} = ed_{active}Bn_{peak-IQE}^2 = ed_{active} \frac{BA_{SRH}}{C_{DL}} = ed_{active}A_{SRH} \frac{p_{p0} \mu_p}{\delta \mu_n}. \quad (22.17)$$

As a numerical example, we use the parameters $d_{active} = 3.0$ nm, $A_{SRH} = 1.0 \times 10^7$ s⁻¹, $p_{p0} = 5.0 \times 10^{17}$ cm⁻³, $\mu_p = 2.5$ cm²/(V s), $\mu_n = 300$ cm²/(V s), and $\delta = 0.1\%$. Using these values in the above equation, we calculate an onset-of-droop current density of 2.0 A/cm². This value is within the range of experimental onset-of-droop current densities, which typically are between 1.0 and 10 A/cm².

In summary, electron drift in the p-type layer of the diode causes a decrease of the injection efficiency and an associated reduction in internal quantum efficiency. The drift-induced leakage function has a strong 3rd-order as well as a 4th-order dependence on the carrier concentration. C_{DL} is found to be approximately 10^{-29} cm⁶/s. The analytic model derived also provides an analytic formula for the onset-of-droop current density.

With respect to the temperature dependence of the C_{DL} coefficient, and considering Eqn. (22.10), p_{p0} depends on the lattice temperature, whereas δ depends on the carrier temperature. Therefore, p_{p0} may exhibit the strongest temperature dependence. As a result, the efficiency droop is expected to increase at low temperatures.

Piprek *et al.* (2010) proposed that the carrier leakage current can be described by

$$I_{Leakage} = (I_{QW})^m \quad (22.18)$$

where I_{QW} is the current flowing into the QW and m is a number that can be assumed to be greater than 1.0. Let us consider the regime where radiative recombination dominates, i.e. in the vicinity of the peak-efficiency point; in this regime $I_{QW} \propto n_{QW}^2$. Since Eqns. (22.9) and (22.12) reveal $I_{Leakage} \propto n_{QW}^3$, and $I_{Leakage} \propto n_{QW}^4$, respectively, one may conclude that $m = 1.5$ near the efficiency peak and rising to $m = 2.0$ in the extreme droop regime.

22.5 Role of “polarization charges and electric fields”

The leakage of electrons out of active region in GaInN LEDs is reduced by the electron-blocking layer (EBL). The EBL is p-doped thereby enabling efficient hole transport while hindering electrons from escaping from the active region. Although EBLs can be efficient in reducing electron leakage, it has been suggested that the strong polarization fields indeed are instrumental in causing electron leakage out of the active region (Kim *et al.*, 2007).

In order for electrons to leak out of the active region, these electrons may either *not be captured* by the active region (e.g. a quantum well), or are captured by the QW but subsequently *escape* from the QW.

The ***non-capture of electrons*** in QW can be due to the following reasons, as illustrated in **Figure 22.6** (Piprek, 2010; Schubert and Schubert, 2010; Zhu *et al.*, 2010): *Firstly*, in GaInN LEDs, the QWs are very thin, usually about 2.5~3.0 nm. The capture probability of a QW is proportional to the thickness of the QW. *Secondly*, due to polarization effects, GaInN QWs are inherently asymmetric, that is, their left-hand “entrance barrier” is much higher than the right-hand “exit barrier” for Ga-face growth. The difference in barrier height can be several 100 meV. Therefore, in order to be captured by a QW, electrons must rapidly lose energy when propagating over the QW.

The ***escape of electrons*** out of the QW can be due to different reasons (Piprek, 2010; Schubert and Schubert, 2010; Zhu *et al.*, 2010): *Firstly*, as illustrated in **Figure 22.6**, a positive sheet charge exists at the spacer-to-EBL interface. This positive sheet charge attracts electrons. When the Al content in an AlGaN EBL is increased, so does the positive sheet charge. Therefore, an increase of the Al content in the EBL to curtail electron leakage has only limited success.

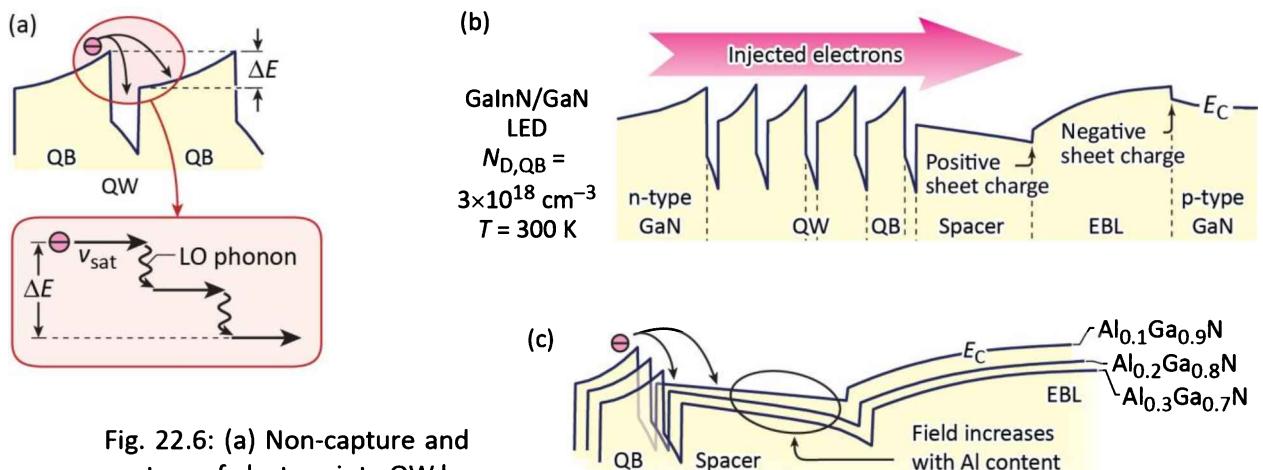


Fig. 22.6: (a) Non-capture and capture of electron into QW by emission of phonons. (b) Simulated conduction-band diagram of GaInN/GaN LED. Due to the positive space charge at the spacer-EBL interface, the electric field in the spacer layer makes electrons drift away from the QWs. (c) As the Al content in the EBL increases, so does the sheet charge at the spacer-EBL interface and the field in the spacer layer (after Zhu *et al.*, 2010).

Secondly, and unrelated to polarization effects but nevertheless relevant are the p-type doping properties of the EBL. In order to enable efficient hole transport in the EBL, commonly made of AlGaN, it is p-doped. However, there are no shallow acceptors available for AlGaN. Therefore, the p-type doping efficiency in AlGaN is quite low. Whereas an EBL with high p-type conductivity blocks electrons, an undoped EBL blocks both, electrons and holes, as illustrated in **Figure 22.7**. This shows that the EBL, to be efficient, must be heavily p-doped.

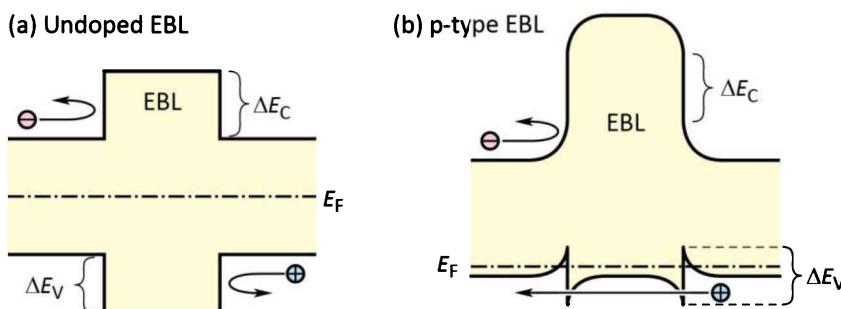


Fig. 22.7: (a) Undoped and (b) p-doped electron-blocking layer (EBL). An undoped EBL blocks electrons as well as holes. A p-doped EBL blocks electrons, but not holes.

For these reasons, the unique polarization charges and electric fields occurring in GaInN LEDs, contribute to electron leakage and the efficiency droop.

22.6 Role of “thermal roll-over”

In III–V phosphides and arsenides, thermal heating is the major cause for the efficiency droop. For this reason, the droop is called ***thermal rollover***. Thermal rollover occurs when the heat

generated in the chip is so great that the light-output power decreases when the injection current exceeds a certain value. Thermal rollover occurs under direct-current (DC) injection conditions, since heating is stronger under DC injection than under pulsed injection conditions.

Figure 22.8 shows the optical emission power of a large-area AlGaNp LED for pulsed and DC injection conditions. Inspection of the figure reveals that the light-output curves are very different for DC and pulsed injection. This large difference indicates that the roll-over is indeed caused by junction heating. Thermal rollover at high junction temperatures is caused by increased non-radiative recombination through traps (SRH recombination) and by carriers no longer being confined to the active region but escaping from the active region (Kish *et al.*, 1994).

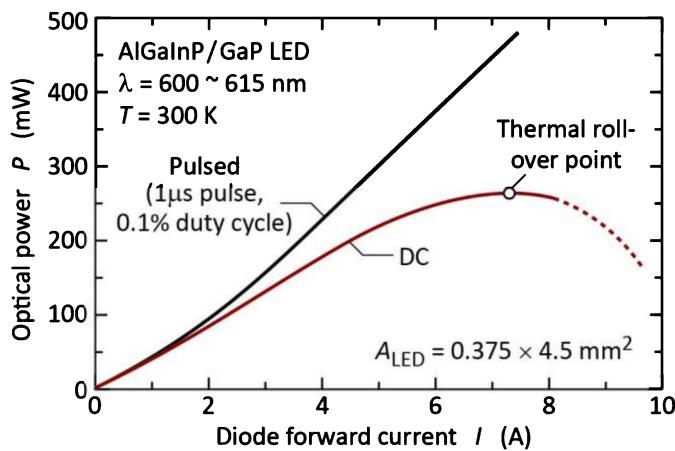


Fig. 22.8: Optical power as a function of injection current for an AlGaNp / GaP LED for DC and pulsed injection conditions. The device has a large area of 1.69 mm^2 . Under DC injection at a current of 7.3 A , the optical power reaches a maximum (“thermal roll-over point”). Increasing the injection current beyond this point results in a decrease of the optical power (after Kish *et al.*, 1994).

Thermal rollover is prominent in III–V phosphides and arsenides and this effect also must be relevant in III–V nitrides. However, in III–V nitrides, comparison between pulsed and DC light-output-versus-current measurements shows that thermal effects play only a minor role in causing the efficiency droop.

22.7 Role of “carrier de-localization”

In the ternary alloy semiconductor GaInN, Ga and In atoms are randomly distributed on the group-III lattice sites (cation sites) of the wurtzite crystal. Therefore, in a random alloy GaInN, the relative concentrations of Ga and In fluctuate. Accordingly, the random alloy GaInN has In-rich regions as well as Ga-rich regions that have a smaller and larger bandgap, respectively.

In addition to random-alloy fluctuations, additional fluctuations of the indium concentrations (“indium clustering”) have been discussed in the literature (Chichibu *et al.*, 1996, 2006; Mukai *et*

al., 1999). In-clustering effects have been attributed to phase separation and spinodal decomposition. Micro-luminescence measurements have indeed revealed non-uniform light emission. The non-uniformity can be due to random-alloy variations as well as In clustering effects.

Free carriers prefer to occupy low-energy states, that is, In-rich regions. **Localization** of electrons in a QW is illustrated in **Figure 22.9**. It has been suggested that carrier-localization effects prevent electrons from diffusing towards dislocations where they would recombine non-radiatively (Chichibu *et al.*, 2006).

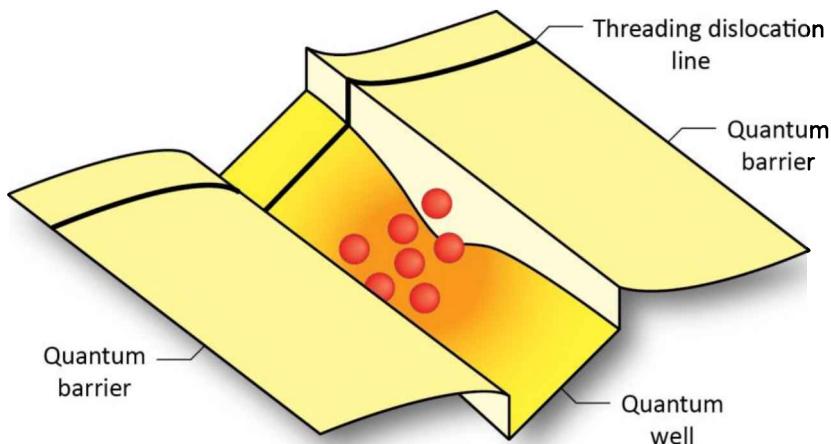


Fig. 22.9: Illustration of electrons localized in a low-energy In-rich domain of a GaInN quantum well. It has been postulated that the low-energy regions attract electrons and keep them from reaching non-radiative centers such as dislocation lines.

It has been suggested that at high carrier densities, carriers are no longer localized and thus diffuse towards dislocations when they recombine radiatively (Chichibu *et al.*, 1996; Mukai *et al.*, 1999; Kim *et al.*, 2001). The **de-localization** of carriers is a possible explanation of droop. However, no evidence was found that the droop would decrease as the dislocation density decreases. In fact, the opposite was found: The efficiency droop increases when the dislocation density decreases (Schubert *et al.*, 2007). Therefore, carriers de-localization is unlikely to be a significant factor in causing the efficiency droop.

22.8 Role of Auger recombination

Auger recombination is a non-radiative recombination process that involves a total of three carriers: Either two electrons and one hole, or, one electron and two holes. Accordingly, Auger recombination rate depends on the carrier concentration according to $R_{Auger} = C_{eeh}n^2p + C_{ehh}np^2$. Under high excitation conditions, where $n = p$, the Auger rate can be written as

$R_{Auger} = C_{Auger} n^3$, where C_{Auger} is the Auger coefficient ($C_{Auger} = C_{eeh} + C_{ehh}$).

It is well known that the Auger coefficient decreases exponentially with the bandgap energy and exhibits a strong temperature dependence as deduced from the formula $C_{Auger} \propto \exp(-E_g/kT)$ (Beattie and Landsberg, 1959). For this reason, the Auger coefficient is expected to be very small in GaN as well as other wide-gap semiconductors. Despite this expectation, Shen *et al.* (2007) and Gardner *et al.* (2007) proposed that the Auger effect causes the efficiency droop in GaInN LEDs.

It has been shown that *inter-band Auger recombination* (i.e. Auger recombination involving multiple conduction bands), and *indirect Auger recombination* (i.e. phonon-assisted Auger recombination) yield higher Auger coefficients than conventional Auger recombination, i.e. Auger recombination involving just two bands (CB and VB) and no phonons. For GaInN having a bandgap wavelength of about 460 nm (blue), the following Auger coefficients were found: 10^{-34} cm⁶/s (Hader *et al.*, 2008), 10^{-32} cm⁶/s (Delaney *et al.*, 2009), 10^{-31} cm⁶/s (Bertazzi *et al.*, 2010), 5×10^{-32} cm⁶/s (Kioupakis *et al.*, 2011), 4.5×10^{-31} cm⁶/s (Scheibenzuber *et al.*, 2011), 1.8×10^{-31} cm⁶/s (Brendel *et al.*, 2011), and 10^{-31} cm⁶/s (Bertazzi *et al.*, 2012).

The quantitative value of the Auger coefficient in GaInN indeed is a key point in answering the question: Is the Auger effect a *significant* effect in GaInN LEDs? Modeling calculations show that if $C_{Auger} \approx 10^{-31}$ cm⁶/s or smaller, then Auger recombination is a marginal effect in typical GaInN LEDs. The analysis by Lin *et al.* (2012) revealed that the C coefficient due to drift-induced-electron leakage, C_{DL} , is much greater than the Auger coefficient, i.e., $C_{DL} \gg C_{Auger}$, with $C_{DL} \approx 10^{-29}$ cm⁶/s and $C_{Auger} \approx 10^{-31}$ cm⁶/s. Therefore, due to the relatively small magnitude of the Auger coefficient, it may be concluded that the Auger effect is an insignificant contribution to the efficiency droop in GaInN LEDs.

In a review, Piprek (2010) discussed the Auger effect and other effects as possible causes of the efficiency droop. The author concluded that although the Auger effect certainly plays a role in GaInN LEDs, it is unlikely that the Auger effect is the dominant cause of the efficiency droop. Bertazzi *et al.* (2010) concluded that although Auger recombination may play a role in GaInN LEDs, it is unlikely the dominant effect in causing the efficiency droop. A similar conclusion was drawn by Brendel *et al.* (2011).

Arguendo, if Auger recombination were indeed the dominant cause for droop, then recombination could be described by the ABC model. This theoretical model predicts an even

symmetry of the *EQE-versus-n* curve (Dai *et al.*, 2011). However, experimental *EQE-versus-n* curves consistently deviate from such even symmetry (Lin *et al.*, 2012).

22.9 Strategies to overcome the efficiency droop

There are multiple strategies to reduce or alleviate the efficiency droop. In the present section, we discuss these strategies.

P-type doping of the p-type confinement layer: Insufficient p-type doping is the root cause of the efficiency droop. The calculated expression for C_{DL} (see Eqn. 22.10) shows the desirability of high hole concentrations, as well as high hole mobilities. Dopant concentrations that are too low will lead to inefficient hole injection. To maximize hole injection efficiency, the p-type doping concentration should be as high as possible but not so high as to generate defects. Very high p-type doping concentrations ($N_A > 10^{19} \text{ cm}^{-3}$) are known to create an abundance of point defects that lead to non-radiative recombination. Furthermore, a defective material is more prone to the diffusion of acceptors that can deteriorate the active region. Therefore, a narrow concentration “window” exists for optimum p-type doping.

Heat sinking: Inadequate heat sinking will increase the junction temperature and decrease the efficiency of the device. Therefore, adequate heat sinking will reduce the droop.

Uniform current injection: For GaN-based lateral LEDs, the current flow through the active region can be crowded (non-uniform) or uniform, as shown in *Figure 22.10* (a) and (b), respectively. A crowded current flow always results in very high current densities in the current-crowded region. Higher current densities result in more droop. Therefore, in order to minimize droop, the current should flow uniformly across the p-n junction.

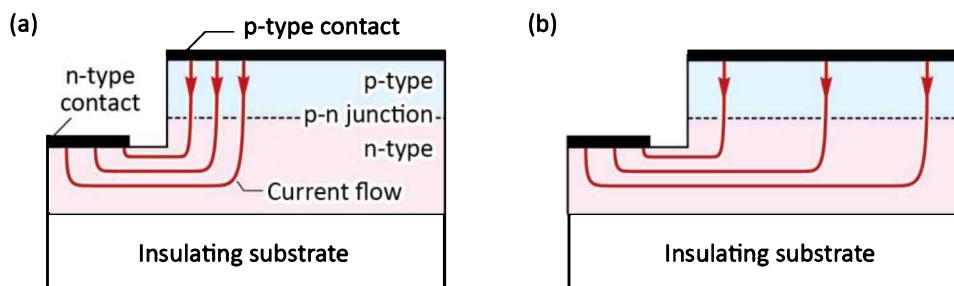


Fig. 22.10: Lateral GaN-based LED grown on an insulating substrate (a) with and (b) without current crowding.

Larger chip area: The magnitude of the droop increases with current density. To decrease the magnitude of the droop, the injection current should be reduced. Alternatively, the device area can be increased, thereby reducing the current density (Edmond *et al.*, 2010). *Figure 22.11* shows the efficiency as a function of current for devices with different areas and a different number of quantum wells. The results show that the droop can be reduced by a larger device area. The figure also shows that an increasing number of quantum wells does

not help to reduce the droop (since carriers mostly recombine in one QW, the one located most closely to the p-type side of the junction).

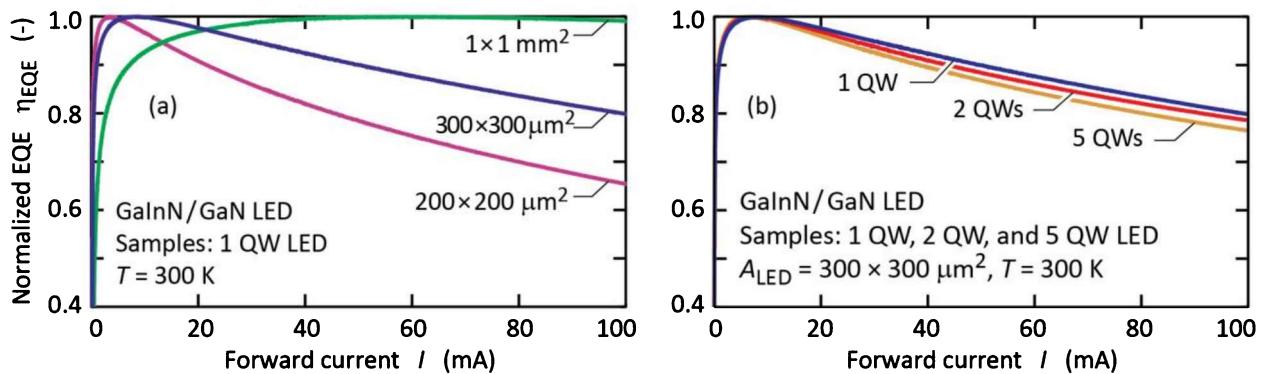


Fig. 22.11: (a) External quantum efficiency (EQE) vs. current for devices with different chip area. (b) EQE vs. current for devices with 1, 2, and 5 quantum wells (after Zhu, 2011).

Thicker QWs: Based on the fact that the efficiency droop has a strong Cn^3 dependence, any decrease in the carrier concentration n , e.g. by an increase in QW thickness, can be very helpful. A doubling of the QW thickness should reduce the carrier concentration n by about a factor of 2 and n^3 by a factor of 8. As a consequence, the Cn^3 term is reduced significantly. However, a wider QW and the concomitant lower n will increase Shockley-Read-Hall recombination. Therefore, very thick QWs (e.g. 20 nm) are prone to decrease the efficiency of LEDs, at least in the low-current regime.

Polarization matching: Given that the droop is related to the strong polarization fields occurring in GaN, there exists a strong motivation to reduce the polarization fields. This can be accomplished in two ways: **(A) Growth on polar substrate orientation:** Epitaxial growth of GaInN LEDs is most commonly carried out on the polar (c-plane) substrate orientation. This substrate orientation results in high polarization fields. It was shown by Kim *et al.*, (2007) that the polarization mismatch between different materials can be reduced or even eliminated by using ternary or quaternary AlGaN semiconductors within the active region. There are two possibilities to reduce the polarization mismatch: *Firstly*, in the MQW region the use of AlGaN QBs instead of GaN QBs reduces polarization mismatch. *Secondly*, in the electron-blocking layer (EBL) region, the use of AlGaN EBLs instead of AlGaN EBLs reduces the polarization mismatch. For both options, appropriate compositions of the $\text{Al}_x\text{Ga}_{1-x}\text{In}_y\text{N}$ need to be selected. The selection can be done by using **Figure 22.12** that shows curves of equal bandgap as well as curves of equal polarization (Kim *et al.*, 2007). Any material composition located on a curve of equal bandgap has the same bandgap. Any material composition located on a curve of equal polarization has the same polarization. Consequently, the figure can be used to “navigate” and choose the AlGaN compositions that are polarization matched (or partially matched). It has been demonstrated that the polarization matching indeed can

reduce the efficiency droop (see, for example, Schubert *et al.*, 2008; Xu *et al.*, 2009; Kuo *et al.*, 2009). A disadvantage of the approach is that the growth of quaternary materials (AlGaN) is more difficult than the growth of ternary materials (AlGaN and GaN).

(B) Growth on non-polar (non-c-plane) substrate orientation: Epitaxial growth of GaN LEDs can also be carried out on the semi-polar and non-polar (non-c-plane) substrate orientations (Walzereit *et al.*, 2000; Chakraborty *et al.*, 2004). This method completely eliminates the polarization mismatch found for c-plane growth. Indeed, reports have confirmed that the efficiency droop is strongly reduced for growth on non-c-plane growth. A disadvantage of the approach is that GaN generally does not grow easily along non-c-plane directions.

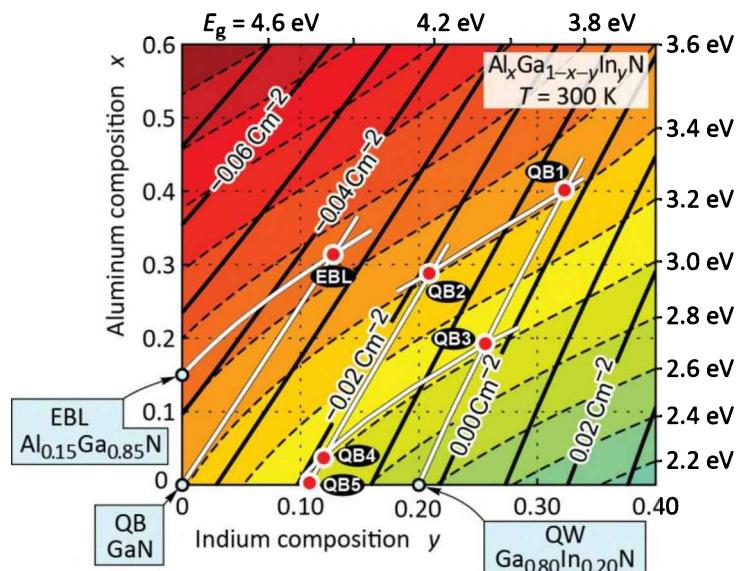


Fig. 22.12: Bandgap energy and polarization sheet charges of $\text{Al}_x\text{Ga}_{1-x-y}\text{In}_y\text{N}$ grown pseudomorphically on GaN . The diagram allows one to identify alloy compositions of AlGaN that are polarization-matched or bandgap-energy-matched to GaN quantum barriers (QB) or GaN quantum wells (QWs) or AlGaN electron-blocking layers (EBLs). As examples, the following compositions are identified: The alloy of point “QB 1” has the same E_g as GaN and no polarization mismatch with $\text{Ga}_{0.80}\text{In}_{0.20}\text{N}$. “QB 2” has the same E_g as GaN and half the polarization mismatch compared with $\text{Ga}_{0.80}\text{In}_{0.20}\text{N}/\text{GaN}$; and so on. “QB 5” is $\text{Ga}_{0.90}\text{In}_{0.10}\text{N}$ which has half

the QB height (half E_g) and half the polarization mismatch. The point “EBL” has the same E_g as $\text{Al}_{0.15}\text{Ga}_{0.85}\text{N}$ and no polarization mismatch with GaN (after Kim *et al.*, 2007).

Tailored doping in the active region: The escape of electrons out of the active region has been identified as a droop-causing loss mechanism. Therefore, if electrons are hindered in their transport across the active region, they are less likely to reach the last QW. Electrons that do not reach the last QW, cannot overshoot or escape from the last QW. Based on this reasoning, one can implement ways to reduce electron transport. In order to reduce electron injection (thereby increasing hole injection), the n-type doping of the QBs in the active region can be reduced. Reduction of the doping in active region hinders electrons and thus leads to a more uniform distribution of carriers among the QWs (Zhu *et al.*, 2010).

Advanced electron-blocking layers (EBLs): The EBL is a key layer, directly adjoining the GaN spacer layer, which is intended to prevent electron leakage out of the active region. Conventional EBLs are made of $\text{Al}_x\text{Ga}_{1-x}\text{N}$. However, there are three problems associated with conventional AlGaN EBLs: *Firstly*, the ability to p-dope AlGaN decreases with increasing Al

content. *Secondly*, a positive polarization-sheet charge is located at the spacer-EBL interface. This sheet charge attracts electrons and thus enhances leakage. As the Al content in the EBL is increased, so does the positive sheet charge (thereby more strongly attracting electrons). *Third*, while EBLs are intended to block electrons and thus are p-doped, they also can block holes, especially if the EBL's p-type doping is not sufficiently high. For these reasons, the ability of conventional EBLs to block electrons is limited. There are several ways that EBLs can be improved over the conventional AlGaN EBLs; Specific ways are: **(A)** Using an AlGaN / GaN superlattice EBL is known to enhance p-type doping and improve hole injection into the active region. **(B)** Using a polarization matched EBL, specifically an AlGaN EBL, has been shown to lower electron leakage out of the active region (Kim *et al.*, 2007; Choi *et al.*, 2010). **(C)** Composition-graded EBLs can introduce an asymmetry between electron and hole transport. This is because electrons and holes cross the EBL in different directions. This asymmetry can be used, to hinder electrons more so than holes (for example by smoothly ramping up the Al composition from the p-side towards the active region).

Thin quantum barriers for reduction of the polarization field: The polarization electric field in a MQW structure does not only depend on the inherent polarization charges at the interfaces, but also on geometrical factors. For a periodic MQW structure, one can derive the following formula for the polarization electric field in the QW (Leroux *et al.*, 1999; Xu, 2011):

$$\mathcal{E}_{QW} = \sigma \frac{L_{QB}}{\epsilon_{QB} L_{QW} + \epsilon_{QW} L_{QB}} \approx \frac{\sigma}{\epsilon_r \epsilon_0} \frac{L_{QB}}{L_{QW} + L_{QB}} \quad (22.19)$$

where σ is the polarization-sheet-charge density occurring at the interfaces of the MQW structure, and ϵ_r is the relative dielectric constant of GaN ($\epsilon_r = 8.9$). The formula shows that MQW structures have smaller electric fields than single QW (SQW) structures (for which $L_{QB} \rightarrow \infty$). For this reason, MQW active regions are preferable over SQW active regions. Furthermore, inspection of the equation shows that the electric field in the QWs can be reduced by employment of *thin* QBs. To summarize, the advantages of thin QBs are: **(A)** Thin QBs reduce the electric field in the QWs and thus reduce the QCSE. **(B)** Thin QBs reduce the electric field in the QWs and thus reduce the likelihood of ballistic transport of electrons over a QW. **(C)** Thin QBs will reduce the transport of electrons *over* the barriers and promote the transport *through* the barriers (by tunneling), thereby reducing the likelihood of ballistic transport of electrons over a QW. Indeed, state-of-the art GaInN MQW structures use QBs that are thin, i.e. on the order of 3~5 nm (while traditionally thicker QBs were used, i.e. on the order of 8~12 nm).

Strain-control combined with thicker QWs: Strain-control through GaInN under-layers allows one to reduce the polarization mismatch in the active region. A reduced QCSE is attained in

the active region. This in turn allows one to increase the QW thickness from typically 25 to 35 Å. Thicker QWs improve electron capture into the QWs and reduce the efficiency droop.

QW profile engineering: It has been proposed that the detailed engineering of the band profiles of QWs will improve efficiency and reduce QSCE and droop. Several different profiles have been proposed (Zhao *et al.*, 2009).

22.10 References

- Beattie A. R. and Landsberg P. T. "Auger effect in semiconductors" *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences* **249**, 16 (1959)
- Bertazzi F., Goano M., and Bellotti E. "A numerical study of Auger recombination in bulk InGaN" *Applied Physics Letters* **97**, 231118 (2010)
- Bertazzi F., Goano M., and Bellotti E. "Numerical analysis of indirect Auger transitions in InGaN" *Applied Physics Letters* **101**, 011111 (2012)
- Brendel M., Kruse A., Jönen H., Hoffmann L., Bremers H., Rossow U., and Hangleiter A. "Auger recombination in GaInN/GaN quantum well laser structures" *Applied Physics Letters* **99**, 031106 (2011)
- Chakraborty A., Haskell B. A., Keller S., Speck J. S., DenBaars S. P., Nakamura S., and Mishraa U. K. "Nonpolar InGaN / GaN emitters on reduced-defect lateral epitaxially overgrown a-plane GaN with drive-current-independent electroluminescence emission peak" *Applied Physics Letters* **85**, 5143 (2004)
- Chichibu S., Azuhata T., Sota T., and Nakamura S. "Spontaneous emission of localized excitons in InGaN single and multiquantum well structures" *Applied Physics Letters* **69**, 30 (1996)
- Chichibu S. F., Uedono A., Onuma T., Haskell B. A., Chakraborty A., Koyama T., Fini P. T., Keller S., DenBaars S. P., Speck J. S., Mishra U. K., Nakamura S., Yamaguchi S., Kamiyama S., Amano H., Akasaki I., Han J., and Sota T. "Origin of defect-insensitive emission probability in In-containing (Al, In, Ga)N alloy semiconductors" *Nature Materials* **5**, 810 (2006)
- Choi S., Kim H. J., Kim S.-S., Liu J., Kim J., Ryou J.-H., Dupuis R. D., Fischer A. M., and Ponce F. A. "Improvement of peak quantum efficiency and efficiency droop in III-nitride visible light-emitting diodes with an AlInN electron-blocking layer" *Applied Physics Letters* **96**, 221105 (2010)
- Dai Q., Shan Q., Wang J., Chhajed S., Cho J., Schubert E. F., Crawford M. H., Koleske D. D., Kim M.-H., and Park Y. "Carrier recombination mechanisms and efficiency droop in GaInN/GaN light-emitting diodes" *Applied Physics Letters* **97**, 133507 (2010)
- Dai Q., Shan Q., Cho J., Schubert E. F., Crawford M. H., Koleske D. D., Kim M.-H., and Park Y. "On the symmetry of efficiency-versus-carrier-concentration curves in GaInN/GaN light-emitting diodes and relation to droop-causing mechanisms" *Applied Physics Letters* **98**, 033506 (2011)
- Delaney K. T., Rinke P., and Van de Walle C. G. "Auger recombination rates in nitrides from first principles" *Applied Physics Letters* **94**, 191109 (2009)
- Edmond J. A., Bergmann M. J., Emerson D. T., and Haberern K. W. "High output group-III Nitride light emitting diodes" US Patent 7,737,459 B2 (2010)
- Gardner N. F., Müller G. O., Shen Y. C., Chen G., Watanabe S., Götz W., and Krames M. R. "Blue-emitting InGaN-GaN double-heterostructure light-emitting diodes reaching maximum quantum efficiency above 200 A/cm²" *Applied Physics Letters* **91**, 243506 (2007)
- Hader J., Moloney J. V., Pasenow B., Koch S. W., Sabathil M., Linder N., and Lutgen S. "On the importance of radiative and Auger losses in GaN-based quantum wells" *Applied Physics Letters* **92**, 261103 (2008)
- Kim A. Y., Götz W., Steigerwald D. A., Wierer J. J., Gardner N. F., Sun J., Stockman S. A., Martin P. S., Krames M. R., Kern R. S., and Steranka F. M. "Performance of high-power AlInGaN light emitting diodes" *Physica Status Solidi A* **188**, 15 (2001)

- Kim M.-H., Schubert M. F., Dai Q., Kim J. K., Schubert E. F., Piprek J., and Park Y. "Origin of efficiency droop in GaN-based light-emitting diodes" *Applied Physics Letters* **91**, 183507 (2007)
- Kish F. A., DeFevere D. A., Vanderwater D. A., Trott G. R., Weiss R. J. and Major J. S. Jr. "High luminous flux semiconductor wafer-bonded AlGaN/P/GaP large-area emitters" *Electronics Letters* **30**, 1790 (1994)
- Kioupkis E., Rinke P., Delaney K. T., and Van de Walle C. G. "Indirect Auger recombination as a cause of efficiency droop in nitride light-emitting diodes" *Applied Physics Letters* **98**, 161107 (2011)
- Kuo Y.-K., Chang J.-Y., Tsai M.-C., and Yen S.-H. "Advantages of blue InGaN multiple-quantum well light-emitting diodes with InGaN barriers" *Applied Physics Letters* **95**, 011116 (2009)
- Leroux M., Grandjean N., Massies J., Gil B., Lefebvre P., and Bigenwald P. "Barrier-width dependence of group-III nitrides quantum-well transition energies" *Physical Review B* **60**, 1496 (1999)
- Lin G.-B., Meynard D., Cho J., Schubert E. F., Shim H., and Sone C. "Analytic model for the efficiency droop in semiconductors with asymmetric carrier-transport properties based on drift-induced reduction of injection efficiency" *Applied Physics Letters* **100**, 161106 (2012)
- Meynard D. S., Lin G.-B., Shan Q., Cho J., Schubert E. F., Shim H., Kim M.-H., and Sone C. "Asymmetry of carrier transport leading to efficiency droop in GaInN based light-emitting diodes" *Applied Physics Letters* **99**, 251115 (2011)
- Mukai T., Yamada M., and Nakamura S. "Characteristics of InGaN-Based UV / blue / green / amber / red light-emitting diodes" *Japanese Journal of Applied Physics* **38**, 3976 (1999)
- Piprek J. "Efficiency droop in nitride-based light-emitting diodes" *Physica Status Solidi A* **207**, 2217 (2010)
- Pope I. A., Smowton P. M., Blood P., Thomson J. D., Kappers M. J., and Humphreys C. J. "Carrier leakage in InGaN quantum well light-emitting diodes emitting at 480 nm" *Applied Physics Letters* **82**, 2755 (2003)
- Scheibenzuber W. G., Schwarz U. T., Sulmoni L., Dorsaz J., Carlin J.-F., and Grandjean N. "Recombination coefficients of GaN-based laser diodes" *Journal of Applied Physics* **109**, 093106 (2011)
- Schubert M. F. provision of figure is gratefully acknowledged (2010)
- Schubert M. F., Xu J., Kim J. K., Schubert E. F., Kim M. H., Yoon S., Lee S. M., Sone C., Sakong T., and Park Y. "Polarization-matched GaInN/AlGaN multi-quantum-well light-emitting diodes with reduced efficiency droop" *Applied Physics Letters* **93**, 041102 (July 2008)
- Shen Y. C., Mueller G. O., Watanabe S., Gardner N. F., Munkholm A., and Krames M. R. "Auger recombination in InGaN measured by photoluminescence" *Applied Physics Letters* **91**, 141101 (2007)
- Waltereit P., Brandt O., Trampert A., Grahn H. T., Menniger J., Ramsteiner M., Reiche M., and Ploog K. H. "Nitride semiconductors free of electrostatic fields for efficient white light-emitting diodes" *Nature* **406**, 865 (2000)
- Xu J. "Efficiency and carrier transport in polarization-matched GaInN-based light-emitting diodes" Dissertation, Rensselaer Polytechnic Institute, Troy, New York (2011)
- Xu J., Schubert M. F., Noemaun A. N., Zhu D., Kim J. K., Schubert E. F., Kim M. H., Chung H. J., Yoon S., Sone C., and Park Y. "Reduction in efficiency droop, forward voltage, ideality factor, and wavelength shift in polarization-matched GaInN/GaN multi-quantum-well light-emitting diodes" *Applied Physics Letters* **94**, 011113 (January 2009)
- Zhao H., Liu G., Li X.-H., Huang G. S., Poplawsky J. D., Penn S. T., Dierolf V., and Tansu N. "Growths of staggered InGaN quantum wells light-emitting diodes emitting at 520~525 nm employing graded growth-temperature profile" *Applied Physics Letters* **95**, 061104 (2009)
- Zhu D. "Investigating and optimizing carrier transport, carrier distribution, and efficiency droop in GaN-based light-emitting diodes" Dissertation, Rensselaer Polytechnic Institute (2011)
- Zhu D., Noemaun A. N., Schubert M. F., Cho J., Schubert E. F., Crawford M. H., and Koleske D. D. "Enhanced electron capture and symmetrized carrier distribution in GaInN light-emitting diodes having tailored barrier doping" *Applied Physics Letters* **96**, 121110 (2010)

23

23 – Human eye sensitivity and photometric quantities

The recipient of the light emitted by most visible-spectrum LEDs is the human eye. In this chapter, the characteristics of human vision and of the human eye are summarized, in particular as these characteristics relate to human eye sensitivity and photometric quantities.

23.1 Light receptors of the human eye

Figure 23.1 (a) shows a schematic illustration of the human eye (Encyclopedia Britannica, 1994). The inside of the eyeball is clad by the retina, which is the light-sensitive part of the eye. The illustration also shows the fovea, a cone-rich central region of the retina which affords the high acuteness of central vision. **Figure 23.1 (b)** shows the cell structure of the retina including the light-sensitive **rod cells** and **cone cells**. Also shown are the ganglion cells and nerve fibers that transmit the visual information to the brain. Rod cells are more abundant and more light sensitive than cone cells. Rods are sensitive over the entire visible spectrum. There are three types of cone cells, namely cone cells sensitive in the red, green, and blue spectral range. The cone cells are therefore denoted as the red-sensitive, green-sensitive, and blue-sensitive cones, or simply as the red, green, and blue cones.

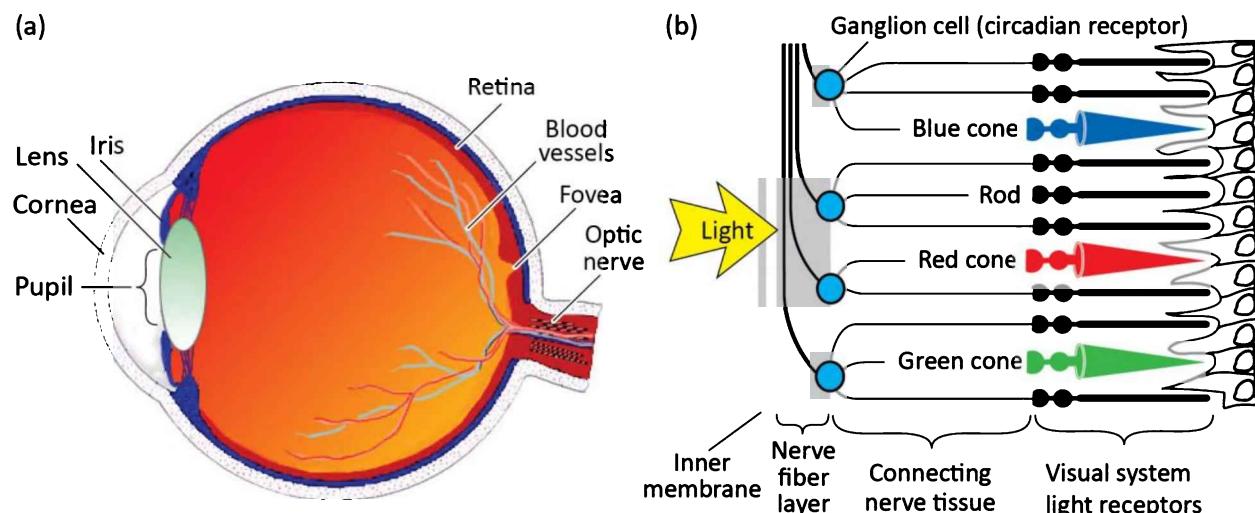


Fig. 23.1: (a) Cross section through a human eye. (b) Schematic view of the retina including rod and cone light receptors (adapted from Encyclopedia Britannica, 1994).

Three different vision regimes are shown in **Figure 23.2** along with the receptors relevant to each of the regimes (Osram Sylvania, 2000). **Photopic vision** relates to human vision at high ambient light levels (e.g. during daylight conditions) when vision is mediated by the cones. The photopic vision regime applies to luminance levels $> 3 \text{ cd/m}^2$. **Scotopic vision** relates to human vision at low ambient light levels (e.g. at night) when vision is mediated by rods. Rods have a much higher sensitivity than the cones. However, the sense of color is essentially lost in the scotopic vision regime. At low light levels such as in a moonless night, objects lose their colors and only appear to have different gray levels. The scotopic vision regime applies to luminance levels $< 0.003 \text{ cd/m}^2$. **Mesopic vision** relates to light levels between the photopic and scotopic vision regime ($0.003 \text{ cd/m}^2 < \text{luminance} < 3 \text{ cd/m}^2$).

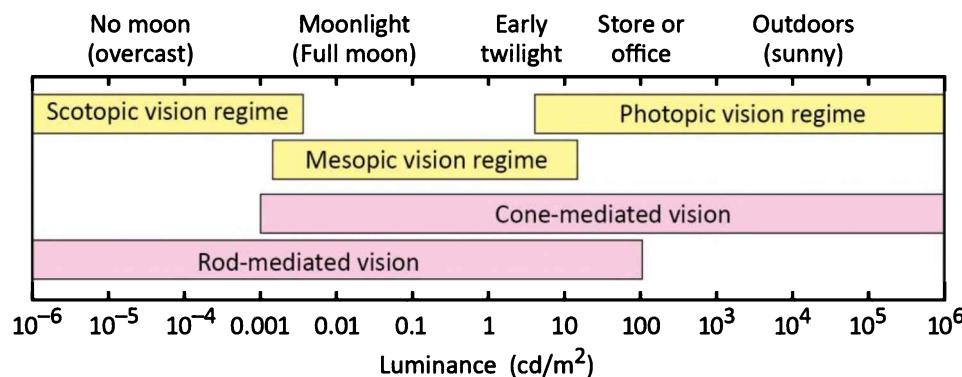


Fig. 23.2: Approximate ranges of vision regimes and receptor regimes (after Osram Sylvania, 2000).

The approximate spectral sensitivity functions of the rods and three types of cones are shown in **Figure 23.3** (Dowling, 1987). Inspection of the figure reveals that night-time vision (scotopic vision) is weaker in the red spectral range and thus stronger in the blue spectral range as compared to day-time vision (photopic vision). The following discussion mostly relates to the photopic vision regime.

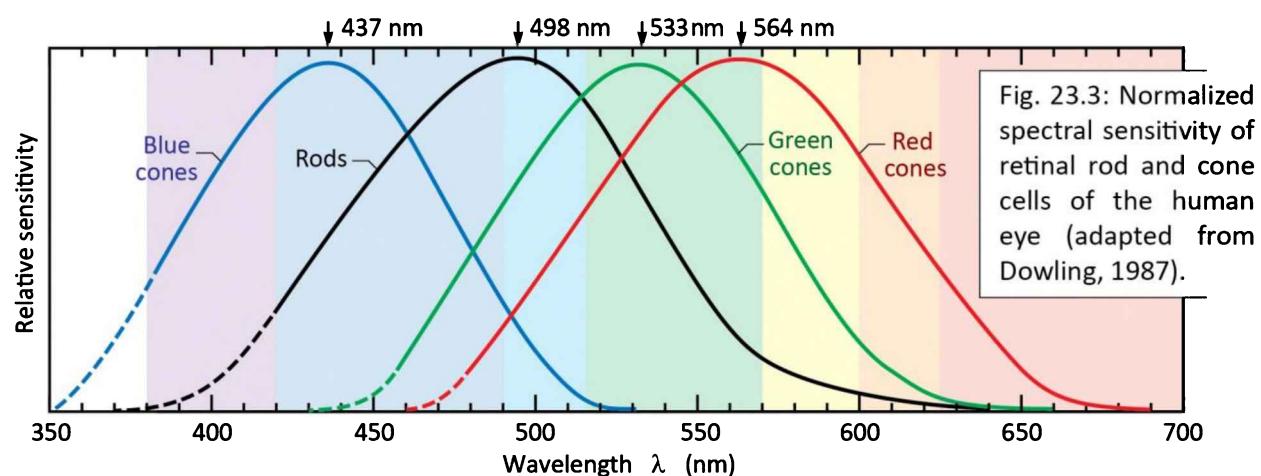


Fig. 23.3: Normalized spectral sensitivity of retinal rod and cone cells of the human eye (adapted from Dowling, 1987).

23.2 Basic radiometric and photometric units

The physical properties of electromagnetic radiation are characterized by **radiometric units**. Using radiometric units, we can characterize light in terms of physical quantities; for example, the number of photons, photon energy, and **optical power** (in the lighting community frequently called the **radiant flux**). However, the radiometric units are irrelevant when it comes to light perception by a human being. For example, infrared radiation causes no luminous sensation in the eye. To characterize the light and color sensation by the **human eye**, different types of units are needed. These units are called **photometric units**.

The **luminous intensity**, which is a photometric quantity, represents the light intensity of an optical source, as perceived by the **human eye**. The luminous intensity is measured in units of **candela** (cd), which is a base unit of the International System of Units (SI unit). The present definition of luminous intensity is as follows: **a monochromatic light source emitting an optical power of (1/683) watt at 555 nm into the solid angle of 1 steradian (sr) has a luminous intensity of 1 candela (cd).**

The unit *candela* has great historical significance. All light intensity measurements can be traced back to the candela. It evolved from an older unit, the **candlepower**, or simply, the **candle**. The original definition of one candela related to the light intensity emitted by a plumber's candle, as shown in **Figure 23.4**, which had a specified construction and dimensions. The original, now obsolete, definition of one candela was:

One standardized candle emits a luminous intensity of 1.0 cd .



Fig. 23.4: Plumber's candle, as used by plumbers in the nineteenth century to melt lead solder when joining water pipes.

The luminous intensity of a light source can thus be characterized by giving the number of standardized candles that, when combined, would emit the same luminous intensity. Note that **candlepower** and **candle** are non-SI units that are no longer current and rarely used at the present time.

The ***luminous flux***, which is also a photometric quantity, ***is the total light power of a source as perceived by the human eye***. The unit of luminous flux is the ***lumen*** (lm). It is defined as follows: ***a monochromatic light source emitting an optical power of (1/683) watt at 555 nm has a luminous flux of 1 lumen (lm)***. The lumen is an SI unit.

A comparison of the definitions for the candela and lumen reveals that 1 candela equals 1 lumen per steradian or $\text{cd} = \text{lm/sr}$. Thus, an isotropically emitting light source with luminous intensity of 1 cd has a luminous flux of $4\pi \text{ lm} = 12.57 \text{ lm}$.

The ***illuminance*** is the luminous flux incident per unit area. The illuminance measured in ***lux*** ($\text{lux} = \text{lm/m}^2$). It is an SI unit used when characterizing illumination conditions. ***Table 23.1*** gives typical values of the illuminance under various lighting conditions.

Table 23.1: Typical illuminance values obtained under various lighting conditions.

Illumination condition	Illuminance
Full moon	1 lux
Street lighting	10 lux
Home lighting	30 to 300 lux
Office desk lighting	100 to 1 000 lux
Surgery lighting	10 000 lux
Direct sunlight	100 000 lux

The ***luminance*** of a ***surface source*** (i.e. a source with a non-zero light-emitting surface area such as a display or an LED) is the ratio of the luminous intensity emitted in a certain direction (measured in cd) divided by the ***projected surface area*** in that direction (measured in m^2). The luminance is measured in units of cd/m^2 . In most cases, the direction of interest is normal to the chip surface. In this case, the luminance is the luminous intensity emitted along the chip-normal direction divided by the chip area.

The ***projected surface area*** mentioned above follows a cosine law, i.e. the projected area is given by $A_{\text{projected}} = A_{\text{surface}} \cos \theta$, where θ is the angle between the direction considered and the surface normal. The light-emitting surface area and the projected area are shown in ***Figure 23.5***. The luminous intensity of LEDs with lambertian emission pattern also depends on the angle θ according to a cosine law. Thus, the luminance of lambertian LEDs is a constant, independent of angle.

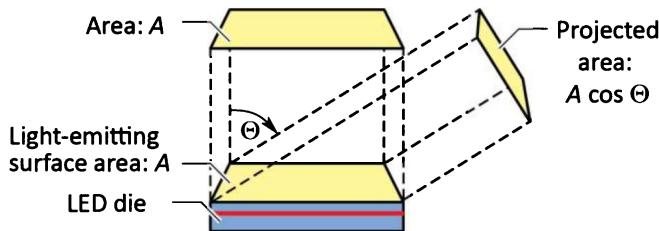


Fig. 23.5: Area of LED, A , and projected area, $A \cos \Theta$, used for the definition of the luminance of an LED.

For LEDs, it is desirable to maximize luminous intensity and luminous flux while keeping the LED chip area minimal. Thus, the luminance is a measure of how efficiently the valuable semiconductor wafer area is used to attain, at a given injection current, a certain luminous intensity.

There are several units that are used to characterize the luminance of a source. The names of these common units are given in **Table 23.2**.

Table 23.2: Conversion between common SI and non-SI units for luminance.

SI Unit	Common name	SI Unit	Common name
1 cd/cm^2	1 stilb	$(1/\pi) \text{ cd/m}^2$	1 apostilb
$(1/\pi) \text{ cd/cm}^2$	1 lambert	Non-SI Unit	Common name
1 cd/m^2	1 nit	$(1/\pi) \text{ cd/ft}^2$	1 foot-lambert

Typical luminances of displays, organic LEDs, and inorganic LEDs are given in **Table 23.3**. The table reveals that displays require a comparatively low luminance because the observer directly views the display from a close distance. This is not the case for high-power inorganic LEDs used for example in traffic light and illumination applications.

Table 23.3: Typical values for the luminance of displays, LEDs fabricated from organic materials, and inorganic LEDs.

Device	Luminance (cd/m^2)	Device	Luminance (cd/m^2)
Display	100 (operation)	Organic LED	$100 \sim 10\,000$
Display	$250 \sim 750$ (max. value)	III-V LED	$1\,000\,000 \sim 10\,000\,000$

Photometric and the corresponding radiometric units are summarized in **Table 23.4**.

Table 23.4: Photometric and corresponding radiometric units.

Photometric unit	Dimension	Radiometric unit	Dimension
Luminous flux	lm	Radiant flux (optical power)	W
Luminous intensity	lm / sr = cd	Radiant intensity	W / sr
Illuminance	lm / m ² = lux	Irradiance (power density)	W / m ²
Luminance	lm / (sr m ²) = cd / m ²	Radiance	W / (sr m ²)

Exercise: Photometric units. A 60 W incandescent light bulb has a luminous flux of 1000 lm. Assume that light is emitted isotropically from the bulb.

- (a) What is the luminous efficacy (i.e. the number of lumens emitted per watt of electrical input power) of the light bulb?
- (b) What number of standardized candles emit the same luminous intensity?
- (c) What is the illuminance, E_{lum} , in units of lux, on a desk located 1.5 m below the bulb?
- (d) Is the illuminance level obtained under (c) sufficiently high for reading?
- (e) What is the luminous intensity, I_{lum} , in units of candela, of the light bulb?
- (f) Derive the relationship between the illuminance at a distance r from the light bulb, measured in lux, and the luminous intensity, measured in candela.
- (g) Derive the relationship between the illuminance at a distance r from the light bulb, measured in lux, and the luminous flux, measured in lumen.
- (h) The definition of the cd involves the optical power of (1/683) W. What, do you suppose, is the origin of this particular power level?

Solution: (a) 16.7 lm/W. (b) 80 candles. (c) $E_{lum} = 35.4 \text{ lm/m}^2 = 35.4 \text{ lux}$. (d) Yes.
 (e) $79.6 \text{ lm/sr} = 79.6 \text{ cd}$. (f) $E_{lum} r^2 = I_{lum}$. (g) $E_{lum} 4\pi r^2 = \Phi_{lum}$.
 (h) Originally, the unit of luminous intensity had been defined as the intensity emitted by a real candle. Subsequently the unit was defined as the intensity of a light source with specified wavelength and optical power. When the power of that light source is (1/683) W, it has the same intensity as the candle. Thus, this particular power level has a historical origin and results from the effort to maintain continuity.

23.3 Eye sensitivity function

The conversion between radiometric and photometric units is provided by the **luminous efficacy function** or **eye sensitivity function**, $V(\lambda)$. In 1924, the CIE introduced the photopic eye sensitivity function $V(\lambda)$ for point-like light sources where the viewer angle is 2° (CIE, 1931). This function is referred to as the **CIE 1931 $V(\lambda)$ function**. It is the current photometric standard in the United States.

A modified $V(\lambda)$ function was introduced by Judd and Vos in 1978 (Vos, 1978; Wyszecki and Stiles, 1982, 2000) and this modified function is here referred to as the **CIE 1978 $V(\lambda)$ function**. The modification was motivated by the underestimation of the human eye sensitivity in the blue

and violet spectral region by the CIE 1931 $V(\lambda)$ function. The modified function $V(\lambda)$ has higher values in the spectral region below 460 nm. The CIE has endorsed the CIE 1978 $V(\lambda)$ function by stating “the spectral luminous efficiency function for a point source may be adequately represented by the Judd modified $V(\lambda)$ function” (CIE, 1988) and “the Judd modified $V(\lambda)$ function would be the preferred function in those conditions where luminance measurements of short wavelengths consistent with color normal observers is desired” (CIE, 1990).

The **photopic** CIE 1931 eye-sensitivity function, $V(\lambda)$, and the CIE 1978 eye-sensitivity function, $V(\lambda)$, are shown in **Figure 23.6**. The photopic eye-sensitivity function has maximum sensitivity in the yellow-green spectral range at 555 nm, where $V(\lambda)$ has a value of unity, i.e. $V(555 \text{ nm}) = 1.0$. Inspection of the figure also reveals that the CIE 1931 $V(\lambda)$ function underestimated the eye sensitivity in the blue spectral range ($\lambda < 460 \text{ nm}$). Numerical values of the CIE 1931 and CIE 1978 $V(\lambda)$ function are tabulated in Appendix 23.1.

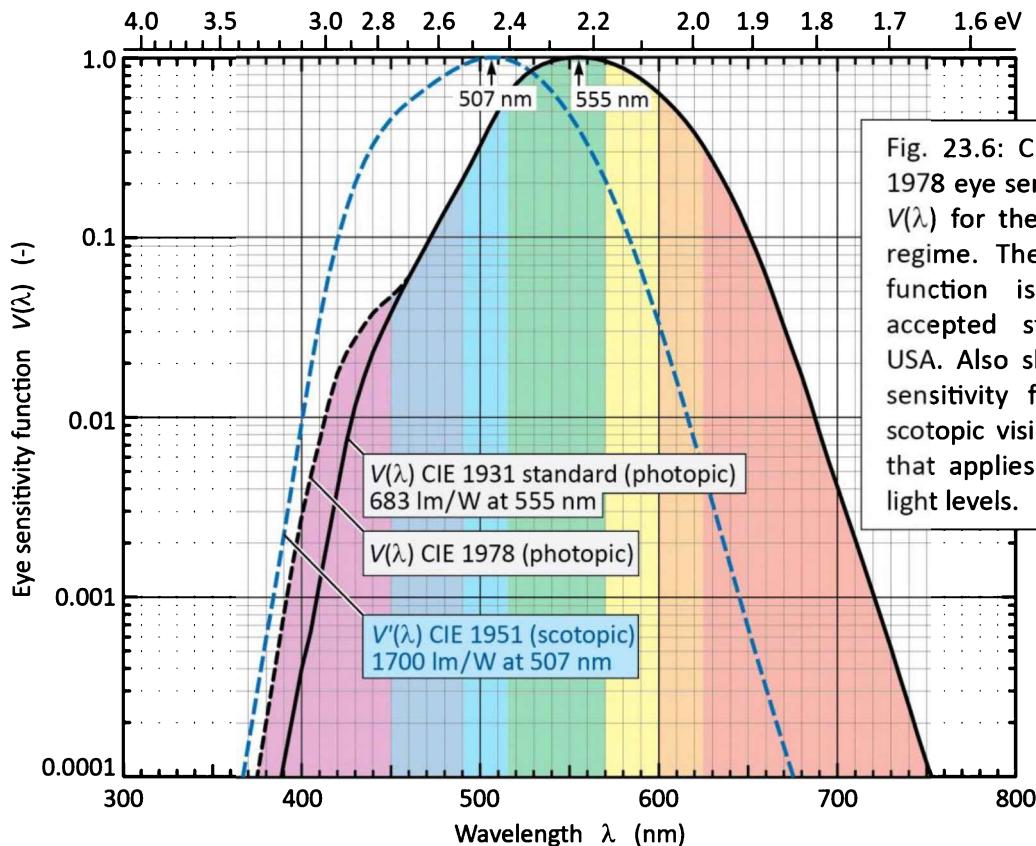
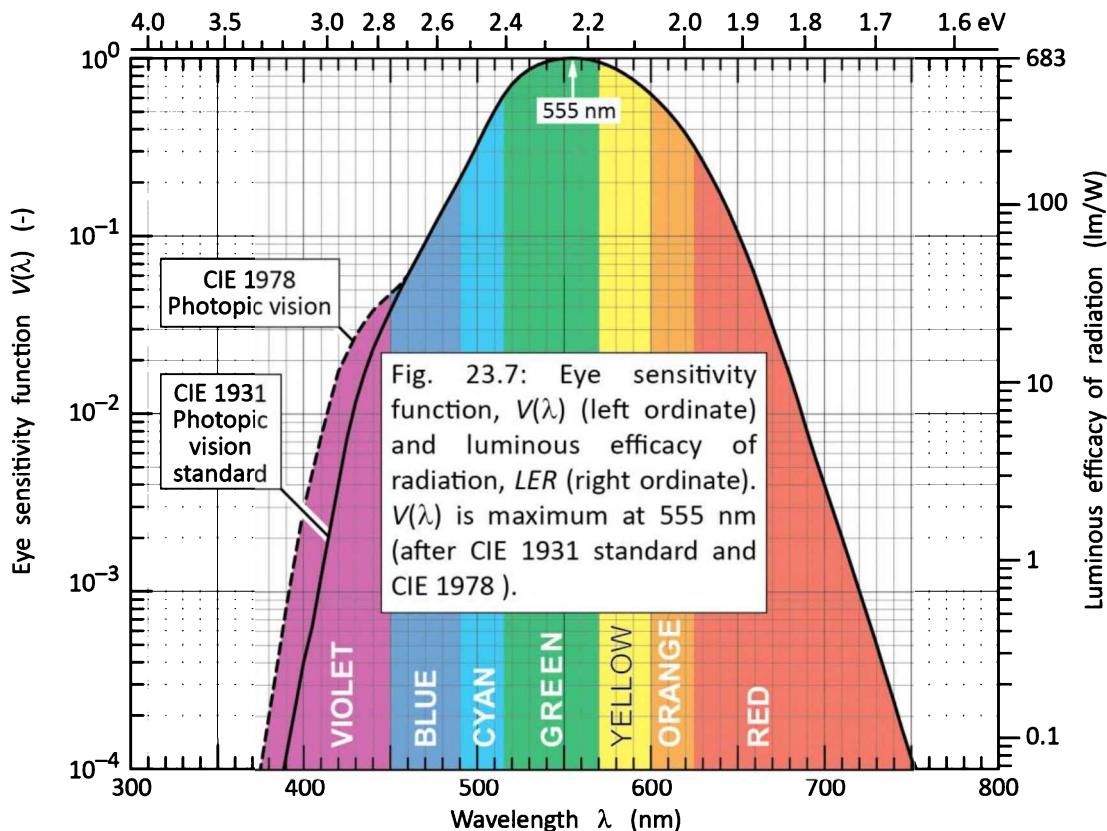


Fig. 23.6: CIE 1931 and CIE 1978 eye sensitivity functions $V(\lambda)$ for the photopic vision regime. The CIE 1931 $V(\lambda)$ function is the currently accepted standard in the USA. Also shown is the eye sensitivity function for the scotopic vision regime, $V'(\lambda)$, that applies to low ambient light levels.

Also shown in **Figure 23.6** is the **scotopic** eye sensitivity function $V'(\lambda)$. The peak sensitivity in the scotopic vision regime, $V'(\lambda) = 1.0$, occurs at $\lambda = 507 \text{ nm}$. This wavelength is markedly shorter than the peak-sensitivity wavelength in the photopic vision regime. Numerical values of the CIE 1951 $V'(\lambda)$ function are tabulated in Appendix 23.1.

The $V(\lambda)$ function (as well as the $V'(\lambda)$ function) is a normalized function, that is its maximum value is 1.0. The eye converts an optical stimulus into a biological sensation that is transmitted to the brain (specifically the visual cortex). The light power, i.e. the magnitude of the optical stimulus, is measured in watt (W), whereas the light power as perceived by the human eye is measured in lumen (lm). Accordingly, a conversion factor, the ***luminous efficacy function*** of the human eye, measured in units of lm/W (lumen per watt) may be introduced. For ***photopic vision***, at the wavelength of maximum sensitivity, 555 nm, the eye's luminous efficacy is 683 lm/W (CIE, 1931). For ***scotopic vision***, at the wavelength of maximum sensitivity, 507 nm, the eye's luminous efficacy is 1700 lm/W (DeCusatis, 1998). Scotopic vision is mediated by rod cells and these cells are more sensitive than cone cells, so that their efficacy is higher.

Note that even though the CIE 1978 $V(\lambda)$ function is preferable, it is not the standard, mostly for practical reasons such as possible ambiguities created by changing standards. Wyszecki and Stiles (2000) note that even though the CIE 1978 $V(\lambda)$ function is not a standard, it has been used in several visual studies. The CIE 1978 $V(\lambda)$ function, which can be considered the most accurate description of the eye sensitivity in the photopic vision regime, is shown in **Figure 23.7** along with the CIE 1931 standard.



The eye sensitivity function has been determined by the **minimum flicker method**, which is the classic method for luminance comparison and for the determination of $V(\lambda)$. The stimulus is a light-emitting small circular area, alternatingly illuminated (with a frequency of 15 Hz) with the standard color and the comparison color. Since the hue-fusion frequency is lower than 15 Hz, the hues fuse. However, the brightness-fusion frequency is higher than 15 Hz and thus if the two colors differ in brightness, then there will be visible flicker. The human subject's task is to adjust the target color until the flicker is minimal.

Any desired chromaticity can be obtained with an infinite variety of spectral power distributions $P(\lambda)$. One of these distributions has the greatest possible luminous efficacy. This limit can be obtained in only one way, namely by the mixture of suitable intensities emitted by two monochromatic sources (MacAdam, 1950). The maximum attainable luminous efficacy obtained with a single monochromatic pair of emitters is shown in **Figure 23.8**. The maximum luminous efficacy of *white* light depends on the color temperature; it is about 420 lm/W for a color temperature of 6500 K and can exceed 500 lm/W for lower color temperatures. The exact value depends on the exact location of the source in the chromaticity diagram.

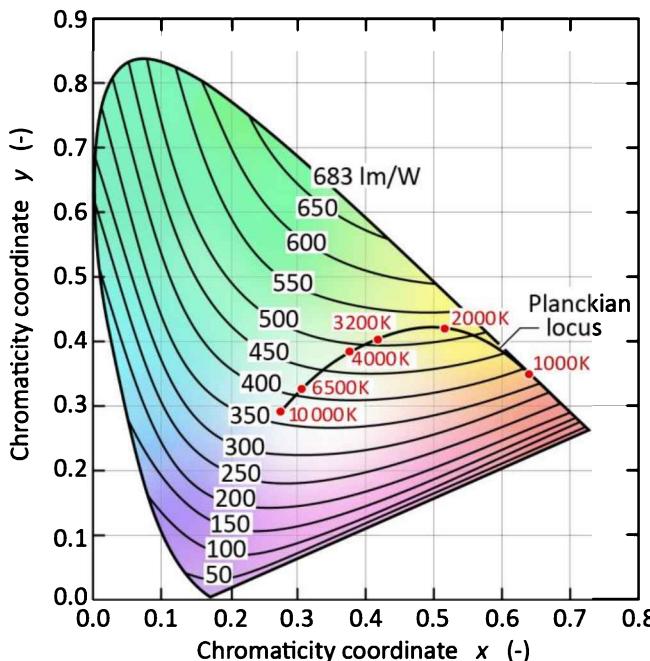


Fig. 23.8: Relation of maximum possible luminous efficacy (lumens per optical watt) and chromaticity in the CIE 1931 (x , y) chromaticity diagram (adapted from MacAdam, 1950).

23.4 Luminous efficacy

The **luminous flux** of a light source, Φ_{lum} , is obtained from the radiometric light power using the equation

$$\Phi_{lum} = 683 \frac{\text{lm}}{\text{W}} \int_{\lambda} V(\lambda) P(\lambda) d\lambda \quad (23.1)$$

where $P(\lambda)$ is the power spectral density, i.e. the light power emitted per unit wavelength, and the prefactor 683 lm/W is a normalization factor. The optical power emitted by a light source is then given by

$$P = \int_{\lambda} P(\lambda) d\lambda . \quad (23.2)$$

For example, single-chip visible LEDs can have a luminous flux of about 10~100 lm at an injection current of 35~350 mA.

The ***luminous efficacy of radiation (LER)***, measured in units of lumens per watt of optical power, is the conversion efficiency from optical power to luminous flux. The luminous efficacy is defined as

$$\begin{aligned} \text{Luminous efficacy of radiation} &= LER = \frac{\Phi_{lum}}{P} \\ &= \frac{683 \frac{\text{lm}}{\text{W}} \int_{\lambda} V(\lambda) P(\lambda) d\lambda}{\int_{\lambda} P(\lambda) d\lambda} . \end{aligned} \quad (23.3)$$

For strictly monochromatic light sources ($\Delta\lambda \rightarrow 0$), the luminous efficacy is equal to the eye sensitivity function $V(\lambda)$ multiplied by 683 lm/W. However, for multicolor light sources and especially for white light sources, the luminous efficacy needs to be calculated by integration over all wavelengths. The luminous efficacy is shown on the right-hand ordinate of **Figure 23.4**.

The ***luminous efficacy of a source (LES)***, also measured in units of lm/W, is the luminous flux of a light source divided by the electrical input power.

$$\text{Luminous efficacy of a source} = LES = \frac{\Phi_{lum}}{IV} \quad (23.4)$$

where the product (IV) is the electrical input power of the device.

Inspection of Eqs. (23.3) and (23.4) reveals that the luminous efficacy is the product of the luminous efficacy of radiation and the electrical-to-optical power conversion efficiency. The luminous efficacy of common light sources is given in **Table 23.5**.

Table 23.5: Luminous source efficacies of different light sources. (a) Incandescent sources. (b) Fluorescent sources. (c) High-intensity discharge (HID) sources.

Light source	Luminous source efficacy
Edison's first light bulb (with C filament) (a)	1.4 lm/W
Tungsten filament light bulbs (a)	10~20 lm/W
Quartz halogen light bulbs (a)	20~25 lm/W
Linear fluorescent lamps (LFL) (b)	50~100 lm/W
Compact fluorescent lamps (CFL) (b)	50~80 lm/W
Mercury vapor light bulbs (c)	50~60 lm/W
Metal halide light bulbs (c)	80~125 lm/W
High-pressure sodium vapor light bulbs (c)	100~140 lm/W

The luminous source efficacy is a highly relevant figure of merit for visible-spectrum LEDs. It is a measure of the light power perceived by a human being normalized to the electrical power expended to operate the LED. For light sources with a perfect electrical-power-to-optical-power conversion, the luminous source efficacy is equal to the luminous efficacy of radiation.

Exercise: Luminous efficacy of radiation and luminous efficacy of LEDs. Consider a red and an amber LED emitting at 625 and 590 nm, respectively. For simplicity, assume that the emission spectra are monochromatic ($\Delta\lambda \rightarrow 0$). What is the luminous efficacy of radiation of the two light sources? Calculate the luminous efficacy of the LEDs, assuming that the red and amber LEDs have an external quantum efficiency of 50%. Assume that the LED voltage is given by $V = E_g/e = h\nu/e$.

Next, assume that the LED spectra are thermally broadened and have a gaussian lineshape with a linewidth of $1.8kT$. Again, calculate the luminous efficacy of radiation and luminous efficacy of the two light sources. How accurate are the results obtained with the approximation of monochromaticity?

Solution:

The LED emitting at 625 nm has a luminous efficacy of radiation of 219.2 lm/W.

The LED emitting at 590 nm has a luminous efficacy of radiation of 517.0 lm/W.

The LED emitting at 625 nm has a power efficiency of 50% and a luminous efficacy of 109.6 lm/W.

The LED emitting at 590 nm has a power efficiency of 50% and a luminous efficacy of 258.5 lm/W.

We use a gaussian lineshape with a full-width at half-maximum (FWHM) of $1.8 kT = 1.8 \times 25.9 \text{ meV} \approx 45 \text{ meV}$. The FWHM and standard deviation of a gaussian function, σ , are related by: $FWHM = 2(2 \ln 2)^{1/2} \sigma = 2.355 \sigma$. Thus, the gaussian function's standard deviation is $\sigma = 19.11 \text{ meV}$ or 6.02 nm (for the 625 nm LED) and 5.37 nm (for the 590 nm LED). A numerical calculation reveals the following:

625 nm LED: Luminous efficacy of radiation = 221.4 lm/W; Luminous source efficacy = 110.7 lm/W.

590 nm LED: Luminous efficacy of radiation = 515.7 lm/W; Luminous source efficacy = 257.8 lm/W.

Comparison of the results obtained with the approximation of monochromaticity with the results obtained by the accurate numerical calculation reveals a difference of only a few percent.

Some LED structures attain excellent power-conversion efficiency by using small light-emitting areas (current injection in a small area of chip) and advanced light-output-coupling structures (see, for example, Schmid *et al.*, 2002). However, such devices have low luminance because only a small fraction of the chip area is injected with current. **Table 23.6** summarizes frequently used figures of merit for light-emitting diodes.

Table 23.6: Summary of photometric, radiometric, and quantum performance measures for LEDs.

Figure of merit	Explanation	Unit
Luminous radiation efficacy	Luminous flux per optical unit power	lm/W
Luminous source efficacy	Luminous flux per input electrical unit power	lm/W
Luminance	Luminous flux per sr per chip unit area	cd/m ²
Power-conversion efficiency	Optical output power per input electrical unit power	%
Internal quantum efficiency	Photons emitted from active region per electron injected	%
External quantum efficiency	Photons emitted from LED per electron injected	%
Light-extraction efficiency	Escape probability of photons emitted from active region	%

23.5 Surface brightness

Although the term **surface brightness** (or simply **brightness**) is frequently used, it lacks a standardized scientific definition. The frequent usage is due to the fact that the general public can more easily relate to the term **brightness** than to the photometric **luminance**.

To quantify the surface brightness of a source, it is useful to differentiate between point and surface-area sources. For a **point source**, the surface brightness (in the photopic vision regime) approaches infinity since the source area approaches zero. For **surface-area sources**, the surface brightness (in the photopic vision regime) can be considered to be equivalent to luminance (measured in cd/m²). That is, the term **surface brightness** may be used as synonym for **luminance**. One may note that for lambertian surface emitters, the surface brightness (luminance) remains constant and is independent of the viewing angle.

23.6 Circadian rhythm and circadian sensitivity

The human wake-sleep rhythm has a period of approximately 24 hours and the rhythm therefore is referred to as the ***circadian rhythm*** or ***circadian cycle***, with the name being derived from the Latin words *circa* and *dies* (and its declination *diem*), meaning *approximately* and *day*, respectively. Light has been known for a long time to be the synchronizing clock (*zeitgeber*) of the human circadian rhythm. For reviews on the development of the understanding of the circadian rhythm including the identification of light as the dominant trigger for the endogenous *zeitgeber*, see Pittendrigh (1993) and Sehgal (2004).

The wake-sleep rhythm of humans is synchronized by the intensity and spectral composition of light. Sunlight is the natural *zeitgeber*. During mid-day hours sunlight has high intensity, a high color temperature, and a high content of blue light. During evening hours, intensity, color temperature, and blue content of sunlight strongly decrease. Humans have adapted to this variation and the circadian rhythm is most likely synchronized by the following three factors: intensity, color temperature, and blue content.

Exposure to inappropriately high intensities of light in the late afternoon or evening can upset the regular wake-sleep rhythm and lead to sleeplessness and even serious illnesses such as cancer (Brainard *et al.*, 2001; Blask *et al.*, 2003). It is therefore highly advisable to limit exposure to high intensity light in the late afternoon and evening hours, to not be counterproductive to the natural circadian rhythm (Schubert, 1997).

It was believed for a long time that rod cells and the three types of cone cells are the only optically sensitive cells in the human eye. However, Brainard *et al.* (2001) postulated that an unknown photoreceptor in the human eye would control the circadian rhythm. Evidence presented by Berson *et al.* (2002) and Hattar *et al.*, (2002) indicates that retinal ganglion cells have an optical sensitivity as well. For a schematic illustration of ganglion cells, see ***Figure 23.1***. The spectral sensitivity of mammalian ganglion cells was measured and the responsivity curve is shown in ***Figure 23.9***. Inspection of the figure reveals a ganglion-cell peak-sensitivity at 484 nm, i.e. in the blue spectral range.

Berson *et al.* (2002) presented evidence that the photosensitive ganglion cells are instrumental in the control of the circadian rhythm. Due to their sensitivity in the blue spectral range, it can be hypothesized that the blue sky occurring near mid-day is a strong factor in synchronizing the endogenous circadian rhythm. The photosensitive ganglion cells have

therefore been referred to as *blue-sky receptors*.

Inspection of the spectral sensitivity of the ganglion cells shown in *Figure 23.9* reveals the huge difference of red light and blue light for circadian efficacy: The efficacy of blue light in synchronizing the circadian rhythm can be three orders of magnitude greater than the efficacy of red light. This particular role of blue light should be taken into account in lighting design and the use of artificial lighting by consumers.

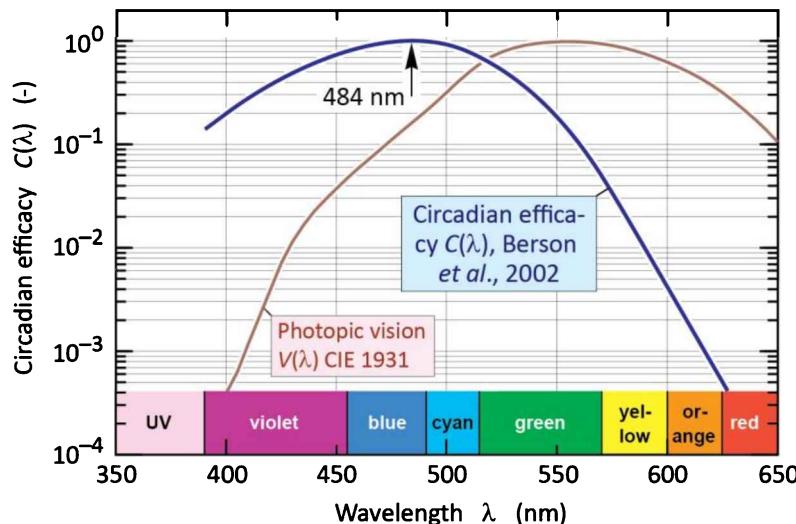


Fig. 23.9: Circadian efficacy curve derived from retinal ganglion cell photoresponse measurements. The ganglion cells on which the measurements were performed originated from mammals. The figure reveals the significant difference between circadian and visual sensitivity (adapted from Berson *et al.*, 2002).

23.7 References

- Berson D. M., Dunn F. A., and Takao M. "Phototransduction by retinal ganglion cells that set the circadian clock" *Science* **295**, 1070 (2002)
- Brainard G. C., Hanifin J. P., Greeson J. M., Byrne B., Glickman G., Gerner E., and Rollag M. D. "Action spectrum for melatonin regulation in humans: Evidence for a novel circadian photoreceptor" *Journal of Neuroscience* **21**, 6405 (2001)
- Blask D. E., Dauchy R. T., Sauer L. A., Krause J. A., Brainard G. C. "Growth and fatty acid metabolism of human breast cancer (MCF-7) xenografts in nude rats: Impact of constant light-induced nocturnal melatonin suppression" *Breast Cancer Research and Treatment* **79**, 313 (2003)
- CIE *Commission Internationale de l'Eclairage Proceedings* (Cambridge University Press, Cambridge, 1931)
- CIE *Proceedings* **1**, Sec. 4; **3**, p. 37; Bureau Central de la CIE, Paris (1951)
- CIE data of 1931 and 1978 available at <http://cvision.ucsd.edu> and <http://www.cvrl.org> (1978). The CIE 1931 $V(\lambda)$ data were modified by D. B. Judd and J. J. Vos in 1978. The Judd–Vos-modified eye-sensitivity function is frequently referred to as $V_M(\lambda)$; see J. J. Vos "Colorimetric and photometric properties of a 2-deg fundamental observer" *Color Research and Applications* **3**, 125 (1978)
- CIE publication 75-1988 *Spectral Luminous Efficiency Functions Based Upon Brightness Matching for Monochromatic Point Sources with 2° and 10° Fields* ISBN 3900734119 (1988)
- CIE publication 86-1990 *CIE 1988 2° Spectral Luminous Efficiency Function for Photopic Vision* ISBN 3900734232 (1990)
- DeCusatis C. (Editor) *Handbook of Applied Photometry* (Springer-Verlag, New York, 1998)
- Dowling J. E. *The retina: An Approachable Part of the Brain* (Harvard University Press, Cambridge, Massachusetts, 1987)

- Encyclopedia Britannica, Inc. Illustration of human eye adopted from 1994 edition of the encyclopedia (1994)
- Hattar S., Liao H.-W., Takao M., Berson D. M., and Yau K.-W. "Melanopsin-containing retinal ganglion cells: Architecture, projections, and intrinsic photosensitivity" *Science* **295**, 1065 (2002)
- Kelly K. L. "Color designations for lights" *Journal of the Optical Society of America* **33**, 627 (1943)
- MacAdam D. L. "Maximum attainable luminous efficiency of various chromaticities" *Journal of the Optical Society of America* **40**, 120 (1950)
- Osram Sylvania Corporation *Lumens and mesopic vision* Application Note FAQ0016-0297 (2000)
- Pittendrigh C. S. "Temporal organization: Reflections of a Darwinian clock-watcher" *Annual Review of Physiology* **55**, 17 (1993)
- Schmid W., Scherer M., Karnutsch C., Plobl A., Wegleiter W., Schad S., Neubert B., and Streubel K. "High-efficiency red and infrared light-emitting diodes using radial outcoupling taper" *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 256 (2002)
- Schubert E. F. The author of this book noticed in 1997 that working after 8 PM under bright illumination conditions in the office allowed him to fall asleep only very late, typically after midnight. The origin of sleeplessness was traced back to high-intensity office lighting conditions. Once the high intensity of the office lighting was reduced, the sleeplessness vanished (1997)
- Sehgal A., editor *Molecular Biology of Circadian Rhythms* (John Wiley and Sons, New York, 2004)
- Vos J. J. "Colorimetric and photometric properties of a 2-deg fundamental observer" *Color Research and Applications* **3**, 125 (1978)
- Wyszecki G. and Stiles W. S. *Color Science – Concepts and Methods, Quantitative Data and Formulae* 2nd edition (John Wiley and Sons, New York, 1982)
- Wyszecki G. and Stiles W. S. *Color Science – Concepts and Methods, Quantitative Data and Formulae* 2nd edition (John Wiley and Sons, New York, 2000)

Appendix 23.1

Values of the 2° degree CIE 1931 photopic eye sensitivity function and the CIE 1978 Judd–Vos-modified photopic eye sensitivity function for point sources (after CIE, 1931 and CIE, 1978). Values of the CIE 1951 eye sensitivity function of the scotopic vision regime, $V'(\lambda)$ (after CIE, 1951).

λ (nm)	CIE 1931 $V(\lambda)$	CIE 1978 $V(\lambda)$	CIE 1951 $V'(\lambda)$
360	3.9170 E-6	0.0000E-4	–
365	6.9650 E-6	0.0000E-4	–
370	1.2390 E-5	0.0000E-4	–
375	2.2020 E-5	0.0000E-4	–
380	3.9000 E-5	2.0000E-4	5.890E-4
385	6.4000 E-5	3.9556E-4	1.108E-3
390	1.2000 E-4	8.0000E-4	2.209E-3
395	2.1700 E-4	1.5457E-3	4.530E-3
400	3.9600 E-4	2.8000E-3	9.290E-3
405	6.4000 E-4	4.6562E-3	1.852E-2
410	1.2100 E-3	7.4000E-3	3.484E-2
415	2.1800 E-3	1.1779E-2	6.040E-2
420	4.0000 E-3	1.7500E-2	9.660E-2
425	7.3000 E-3	2.2678E-2	1.436E-1
430	1.1600 E-2	2.7300E-2	1.998E-1
435	1.6840 E-2	3.2584E-2	2.625E-1
440	2.3000 E-2	3.7900E-2	3.281E-1
445	2.9800 E-2	4.2391E-2	3.931E-1
450	3.8000 E-2	4.6800E-2	4.550E-1
455	4.8000 E-2	5.2122E-2	5.130E-1
460	6.0000 E-2	6.0000E-2	5.670E-1
465	7.3900 E-2	7.2942E-2	6.200E-1
470	9.0980 E-2	9.0980E-2	6.760E-1
475	0.11260	0.11284	7.340E-1
480	0.13902	0.13902	7.930E-1
485	0.16930	0.16987	8.510E-1
490	0.20802	0.20802	9.040E-1
495	0.25860	0.25808	9.490E-1
500	0.32300	0.32300	9.820E-1
505	0.40730	0.40540	9.980E-1
510	0.50300	0.50300	9.970E-1
515	0.60820	0.60811	9.750E-1
520	0.71000	0.71000	9.350E-1
525	0.79320	0.79510	8.800E-1
530	0.86200	0.86200	8.110E-1
535	0.91485	0.91505	7.330E-1
540	0.95400	0.95400	6.500E-1
545	0.98030	0.98004	5.640E-1
550	0.99495	0.99495	4.810E-1
555	1.00000	1.00000	4.020E-1
560	0.99500	0.99500	3.288E-1
565	0.97860	0.97875	2.639E-1
570	0.95200	0.95200	2.076E-1
575	0.91540	0.91558	1.602E-1

580	0.87000	0.87000	1.212E–1
585	0.81630	0.81623	8.990E–2
590	0.75700	0.75700	6.550E–2
595	0.69490	0.69483	4.690E–2
600	0.63100	0.63100	3.315E–2
605	0.56680	0.56654	2.312E–2
610	0.50300	0.50300	1.593E–2
615	0.44120	0.44172	1.088E–2
620	0.38100	0.38100	7.370E–3
625	0.32100	0.32052	4.970E–3
630	0.26500	0.26500	3.335E–3
635	0.21700	0.21702	2.235E–3
640	0.17500	0.17500	1.497E–3
645	0.13820	0.13812	1.005E–3
650	0.10700	0.10700	6.770E–4
655	8.1600 E–2	8.1652E–2	4.590E–4
660	6.1000 E–2	6.1000E–2	3.129E–4
665	4.4580 E–2	4.4327E–2	2.146E–4
670	3.2000 E–2	3.2000E–2	1.480E–4
675	2.3200 E–2	2.3454E–2	1.026E–4
680	1.7000 E–2	1.7000E–2	7.150E–5
685	1.1920 E–2	1.1872E–2	5.010E–5
690	8.2100 E–3	8.2100E–3	3.533E–5
695	5.7230 E–3	5.7723E–3	2.501E–5
700	4.1020 E–3	4.1020E–3	1.780E–5
705	2.9290 E–3	2.9291E–3	1.273E–5
710	2.0910 E–3	2.0910E–3	9.140E–6
715	1.4840 E–3	1.4822E–3	6.600E–6
720	1.0470 E–3	1.0470E–3	4.780E–6
725	7.4000 E–4	7.4015E–4	3.482E–6
730	5.2000 E–4	5.2000E–4	2.546E–6
735	3.6110 E–4	3.6093E–4	1.870E–6
740	2.4920 E–4	2.4920E–4	1.379E–6
745	1.7190 E–4	1.7231E–4	1.022E–6
750	1.2000 E–4	1.2000E–4	7.600E–7
755	8.4800 E–5	8.4620E–5	5.670E–7
760	6.0000 E–5	6.0000E–5	4.250E–7
765	4.2400 E–5	4.2446E–5	3.196E–7
770	3.0000 E–5	3.0000E–5	2.413E–7
775	2.1200 E–5	2.1210E–5	1.829E–7
780	1.4990 E–5	1.4989E–5	1.390E–7
785	1.0600 E–5	1.0584E–5	–
790	7.4657 E–6	7.4656E–6	–
795	5.2578 E–6	5.2592E–6	–
800	3.7029 E–6	3.7028E–6	–
805	2.6078 E–6	2.6076E–6	–
810	1.8366 E–6	1.8365E–6	–
815	1.2934 E–6	1.2950E–6	–
820	9.1093 E–7	9.1092E–7	–
825	6.4153 E–7	6.3564E–7	–

24

24 – Colorimetry

The assessment and quantification of color is referred to as **colorimetry** or the “science of color”. Colorimetry is closely associated with human color vision. Both colorimetry and human vision have attracted a great deal of interest that spans many centuries. For a thorough and entertaining review of the history of colorimetry including early attempts to understand color, we recommended the collection of historical reprints complied by MacAdam (1993).

The human sense of vision is very different from the human sense of hearing. If we hear two frequencies simultaneously, e.g. two frequencies generated by a musical instrument, we are able to recognize two distinct frequencies. This is not the case for optical signals and the sense of vision. Mixing two monochromatic optical signals will appear to us as one color and we are unable to recognize the original dichromatic composition of that color.

Visible wavelengths range from 380 to 770 nm. At the two boundaries of this range, the eye sensitivity is, by definition, very low. For this reason, the two boundaries of the visible range are not highly accurate quantities but somewhat vague. The technical literature puts the short-wavelength boundary at 380 to 400 nm and the long-wavelength boundary at 700 to 780 nm. The colors of monochromatic light sources ($\Delta\lambda \rightarrow 0$), emitting within the visible wavelength range at wavelength λ , are shown in **Figure 24.1**. The condition $\Delta\lambda \rightarrow 0$ is closely met by lasers but to a lesser extent by LEDs, since LEDs have a spectral linewidth of $\Delta E = 1.8 kT$.

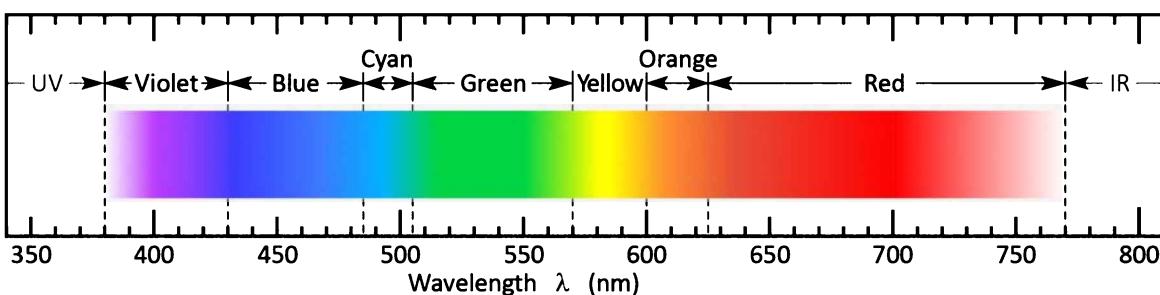


Fig. 24.1: Colors of a strictly monochromatic light source versus wavelength. The light source has a spectral linewidth approaching zero, i.e. $\Delta\lambda \rightarrow 0$.

24.1 Colors of near-monochromatic emitters

For wavelengths ranging from 380 to 770 nm, the eye sensitivity function $V(\lambda)$ is greater than about 10^{-4} . Therefore, the wavelength range $380 \text{ nm} \leq \lambda \leq 770 \text{ nm}$ is considered the *visible wavelength range*. The relationship between color and wavelength for near-monochromatic emitters is given in **Table 24.1**. This relationship is valid for monochromatic or near-monochromatic light sources such as LEDs. The relationship between the color names and dominant wavelength was first published by Kelly (1943) and the table is overall consistent with Kelly (1943). Note that color is, to some extent, a subjective quantity and that the transitions between different colors are continuous so that the boundaries are not as abrupt as the table may suggest.

Table 24.1: Color names and associated LED peak wavelength ranges. The “Royal blue” range was termed “blue violet” by Kelly (1943). “Cyan” may also be termed “blue green”. “Amber” may also be termed “yellowish orange”.

Color	Wavelength	Color	Wavelength
Ultraviolet (UV)	< 380 nm	Yellow green	550~570 nm
Violet	380~430 nm	Yellow	570~600 nm
Royal blue	430~450 nm	Amber yellow	590~600 nm
Blue	430~485 nm	Orange	600~625 nm
Cyan	485~505 nm	Red	625~770 nm
Green	505~550 nm	Infrared (IR)	> 770 nm

24.2 Color-matching functions and chromaticity diagram

Light causes different levels of excitation of the red, green, and blue cones. However, the sensation of color and luminous flux caused by a particular light source varies slightly among different individuals. Furthermore, the sensation of color is, to some extent, a subjective quantity. For these reasons, *The International Commission for Illumination (Commission Internationale de l'Eclairage, CIE)* has standardized the measurement of color by means of **color-matching functions** and the **chromaticity diagram** (CIE, 1931).

How are color-matching functions obtained? Consider two lighted fields lying side by side: One field being monochromatic and the other field being a mixture of three primary lights with colors red, green, and blue, as shown in **Figure 24.2**. A human subject will be able to make the two lighted fields appear identical (i.e. “match” them) by adjusting the relative intensities of the

red, green, and blue light. The three color-matching functions are obtained from a series of such matches, in which the subject sets the intensities of the three primary lights required to match a series of monochromatic lights across the visible spectrum.

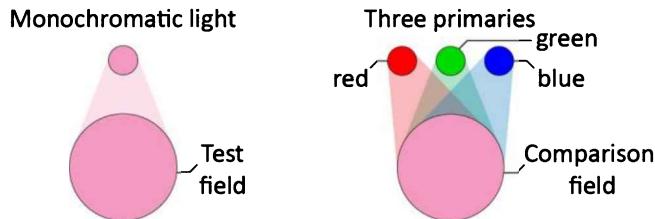


Fig. 24.2: Principle of color matching: A monochromatic test light (imaged on the “test field”) is color-matched by mixing three adjustable primary lights, red, green, and blue (imaged on the “comparison field”).

Subsequently, the measured set of color-matching functions is mathematically transformed into a new set of color-matching functions for which the green color-matching function, $\bar{y}(\lambda)$, is chosen to be identical to the eye sensitivity function, $V(\lambda)$, i.e.

$$\bar{y}(\lambda) = V(\lambda). \quad (24.1)$$

The CIE 1931 and CIE 1978 color-matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are shown in **Figure 24.3**. The numerical values of these color-matching functions are tabulated in Appendices 24.1 and 24.2, respectively. The three color-matching functions reflect the fact that human color vision possesses **trichromacy**, that is, the color of any light source can be described by just three variables. Note that $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are dimensionless quantities. Also note that neither the color-matching functions nor the chromaticity diagram is unique (see, for example, Judd, 1951 or Vos, 1978). In fact there have been several different versions of the color-matching functions and of the chromaticity diagram.

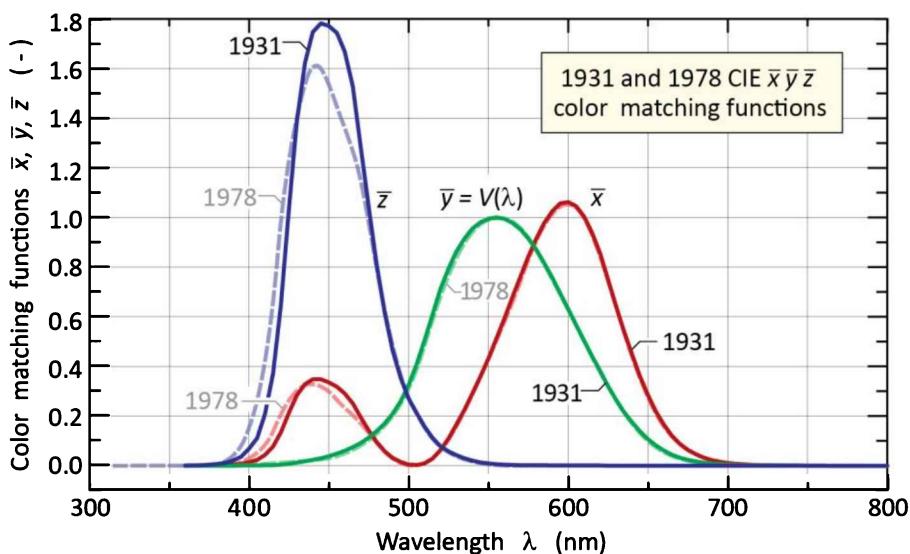


Fig. 24.3: CIE (1931) and CIE (1978) $\bar{x}\bar{y}\bar{z}$ color-matching functions. The \bar{y} color-matching function is identical to the eye sensitivity function $V(\lambda)$. The CIE 1931 color-matching functions are the currently valid official standard in the United States.

For a given power-spectral density $P(\lambda)$, the degree of stimulation required to match the color of $P(\lambda)$ is given by

$$X = \int_{\lambda} \bar{x}(\lambda) P(\lambda) d\lambda \quad (24.2)$$

$$Y = \int_{\lambda} \bar{y}(\lambda) P(\lambda) d\lambda \quad (24.3)$$

$$Z = \int_{\lambda} \bar{z}(\lambda) P(\lambda) d\lambda \quad (24.4)$$

where X , Y , and Z are the **tristimulus values** that give the stimulation (i.e. power) of each of the three primary red, green, and blue lights needed to match the color of $P(\lambda)$. Comparatively large values of X , Y , and Z indicate that the spectrum $P(\lambda)$ has a predominance of the red, green, and blue color, respectively.

Because of the distinct similarity of the three retinal-cone-sensitivity functions on one hand, and the color-matching functions on the other hand (both groups of functions have three peaks), each tristimulus value represents the *approximate* (but not *exact*) degree of stimulation that each type of retinal cone experiences when illuminated by a light source with spectrum $P(\lambda)$.

Inspection of Eqns. (24.2) ~ (24.4) suggests that the unit of the tristimulus values is “watt”. However, the tristimulus values are usually given as dimensionless quantities. The prefactor “watt⁻¹” in front of the integral can be included so that the tristimulus values become dimensionless. If only *ratios* of tristimulus values are employed, as below, the prefactors and units cancel and thus become irrelevant.

The **chromaticity coordinates** x and y are calculated from the tristimulus values using

$$x = \frac{X}{X + Y + Z} \quad (24.5)$$

$$y = \frac{Y}{X + Y + Z} \quad (24.6)$$

Thus, the value of a chromaticity coordinate is the stimulation of each primary light (or of each type of retinal cone) divided by the entire stimulation ($X + Y + Z$). The value of the z chromaticity coordinate is calculated analogously, that is

$$z = \frac{Z}{X+Y+Z} = 1 - x - y. \quad (24.7)$$

Note that the z chromaticity value can be obtained from x and y , so that there is no new information provided by the z chromaticity coordinate. Therefore, the z coordinate is redundant and, for this reason, does not need to be used.

The (x, y) chromaticity diagram is shown in **Figure 24.4**. Reddish and greenish colors are found for large values of x and y , respectively. Bluish colors are found for large values of z , which is, according to Eqn. (24.7), for low values of x and y , or near the origin of the chromaticity diagram.

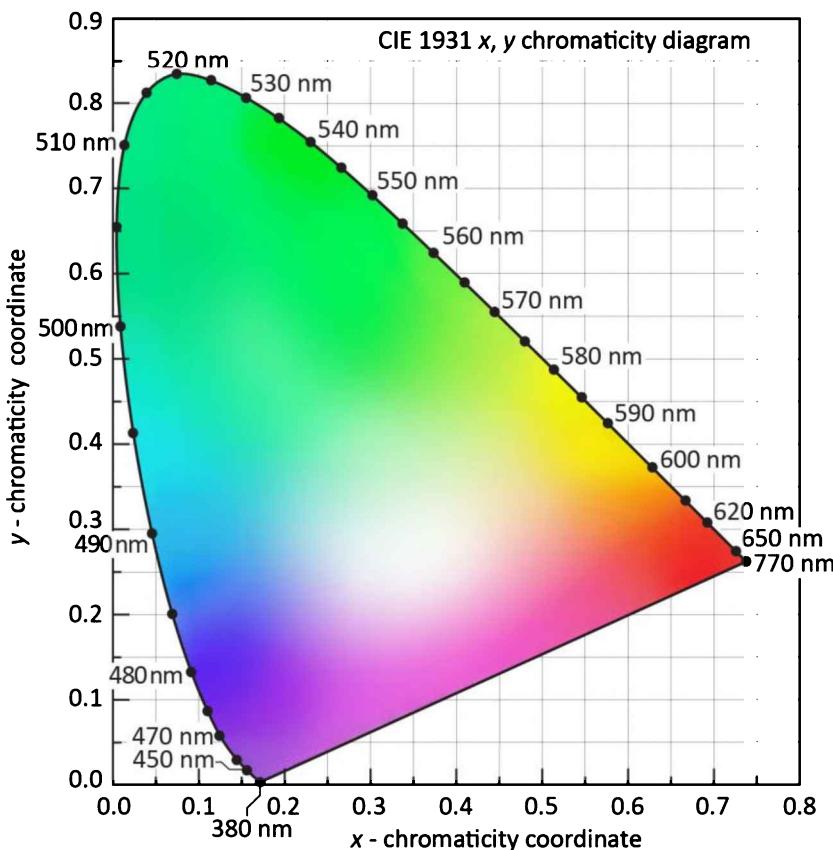


Fig. 24.4: CIE 1931 (x, y) chromaticity diagram. Monochromatic colors are located on the perimeter and white light is located in the center of the diagram.

A detailed assignment of colors to specific regions within the CIE (x, y) chromaticity diagram is shown in **Figure 24.5**. The assignment is due to Kelly (1943) and Gage *et al.* (1977).

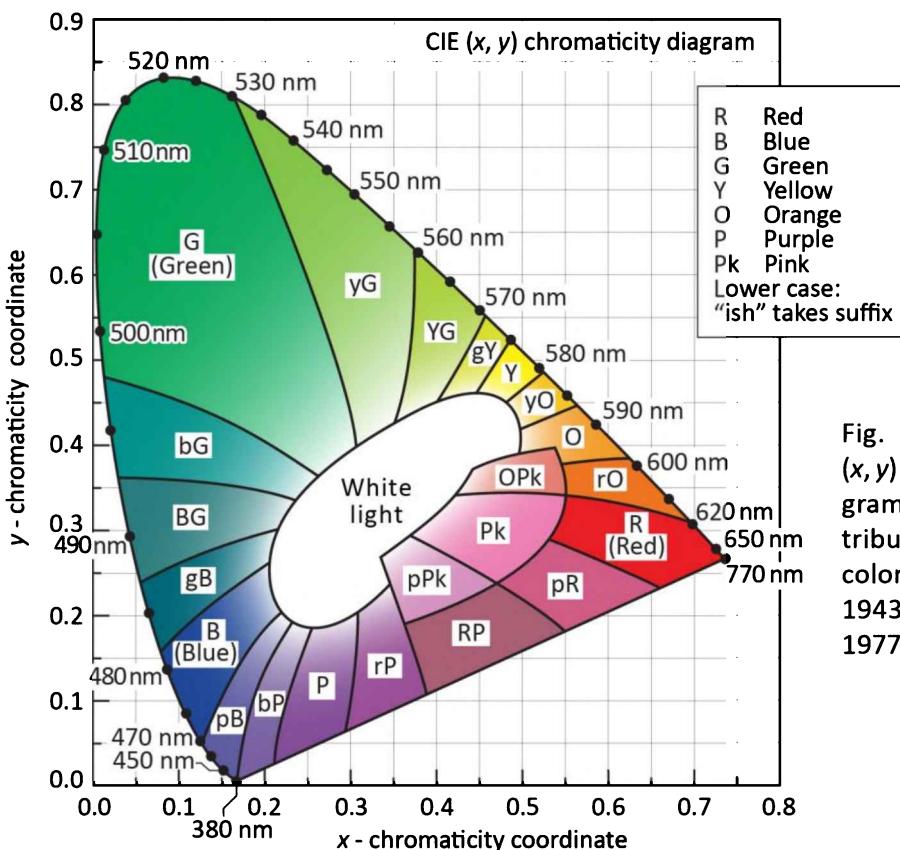


Fig. 24.5: 1931 CIE (x, y) chromaticity diagram with areas attributed to distinct colors (after Kelly, 1943; Gage *et al.*, 1977).

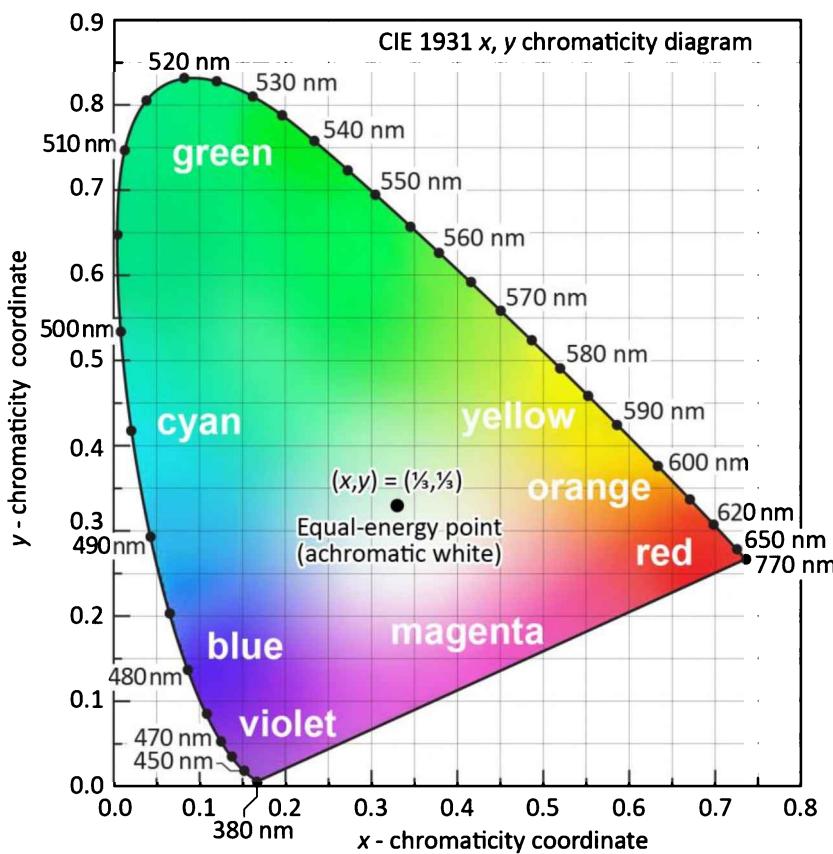


Fig. 24.6: CIE 1931 (x, y) chromaticity diagram. Monochromatic colors are located on the perimeter. Color saturation decreases towards the center of the diagram. White light is located in the center. Also shown are the regions of distinct colors. The equal-energy point is located at the center and has the coordinates $(x, y) = (1/3, 1/3)$. It can be considered as the achromatic white point.

A simple assignment of common colors in the chromaticity diagram is given in **Figure 24.6**. The figure also shows the ***equal-energy point*** located in the center of the chromaticity diagram at $(x, y, z) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. At this point, the tristimulus values are equal, i.e. $X = Y = Z$. An optical spectrum corresponding to the equal-energy point can have an about constant spectral distribution, i.e. the optical power per wavelength interval $d\lambda$ is about constant across the visible spectrum. Light of the equal-energy point is colorless and has no hue, i.e. it is not yellowish or bluish. Therefore, light of the equal-energy point is termed ***achromatic white*** (colorless white). Monochromatic or pure colors are found on the perimeter of the chromaticity diagram. White light is found in the center of the chromaticity diagram. All chromaticities can be characterized in terms of their location in the chromaticity diagram.

MacAdam (1943) analyzed the color differences of closely spaced points in the chromaticity diagram. MacAdam concluded that two chromaticity points must have a *minimum geometrical distance* to be perceptibly different in color. That is, colors within a limited small area in the chromaticity diagram appear identical to the average human observer. MacAdam showed that these areas have the shape of an ellipsis and these ellipses are called ***1-step MacAdam ellipses***. Examples of ***MacAdam ellipses*** are shown in **Figure 24.7** (MacAdam, 1943, 1993; Wright, 1943). Multi-step MacAdam ellipses are defined as follows: An N -step MacAdam ellipsis has axes that are N times longer than those of a 1-step MacAdam ellipsis. The following classification is valid for the average human observer:

- The colors on the boundary of a 1-step MacAdam ellipse are *indistinguishable* from the color in the center of the ellipse.
- Colors on the boundary of a 2- to 4-step MacAdam ellipse are *barely distinguishable* from the color in the center of the ellipse.
- Colors on the boundary of a 5- to 6-step MacAdam ellipse are *readily distinguishable* from the color in the center of the ellipse.

Inspection of **Figure 24.7** reveals that ellipses in the blue and green region are very different in size. Thus, the geometric distance between two points in the (x, y) chromaticity diagram does *not* scale linearly with the color difference. We also note that white LEDs are frequently binned at the 3- to 7-step MacAdam level so that minor differences in color are discernable for white LEDs from the same bin.

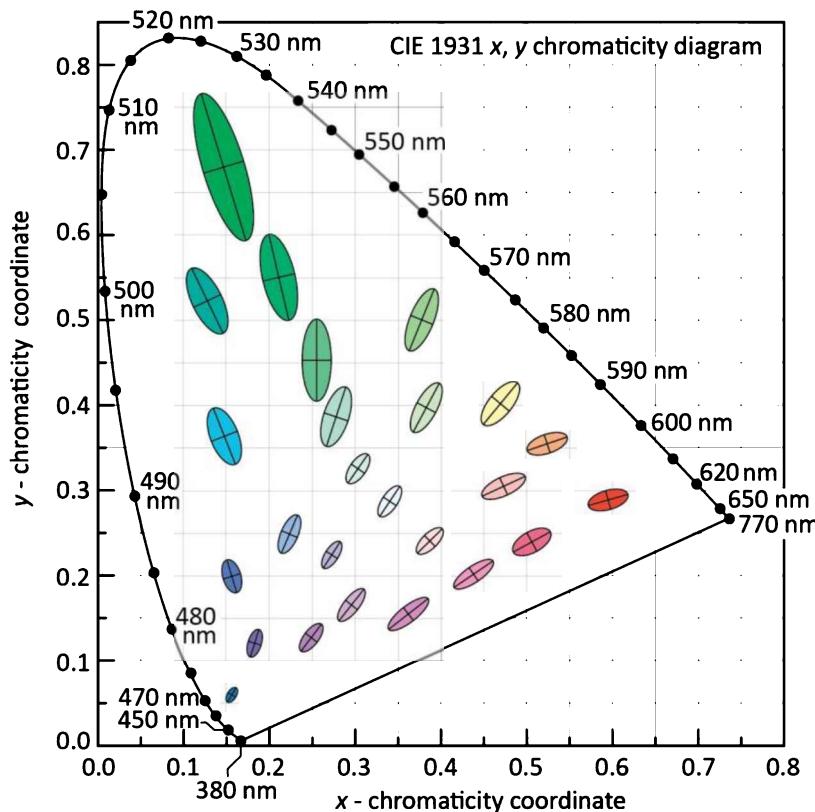


Fig. 24.7: MacAdam ellipses plotted in the CIE 1931 (x , y) chromaticity diagram. The axes of the ellipses shown are ten times the length of 1-step MacAdam ellipses. In other words, the ellipses shown are 10-step MacAdam ellipses (after MacAdam, 1943; Wright, 1943; MacAdam, 1993).

The total number of differentiable chromaticities can be obtained by dividing the area of the chromaticity diagram through the average area of the MacAdam ellipses. This calculation yields the result that humans can discern approximately 50 000 distinct chromaticities. If variations in lightness (or brightness, i.e., luminance) are taken into account, the number of discernable colors increases to a value greater than 10^6 . While the exact number of discernable colors is not known, estimates vary between 10^6 and 10^7 , i.e. between 1 million and 10 million.

In the chromaticity diagram, it is very desirable for the color difference to be proportional to the geometric difference. This has motivated the *uniform* chromaticity diagram. In 1960, the CIE introduced the (u, v) and in 1976 the (u', v') **uniform chromaticity coordinates** (Wyszecki and Stiles, 2000). These coordinates form the **uniform chromaticity diagram**. The uniform chromaticity coordinates are calculated from the tristimulus values according to

$$u = \frac{4X}{X + 15Y + 3Z} \quad v = \frac{6Y}{X + 15Y + 3Z} \quad (\text{CIE, 1960}) \quad (24.8)$$

and

$$u' = \frac{4X}{X + 15Y + 3Z} \quad v' = \frac{9Y}{X + 15Y + 3Z} \quad (\text{CIE, 1976}). \quad (24.9)$$

The CIE 1976 (u', v') uniform chromaticity diagram is shown in **Figure 24.8**.

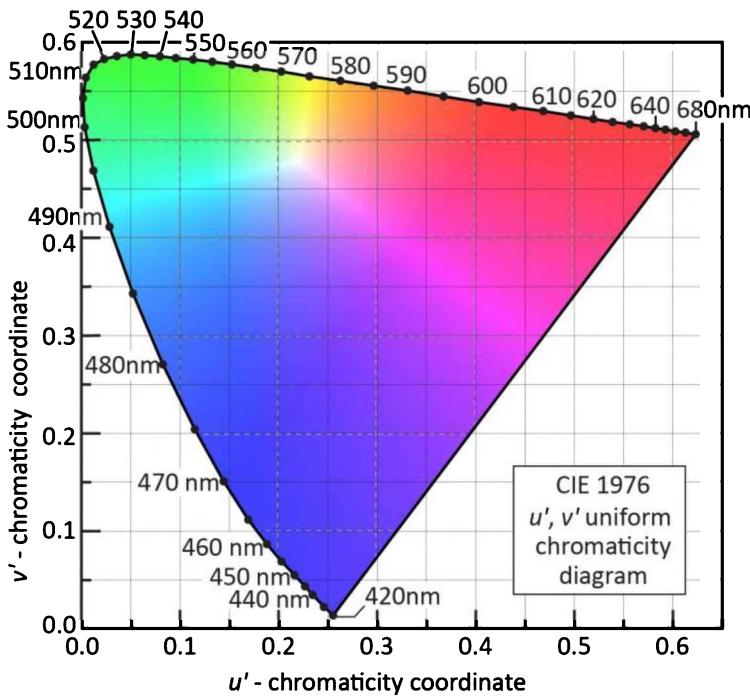


Fig. 24.8: CIE 1976 (u' , v') uniform chromaticity diagram calculated using the CIE 1931 2° standard observer.

The (u, v) and (u', v') uniform chromaticity coordinates can be calculated from the (x, y) chromaticity coordinates according to

$$u = u' = \frac{4x}{-2x + 12y + 3} \quad (24.10)$$

and

$$v = \frac{6y}{-2x + 12y + 3} \quad v' = \frac{9y}{-2x + 12y + 3}. \quad (24.11)$$

Conversely, one obtains

$$x = \frac{9u'}{6u' - 16v' + 12} \quad y = \frac{2v'}{3u' - 8v' + 6} \quad (24.12)$$

and

$$x = \frac{3u}{2u - 8v + 4} \quad y = \frac{2v}{2u - 8v + 4}. \quad (24.13)$$

The color differences between two points in the (x, y) chromaticity diagram are spatially very non-uniform, that is, the color changes much more rapidly in one direction, e.g. the x -direction, compared with the other direction, e.g. the y -direction. This deficiency of the (x, y) chromaticity diagram is strongly reduced, although not eliminated, in the (u, v) and (u', v') uniform chromaticity diagrams. As a result, the *color difference* between two locations in the uniform chromaticity diagram is (approximately) directly proportional to the *geometrical distance*

between these points.

The 1943 MacAdam ellipses in the (x, y) chromaticity diagram can be transformed to the uniform (u', v') chromaticity diagram. This transformation is shown in **Figure 24.9**, which reveals that the areas of non-discernable colors are much more uniform in shape and area than the MacAdam ellipses in the (x, y) chromaticity diagram.

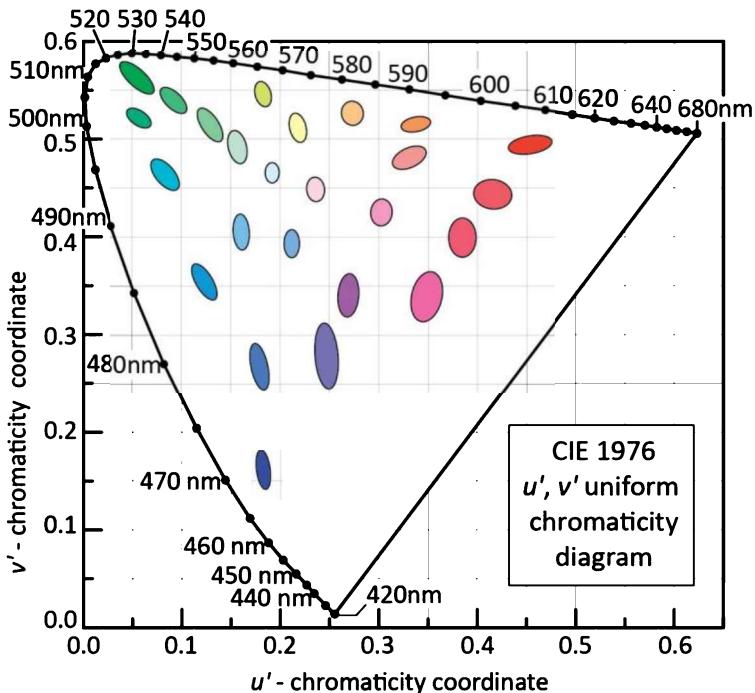


Fig. 24.9: MacAdam ellipses transformed to uniform CIE 1976 (u', v') chromaticity coordinates. For clarity, the axes of the transformed ellipses are ten times their actual lengths. That is, the ellipses shown are 10-step MacAdam ellipses. Transformed ellipses are not ellipses in a strict mathematical sense, but their shapes closely resemble those of ellipses. The areas of the transformed ellipses in the (u', v') diagram are much more similar in area than the MacAdam ellipses in the (x, y) diagram.

Note that the transformation from the (x, y) to the (u', v') coordinate system is mathematically non-linear so that ellipses in the (x, y) chromaticity diagram do *not* transform into ellipses in the (u', v') chromaticity diagram. However, if the ellipses in the (x, y) chromaticity diagram are sufficiently small in size, non-linear distortions are small as well, so that the transformed regions are very close to ellipses.

24.3 Color purity

Monochromatic sources ($\Delta\lambda \rightarrow 0$) are located on the perimeter of the chromaticity diagram. However, as the spectral linewidth of a light source gets broader, the color location in the chromaticity diagram moves towards the center of the chromaticity diagram. If the spectral width of a light source becomes comparable to the entire visible range, the light source is *white* and thus located near the center of the chromaticity diagram.

The **dominant wavelength** of a test light source is defined as the wavelength (i.e. monochromatic color) located on the perimeter of the chromaticity diagram that appears to be closest to the color of the test light source. The dominant wavelength is determined by drawing a straight line from the equal-energy point to the (x, y) chromaticity coordinate of the test light source, and by extending the straight line to the perimeter of the chromaticity diagram. The intersection point is the dominant wavelength of the light source. The procedure is schematically shown in **Figure 24.10**.

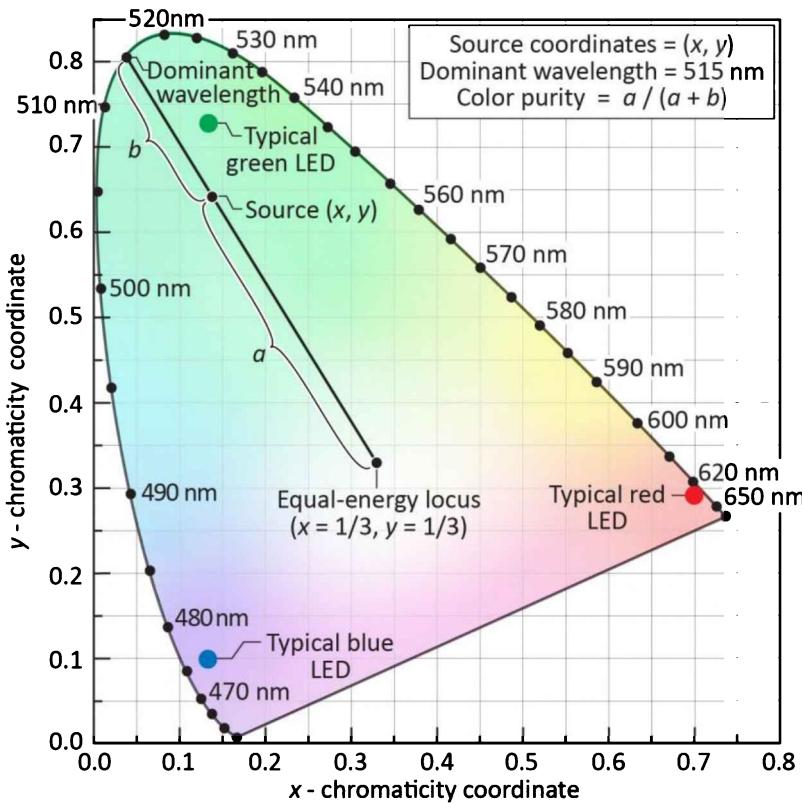


Fig. 24.10: Chromaticity diagram showing the determination of the **dominant wavelength** and **color purity** of a light source with chromaticity coordinates (x, y) using the equal-energy locus ($x = 1/3, y = 1/3$) as the white-light reference. Also shown are typical locations of blue, green, and red LEDs.

The **color purity** or **color saturation** of a light source is the distance in the chromaticity diagram between the (x, y) color-coordinate point of the test source and the coordinate of the equal-energy point divided by the distance between the equal-energy point and the dominant-wavelength point. The color purity is thus given by

$$\text{Color purity} = \frac{a}{a + b} = \frac{\sqrt{(x - x_{ee})^2 + (y - y_{ee})^2}}{\sqrt{(x_d - x_{ee})^2 + (y_d - y_{ee})^2}} \quad (24.14)$$

where a and b are shown in **Figure 24.10** and (x, y) , (x_{ee}, y_{ee}) , and (x_d, y_d) represent the chromaticity coordinates of the light source under test, of the equal-energy reference illuminant,

and of the dominant-wavelength point, respectively. Thus, the color purity is the relative distance of a light source under test from the center of the chromaticity diagram. Generally, the color purity is 100% for monochromatic light sources ($\Delta\lambda \rightarrow 0$) located on the perimeter of the chromaticity diagram and near 0% for white illuminants.

Note that the dominant wavelength and color purity are an alternative way to uniquely characterize the location of an emitter on the chromaticity diagram. Dominant wavelength and color purity are quite intuitive quantities (more so than the numerical (x, y) chromaticity coordinates) and they are therefore frequently preferred.

24.4 LEDs in the chromaticity diagram

Monochromatic light sources ($\Delta\lambda \rightarrow 0$) are located on the perimeter of the chromaticity diagram. Light emission from LEDs is monochromatic (single color) to the eye but LEDs are not monochromatic in the strict physical sense since LEDs have a spectral linewidth of about $1.8kT$. Owing to the finite spectral linewidth of LEDs, they are not located on the very perimeter of the chromaticity diagram but are located *close* to the perimeter. When a source emits light distributed over a range of wavelengths, then the chromaticity location moves towards the center of the diagram.

The location of different LEDs on the chromaticity diagram is shown in *Figure 24.11*. Inspection of the figure reveals that the location of red and blue LEDs is on the perimeter of the chromaticity diagram. That is, their color purity is very high, close to 100%. However, blue-green and green LEDs are located off the perimeter closer to the center of the diagram due to the finite linewidth of the emission spectrum and the strong curvature of the chromaticity diagram in the green wavelength range.

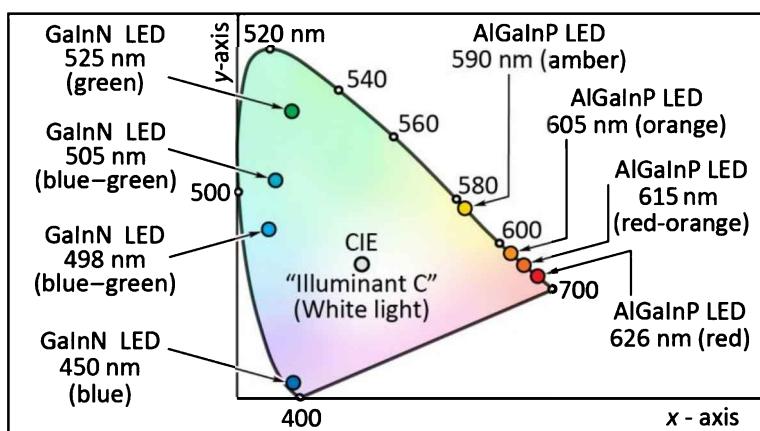


Fig. 24.11: Location of LED light emission on the chromaticity diagram (adapted from Schubert and Miller, 1999).

24.5 Relationship between chromaticity and color

A limitation of chromaticity, as discussed in prior sections of this chapter, is that it does not include the lightness of an object. As an example, *Figure 24.12* (a) shows the hues green and magenta, as well as gray. Within each row, the chromaticity does not change. However, the lightness of the color samples changes within each row thereby changing its appearance quite dramatically.

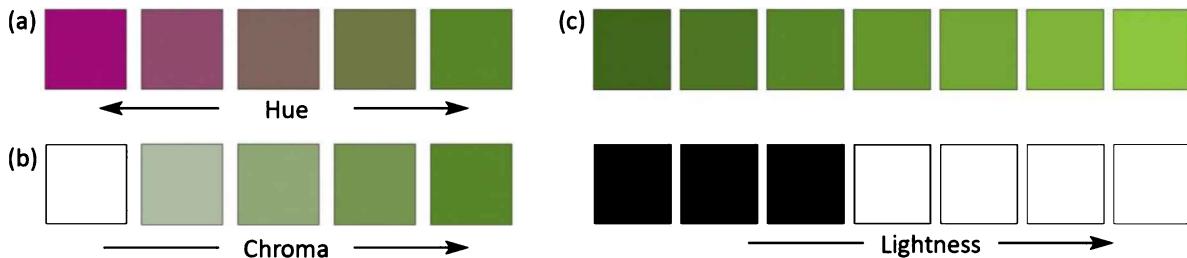


Fig. 24.12: (a) Color samples with changing hue. (b) Color samples with changing chroma. (c) Color samples having, in each row, different lightness.

The combination of chromaticity and lightness constitutes the **color** of an object. Whereas chromaticity forms a two-dimensional plane, the addition of lightness adds a third dimension, thereby opening up a three-dimensional **color space**.

Although color spaces could be defined by chromaticity and lightness, color spaces are usually defined by **hue**, **chroma**, and **lightness**. The meaning of these terms may be gained by inspection of *Figure 24.12*. Alternatively, the meaning of these terms can be articulated as follows:

Hue: The hue of a color sample relates the appearance of the sample to the colors of (at least) red, green, and blue, including the similarity of the sample to these colors or the differences from these colors. Hue refers to a pure color and is an element of the color wheel (for a discussion of the color wheel, the reader is referred to Wilhelm von Bezold's 1874 color wheel, as discussed under the key work "color wheel" by Wikipedia, 2013).

Chroma: The chroma of a color sample relates to the degree of color saturation or the degree of color purity. Color samples with high chroma exhibit saturated, pure, strong colors.

Lightness: The lightness of a color sample relates to the degree of the sample's whiteness, grayness, or blackness. The term *lightness* is used for non-light-emitting color samples, whereas the term *brightness* is used for light sources. Recall that the brightness of a light source is synonymous with luminance.

A common color space is the **CIE Lab color space**, which will be discussed next.

24.6 The CIE Lab color space

The 1976 CIE *Lab* color space (CIE L^* , a^* , b^*) is the most commonly used color space. It was established by the CIE in 1976 (CIE, 1976). It bears its name from the three coordinates of the space, i.e. L^* , a^* , and b^* . The 1976 CIE L^* , a^* , b^* color space is a modification and refinement of the 1948 Hunter color space, developed by Hunter (1948a; 1949b), which had that coordinates L , a , and b . The star symbol ("*") was appended to the three coordinates in order to differentiate the 1976 CIE L^* , a^* , b^* color space from the 1948 Hunter L , a , b color space. The 1976 CIE L^* , a^* , b^* color space is frequently referred to as the 1976 CIE *Lab* color space. The CIE *Lab* color space allows one to quantify colors, including the differences of colors.

The CIE *Lab* color space is an independent standard that does not depend on a specific input device (such as a camera or a scanner) and output device (such as a color printer or a display). This characteristic, i.e. being a device-independent standard, facilitates the transmittance of colors between input and output devices from different manufacturers while maintaining a high fidelity in colors.

The CIE *Lab* color space is based on human perception; this means that the geometric distance between color points is equivalent to the color differences perceived by a human subject, or specifically, by the human visual system (eye and brain).

The variables L^* , a^* , b^* form a coordinate system with the coordinate axes intersecting at right angles, as illustrated in the schematic of *Figure 24.13*. The coordinates are defined as:

- L^* , lightness, with $L^* = 0$ being pure black and $L^* = 100$ being pure white
- a^* , the red–green axis, with positive values being red, and negative values being green
- b^* , the yellow–blue axis, with positive values being yellow, and negative values being blue.

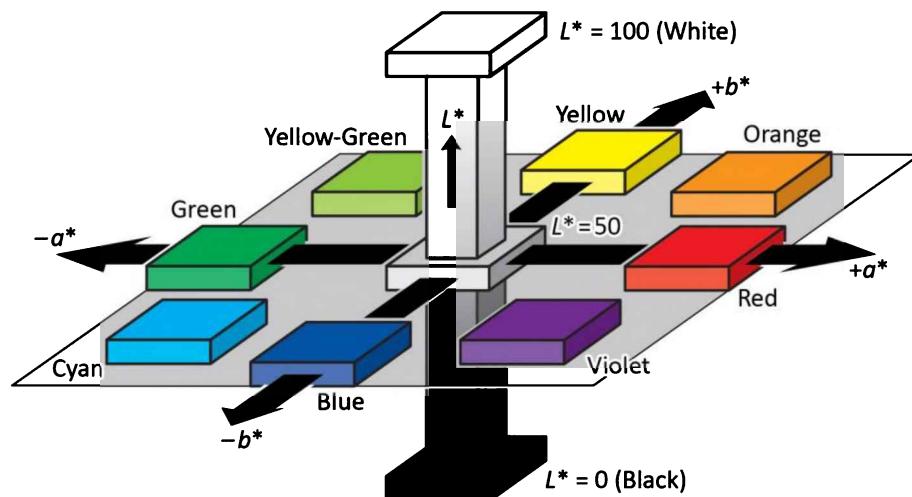


Fig. 24.13: CIE *Lab* color space defined by the coordinates L^* (representing lightness), a^* , and b^* . Lightness is represented by the vertical direction; Hue by the position on the perimeter of the horizontal plane (“color wheel”); Chroma (or color saturation) by the radial distance from the center.

The center point of the color space, i.e. $L^* = 50$ and $a^* = b^* = 0$, is an achromatic gray. This point is chosen to provide normalization, e.g. chosen to be a reference white point such as the D₆₅ point, so that color points can be described for a specific illumination condition. Moving around the vertical axis in a circular motion describes what is frequently called the “**color wheel**”, which describes the basic colors of the visible spectrum.

The 1976 CIE *Lab* color space is illustrated with colors displayed in a realistic, i.e. continuous manner, in **Figure 24.14**. Hue and chroma may be defined as (Schanda, 2007):

Hue = $\arctan(b^*/a^*)$. Accordingly, red hue is described by the angle 0°, yellow hue by 90°, green hue by 180°, and blue hue by 270°.

Chroma = $(a^{*2} + b^{*2})^{1/2}$. Accordingly, the center vertical axis has a chroma of zero.

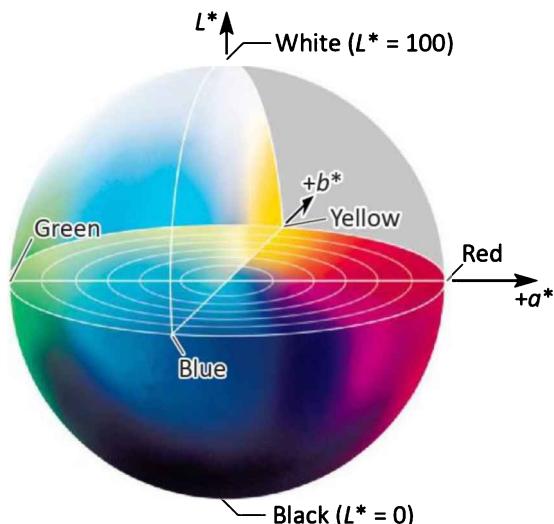


Fig. 24.14: CIE *Lab* color space defined by the three coordinates L^* , a^* and b^* . Lightness is represented by the vertical direction; chroma by the radial direction originating from the center; and hue by the position around the equator of the sphere. A color point can be expressed by its three coordinates. The center of the sphere (i.e. $L^* = 50$ and $a^* = b^* = 0$) is a reference achromatic point (or gray or white point), e.g., the D₆₅ chromaticity point. Color difference, ΔE , is the geometric distance between two color points.

The variables L^* , a^* , b^* are calculated from the CIE *X*, *Y*, and *Z* tristimulus values according to (CIE, 1976)

$$L^* = 116 f(Y/Y_n) - 16 \quad (24.15)$$

$$a^* = 500 [f(X/X_n) - f(Y/Y_n)] \quad (24.16)$$

$$b^* = 200 [f(Y/Y_n) - f(Z/Z_n)] \quad (24.17)$$

where

$$f(\alpha) = \alpha^{1/3} \quad \text{for} \quad \alpha > (6/29)^3 \quad (24.18a)$$

$$f(\alpha) = \frac{841}{108} \alpha + \frac{4}{29} \quad \text{for} \quad \alpha \leq (6/29)^3 \quad (24.18a)$$

and X_n , Y_n , and Z_n are the tristimulus values of the reference white point.

The CIE *Lab* color space allows for the quantitative determination of **color differences**. Assuming that the geometric distance is proportional to the color difference, the color difference between two color points with coordinates L_1^* , a_1^* , b_1^* and L_2^* , a_2^* , b_2^* is given by (CIE, 1976)

$$\begin{aligned} \Delta E_{1976}^* &= \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2} \\ &= \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \end{aligned} \quad (24.19)$$

The question arises as to what numerical value of ΔE constitutes a **just noticeable difference** in color? It is generally accepted that $\Delta E = 2.3$ constitutes a just noticeable difference. Color differences are highly relevant for determination of the color rendering index, to be discussed in a subsequent chapter.

Note that Eqn. (24.19) gives the 1976 CIE color difference. Additional modifications to the formula calculating ΔE have been made by the CIE in 1994 (ΔE_{1994}) and 2000 (ΔE_{2000}) in order to more closely emulate human color vision. The reader is referred to the technical literature for the formulas giving ΔE_{1994} and ΔE_{2000} (see e.g. Wikipedia under “color difference”).

24.7 The CIE UVW and CIE Luv color spaces

The CIE has introduced additional color spaces: Most notably the 1964 CIE U^* , V^* , W^* color space, also called the **1964 CIE UVW** color space. This color space is frequently used for the calculation of the color rendering index (to be discussed in a subsequent chapter). Furthermore, the 1976 CIE L^* , u^* , v^* uniform color space, also called the **1976 CIE Luv** uniform color space, which can be considered a refinement of the 1964 CIE *UVW* color space (Schanda, 2007). The interested reader is referred to the literature for these color spaces (e.g., Schanda, 2007).

24.8 References

- CIE *Commission Internationale de l'Eclairage Proceedings* (Cambridge University Press, Cambridge, 1931)
- CIE data of 1931 and 1978 available at <http://www.cvrl.org> (1931)
- CIE data of 1960 relating to the (u, v) chromaticity coordinates can be found in CIE, 1986 (1960)
- CIE data of 1976 relating to the (u', v') chromaticity coordinates can be found in CIE, 1986 (1976)
- CIE *Lab* color space of 1976 can be found (i) at the CIE Internet site <<http://www.cie.co.at/>>, (ii) in the

- book “Colorimetry – Understanding the CIE system” edited by J. Schanda (Wiley-Interscience, Hoboken NJ, 2007), and (iii) on Wikipedia under the term “CIE Lab” (accessed 2013) (1976)
- CIE data of 1931 and 1978 available at <http://www.cvrl.org> (1978)
- CIE publication No. 15 (E.1.3.1) 1971: *Colorimetry*; this publication was updated in 1986 to CIE Publication 15.2 *Colorimetry* (CIE, Vienna Austria, 1986)
- Gage S., Evans D., Hodapp M. W., and Sorensen H. *Optoelectronics Applications Manual* 1st edition (McGraw Hill, New York, 1977)
- Hunter R. S. “Photoelectric color-difference meter” *Journal of the Optical Society of America* **38**, 661 (1948a)
- Hunter “Accuracy, precision, and stability of new photo-electric color-difference meter” *Journal of the Optical Society of America* **38**, 1094 (1948b)
- Judd D. B. “Report of US Secretariat Committee on Colorimetry and Artificial Daylight” in *Proceedings of the 12th Session of the CIE* **1**, p. 11 (Bureau Central de la CIE, Paris, 1951)
- Kelly K. L. “Color designations for lights” *Journal of the Optical Society of America* **33**, 627 (1943)
- MacAdam D. L. “Specification of small chromaticity differences” *Journal of the Optical Society of America* **33**, 18 (1943)
- MacAdam D. L. (Editor) *Colorimetry – Fundamentals* (SPIE Optical Engineering Press, Bellingham, Washington, 1993)
- Schanda J. (Editor) *Colorimetry – Understanding the CIE system* (Wiley Interscience, Hoboken, New Jersey, 2007)
- Schubert E. F. and Miller J. N “Light-emitting diodes - An introduction” *Encyclopedia of Electrical Engineering*, edited by John G. Webster, Vol. **11**, p. 326 (John Wiley and Sons, New York, March 1999)
- Vos J. J. “Colorimetric and photometric properties of a 2-degree fundamental observer” *Color Research and Applications* **3**, 125 (1978)
- Wikipedia “Color wheel” (accessed 2013)
- Wyszecki G. and Stiles W. S. *Color Science – Concepts and Methods, Quantitative Data and Formulae* 2nd edition (John Wiley and Sons, New York, 2000)
- Wright W. D. “The graphical representation of small color differences” *Journal of the Optical Society of America* **33**, 632 (1943)

Appendix 24.1

Tabulated CIE 1931 2° color-matching and eye sensitivity function for point sources (after CIE, 1931).

λ (nm)	$\bar{x}(\lambda)$ <i>red</i>	$\bar{y}(\lambda) = V(\lambda)$ <i>green</i>	$\bar{z}(\lambda)$ <i>blue</i>
360	1.2990 E-4	3.9170 E-6	6.0610 E-4
365	2.3210 E-4	6.9650 E-6	1.0860 E-3
370	4.1490 E-4	1.2390 E-5	1.9460 E-3
375	7.4160 E-4	2.2020 E-5	3.4860 E-3
380	1.3680 E-3	3.9000 E-5	6.4500 E-3
385	2.2360 E-3	6.4000 E-5	1.0550 E-2
390	4.2430 E-3	1.2000 E-4	2.0050 E-2
395	7.6500 E-3	2.1700 E-4	3.6210 E-2
400	1.4310 E-2	3.9600 E-4	6.7850 E-2
405	2.3190 E-2	6.4000 E-4	0.11020
410	4.3510 E-2	1.2100 E-3	0.20740
415	7.7630 E-2	2.1800 E-3	0.37130
420	0.13438	4.0000 E-3	0.64560
425	0.21477	7.3000 E-3	1.03905
430	0.28390	1.1600 E-2	1.38560
435	0.32850	1.6840 E-2	1.62296
440	0.34828	2.3000 E-2	1.74706
445	0.34806	2.9800 E-2	1.78260
450	0.33620	3.8000 E-2	1.77211
455	0.31870	4.8000 E-2	1.74410
460	0.29080	6.0000 E-2	1.66920
465	0.25110	7.3900 E-2	1.52810
470	0.19536	9.0980 E-2	1.28764
475	0.14210	0.11260	1.04190
480	9.5640 E-2	0.13902	0.81295
485	5.7950 E-2	0.16930	0.61620
490	3.2010 E-2	0.20802	0.46518
495	1.4700 E-2	0.25860	0.35330
500	4.9000 E-3	0.32300	0.27200
505	2.4000 E-3	0.40730	0.21230
510	9.3000 E-3	0.50300	0.15820
515	2.9100 E-2	0.60820	0.11170
520	6.3270 E-2	0.71000	7.8250 E-2
525	0.10960	0.79320	5.7250 E-2
530	0.16550	0.86200	4.2160 E-2
535	0.22575	0.91485	2.9840 E-2
540	0.29040	0.95400	2.0300 E-2
545	0.35970	0.98030	1.3400 E-2
550	0.43345	0.99495	8.7500 E-3
555	0.51205	1.00000	5.7500 E-3
560	0.59450	0.99500	3.9000 E-3
565	0.67840	0.97860	2.7500 E-3
570	0.76210	0.95200	2.1000 E-3
575	0.84250	0.91540	1.8000 E-3
580	0.91630	0.87000	1.6500 E-3

585	0.97860	0.81630	1.4000 E-3
590	1.02630	0.75700	1.1000 E-3
595	1.05670	0.69490	1.0000 E-3
600	1.06220	0.63100	8.0000 E-4
605	1.04560	0.56680	6.0000 E-4
610	1.00260	0.50300	3.4000 E-4
615	0.93840	0.44120	2.4000 E-4
620	0.85445	0.38100	1.9000 E-4
625	0.75140	0.32100	1.0000 E-4
630	0.64240	0.26500	5.0000 E-5
635	0.54190	0.21700	3.0000 E-5
640	0.44790	0.17500	2.0000 E-5
645	0.36080	0.13820	1.0000 E-5
650	0.28350	0.10700	0.0000 E-5
655	0.21870	8.1600 E-2	0.0000 E-5
660	0.16490	6.1000 E-2	0.0000 E-5
665	0.12120	4.4580 E-2	0.0000 E-5
670	8.7400 E-2	3.2000 E-2	0.0000 E-5
675	6.3600 E-2	2.3200 E-2	0.0000 E-5
680	4.6770 E-2	1.7000 E-2	0.0000 E-5
685	3.2900 E-2	1.1920 E-2	0.0000 E-5
690	2.2700 E-2	8.2100 E-3	0.0000 E-5
695	1.5840 E-2	5.7230 E-3	0.0000 E-5
700	1.1359 E-2	4.1020 E-3	0.0000 E-5
705	8.1109 E-3	2.9290 E-3	0.0000 E-5
710	5.7903 E-3	2.0910 E-3	0.0000 E-5
715	4.1065 E-3	1.4840 E-3	0.0000 E-5
720	2.8993 E-3	1.0470 E-3	0.0000 E-5
725	2.0492 E-3	7.4000 E-4	0.0000 E-5
730	1.4400 E-3	5.2000 E-4	0.0000 E-5
735	9.9995 E-4	3.6110 E-4	0.0000 E-5
740	6.9008 E-4	2.4920 E-4	0.0000 E-5
745	4.7602 E-4	1.7190 E-4	0.0000 E-5
750	3.3230 E-4	1.2000 E-4	0.0000 E-5
755	2.3483 E-4	8.4800 E-5	0.0000 E-5
760	1.6615 E-4	6.0000 E-5	0.0000 E-5
765	1.1741 E-4	4.2400 E-5	0.0000 E-5
770	8.3075 E-5	3.0000 E-5	0.0000 E-5
775	5.8707 E-5	2.1200 E-5	0.0000 E-5
780	4.1510 E-5	1.4990 E-5	0.0000 E-5
785	2.9353 E-5	1.0600 E-5	0.0000 E-5
790	2.0674 E-5	7.4657 E-6	0.0000 E-5
795	1.4560 E-5	5.2578 E-6	0.0000 E-5
800	1.0254 E-5	3.7029 E-6	0.0000 E-5
805	7.2215 E-6	2.6078 E-6	0.0000 E-5
810	5.0859 E-6	1.8366 E-6	0.0000 E-5
815	3.5817 E-6	1.2934 E-6	0.0000 E-5
820	2.5225 E-6	9.1093 E-7	0.0000 E-5
825	1.7765 E-6	6.4153 E-7	0.0000 E-5

Appendix 24.2

Tabulated values of the CIE 1978 two-degree color-matching functions and eye sensitivity function for point sources (after CIE, 1978). The functions are also called the Judd–Vos-modified color-matching functions.

λ (nm)	$\bar{x}(\lambda)$ <i>red</i>	$\bar{y}(\lambda) = V(\lambda)$ <i>green</i>	$\bar{z}(\lambda)$ <i>blue</i>
380	2.6899E-3	2.0000E-4	1.2260E-2
385	5.3105E-3	3.9556E-4	2.4222E-2
390	1.0781E-2	8.0000E-4	4.9250E-2
395	2.0792E-2	1.5457E-3	9.5135E-2
400	3.7981E-2	2.8000E-3	1.7409E-1
405	6.3157E-2	4.6562E-3	2.9013E-1
410	9.9941E-2	7.4000E-3	4.6053E-1
415	1.5824E-1	1.1779E-2	7.3166E-1
420	2.2948E-1	1.7500E-2	1.0658
425	2.8108E-1	2.2678E-2	1.3146
430	3.1095E-1	2.7300E-2	1.4672
435	3.3072E-1	3.2584E-2	1.5796
440	3.3336E-1	3.7900E-2	1.6166
445	3.1672E-1	4.2391E-2	1.5682
450	2.8882E-1	4.6800E-2	1.4717
455	2.5969E-1	5.2122E-2	1.3740
460	2.3276E-1	6.0000E-2	1.2917
465	2.0999E-1	7.2942E-2	1.2356
470	1.7476E-1	9.0980E-2	1.1138
475	1.3287E-1	1.1284E-1	9.4220E-1
480	9.1944E-2	1.3902E-1	7.5596E-1
485	5.6985E-2	1.6987E-1	5.8640E-1
490	3.1731E-2	2.0802E-1	4.4669E-1
495	1.4613E-2	2.5808E-1	3.4116E-1
500	4.8491E-3	3.2300E-1	2.6437E-1
505	2.3215E-3	4.0540E-1	2.0594E-1
510	9.2899E-3	5.0300E-1	1.5445E-1
515	2.9278E-2	6.0811E-1	1.0918E-1
520	6.3791E-2	7.1000E-1	7.6585E-2
525	1.1081E-1	7.9510E-1	5.6227E-2
530	1.6692E-1	8.6200E-1	4.1366E-2
535	2.2768E-1	9.1505E-1	2.9353E-2
540	2.9269E-1	9.5400E-1	2.0042E-2
545	3.6225E-1	9.8004E-1	1.3312E-2
550	4.3635E-1	9.9495E-1	8.7823E-3
555	5.1513E-1	1.0000	5.8573E-3
560	5.9748E-1	9.9500E-1	4.0493E-3
565	6.8121E-1	9.7875E-1	2.9217E-3
570	7.6425E-1	9.5200E-1	2.2771E-3
575	8.4394E-1	9.1558E-1	1.9706E-3
580	9.1635E-1	8.7000E-1	1.8066E-3
585	9.7703E-1	8.1623E-1	1.5449E-3
590	1.0230	7.5700E-1	1.2348E-3

595	1.0513	6.9483E–1	1.1177E–3
600	1.0550	6.3100E–1	9.0564E–4
605	1.0362	5.6654E–1	6.9467E–4
610	9.9239E–1	5.0300E–1	4.2885E–4
615	9.2861E–1	4.4172E–1	3.1817E–4
620	8.4346E–1	3.8100E–1	2.5598E–4
625	7.3983E–1	3.2052E–1	1.5679E–4
630	6.3289E–1	2.6500E–1	9.7694E–5
635	5.3351E–1	2.1702E–1	6.8944E–5
640	4.4062E–1	1.7500E–1	5.1165E–5
645	3.5453E–1	1.3812E–1	3.6016E–5
650	2.7862E–1	1.0700E–1	2.4238E–5
655	2.1485E–1	8.1652E–2	1.6915E–5
660	1.6161E–1	6.1000E–2	1.1906E–5
665	1.1820E–1	4.4327E–2	8.1489E–6
670	8.5753E–2	3.2000E–2	5.6006E–6
675	6.3077E–2	2.3454E–2	3.9544E–6
680	4.5834E–2	1.7000E–2	2.7912E–6
685	3.2057E–2	1.1872E–2	1.9176E–6
690	2.2187E–2	8.2100E–3	1.3135E–6
695	1.5612E–2	5.7723E–3	9.1519E–7
700	1.1098E–2	4.1020E–3	6.4767E–7
705	7.9233E–3	2.9291E–3	4.6352E–7
710	5.6531E–3	2.0910E–3	3.3304E–7
715	4.0039E–3	1.4822E–3	2.3823E–7
720	2.8253E–3	1.0470E–3	1.7026E–7
725	1.9947E–3	7.4015E–4	1.2207E–7
730	1.3994E–3	5.2000E–4	8.7107E–8
735	9.6980E–4	3.6093E–4	6.1455E–8
740	6.6847E–4	2.4920E–4	4.3162E–8
745	4.6141E–4	1.7231E–4	3.0379E–8
750	3.2073E–4	1.2000E–4	2.1554E–8
755	2.2573E–4	8.4620E–5	1.5493E–8
760	1.5973E–4	6.0000E–5	1.1204E–8
765	1.1275E–4	4.2446E–5	8.0873E–9
770	7.9513E–5	3.0000E–5	5.8340E–9
775	5.6087E–5	2.1210E–5	4.2110E–9
780	3.9541E–5	1.4989E–5	3.0383E–9
785	2.7852E–5	1.0584E–5	2.1907E–9
790	1.9597E–5	7.4656E–6	1.5778E–9
795	1.3770E–5	5.2592E–6	1.1348E–9
800	9.6700E–6	3.7028E–6	8.1565E–10
805	6.7918E–6	2.6076E–6	5.8626E–10
810	4.7706E–6	1.8365E–6	4.2138E–10
815	3.3550E–6	1.2950E–6	3.0319E–10
820	2.3534E–6	9.1092E–7	2.1753E–10
825	1.6377E–6	6.3564E–7	1.5476E–10

25

25 – Black-body radiation and color temperature

Incandescence (heat glow), black-body radiation, and its associated color temperature are intimately related to white light sources, including natural white light sources (such as the sun) and man-made white light sources (such as an LED light bulb). Therefore, we will discuss these concepts and show their usefulness for the characterization of white light sources used for general lighting applications.

25.1 Incandescence (heat glow)

We are all familiar with the concept of **heat glow**, that is, the emission of light by a hot object. Consider an object, such as a horse shoe, that is placed in a fire thereby heating up the object. Assume that the object will not undergo a chemical reaction when being heated up. Upon reaching a temperature of about 700°C , the object starts to glow, i.e. emit radiation within the visible spectrum. The color of the glow starts with red; as the temperature increases, the glow becomes orange, yellow, and ultimately the glow becomes purely white. The heat glow of a black object, heated to different temperatures, is schematically shown in *Figure 25.1*.

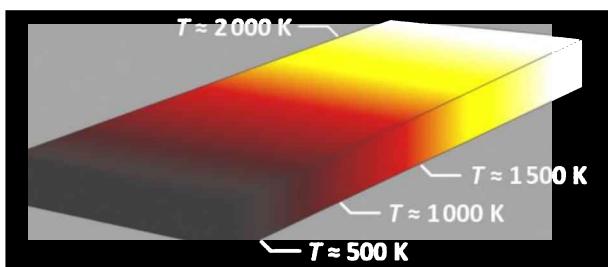


Fig. 25.1: Schematic illustration of an object, for example a metal bar, heated to the temperatures indicated. The object emits heat glow, also known as incandescence. At lower temperatures, the heat glow is red, changing to orange, yellow, and white, as the temperature is increased.

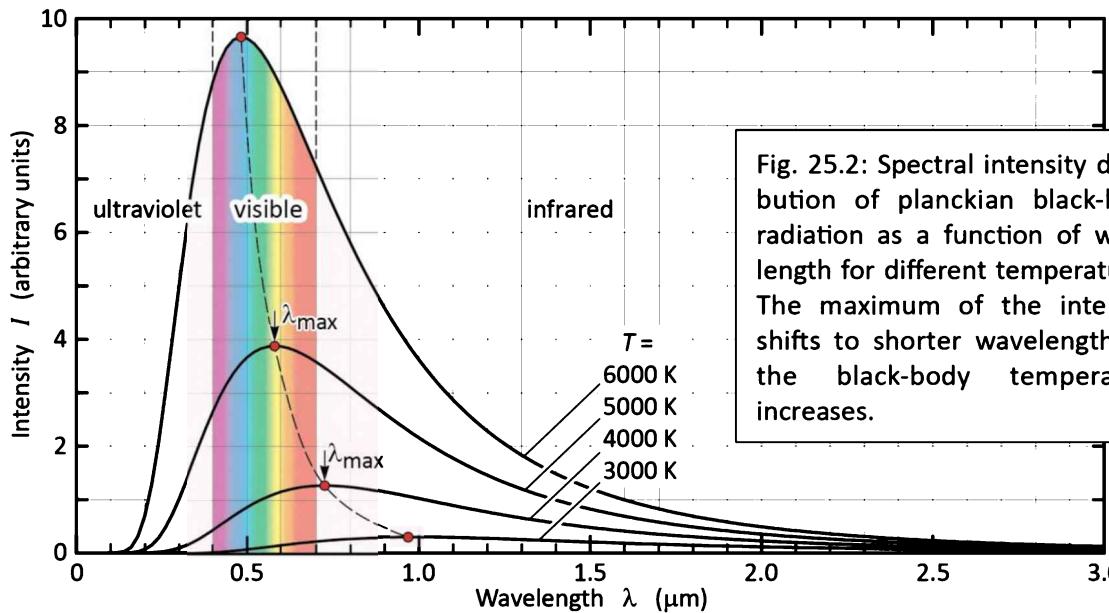
Incandescence is of Latin origin and has, in the present context, the same meaning as heat glow. Accordingly, light sources based on incandescence, are called **incandescent sources**. For example, conventional light bulbs use a tungsten filament that is heated by an electrical current so that the filament reaches temperatures of approximately 3 000 K. As a result, the filament emits yellowish white light, i.e. incandescence. Conventional light bulbs based on a hot filament, are therefore incandescent sources.

25.2 The planckian spectrum

The theory of electromagnetic radiation emanating from an incandescent object is quite complex and was first developed by Max Planck (1900). He assumed that the incandescent object is black so that any light emitted by the object can only be due to incandescence, or black-body radiation. The **black-body radiation spectrum** (or simply **planckian spectrum**) depends on only one parameter, T , the temperature of the body. The black-body spectrum is given by (Planck, 1900)

$$I(\lambda) = \frac{2 h c^2}{\lambda^5 \left[\exp\left(\frac{hc}{\lambda kT}\right) - 1 \right]} . \quad (25.1)$$

This formula allows us to calculate the intensity, I , of the light emitted from the black body as a function its wavelength, λ . The planckian spectrum is shown for different black-body temperatures in **Figure 25.2**. Inspection of the figure reveals that at “low” black-body temperatures, e.g. 3 000 K, the radiation occurs mostly in the infrared. As the temperature increases, the maximum of the radiation shifts into the visible wavelength range.



The maximum intensity of radiation emanating from a black body of temperature T occurs at a specific wavelength that shifts to shorter wavelengths as the black-body temperature increases. The wavelength of maximum intensity, λ_{max} , is given by **Wien's displacement law**, i.e.

$$\lambda_{max} = \frac{2880 \mu\text{m K}}{T} . \quad (25.2)$$

The development of the theory of black-body radiation was a major scientific development (Planck, 1900). The development of the theory was based on the assumption, that charged particles (electrons and atom nuclei), whose mechanical oscillation becomes stronger under the influence of heat, can only have specific oscillatory energies, specifically multiples of $h\nu$, where h is Planck's constant, and ν is the vibrational frequency of the charged particles. Therefore, the theory of black-body radiation is also associated with the origination of quantum mechanics. A photograph of Max Planck is shown in *Figure 25.3*.

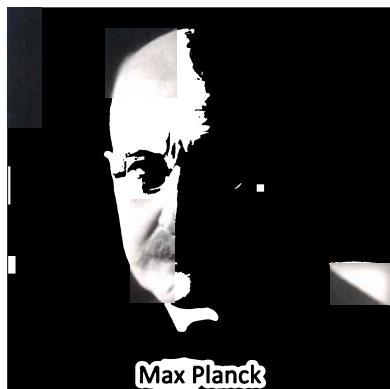


Fig. 25.3: Max Planck (1858–1947), a German theoretical physicist, who developed the theory of black body radiation. Planck also originated quantum theory, for which he won the Nobel Prize in Physics in 1918. Quantum theory revolutionized the understanding of atomic and subatomic processes.

As we will see, the planckian black-body spectrum is highly relevant to our natural light source, the sun, as well as to lighting in general. Indeed, the black-body spectrum is used to characterize white light. White light is a unique color. There are a very large number of optical spectra that can be used to generate white light. Among these spectra, the black-body spectrum forms a unique and very useful standard because (i) daylight strongly resembles the black-body spectrum and (ii) it allows us to describe the spectrum with only one parameter, namely the black-body temperature.

25.3 The solar spectrum

White light usually has a broad spectrum extending over the entire visible spectral range. An instructive model for white light is sunlight. The sun's optical spectrum is shown in *Figure 25.4*, including the spectrum at sea level with the sun at zenith, incident above the earth's atmosphere, and at sunset and sunrise (Jackson, 1975). The solar spectrum extends over the entire visible region. However, the sun's spectrum, to a certain degree, depends on the time of day, season, altitude, and weather.

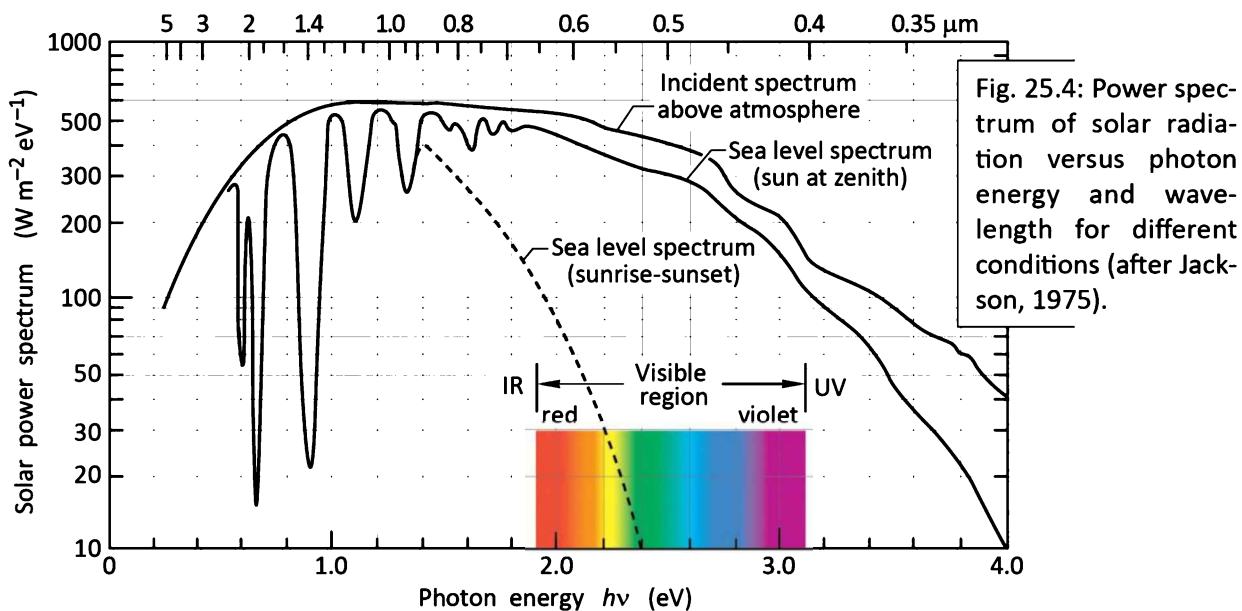


Fig. 25.4: Power spectrum of solar radiation versus photon energy and wavelength for different conditions (after Jackson, 1975).

Exact replication of the solar spectrum for white-light illumination sources would not yield an efficient source due to the large infrared (IR) and ultraviolet (UV) components of the solar spectrum. Thus, the exact replication of the spectrum of “mother nature” would not result in an efficient white source. Even if the IR and UV components of the spectrum were to be eliminated, the solar spectrum still would not be the most energy efficient one, due to its high intensity at the visible–IR and visible–UV boundaries.

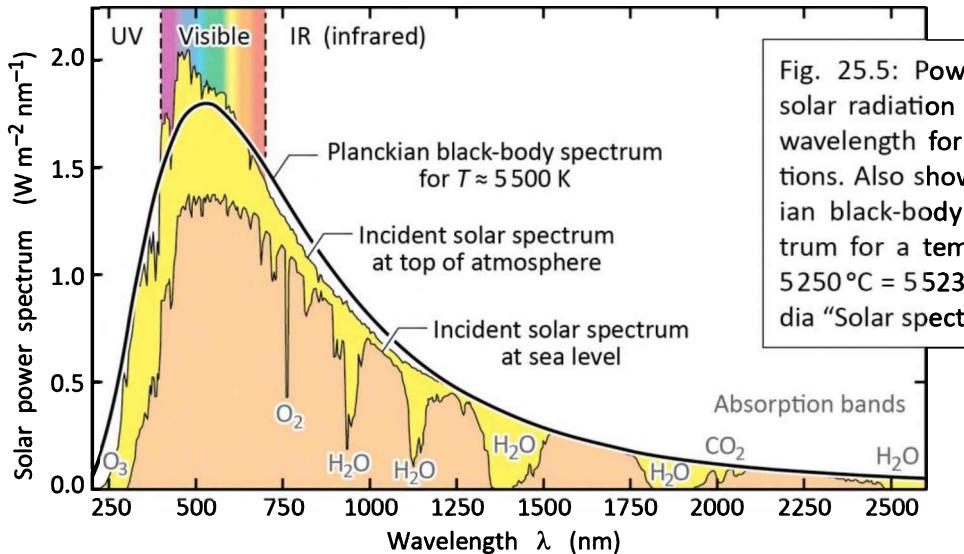


Fig. 25.5: Power spectrum of solar radiation as a function of wavelength for different conditions. Also shown is the planckian black-body radiation spectrum for a temperature of $T = 5250^\circ\text{C} = 5523 \text{ K}$ (after Wikipedia “Solar spectrum”, 2013).

Figure 25.5 shows the solar spectrum along with the planckian black-body spectrum for $T \approx 5500 \text{ K}$ (Wikipedia, 2013). Inspection of the figure reveals two notable points: *First*, the black-body spectrum closely replicates the solar spectrum! This legitimizes the use of the black-body spectrum as an exemplary spectrum for white light. *Second*, human vision has evolved to near-

perfectly match the radiation spectrum of solar radiation! That is, the spectral region of maximum solar radiation falls into the visible spectrum.

25.4 Planckian radiation in the chromaticity diagram

The location of the black-body radiation in the (x, y) chromaticity diagram, called **planckian locus**, is shown in *Figure 25.6*. As the temperature of the black body increases, the chromaticity location moves from the red wavelength range towards the center of the diagram. Typical black-body temperatures in the white region of the chromaticity diagram range between 2 500 and 10 000 K.

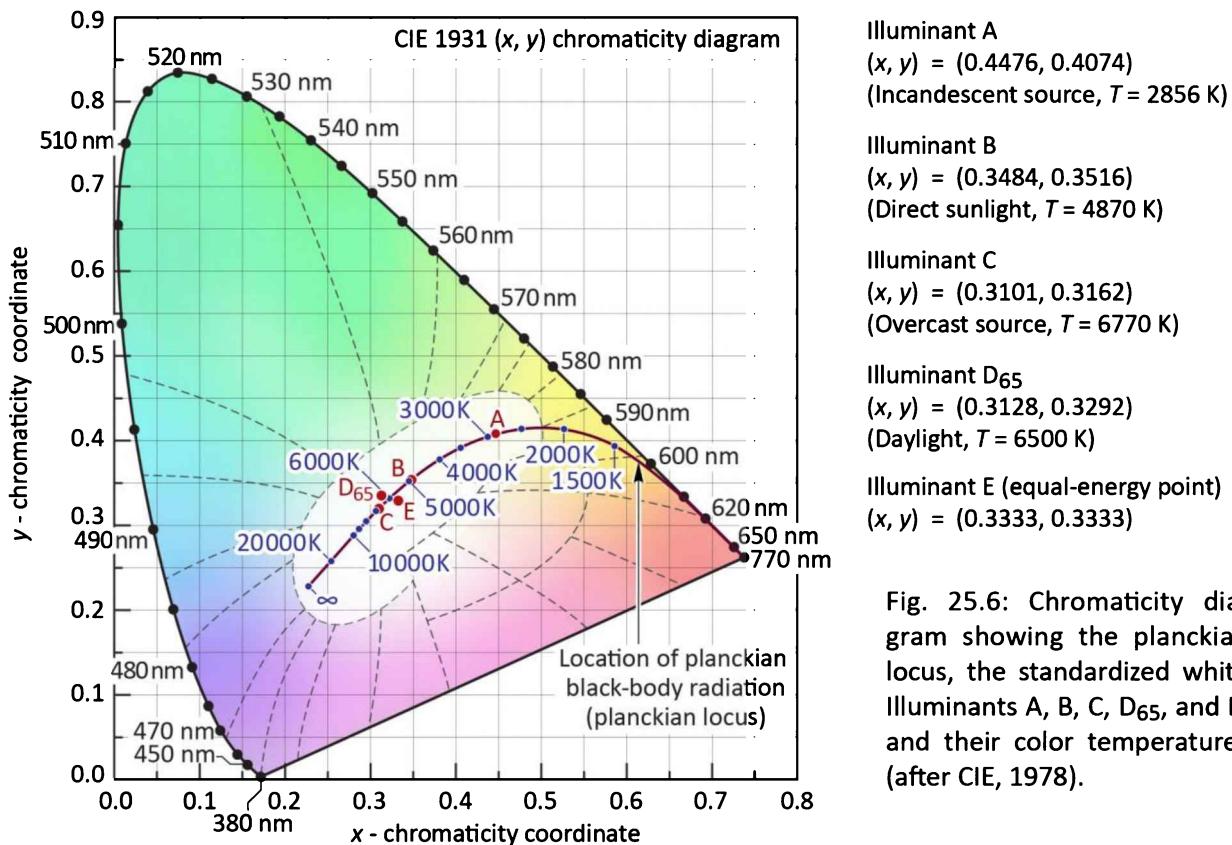


Fig. 25.6: Chromaticity diagram showing the planckian locus, the standardized white Illuminants A, B, C, D₆₅, and E, and their color temperatures (after CIE, 1978).

Also shown in *Figure 25.6* are the locations of several illuminants standardized by the CIE. These standard illuminants include Illuminants A, B, C, D₆₅, and E. We see that these standard illuminants are all located on the planckian locus or exceedingly close to the planckian locus.

The planckian locus and the locations of the black-body temperatures in the (u', v') uniform chromaticity diagram are shown in *Figure 25.7*. The (x, y) and in (u', v') chromaticity coordinates of the planckian radiator are tabulated in Appendix 25.1.

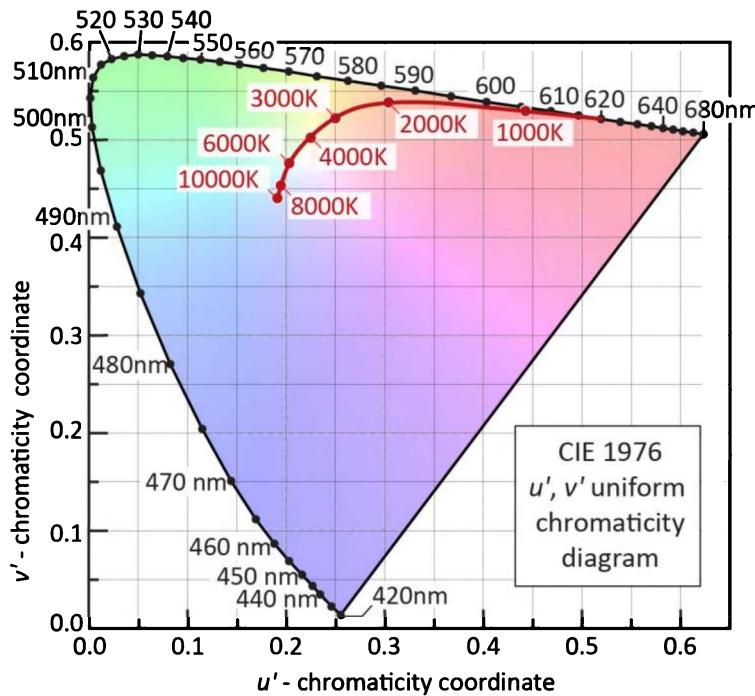


Fig. 25.7: CIE 1976 (u' , v') uniform chromaticity diagram calculated using the CIE 1931 2° standard observer and planckian locus.

In both the (x, y) and the (u', v') chromaticity diagrams, the planckian locus starts out in the red, then moves through the orange and yellow, to finally the white region. This sequence of colors is reminiscent of the colors of a real object (e.g. a piece of metal) heated to high temperatures, indicating that real objects closely follow the chromaticity of Planck's idealized black body.

25.5 Color temperature and correlated color temperature

Color temperature may appear to be a somewhat surprising quantity because the two quantities, *color* and *temperature*, don't seem to have a direct relationship with each other. However, the relationship is derived from the black-body radiator. With increasing temperatures, it glows in the red, orange, yellowish white, white, and ultimately bluish white.

When plotting the chromaticity point of a white light source into the chromaticity diagram, the point either (i) falls onto the planckian locus or (ii) does not fall onto the planckian locus; then, the white source has a (i) *color temperature* or a (ii) *correlated color temperature*, respectively.

The **color temperature** (CT) of a white light source, is the temperature of the planckian black-body radiator that has the same chromaticity location as the white light source considered.

Figure 25.8 shows the planckian locus and associated CTs on the planckian locus in the $(u', 2/3 v')$ chromaticity diagram.

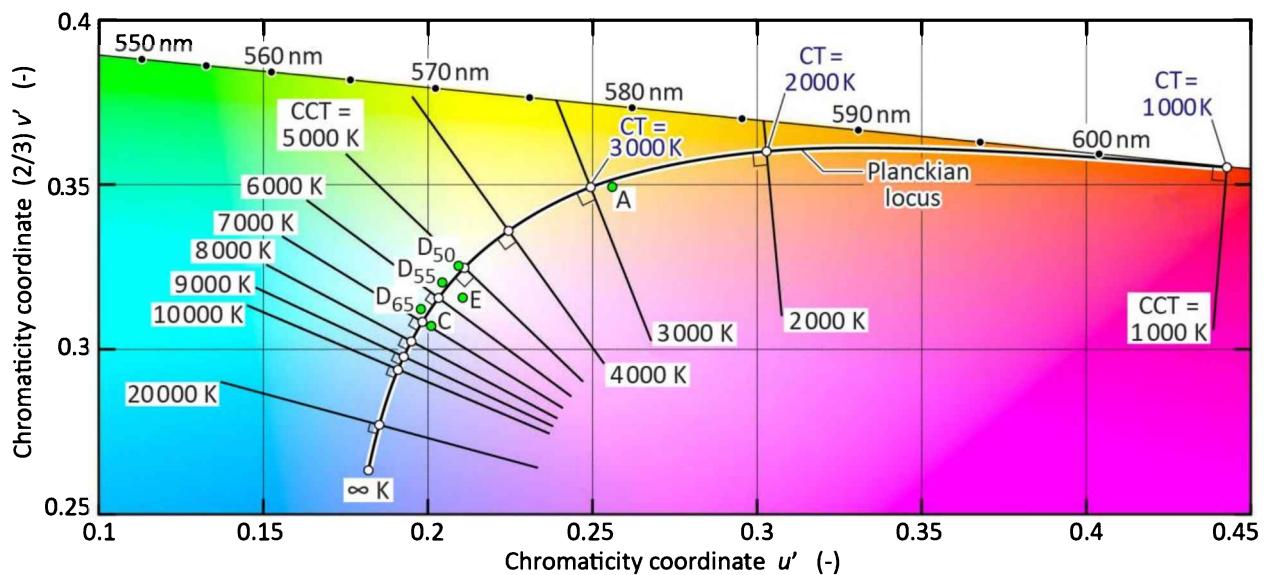


Fig. 25.8: Planckian locus and right-angle intersecting lines of constant correlated color temperature (CCT) in the $(u', 2/3 v')$ chromaticity diagram. CCT points have the minimum distance from planckian color temperature (CT) points. The CCT is defined for white sources that are no more than 0.05 units away from the planckian locus.

If the color of a white light source does not fall on the planckian locus, the **correlated color temperature** (CCT) is used. The correlated color temperature of a white light source is understood as the temperature of a planckian black-body radiator whose chromaticity is closest to the chromaticity of the white light source. The CIE considers the CCT to be a valid concept for light sources that are within a distance of 0.05 from the planckian locus, as indicated in **Figure 25.8**.

The correlated color temperature of a light source is determined according to the following specific procedure (Robertson, 1968; CIE publication No. 17.4, of 1987; and Schanda, 2007): On the $(u', 2/3 v')$ chromaticity diagram, the point on the planckian locus that is *closest* to the chromaticity location of the light source, is determined (i.e. shortest geometrical distance). The correlated color temperature is the temperature of the planckian black-body radiator at that point. The determination of the correlated color temperature was discussed by. **Figure 25.8** shows the $(u', 2/3 v')$ chromaticity diagram along with the planckian locus, CTs on the planckian locus, as well as the locations of constant CCTs off the planckian locus.

On the (x, y) chromaticity diagram, the correlated color temperature *cannot* be determined by using the shortest distance to the planckian locus due to the non-uniformity of the (x, y) chromaticity diagram. The lines of constant correlated color temperature in the (x, y) chromaticity diagram are shown in **Figure 25.9** (Duggal, 2005).

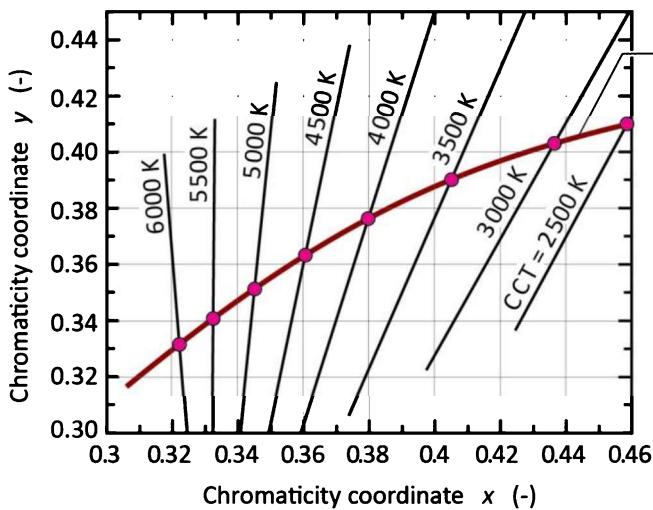


Fig. 25.9: Lines of constant correlated color temperature (CCT) in the (x, y) chromaticity diagram. Whereas the CCT follows from the minimum distance to the planckian in the (u', v') diagram, this is not the case in the (x, y) diagram. The CCT is defined for white light sources that are a distance of no more than 0.05 away from the planckian locus (after Duggal, 2005).

The chromaticity locations of incandescent light sources are close to, but not on the planckian locus (Ohno, 2001). For such incandescent light sources, the *color temperature* can be specified. Standard incandescent sources have color temperatures ranging from about 2 000 to 2 900 K. Conventional incandescent light sources (W filament light bulbs) have a color temperature of about 2 800 K. Quartz halogen incandescent lamps have a color temperature ranging from 2 800 to 3 200 K (Ohno, 1997). Other light sources, such as metal-halide sources, are further removed from the planckian locus. For such light sources, the correlated color temperature should be used. Bluish white lights have a correlated color temperature of greater than 6 500 K. Color temperatures of common man-made and natural light sources are given in **Table 25.1**.

Table 25.1: Correlated color temperatures of artificial and natural light sources.

Light source	Correlated color temperature (K)
Wax candle flame / CIE standard candle flame	1 500 to 2 000 / 2 000
W filament incandescent lamp: 40 / 60 / 100 W	2 700 / 2 800 / 2 850
W filament halogen incandescent lamp	2 800 to 3 200
“Warm white” fluorescent tube	3 000
“Daylight white” fluorescent tube	4 300
“True daylight” color match fluorescent tube	6 500
Carbon arc white flame	5 000
Xenon arc (unfiltered)	6 000
Summer sunlight (before 9.00 or after 15.00 h)	4 900 to 5 600
Summer sunlight (9.00 to 15.00 h)	5 400 to 5 700
Direct sun	5 700 to 6 500
Overcast daylight	6 500 to 7 200
Cloudy blue sky to cloudless blue sky	8 000 to 27 000
Equal energy point	5 500 K

The ranges of color temperatures of man-made light sources (incandescent lamp and wax candle) as well as natural light sources (sun and moon under different conditions) are illustrated in **Figure 25.10**. Achromatic white light (also called ‘pure’ white light) is located close to the equal-energy point in the chromaticity diagram. Achromatic white has no hue and is used as the white point of color displays. It may be noted that LEDs can generate light of any color temperature by mixing the light from multiple LED chips or by the use of phosphors.

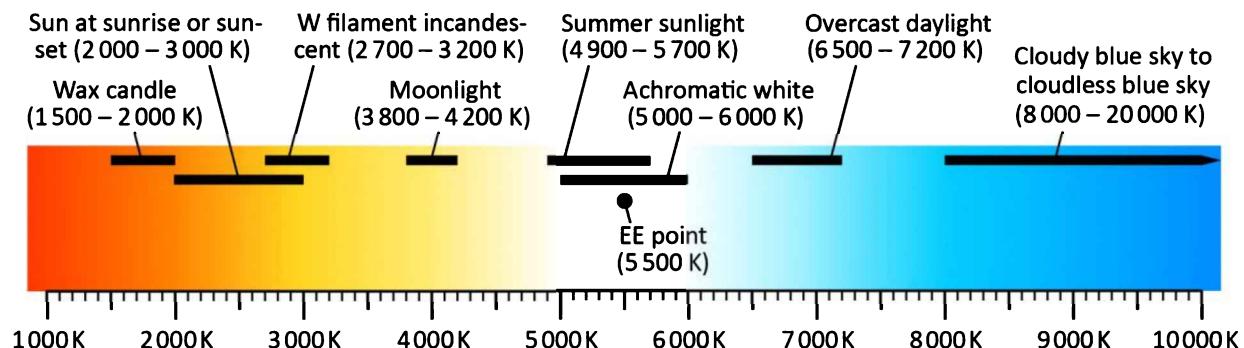


Fig. 25.10: Color temperatures of natural and man-made light sources. The equal-energy point is denoted as ‘EE point’.

25.6 Standardization of white light by means of color-temperature binning

The definition of “white light” is guided by natural white light such as sunlight (under various conditions). As pointed out above in this chapter, sunlight resembles black-body radiation and therefore has a chromaticity on or close to the planckian locus.

What is the range of color temperatures that qualifies as “white light”? A narrower range of color temperatures qualifying for white light is 2500 K to 6500 K. This range is generally acceptable to those working in the field of lighting. A wider range of color temperatures is about 1750 K to 8000 K, applicable when identifying the broadest possible range. Note that in comparison with neutral white (achromatic), 1750 K and 8000 K is orangish and bluish, respectively. However, given chromatic adaptation, this broad range can still qualify as white light.

What is the distance from the planckian locus that qualifies as “white light”? A narrower range of distances is ± 0.03 in the (x, y) chromaticity diagram. A wider range of distances is about ± 0.05 . Whereas the narrower range is generally acceptable to those working in the field of lighting, the wider range may be applied when identifying the broadest possible range.

Given these general boundaries for white light, the American National Standards Institute (ANSI) defined a white-color-bin standard for LEDs that is shown in *Figure 25.11*. The standard is known as the ANSI Standard C78.377-2008 (ANSI, 2008). The center points of the bins are oriented along a line that is quasi-parallel to the planckian locus and intersects the CIE Illuminant A and D65 chromaticity points. The bins are identified according to their correlated color temperature (CCT) and are therefore referred as “**CCT bins**”. A total of eight bins are defined with the CCTs of the bins being 2 700 K, 3 000 K, 3 500 K, 4 000 K, 4 500 K, 5 000 K, 5 700 K, and 6 500 K. The sizes of the ANSI bins are close to 7-step Macadam ellipses; this implies that color differences within one bin are recognizable.

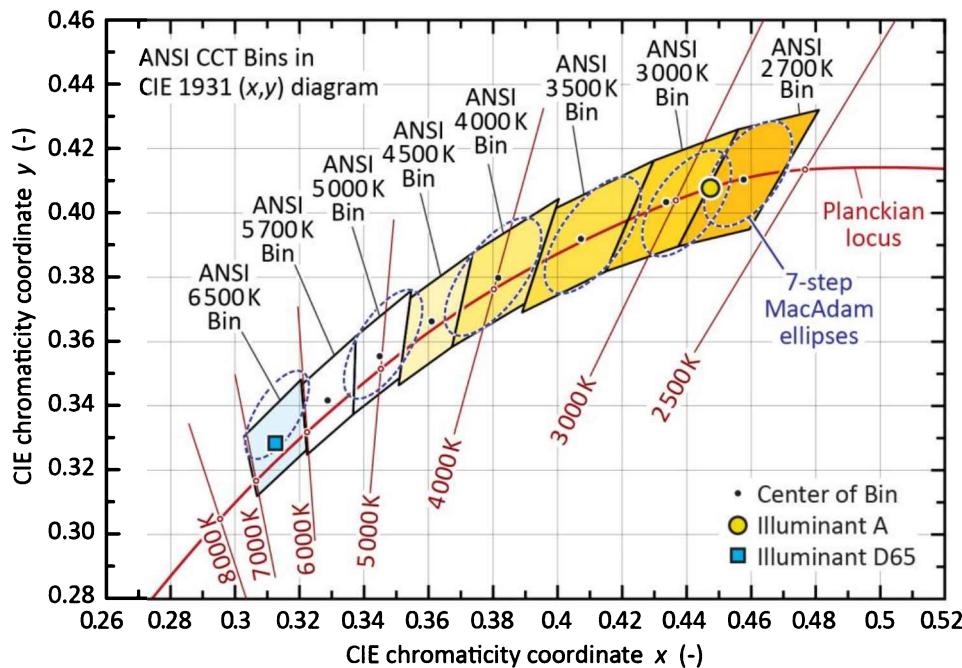


Fig. 25.11: ANSI bins of white LEDs organized according to correlated color temperature (CCT). The center points of the quadrilateral CCT bins are chosen to align with Illuminant A and D65. The size of the CCT bins is chosen to be comparable to 7-step MacAdam ellipses. The total number of CCT bins is eight (after ANSI, 2008).

A finer CCT binning, based on 3-step Macadam ellipses, is shown in *Figure 25.12* (Osram, 2008). The “**CCT fine bins**” are obtained by further developing the 2008 ANSI standard. That is, the CCT fine bins are obtained by sub-dividing and expanding the ANSI CCT bins along directions away from the planckian locus. A total of 120 CCT fine bins are defined by the combination of one letter and one number. For example, the “H6 bin” is neutral (achromatic) white, has a CCT of slightly more than 5000 K, and is located above yet close to the planckian locus, as can be verified by inspecting *Figure 25.12*.

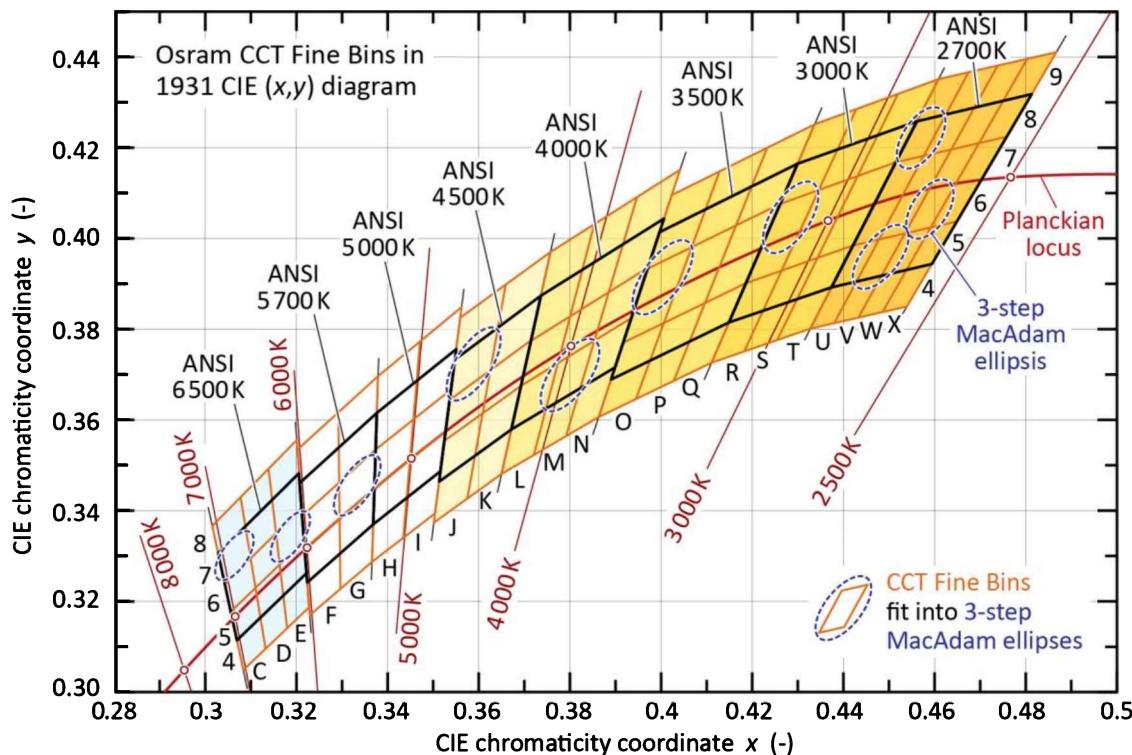


Fig. 25.12: Correlated-color-temperature (CCT) fine bins for white LEDs based on 3-step MacAdam ellipses. The CCT fine bins are obtained by sub-dividing and expanding the 2008 ANSI CCT bins. The total number of the CCT fine bins is 120 (after Osram Company, 2008).

25.7 References

- ANSI “Specifications for the chromaticity of solid state lighting products” ANSI Standard C78.377-2008; American National Standard Institute, Washington, D.C. (2008)
- CIE publication No. 17.4 *International Lighting Vocabulary*, see <http://www.cie.co.at> (CIE, Vienna, Austria, 1987)
- Duggal A. R. “Organic electroluminescent devices for solid-state lighting” in *Organic Electroluminescence* edited by Z. H. Kafafi (Taylor & Francis Group, Boca Raton, Florida, 2005)
- Jackson J. D. *Classical Electrodynamics* (John Wiley and Sons, New York, 1975)
- Ohno Y. “Photometric standards” Chapter 3 in *OSA/AIP Handbook of Applied Photometry*, 55 (Optical Society of America, Washington DC, 1997)
- Ohno Y. “Photometry and radiometry” Chapter 14 in *OSA Handbook of Optics, Volume III Review for Vision Optics, Part 2, Vision Optics* (McGraw-Hill, New York, 2001)
- Osram Company “Osram Opto creates 120 bins for white LEDs” *2008 Strategies in Light Technical Conference*, as published on Internet site <<http://www.ledsmagazine.com/articles/2008/02/osram-opto-creates-120-bins-for-white-leds.html>> (2008)
- Planck M. “On the theory of the law on energy distribution in the normal spectrum (translated from German)” *Verhandlungen der Deutschen Physikalischen Gesellschaft* 2, 237 (1900)
- Robertson R. “Computation of correlated color temperature and distribution temperature” *Journal of the Optical Society of America* 58, 1528 (1968)
- Schanda J. (Editor) *Colorimetry – Understanding the CIE system* (Wiley Interscience, Hoboken, New Jersey, 2007)
- Wikipedia <<http://www.wikipedia.org/>> under key word “Solar spectrum” (accessed 2013)

Appendix 25.1

Color temperature T and (x, y) and (u', v') chromaticity coordinates of a planckian emitter.

T	x	y	u'	v'
1 000 K	0.649	0.347	0.443	0.533
1 200 K	0.623	0.370	0.402	0.538
1 400 K	0.597	0.389	0.369	0.541
1 600 K	0.572	0.402	0.342	0.542
1 800 K	0.549	0.412	0.321	0.542
2 000 K	0.527	0.417	0.303	0.540
2 200 K	0.506	0.420	0.288	0.538
2 400 K	0.487	0.419	0.276	0.535
2 600 K	0.470	0.417	0.266	0.531
2 800 K	0.454	0.414	0.257	0.528
3 000 K	0.439	0.409	0.250	0.524
3 200 K	0.425	0.404	0.243	0.520
3 400 K	0.413	0.399	0.237	0.516
3 600 K	0.402	0.393	0.233	0.512
3 800 K	0.392	0.388	0.228	0.508
4 000 K	0.383	0.382	0.225	0.504
4 200 K	0.374	0.376	0.221	0.500
4 400 K	0.367	0.371	0.218	0.497
4 600 K	0.360	0.366	0.216	0.494
4 800 K	0.353	0.361	0.213	0.490
5 000 K	0.347	0.356	0.211	0.487
5 200 K	0.342	0.351	0.209	0.484
5 400 K	0.337	0.347	0.208	0.481
5 600 K	0.332	0.343	0.206	0.479
5 800 K	0.328	0.339	0.205	0.476
6 000 K	0.324	0.335	0.203	0.473
6 200 K	0.321	0.332	0.202	0.471
6 400 K	0.317	0.328	0.201	0.469
6 500 K	0.315	0.327	0.201	0.468
6 600 K	0.314	0.325	0.200	0.466
6 800 K	0.311	0.322	0.199	0.464
7 000 K	0.308	0.319	0.198	0.462
7 200 K	0.306	0.317	0.198	0.460
7 400 K	0.303	0.314	0.197	0.459
7 600 K	0.301	0.312	0.196	0.457
7 800 K	0.299	0.309	0.196	0.455
8 000 K	0.297	0.307	0.195	0.454
8 500 K	0.292	0.301	0.194	0.450
9 000 K	0.289	0.297	0.193	0.447
9 500 K	0.285	0.293	0.192	0.444
10 000 K	0.282	0.290	0.191	0.441

26

26 – Color mixing

There are two types of color mixing: Subtractive color mixing, which applies to the mixing of paint, and additive color mixing, which applies to the mixing of light. This chapter is mostly concerned with additive color mixing.

26.1 Subtractive color mixing

The mixing of two or more types of paint, generally referred to as *subtractive color mixing*, allows one to create paint having a new color. A schematic of color mixing for magenta, yellow, and cyan paint is shown in **Figure 26.1**.

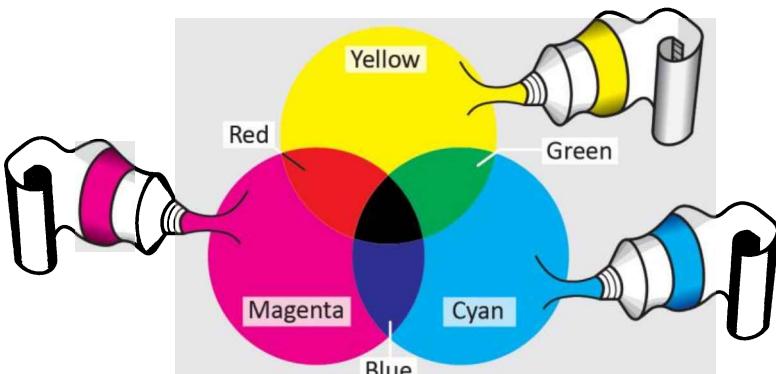


Fig. 26.1: Subtractive color mixing of three primary colors yellow, magenta, and cyan. Subtractive color mixing applies to the mixing of paint. Note that the mixture of the three colors yields a black paint.

The figure reveals that the mixing of two colors gives a new color, e.g. the mixing of yellow and cyan gives green. The figure also reveals that the mixing of all three colors gives black. How can this be understood? Each of the three paints contains pigments that absorb light over nearly the entire spectrum except for a certain wavelength range (e.g. cyan is non-reflective from the blue to the green range). When the three colors are mixed, the mixture will contain pigments that, collectively, absorb light at any wavelength. As a result, the mixed paint is black.

26.2 Additive color mixing

The mixing of two or more types of light, generally referred to as *additive color mixing*, allows one to create light having a new color. A schematic of color mixing for red, green, and blue light is shown in **Figure 26.2**.

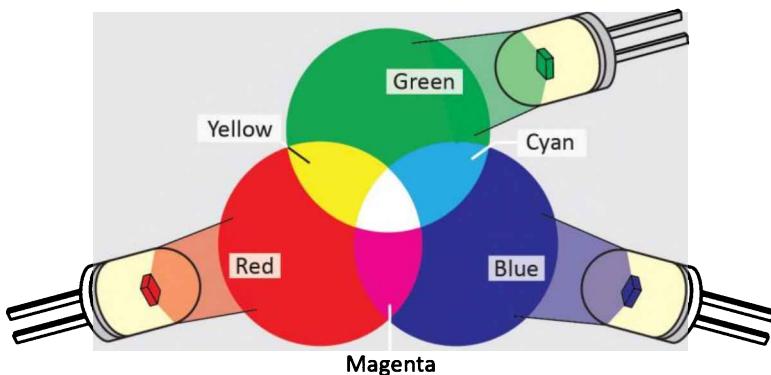


Fig. 26.2: Additive color mixing of three primary colors red, green, and blue. Additive color mixing applies to the mixing of light. Note that the mixture of the three colors yields white light.

The mixing of two or more light sources is employed in a number of applications. For example, in LED displays, three different types of LEDs, usually emitting in the red, green, and blue, are used. The three colors are mixed so that the observer can see a wide range of colors. Another application of color mixing is the generation of white light by the mixing of multiple colors, usually two complementary colors (such as blue and yellow), three colors (such as red, green, and blue), or more than three colors.

A photograph of an experiment visually showing the mixing of light coming from a group of LEDs is shown in **Figure 26.3** (Schubert, 2006). The differently colored LEDs are directed towards a center region where the light is white.

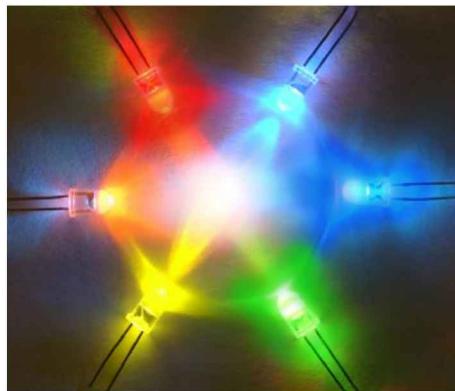


Fig. 26.3: Photograph of additive color mixing of light by using LEDs (Schubert, 2006).

26.3 Calculation of color mixing

In this section, we determine the chromaticity coordinates of the mixture of three discrete emission bands. Assume that the three emission bands have spectral power densities $P_1(\lambda)$, $P_2(\lambda)$, and $P_3(\lambda)$ with peak wavelengths of λ_1 , λ_2 , λ_3 , respectively. We assume that each emission band is much narrower than any of the three color-matching functions. We further assume that the three light sources have the chromaticity coordinates (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) . Then the tristimulus values are given by

$$\begin{aligned}
X &= \int_{\lambda} \bar{x}(\lambda) P_1(\lambda) d\lambda + \int_{\lambda} \bar{x}(\lambda) P_2(\lambda) d\lambda + \int_{\lambda} \bar{x}(\lambda) P_3(\lambda) d\lambda \\
&\approx \bar{x}(\lambda_1) P_1 + \bar{x}(\lambda_2) P_2 + \bar{x}(\lambda_3) P_3
\end{aligned} \tag{26.1}$$

$$\begin{aligned}
Y &= \int_{\lambda} \bar{y}(\lambda) P_1(\lambda) d\lambda + \int_{\lambda} \bar{y}(\lambda) P_2(\lambda) d\lambda + \int_{\lambda} \bar{y}(\lambda) P_3(\lambda) d\lambda \\
&\approx \bar{y}(\lambda_1) P_1 + \bar{y}(\lambda_2) P_2 + \bar{y}(\lambda_3) P_3
\end{aligned} \tag{26.2}$$

$$\begin{aligned}
Z &= \int_{\lambda} \bar{z}(\lambda) P_1(\lambda) d\lambda + \int_{\lambda} \bar{z}(\lambda) P_2(\lambda) d\lambda + \int_{\lambda} \bar{z}(\lambda) P_3(\lambda) d\lambda \\
&\approx \bar{z}(\lambda_1) P_1 + \bar{z}(\lambda_2) P_2 + \bar{z}(\lambda_3) P_3
\end{aligned} \tag{26.3}$$

where P_1 , P_2 , and P_3 are the optical powers emitted by the three sources. Using the abbreviations

$$L_1 = \bar{x}(\lambda_1) P_1 + \bar{y}(\lambda_1) P_1 + \bar{z}(\lambda_1) P_1 \tag{26.4}$$

$$L_2 = \bar{x}(\lambda_2) P_2 + \bar{y}(\lambda_2) P_2 + \bar{z}(\lambda_2) P_2 \tag{26.5}$$

$$L_3 = \bar{x}(\lambda_3) P_3 + \bar{y}(\lambda_3) P_3 + \bar{z}(\lambda_3) P_3 \tag{26.6}$$

the chromaticity coordinates of the mixed light can be calculated from the tristimulus values to yield

$$x = \frac{x_1 L_1 + x_2 L_2 + x_3 L_3}{L_1 + L_2 + L_3} \tag{26.7}$$

$$y = \frac{y_1 L_1 + y_2 L_2 + y_3 L_3}{L_1 + L_2 + L_3} . \tag{26.8}$$

Thus, the chromaticity coordinate of the multi-component light is a linear combination of the individual chromaticity coordinates weighted by the L_i factors.

The principle of color mixing in the chromaticity diagram is shown in **Figure 26.4**. The figure shows the mixing of two colors with chromaticity coordinates (x_1, y_1) and (x_2, y_2) . For the case of two colors, $L_3 = P_3 = 0$. The mixed color will be located on the straight line connecting the chromaticity coordinates of the two light sources. Thus, any color (including white) located between the two chromaticity points can be created by mixing the two colors.

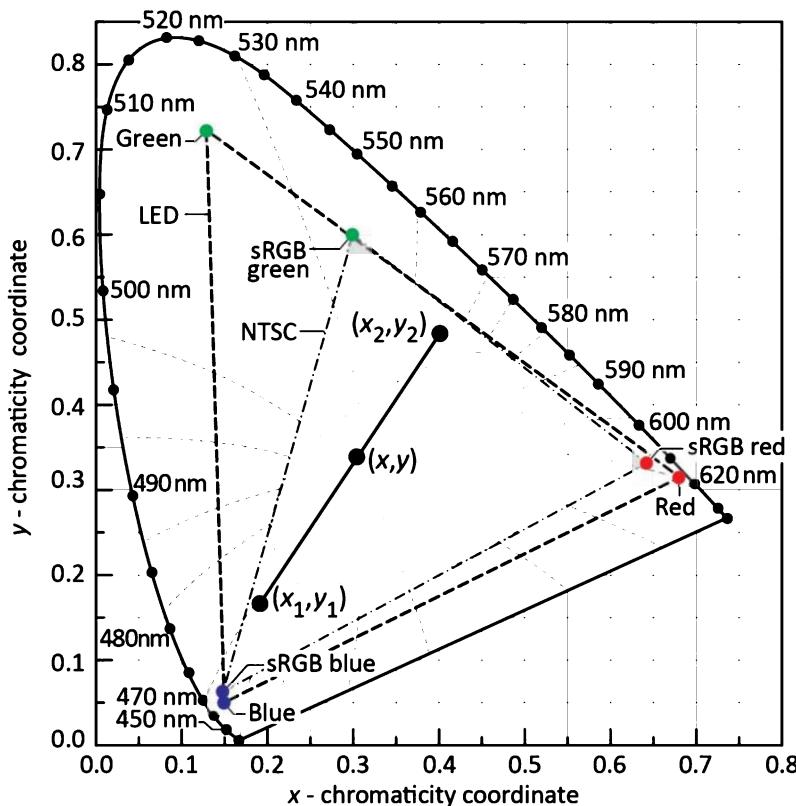


Fig. 26.4: Principle of color mixing illustrated with two light sources with chromaticity coordinates (x_1, y_1) and (x_2, y_2) . The resulting color has the coordinates (x, y) . Also shown is the triangular area of the chromaticity diagram (color gamut) accessible by additive mixing of a red, green, and blue LED. The locations of the red, green, and blue phosphors of the sRGB display standard ($x_r = 0.64$, $y_r = 0.33$, $x_g = 0.30$, $y_g = 0.60$, $x_b = 0.15$, $y_b = 0.06$) are also shown. The sRGB standard is similar to the NTSC standard.

Figure 26.4 also shows the mixing of *three* colors, located in the red, green, and blue regions of the chromaticity diagram. The three chromaticity points, connected by a dashed line, are typical points for red, green, and blue LEDs. The area located within the dashed line, called the **color gamut**, represents all colors that can be created by mixing the three primary colors red, green, and blue. The ability to create a great variety of colors is an important quality for displays. It is desirable that the color gamut provided by the three light sources is as large as possible to create displays able to show brilliant, saturated colors.

The color gamut represents the entire range of colors that can be created from a set of primary sources. Color gamuts are polygons positioned within the perimeter of the chromaticity diagram. For the case of *three* primary colors, the color gamut is a *triangle*, as shown in **Figure 26.4**. All colors created by additive mixtures of the vertex points (primary colors) of a gamut, are necessarily located inside the gamut.

The insight now gained on color mixing allows one to understand the location of different LEDs in the chromaticity diagram. The perimeter of the chromaticity diagram in the red spectral region is approximately a *straight line*, so that red LEDs, despite their thermal broadening, are located directly on the perimeter of the chromaticity diagram (or very closely to the perimeter of the chromaticity diagram). In contrast, the perimeter is strongly curved in the green region, so

that green LEDs, due to their spectral broadening, are displaced from the perimeter towards the center of the chromaticity diagram. For further discussion of color mixing, the reader is referred to Wyszecki and Stiles (2000).

26.4 Implementation of color mixing

It would be undesirable for a white light source to emit bluish white light along one direction and yellowish white light along another direction. Such white LED would have an incomplete mixture of colors. Therefore, *complete mixture* of light is highly desirable. If light coming from multiple sources is *completely* mixed, then the emitted color is independent of the emission direction. How can such complete mixture be accomplished? It can be accomplished by (i) co-locating the differently colored sources and (ii) completely randomizing their light propagation direction, i.e. by diffusing the light. The randomization of the light-propagation direction can be accomplished by (1) light-diffusing volumes or (2) light-diffusing surfaces.

Light-diffusing volumes

Light can be diffused (scattered) in a volume of a material if this volume contains ***light-scattering particles*** (or light-diffusing particles). The light-scattering particles should be randomly distributed and located in a host material such as a transparent encapsulation material. In order for the particles to scatter light, the refractive indices of the particles (e.g. SiO₂) and of the transparent encapsulation material (e.g. silicone) must be different. The higher the difference in refractive index, the more efficiently the particles will scatter the light. SiO₂ (silicon dioxide, also called glass), Al₂O₃ (aluminum oxide, also called alumina), TiO₂ (titanium dioxide, also called titania), CaF₂ (calcium fluoride), CaCO₃ (calcium carbonate), BaTiO₃ (barium titanate), and BaSO₄ (barium sulfate) are suitable materials for light-scattering particles (Haitz, 1973; Haitz *et al.*, 1995; Shimizu *et al.*, 1999; Reeh *et al.*, 2006). Light-scattering particles are also called “light-diffusing particles”, “light-diffusion agents”, “diffusants”, or “mineral diffusers”. The light-scattering particles are mostly oxides; oxides are very stable chemical compounds, have a large bandgap energy, and thus are transparent.

The addition of light-scattering particles results in a device of uniformly appearance that does not reveal the point-like nature of an LED chip. As a result, the transparency of the encapsulation material changes from specular transparency to diffuse transparency. Since the 1960s (i.e. since the infancy of commercial LED technology), to accomplish a uniform appearance of LED indicator-

lights, light-scattering particles have been added to the encapsulation material. In the 1960s, the light-scattering particles were mostly ground glass (SiO_2).

An LED display package employing light-diffusing particles is shown in *Figure 26.5 (a)* (Haitz, 1973). Such packages are used for alphanumeric displays, such as 7-, 8-, 14-, or 16-segment displays, as shown in *Figure 26.5 (b)* (www.sparkfun.com, 2013). The light-diffusing particles will cause the surface brightness of each segment to be uniform.

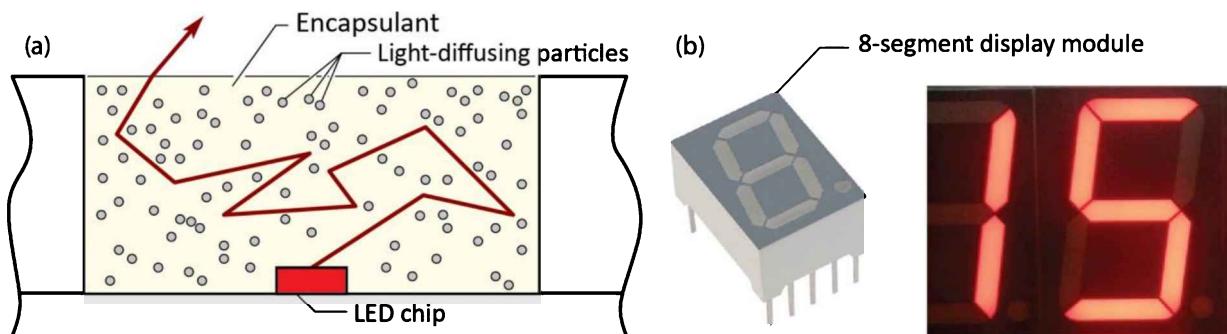


Fig. 26.5: (a) LED display package having an encapsulant filled with light-diffusing particles. Light-diffusing particles may consist of optically transparent materials such as SiO_2 , TiO_2 , CaF_2 , CaCO_3 , or BaSO_4 (after Haitz, 1973). (b) Examples of 8-segment LED display modules having uniform surface brightness (after www.sparkfun.com, 2013).

A comparison of 5-mm-radial LED packages with and without an optical diffusant distributed within the encapsulant body is shown in *Figure 26.6*. Also shown are the associated emission patterns. The emission pattern of the device having the clear encapsulant is determined by ray optics and depends on the shape of the reflector cup and encapsulant body. The emission pattern of the device having an optical diffusant is strongly influenced by the diffusant. In the limit of strong light diffusion, all directions have equal emission probability.

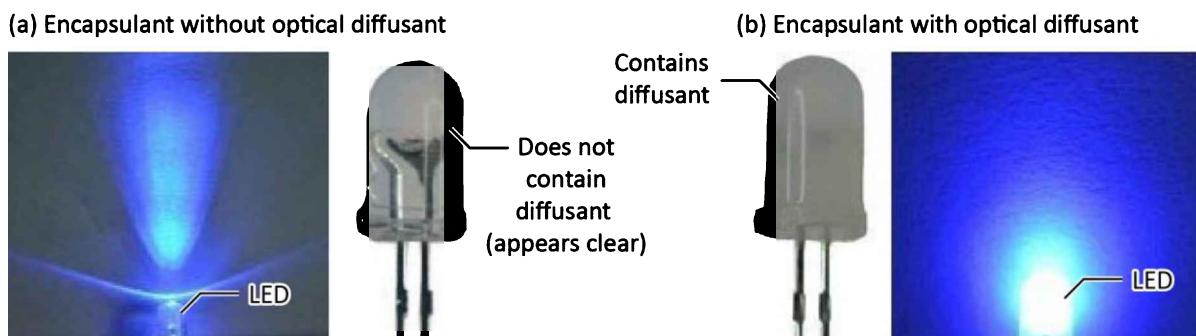


Fig. 26.6: LED package (a) not containing and (b) containing an optical diffusant (light-diffusing material) in encapsulant and associated emission patterns.

Light-scattering particles are also used in white LEDs that employ phosphor-based wavelength conversion. Such LEDs frequently consist of a blue LED chip coated with a yellow-emitting phosphor. In order to thoroughly mix the blue and yellow light, light-scattering particles can be used to attain a uniform spatial distribution of the emission color (see, for example, Shimizu *et al.*, 1999, 2009; Reeh *et al.*, 2006). Note that the diffusion of light in phosphor-based LEDs is in part accomplished by phosphor particles but can be complemented by the addition of light-scattering particles.

The optical scattering of light can be classified into three categories that represent the physical origin of the scattering process:

Geometric-optics scattering is relevant for particle sizes $D_{50} \gg \lambda$ (where D_{50} is the median particle diameter). Irregularly shaped particles with diameters much greater than the wavelength of light (λ) refract light (as long as their refractive index differs from the one of their host material). For irregularly shaped particles, the direction of the scattered light is irregular and, after a sufficient number of refraction events, completely randomized. Optical scattering by a particle can be described by means of the **scattering cross section** (*SCS*) of the particle, which is a hypothetical area that, when light is incident within the *SCS*, scatters the light. For geometric-optics scattering, $SCS = (D_{50}/2)^2 \pi$, i.e. the *SCS* is equal to the cross-sectional area of the particle. When decreasing D_{50} (while keeping the mass density of the scattering substance constant), the number of particles (N) increases ($N \propto (D_{50}/2)^{-3}$) but their *SCS* decreases ($SCS \propto (D_{50}/2)^2$) so that the **scattering strength** ($\propto N \times SCS \propto D_{50}^{-1}$) increases. Therefore, in the geometric-optics regime, a larger number of smaller particles will scatter light more effectively.

Mie scattering is relevant for particle sizes $D_{50} \approx \lambda$. For Mie scattering, the *SCS* of a scattering particle is a constant, i.e. independent of D_{50} and independent of λ . When decreasing D_{50} (while keeping the mass density of the scattering substance constant), the number of particles increases ($N \propto (D_{50}/2)^{-3}$) so that the scattering strength ($\propto N \times SCS \propto D_{50}^{-3}$) increases. Therefore, in the Mie regime, a larger number of smaller particles will scatter light more effectively.

Rayleigh scattering is relevant for particle sizes $D_{50} \ll \lambda$. For Rayleigh scattering, the *SCS* of a scattering particle depends on λ and obeys the proportionality $SCS \propto (D_{50}/\lambda)^4$. This implies that short-wavelength light is more strongly scattered than long-wavelength light. Furthermore, when decreasing D_{50} (while keeping the mass density of the scattering substance constant), the number of particles increases ($N \propto (D_{50}/2)^{-3}$) but their *SCS* decreases ($\propto (D_{50}/2)^4$) so that the scattering strength ($\propto N \times SCS \propto D_{50}$) decreases.

Therefore, in the Rayleigh regime, a larger number of smaller particles will scatter light less effectively. That is, in the limit of small particles (i.e. nano-particles), Rayleigh scattering becomes negligible. Accordingly, the decrease of D_{50} to values $\ll \lambda$ would be detrimental for the effective scattering of light.

To conclude, for geometric-optics scattering and Mie scattering, light scattering increases as D_{50} decreases (assuming that the mass density of the light-scattering substance, measured in g/cm³, is kept constant so that the number of particles increases when D_{50} decreases). In contrast, for Rayleigh scattering, light scattering decreases as D_{50} decreases and becomes negligible for nano-sized particles for which $D_{50} \ll \lambda$ (again assuming that the mass density of the light-scattering substance is kept constant).

The **scattering length** has been used in the context of light-scattering particles as well as phosphors (Bechtel *et al.*, 2010). The scattering length is the distance that the light travels before its propagation direction is randomized by scattering.

Light-diffusing surfaces

Surfaces can transmit light in a **specular** or a **diffuse manner**. Specular surfaces follow Snell's law of refraction and are thus deterministic. Diffuse surfaces diffuse transmitted light, thereby mixing transmitted light. Diffusive surfaces are rough, random, and thus non-deterministic. Such surfaces are frequently called "light-diffusing surface", "frosted-glass surface", or "opaque-glass surface". Specular and diffuse surfaces are contrasted in **Figure 26.7** (photographs after www.eShowerDoor.com, 2013).

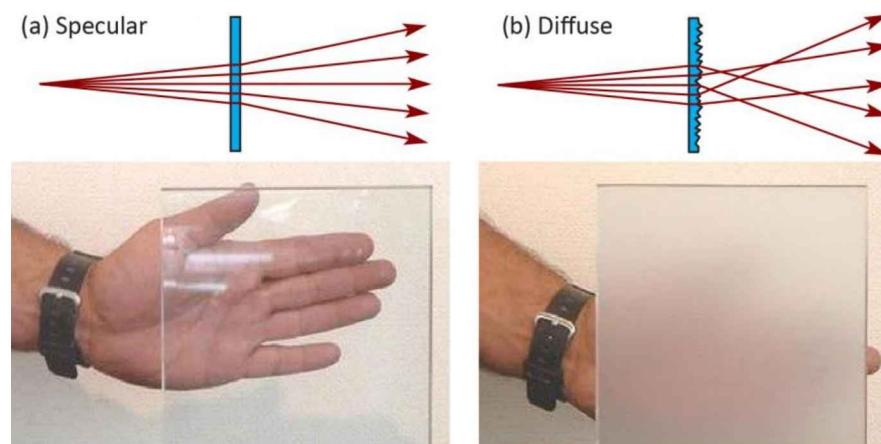


Fig. 26.7: (a) Specular and (b) diffuse transparency. Specular transparency is deterministic so that information (such as the contour of a hand) is not lost during the transmittance. In contrast, diffuse transparency is non-deterministic so that information is lost. (photos after www.eShowerDoor.com, 2013).

A light-diffusing glass envelope (“bulb”) of an LED light bulb is shown in *Figure 26.8*. The light-diffusing bulb makes the emission pattern more uniform and enhances the mixing of the different components of the light emission (i.e. LED chip emission and phosphor emission).

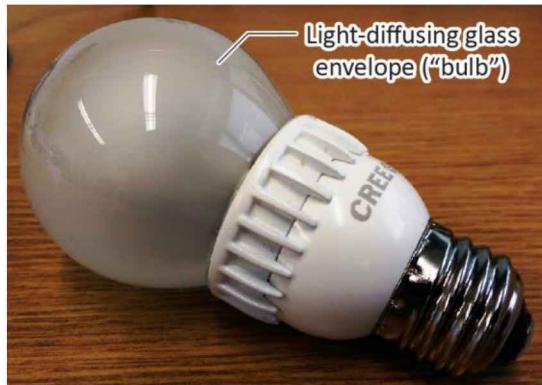


Fig. 26.8: LED light bulb having a light-diffusing glass envelope (“bulb”). The bulb has an opaque-glass appearance that does not allow one to see through the glass.

26.5 References

- Bechtel H., Busselt W., and Schmidt P. “Electroluminescence device” US Patent 7,800,123 (2010)
 Haitz R. H. “Electroluminescent semiconductor display apparatus and method of fabricating the same” US Patent 3,780,357 (1973)
 Haitz R. H., Craford M. G., and Weissman R. H. “Light-emitting diodes” Chapter 12 in “Handbook of Optics” edited by Michael Bass (McGraw-Hill, New York, 1995)
 Reeh U., Hohn K., Stath N., Waitl G., Schlotter P., Schneider J., and Schmidt R. “Light-radiating semiconductor component with a luminescence conversion element” US Patent 7,151,283 (2006)
 Schubert E. F. “Light-Emitting Diodes” 2nd edition (Cambridge University Press, Cambridge UK, 2006)
 Shimizu Y., Sakano K., Noguchi Y., and Moriguchi T. “Light-emitting device having a nitride compound semiconductor and a phosphor containing a garnet fluorescent material” US Patent 5,998,925 (1999)
 Shimizu Y., Sakano K., Noguchi Y., and Moriguchi T. “Light-emitting device with blue light LED and phosphor components” US Patent 7,531,960 (2009)
www.eShowerDoor.com, accessed July 2013 (2013)
www.sparkfun.com, accessed July 2013 (2013)
 Wyszecki G. and Stiles W. S. “Color Science – Concepts and Methods, Quantitative Data and Formulae” 2nd edition “Wiley Classics Library” (John Wiley and Sons, New York, 2000)

27

27 – Color rendering

27.1 Color rendering of physical objects

An important characteristic of a white light source is its ability to show (i.e. render) the true colors of physical objects, e.g. fruits, plants, or toys, that are being illuminated by the source. The ability to render the colors of an object is measured in terms of the **color-rendering index** or **CRI** (Wyszecki and Stiles, 1982, 2000; MacAdam, 1993; Berger-Schunn, 1994; CIE, 1974, 1995, 1999, 2007). It is a measure of the ability of the *illuminant* (i.e. a white illumination source) to accurately render the colors of physical objects illuminated by the source.

Figure 27.1 shows an example of a physical object (here a painting by the expressionist Franz Marc) under illumination with a high-CRI source and with a low-CRI source. Colors appear richer and more vivid under illumination with a high-CRI source. Whereas high color rendering is important in locations such as museums, homes, and offices, it is less so in locations such as streets and parking lots. The color-rendering index is irrelevant for white light sources used in indicator lamp and signage applications.



Fig. 27.1: Artwork titled “Blue horse” illuminated with (a) high-CRI illumination source and (b) low-CRI illumination source (Franz Marc, German expressionist, 1880–1916).

The color-rendering ability of a **test light source** (“test source”) is evaluated by comparing it with the color-rendering ability of a **reference light source** (“reference source”). For the calculation of the **CRI**, the reference source is chosen as follows (CIE, 1974, 1995, 1999, 2007): (i) If the chromaticity point of the test source is located *on* the planckian locus, the reference source is a

planckian black-body radiator with the same color temperature as the test source. (ii) If the chromaticity point of the test source is located *off* the planckian locus, the reference source is a planckian black-body radiator with the same *correlated* color temperature as the test source. (iii) Alternatively, for test sources with a color temperature exceeding 5 000 K, one of the standardized CIE illuminants (e.g. Illuminant D₆₅) can be used as a reference source (CIE, 1995, 1999, 2007). Ideally, the test source and reference source have the same chromaticity coordinates.

By convention, the planckian black-body reference source is assumed to have perfect color-rendering properties and thus its color-rendering index is *CRI* = 100. This convention was agreed upon because natural daylight closely resembles a planckian black-body source and thus rightfully deserves to be established as the standard reference source. Illuminants other than the reference source necessarily have a color-rendering index lower than 100. Because the *CRI* depends sensitively on the choice of the reference source, the selection of the reference source is of critical importance when calculating the *CRI*s of test sources.

Because the emission spectrum of an incandescent lamp closely follows that of a planckian black-body radiator, such lamps have the highest possible *CRI* and thus the best color-rendering properties of all artificial light sources. Incandescent quartz-halogen lamps have been used in locations where color rendering is of prime importance, such as in museums, art galleries, and clothing shops. The drawback of quartz-halogen lamps is high power consumption.

In addition to the test source and the reference source, ***test-color samples*** are instrumental in determining the *CRI* of a test source. Test-color samples can be derived from real objects, e.g. fruit, flowers, wood, furniture, and clothes. However, in the interest of international standardization, a specific set of 14 test-color samples has become an agreed-upon standard. These 14 test-color samples are a subset of a larger collection of test-color samples initially introduced by Albert H. Munsell, a university professor who taught at Rochester Institute of Technology (Rochester, NY) in the late 1800s and early 1900s (Munsell, 1905; 2005; Billmeyer, 1987; Long and Luke, 2001). Munsell introduced a color notation – the ***Munsell color system*** – which is a notation for defining a very wide range of colors.

The *CRI* calculation has been discussed in detail by Wyszecki and Stiles (1982; 2000) and by CIE (1974, 1995, 1999, 2007). The ***CIE general CRI*** is an average calculated according to

$$CRI_{general} = \frac{1}{8} \sum_{i=1}^8 CRI_i \quad (27.1)$$

where the CRI_i are the **special CRIs** for a set of eight test-color samples. The special color-rendering indices are calculated according to

$$CRI_i = 100 - 4.6 \Delta E_i^* \quad (27.2)$$

where ΔE_i^* represents the quantitative color change that occurs when a test-color sample is illuminated with, first, the reference source, and subsequently with the test source. The special color-rendering indices are calculated in such a way that they have a value of 100 if there is no difference in color. The quantitative color change, ΔE_i^* , plays a key role in the calculation of the CRI and the determination of ΔE_i^* will be discussed in subsequent sections of this chapter where we will differentiate between on-planckian-locus and off-planckian-locus test sources.

At the time Eqn. (27.2) was established, the pre-factor 4.6 was chosen in such a way that the general CRI equals 60 when a “standard warm white” fluorescent lamp was used as a test source and a planckian black-body radiator was used as a reference source. Current fluorescent light sources have higher CRI s, typically in the range 60~85 (Kendall and Scholand, 2001).

The CIE test-color samples (TCS) are defined in terms of their spectral reflectivity. The colors of the first eight CIE test-color samples are shown in **Figure 27.2**. The first group, (TCS-01 to TCS-08), is used to calculate the general CRI . The second group, (TCS-09 to TCS-14), is used to calculate the special CRI s. As can be noticed, whereas TCS-01 to TCS-08 are weakly saturated colors, TCS-09 to TCS-12 are strongly saturated colors. TCS-13 is light yellowish pink (originally called the complexion of a person of European decent). TCS-14 represents a tree leaf. TCS-15, which is optional, represents white.



Fig. 27.2: CIE test-color samples (TCSs) used to calculate the CRI. TCS-01 to 08 have low-saturation colors (pastel colors) used for calculation of the general CRI, $CRI_{general}$. TCS-09 to 12 have high-saturation colors used for calculation of the special CRIs, CRI_i .

The spectral reflectivity curves of the test-color samples are shown in **Figure 27.3**. The numerical values of the reflectivity of the eight test-color samples are listed in Appendix 27.1. The general color-rendering index is calculated from these eight test-color samples ($i = 1 \sim 8$).

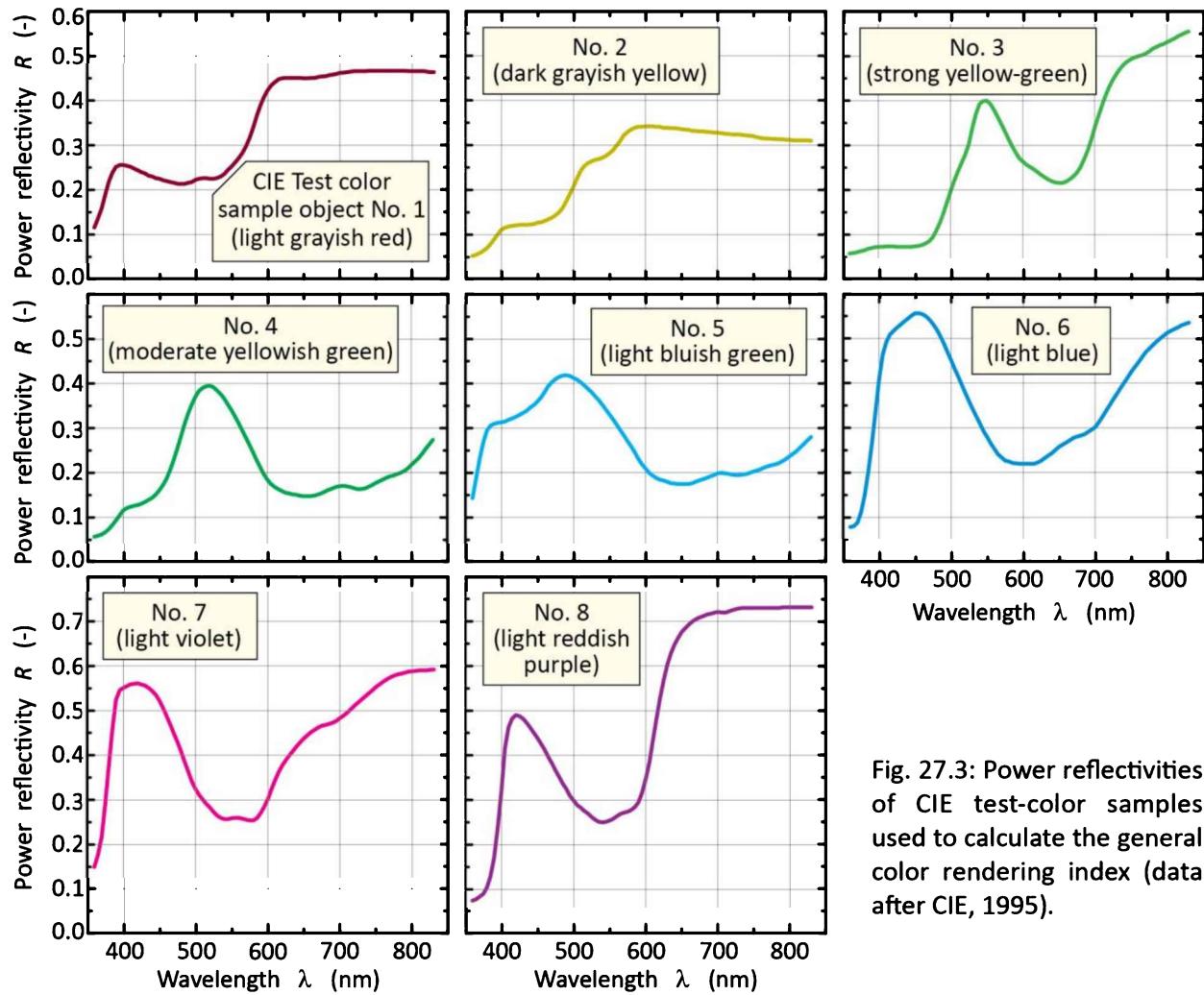
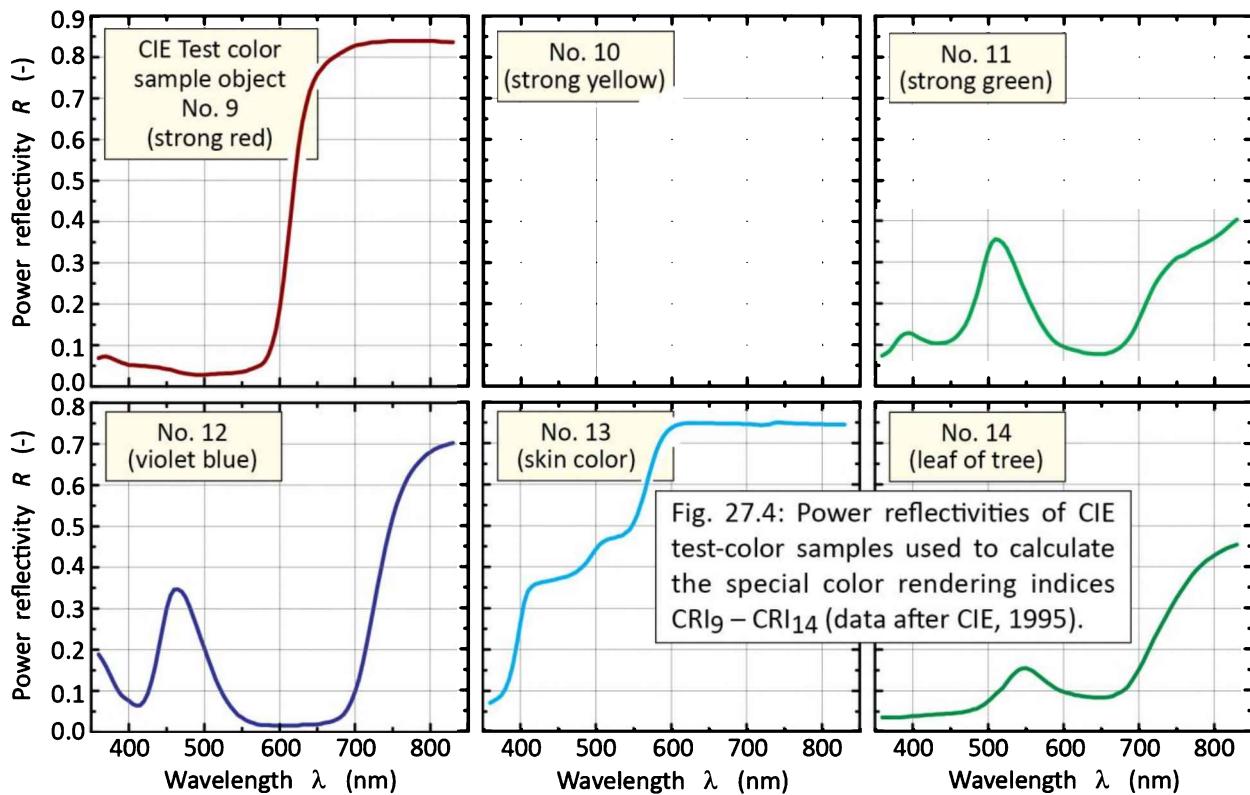


Fig. 27.3: Power reflectivities of CIE test-color samples used to calculate the general color rendering index (data after CIE, 1995).

In addition to the test-color samples (with numbers 1~8) used to calculate the *general* color rendering index, six supplemental test-color samples (with numbers 9~14) are used to further assess the color rendering capabilities of test sources. These supplemental test-color samples have the following colors: 9 – strong red; 10 – strong yellow; 11 – strong green; 12 – strong purplish blue; 13 – yellowish pink (originally called the complexion of a person of European decent); 14 – leaf of tree. The reflectivity spectra and the numerical values of the reflectivity of the supplemental test-color samples are given in **Figure 27.4** and in Appendix 27.2, respectively. Inspection of the reflectivity curves reveals that the colors of the test-color samples 9~14 have particularly strong colors with relatively narrow peaks. CRI_9 to CRI_{14} are referred to as the *special color-rendering indices* 9~14.



An example of a color-rendering-index determination is shown in **Figure 27.5**. The figure shows the test-color samples when illuminated by (i) the reference source (upper rows) and (ii) the test source (lower rows). A comparison of the paired color patches reveals that the colors of the patches illuminated by the test source are slightly different from the colors of the patches when illuminated by the reference source (with TCS 09 showing the strongest difference among the pairs). The differences in color are used to calculate the general *CRI* as well as the special *CRI*s.

TCS colors rendered by:



Fig. 27.5: Example of TCS color rendition when TCSs are illuminated with (a, c) a reference source and (b, d) test source. A small color difference can be observed for several TCSs.

The meaning of *chromaticity difference* of a test and a reference illumination source and the rendered colors of a test-color sample, when illuminated with the test and reference illumination sources, is illustrated in **Figure 27.6**. In the example shown in the figure, the test source is located slightly off the planckian locus. The reference source is a planckian source with the least possible distance from the test-source chromaticity point. As a result, the color temperature of the reference source is equal to the correlated color temperature of the test source. The four chromaticity points shown in **Figure 27.6** enter the calculation of the *CRI*.

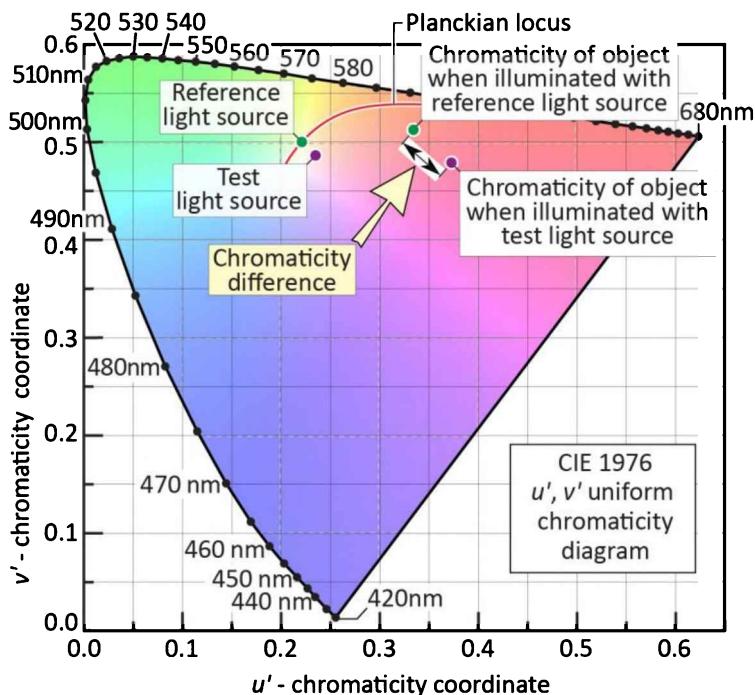


Fig. 27.6. Chromaticity difference resulting from the illumination of an object with a reference and a test light source. In the CIE 1976 u' , v' uniform chromaticity diagram, the chromaticity difference is directly proportional to the geometric distance. The reference light source is located on the planckian locus at the correlated color temperature of the test light source.

Note, however, that the term “color” as used by the CIE, is not equal to “chromaticity”. The broad CIE definition for color includes hue, saturation, and additionally, brightness (for light) or lightness (for physical objects). Whereas hue and saturation are fully defined by location in the chromaticity coordinate system, brightness and lightness are not.

27.2 CIE uniform color space

A **uniform color space** (CIE, 1986) is motivated by the need for a quantitative color space that includes chromaticity and brightness/lightness. This uniform color space provides direct proportionality between color difference and geometrical distance. Thus, a color difference is directly related to the geometric distance between two points in the uniform color space. The

CIE has introduced two three-dimensional uniform color spaces, namely the (L^*, u^*, v^*) and (L^*, a^*, b^*) spaces (CIE, 1986; Wyszecki and Stiles, 2000). For our purposes, it will be sufficient to consider the (L^*, u^*, v^*) uniform color space and we will therefore restrict our considerations to this space. The CIE (L^*, u^*, v^*) uniform color space is a three-dimensional space with Cartesian coordinates, two coordinates being associated primarily with the chromaticity (u^*, v^*) , and the third coordinate, L^* , representing the brightness (of a source) or lightness (of a physical object). The uniform color space is particularly suited for quantifying color differences.

To become familiar with the CIE (L^*, u^*, v^*) color space, we consider a test-color sample i illuminated with a white *reference source*. The coordinates, L^* , u^* , and v^* , which give the color difference between the object color and the white reference-source color, are defined as

$$L^*|_{ref} = 116 \left(Y_{ref,i} / Y_{ref} \right)^{1/3} - 16 \quad (27.3a)$$

$$u^*|_{ref} = 13 L^* (u_{ref,i} - u_{ref}) \quad (27.3b)$$

$$v^*|_{ref} = 13 L^* (v_{ref,i} - v_{ref}) \quad (27.3c)$$

where $Y_{ref,i}$, $u_{ref,i}$, and $v_{ref,i}$ describe the color stimulus of the test-color sample i when illuminated with the *reference source*, and Y_{ref} , u_{ref} , and v_{ref} describe the color stimulus of the reference source.

Next consider the test-color sample illuminated with a white *test source*. The coordinates, L^* , u^* , and v^* , which give the color difference between object color and test-source color, are then given by

$$L^*|_{test} = 116 \left(Y_{test,i} / Y_{test} \right)^{1/3} - 16 \quad (27.4a)$$

$$u^*|_{test} = 13 L^* (u_{test,i} - u_{test}) \quad (27.4b)$$

$$v^*|_{test} = 13 L^* (v_{test,i} - v_{test}) \quad (27.4c)$$

where $Y_{test,i}$, $u_{test,i}$, and $v_{test,i}$ describe the color stimulus of the test-color sample i when illuminated with the *test source*, and Y_{test} , u_{test} , and v_{test} describe the color of the test source.

The difference in color between the two points located in the (L^*, u^*, v^*) space given by Eqns. (27.3) and (27.4) is equal to the euclidean distance between the points, that is

$$\Delta E^* = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2} \quad (27.5)$$

where

$$\Delta L^* = L^*|_{test} - L^*|_{ref} \quad (27.6a)$$

$$\Delta u^* = u^*|_{test} - u^*|_{ref} \quad (27.6b)$$

$$\Delta v^* = v^*|_{test} - v^*|_{ref}. \quad (27.6c)$$

The difference in color, ΔE^* , forms the basis of the *CRI* calculation, as will be described in the subsequent sections.

An overview of the general color-rendering indices of common light sources is given in **Table 27.1**. The table includes several types of LED sources including dichromatic white LEDs, trichromatic white LEDs, and phosphor-based white LEDs. A *CRI* between 85 and 100 is suitable for virtually all illumination applications. A *CRI* between 70 and 85 is suitable for many illumination applications. Light sources with a *CRI* below 70 are considered to be of lower quality.

Table 27.1: General color-rendering indices (*CRI*s) of different light sources. (a) Using sunlight as reference source. (b) Using incandescent light with the same correlated color temperature as the reference source. (c) Using Illuminant D₆₅ as the reference source (some data after Kendall and Scholand, 2001).

Light source	Color-rendering index
Sunlight	100 (a)
Quartz halogen W filament incandescent light	100 (b)
W filament incandescent light	100 (b)
Fluorescent light	60~95 (b)
Trichromatic white LED	60~95 (b, c)
Tetrachromatic white LED	70~95 (b, c)
Phosphor-based LED	55~95 (b, c)
Broadened dichromatic white LED	10~60 (b, c)
Hg vapor light coated with phosphor	50 (b)
Hg vapor light	33 (b)
Low and high-pressure Na vapor light	10 and 22 (b)
Green monochromatic light	-50 (c)

27.3 Color-rendering index for planckian-locus illumination sources

The following calculation gives the value of the quantity ΔE_i^* , which, according to Eqn. (27.2), is needed to determine the *CRI* of a test source. The calculation is suited for test sources that are located on the planckian locus or extremely close to the planckian locus. The calculation described here follows Wyszecki and Stiles (1982, 2000) and CIE (1995). ΔE_i^* , which is the difference in appearance of a test-color sample when illuminated with the test and reference sources, is calculated according to

$$\Delta E_i^* = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2} \quad (27.7)$$

where

$$\Delta L^* = L_{ref}^* - L_{test}^* = \left[116 \left(\frac{Y_{ref,i}}{Y_{ref}} \right)^{1/3} - 16 \right] - \left[116 \left(\frac{Y_{test,i}}{Y_{ref}} \right)^{1/3} - 16 \right] \quad (27.8)$$

$$\Delta u^* = u_{ref}^* - u_{test}^* = 13 L_{ref}^* (u_{ref,i} - u_{ref}) - 13 L_{test}^* (u_{test,i} - u_{ref}) \quad (27.9)$$

$$\Delta v^* = v_{ref}^* - v_{test}^* = 13 L_{ref}^* (v_{ref,i} - v_{ref}) - 13 L_{test}^* (v_{test,i} - v_{ref}) \quad (27.10)$$

and

$$u = \frac{4X}{X + 15Y + 3Z} \quad \text{and} \quad v = \frac{6Y}{X + 15Y + 3Z} . \quad (27.11)$$

Note that u and v are calculated from the tristimulus values of the reference source spectrum (subscript “ref”), from the reference source spectrum reflected off the test-color samples (subscript “ref, i”), and from the test source spectrum reflected off the test-color samples (subscript “test, i”).

When calculating the *CRI* using the equations given above, the chromaticity coordinates and luminous flux of the test and reference sources should be identical in order to get the highest possible *CRI* for the test source. That is, the conditions $u_{test} = u_{ref}$, $v_{test} = v_{ref}$, and $Y_{test} = Y_{ref}$ should be satisfied.

The calculation of the color-rendering index in terms of Eqns. (27.1) ~ (27.11) illustrates that it is calculated from the ability of a test source to render the chromaticity of physical objects (taken into account by Δu^* and Δv^*) but also from the ability of the test source to render the lightnesses of the physical objects (taken into account by ΔL^*). The *CRI* calculation is based on

the premise that the reference source renders the true chromaticity and lightness, i.e. the true color, of physical objects.

The choice of the numerical prefactors in Eqns. (27.8) ~ (27.11) is somewhat arbitrary. These prefactors have been determined in extensive experiments with human subjects. Evidence exists, however, that the current prefactors may not be optimal (Wyszecki and Stiles, 1982, 2000).

27.4 Color-rendering index for non-planckian-locus illumination sources

The following calculation of ΔE_i^* is suited for test sources that are located *off the planckian locus*. The calculation described here follows the procedure developed by CIE (1995, 1999) and takes into account the ***adaptive color shift*** that follows from the human ability of ***chromatic adaptation***.

The ***human visual system*** consists of the eyes, the optic nerve, and the visual cortex (the visual cortex is the part of the brain dedicated to vision). Chromatic adaptation is the well-known ability of the human visual system to adapt to certain illumination conditions without a substantial loss in color perception. For example, colorful objects (e.g. apples, bananas, or color prints) can be viewed under vastly different illumination conditions such as in candle light ($CT \approx 2\,000\text{ K}$), incandescent light ($CT \approx 2\,800\text{ K}$), direct sunlight ($CT \approx 5\,000\text{ K}$), or overcast conditions ($CT \approx 6\,500\text{ K}$). Nevertheless, under any of these conditions, the colors of the objects appear very similar (or the same). This is the result of chromatic adaptation.

The ability to chromatically adapt exists even for non-planckian-locus sources. For example, yellow light sources that do not contain short-wavelength light (violet, blue and cyan) are frequently used in semiconductor-manufacturing clean rooms. Such yellow sources are located off the planckian locus. Nevertheless, after having adapted to the clean-room illumination conditions (which typically takes several tens of minutes), the color-distortions caused by the yellow lights are no longer apparent and the lighting conditions are perceived as natural.

When calculating the *CRI* for off-planckian-locus sources according to the method described in the previous section, the *CRI* is very low. However, such low *CRI* values are not supported by experiments with human subjects: Due to chromatic adaptation, colors can appear vivid and natural, even for illumination sources slightly off the planckian locus. To overcome this discrepancy and to allow for a more realistic calculation of the *CRI*, the CIE (1995, 1999) introduced an alternative method to calculate the *CRI*. This alternative method takes into account

the human ability of chromatic adaptation by introducing an adaptive color shift of the test source towards the planckian reference source.

As a rule of thumb, the pleasantness and quality of white illumination sources decreases rapidly if the chromaticity point of the illumination source deviates from the planckian locus by a distance greater than 0.01 in the (x, y) chromaticity system. This corresponds to the distance of about 4 MacAdam ellipses, a standard employed by the lighting industry (Duggal, 2005). Note however, that the 0.01-rule-of-thumb is a necessary but not a sufficient condition for high quality of illumination sources.

The calculation starts with the uniform chromaticity coordinates of the reference and test sources and the chromaticity coordinates of the test-color samples when illuminated with the reference and test sources, i.e. (u_{ref}, v_{ref}) , (u_{test}, v_{test}) , $(u_{ref,i}, v_{ref,i})$, and $(u_{test,i}, v_{test,i})$.

To account for the adaptive color shift, the (u, v) coordinates of (u_{ref}, v_{ref}) , (u_{test}, v_{test}) , and $(u_{test,i}, v_{test,i})$ are transformed into (c, d) coordinates using the formulae

$$c = (4 - u - 10v) / v \quad (27.12)$$

$$d = (1.708v + 0.404 - 1.481u) / v . \quad (27.13)$$

Note that these two equations correspond to six equations when transforming (u_{ref}, v_{ref}) , (u_{test}, v_{test}) , and $(u_{test,i}, v_{test,i})$ into (c_{ref}, d_{ref}) , (c_{test}, d_{test}) , and $(c_{test,i}, d_{test,i})$, respectively. Subsequently the adaptive-color-shifted chromaticity coordinates of the test-color samples are calculated according to

$$u_{test,i}^{**} = \frac{10.872 + 0.404 \frac{c_{ref}}{c_{test}} c_{test,i} - 4 \frac{d_{ref}}{d_{test}} d_{test,i}}{16.518 + 1.481 \frac{c_{ref}}{c_{test}} c_{test,i} - \frac{d_{ref}}{d_{test}} d_{test,i}} \quad (27.14)$$

$$v_{test,i}^{**} = \frac{5.520}{16.518 + 1.481 \frac{c_{ref}}{c_{test}} c_{test,i} - \frac{d_{ref}}{d_{test}} d_{test,i}} . \quad (27.15)$$

Correspondingly, the adaptive-color-shifted chromaticity coordinates of the test source are calculated according to

$$u_{test}^{**} = \frac{10.872 + 0.404 c_{ref} - 4 d_{ref}}{16.518 + 1.481 c_{ref} - d_{ref}} = u_{ref} \quad (27.16)$$

$$v_{test}^{**} = \frac{5.520}{16.518 + 1.481 c_{ref} - d_{ref}} = v_{ref}. \quad (27.17)$$

The values of u_{test}^{**} and v_{test}^{**} are the chromaticity coordinates of the light source to be tested after the adaptive color shift has been performed (note that $u_{test}^{**} = u_{ref}$ and $v_{test}^{**} = v_{ref}$). Finally, the color difference is calculated in terms of the uniform color space coordinates

$$\Delta E_i^* = \sqrt{(\Delta L^{**})^2 + (\Delta u^{**})^2 + (\Delta v^{**})^2} \quad (27.18)$$

where

$$\Delta L^{**} = L_{ref,i}^{**} - L_{test,i}^{**} = [25(Y_{ref,i})^{1/3} - 17] - [25(Y_{test,i})^{1/3} - 17] \quad (27.19)$$

$$\Delta u^{**} = u_{ref}^{***} - u_{test}^{***} = 13 L_{ref,i}^{**} (u_{ref,i} - u_{ref}) - 13 L_{test,i}^{**} (u_{test,i}^{**} - u_{test}^{**}) \quad (27.20)$$

$$\Delta v^{**} = v_{ref}^{***} - v_{test}^{***} = 13 L_{ref,i}^{**} (v_{ref,i} - v_{ref}) - 13 L_{test,i}^{**} (v_{test,i}^{**} - v_{test}^{**}). \quad (27.21)$$

Note that the calculation requires that $Y_{ref} = Y_{test} = 100$ (CIE, 1995, 1999). Using the calculated values of ΔE_i^* , the general CRI is calculated using Eqns. (27.1) and (27.2). The special CRI_i for $i = 9$ to 14 may be of interest for a complete assessment of the color rendering properties of an illumination source.

Exercise: Color rendering. The color of a physical object, as seen by a human being, is not just a function of the object but also a function of the light source illuminating the object! In fact, the color of an object can depend very strongly on the light source illuminating the object. Some light sources do render the natural colors of an object (true color rendering) while some light sources do not (false color rendering).

- (a) What is the color of a yellow banana when illuminated with a red LED?
- (b) What is the color of a green banana when illuminated with a yellow LED?
- (c) Could it be advantageous for a grocer to illuminate meat with red LEDs, bananas with yellow LEDs, and oranges with orange LEDs?
- (d) Is it possible for two physical objects of different colors to appear to have the same color under certain illumination conditions?
- (e) Why are low-pressure Na vapor lights used despite their low color-rendering index?
- (f) What would be the advantage and disadvantage of using green LEDs for illumination?

Solution:

(a) Red. (b) Yellow. (c) Yes – but his truthfulness in displaying fruit could be questioned. (d) Yes. (e) Because of their high luminous efficiency (and thus low electricity consumption). (f) High luminous efficacy would be an advantage but low color-rendering properties would be a disadvantage.

27.5 Criticism of the CRI and alternative approaches

In addition to the general *CRI*, the CIE has established the special *CRI_{special}* (*CRI₉* ~ *CRI₁₅*). These special *CRI* values should be considered when fully evaluating the color rendering capabilities of a light source. In the early days of phosphor-based white LEDs, much attention was paid to the *CRI₉* value, since it measures the rendering of saturated red, which early phosphor-based white LEDs had difficulties in rendering. This shortcoming can be corrected by adding a red LED or by adding a red-emitting phosphor.

The *CRI* has been criticized by pointing out that a black-body-radiation (planckian) reference light source having a very low color temperature, e.g. 2 000 K, has a favorable *CRI*, 100, even though colors are not rendered well (Ohno, 2005). The same can be said for a planckian reference light source having a very high color temperature, e.g. 8 000 K (Ohno, 2005).

The *CRI* metric should be employed with caution for white-light sources with a “spiky” emission spectrum, i.e. a spectrum consisting of multiple narrow emission bands (usually three to four narrow emission bands). Such light sources include RGB-LED lamps and the now-obsolete tri-phosphor fluorescent lamps. The *CRI* depends strongly on the specific emission wavelength of these “spikes”. Furthermore, although the *CRI* of such sources may have acceptably high values, the actual color rendition, when inspected visually and compared to a light source with a continuous-spectrum, can be unsatisfactory. Therefore, caution must be exercised when applying the *CRI* metric to “spiky” light sources.

On the other hand, white-light sources with a “spiky” emission spectrum, particularly RGB sources with the emission spikes located in the red, green and blue, can increase saturation (chroma) of object colors. This can enhance the visual experience, specifically “visual clarity” and the visual “feeling of contrast” (Hashimoto *et al.*, 2007). The authors pointed out that LED light sources with emission spectra consisting of multiple narrow lines (e.g. an RGB light source) tend to increase the saturation (chroma) of object colors thereby increasing visual clarity. To quantify this positive effect, the authors proposed a new metric, the “feeling contrast index” (FCI) that is

based on “visual clarity” and the visual “feeling of contrast”. The authors proposed that the FCI be used in addition to the *CRI* to provide a more comprehensive metric for light sources.

An increase the number of test color samples from 15 to more than 1000 was proposed to reduce the inherent variation caused by the limited number of test-color samples (Zukauskas *et al.*, 2009). Furthermore, the authors categorized the rendered colors into three groups — (i) colors rendered with high fidelity, (ii) colors rendered with increased saturation, and (iii) those rendered with distorted hue — and introduced three corresponding indexes to assess the color quality of light sources.

Davis and Ohno (2010) proposed a “color-quality scale” (CQS). The CQS uses a larger set of reflective samples and combines the color differences of the samples with a root mean square. The *CRI* penalizes any light source that causes any changes in object colors (upon illumination with the test source). However, Davis and Ohno (2010) suggested that while reduction in chroma (desaturation or dulling) was undesirable, enhancement in chroma was desirable. For this reason, the CQS does not penalize light sources that increase chroma (saturation) of object colors. The authors suggested that in practical applications, increases in chroma would be desirable due to higher visual clarity. However, general validity of such desirability is not assured. There are very good reasons, including the strong desire for color-fidelity, to penalize *any* deviation in color, including reduction in chroma (more dull) and enhancement of chroma (more colorful).

A test light source can make object colors more dull (lower chroma) or more saturated (higher chroma). To convey these two possibilities to the users of the light source, Zukauskas *et al.* (2013) proposed that the *CRI* value be appended by either a “D” (for dulling) or an “S” (for saturating). A third possibility, neither dulling nor saturating, may also be considered.

In conclusion, the conventional general *CRI* and the special *CRI*s are widely used. The *CRI* is a very helpful and valuable standard for assessing the color-rendering capability of a light source. Furthermore, the *CRI* is a well-established thereby providing historical continuity. However, the *CRI* is not a perfect metric.

27.6 References

- Berger-Schunn A. *Practical Color Measurement* (John Wiley and Sons, New York, 1994)
- Billmeyer Jr. F. W. “Survey of color order systems” *Color Research and Applications* **12**, 173 (1987)
- CIE publication No. 13.2, *Method of measuring and specifying colour rendering properties of light sources* (CIE, Vienna, Austria, 1974 and CIE, Paris, France, Bureau Central de la CIE, 1974)
- CIE publication No. 15 (E.1.3.1) 1971: *Colorimetry* this publication was updated in 1986 to CIE Publication

- 15.2 *Colorimetry* (CIE, Vienna, Austria, 1986)
- CIE publication No. 13.3, *Method of Measuring and Specifying Color-Rendering of Light Sources* (see also www.cie.co.at) (CIE, Vienna, Austria, 1995 and CIE, Paris, France, Bureau Central de la CIE, 1995)
- CIE publication No. 135-1999 and CIE 135.1-2000 Revision 1, *Physical Measurement of Light and Radiation*, ISBN 978-3-900734-97-8 (CIE, Vienna, Austria, 1999)
- CIE publication No. 177:2007, *Colour Rendering of White LED Light Sources*, ISBN 978-3-901906-57-2 (CIE, Vienna, Austria, 2007)
- Duggal A. R. "Organic electroluminescent devices for solid-state lighting" in *Organic Electro-luminescence* edited by Z. H. Kafafi (Taylor and Francis Group, Boca Raton, Florida, 2005)
- Hashimoto K., Yano T., Shimizu M., and Nayatani Y. "New method for specifying color-rendering properties of light sources based on feeling of contrast" *Color Research and Application* **32**, 361 (2007)
- Kendall M. and Scholand M. *Energy Savings Potential of Solid State Lighting in General Lighting Applications* available to the public from National Technical Information Service (NTIS), US Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161 (2001)
- Long J. and Luke J. T. *The New Munsell Student Color Set* 2nd ring-bound edition (Fairchild Books and Visuals, New York, 2001)
- MacAdam D. L. (Editor) *Colorimetry – Fundamentals* (SPIE Optical Engineering Press, Bellingham, Washington, 1993)
- Munsell A. H. *A Color Notation – An Illustrated System Defining All Colors and Their Relations by Measured Scales of Hue, Value, and Chroma* (G. H. Ellis, Boston, 1905)
- Munsell *Munsell Book of Color, Matte Edition* is available through GretagMacbeth Corporation, gretagmacbeth.com and munsell.com (Regensdorf, Switzerland, 2005)
- Ohno Y. "Spectral design considerations for white LED color rendering" *SPIE Optical Engineering* **44**, 111302 (2005)
- Davis W. and Ohno Y. "Color quality scale" *SPIE Optical Engineering* **49**, 033602 (2010)
- Wyszecki G. and Stiles W. S. *Color Science – Concepts and Methods, Quantitative Data and Formulae* 2nd edition (John Wiley and Sons, New York, 1982)
- Wyszecki G. and Stiles W. S. *Color Science – Concepts and Methods, Quantitative Data and Formulae* 2nd edition "Wiley Classics Library" (John Wiley and Sons, New York, 2000)
- Zukauskas A., Vaicekauskas R., Ivanauskas F., Vaitkevicius H., Vitta P., and Shur M. S. "Statistical approach to color quality of solid-state lamps" *IEEE Journal of Selected Topics in Quantum Electronics* **15**, 1189 (2009)
- Zukauskas A., Vaicekauskas R., and Shur M. S. "Resolving the ambiguity of colour fidelity indices" AIC Colour 2013, 12th Congress of the International Colour Association, Vol. **3**, 1129 (2013)

Appendix 27.1

Spectral reflectivity $R_i(\lambda)$ of the CIE 1974 Test-Color Samples (TCS) Nos. 1~8 to be used in calculating the General Color-Rendering Index (General CRI) (after CIE, 1995)

λ (nm)	R_1 (-)	R_2 (-)	R_3 (-)	R_4 (-)	R_5 (-)	R_6 (-)	R_7 (-)	R_8 (-)
360	0.116	0.053	0.058	0.057	0.143	0.079	0.150	0.075
365	0.136	0.055	0.059	0.059	0.187	0.081	0.177	0.078
370	0.159	0.059	0.061	0.062	0.233	0.089	0.218	0.084
375	0.190	0.064	0.063	0.067	0.269	0.113	0.293	0.090
380	0.219	0.070	0.065	0.074	0.295	0.151	0.378	0.104
385	0.239	0.079	0.068	0.083	0.306	0.203	0.459	0.129
390	0.252	0.089	0.070	0.093	0.310	0.265	0.524	0.170
395	0.256	0.101	0.072	0.105	0.312	0.339	0.546	0.240
400	0.256	0.111	0.073	0.116	0.313	0.410	0.551	0.319
405	0.254	0.116	0.073	0.121	0.315	0.464	0.555	0.416
410	0.252	0.118	0.074	0.124	0.319	0.492	0.559	0.462
415	0.248	0.120	0.074	0.126	0.322	0.508	0.560	0.482
420	0.244	0.121	0.074	0.128	0.326	0.517	0.561	0.490
425	0.240	0.122	0.073	0.131	0.330	0.524	0.558	0.488
430	0.237	0.122	0.073	0.135	0.334	0.531	0.556	0.482
435	0.232	0.122	0.073	0.139	0.339	0.538	0.551	0.473
440	0.230	0.123	0.073	0.144	0.346	0.544	0.544	0.462
445	0.226	0.124	0.073	0.151	0.352	0.551	0.535	0.450
450	0.225	0.127	0.074	0.161	0.360	0.556	0.522	0.439
455	0.222	0.128	0.075	0.172	0.369	0.556	0.506	0.426
460	0.220	0.131	0.077	0.186	0.381	0.554	0.488	0.413
465	0.218	0.134	0.080	0.205	0.394	0.549	0.469	0.397
470	0.216	0.138	0.085	0.229	0.403	0.541	0.448	0.382
475	0.214	0.143	0.094	0.254	0.410	0.531	0.429	0.366
480	0.214	0.150	0.109	0.281	0.415	0.519	0.408	0.352
485	0.214	0.159	0.126	0.308	0.418	0.504	0.385	0.337
490	0.216	0.174	0.148	0.332	0.419	0.488	0.363	0.325
495	0.218	0.190	0.172	0.352	0.417	0.469	0.341	0.310
500	0.223	0.207	0.198	0.370	0.413	0.450	0.324	0.299
505	0.225	0.225	0.221	0.383	0.409	0.431	0.311	0.289
510	0.226	0.242	0.241	0.390	0.403	0.414	0.301	0.283
515	0.226	0.253	0.260	0.394	0.396	0.395	0.291	0.276
520	0.225	0.260	0.278	0.395	0.389	0.377	0.283	0.270
525	0.225	0.264	0.302	0.392	0.381	0.358	0.273	0.262
530	0.227	0.267	0.339	0.385	0.372	0.341	0.265	0.256
535	0.230	0.269	0.370	0.377	0.363	0.325	0.260	0.251
540	0.236	0.272	0.392	0.367	0.353	0.309	0.257	0.250
545	0.245	0.276	0.399	0.354	0.342	0.293	0.257	0.251
550	0.253	0.282	0.400	0.341	0.331	0.279	0.259	0.254
555	0.262	0.289	0.393	0.327	0.320	0.265	0.260	0.258
560	0.272	0.299	0.380	0.312	0.308	0.253	0.260	0.264
565	0.283	0.309	0.365	0.296	0.296	0.241	0.258	0.269
570	0.298	0.322	0.349	0.280	0.284	0.234	0.256	0.272
575	0.318	0.329	0.332	0.263	0.271	0.227	0.254	0.274
580	0.341	0.335	0.315	0.247	0.260	0.225	0.254	0.278
585	0.367	0.339	0.299	0.229	0.247	0.222	0.259	0.284

590	0.390	0.341	0.285	0.214	0.232	0.221	0.270	0.295
595	0.409	0.341	0.272	0.198	0.220	0.220	0.284	0.316
600	0.424	0.342	0.264	0.185	0.210	0.220	0.302	0.348
605	0.435	0.342	0.257	0.175	0.200	0.220	0.324	0.384
610	0.442	0.342	0.252	0.169	0.194	0.220	0.344	0.434
615	0.448	0.341	0.247	0.164	0.189	0.220	0.362	0.482
620	0.450	0.341	0.241	0.160	0.185	0.223	0.377	0.528
625	0.451	0.339	0.235	0.156	0.183	0.227	0.389	0.568
630	0.451	0.339	0.229	0.154	0.180	0.233	0.400	0.604
635	0.451	0.338	0.224	0.152	0.177	0.239	0.410	0.629
640	0.451	0.338	0.220	0.151	0.176	0.244	0.420	0.648
645	0.451	0.337	0.217	0.149	0.175	0.251	0.429	0.663
650	0.450	0.336	0.216	0.148	0.175	0.258	0.438	0.676
655	0.450	0.335	0.216	0.148	0.175	0.263	0.445	0.685
660	0.451	0.334	0.219	0.148	0.175	0.268	0.452	0.693
665	0.451	0.332	0.224	0.149	0.177	0.273	0.457	0.700
670	0.453	0.332	0.230	0.151	0.180	0.278	0.462	0.705
675	0.454	0.331	0.238	0.154	0.183	0.281	0.466	0.709
680	0.455	0.331	0.251	0.158	0.186	0.283	0.468	0.712
685	0.457	0.330	0.269	0.162	0.189	0.286	0.470	0.715
690	0.458	0.329	0.288	0.165	0.192	0.291	0.473	0.717
695	0.460	0.328	0.312	0.168	0.195	0.296	0.477	0.719
700	0.462	0.328	0.340	0.170	0.199	0.302	0.483	0.721
705	0.463	0.327	0.366	0.171	0.200	0.313	0.489	0.720
710	0.464	0.326	0.390	0.170	0.199	0.325	0.496	0.719
715	0.465	0.325	0.412	0.168	0.198	0.338	0.503	0.722
720	0.466	0.324	0.431	0.166	0.196	0.351	0.511	0.725
725	0.466	0.324	0.447	0.164	0.195	0.364	0.518	0.727
730	0.466	0.324	0.460	0.164	0.195	0.376	0.525	0.729
735	0.466	0.323	0.472	0.165	0.196	0.389	0.532	0.730
740	0.467	0.322	0.481	0.168	0.197	0.401	0.539	0.730
745	0.467	0.321	0.488	0.172	0.200	0.413	0.546	0.730
750	0.467	0.320	0.493	0.177	0.203	0.425	0.553	0.730
755	0.467	0.318	0.497	0.181	0.205	0.436	0.559	0.730
760	0.467	0.316	0.500	0.185	0.208	0.447	0.565	0.730
765	0.467	0.315	0.502	0.189	0.212	0.458	0.570	0.730
770	0.467	0.315	0.505	0.192	0.215	0.469	0.575	0.730
775	0.467	0.314	0.510	0.194	0.217	0.477	0.578	0.730
780	0.467	0.314	0.516	0.197	0.219	0.485	0.581	0.730
785	0.467	0.313	0.520	0.200	0.222	0.493	0.583	0.730
790	0.467	0.313	0.524	0.204	0.226	0.500	0.585	0.731
795	0.466	0.312	0.527	0.210	0.231	0.506	0.587	0.731
800	0.466	0.312	0.531	0.218	0.237	0.512	0.588	0.731
805	0.466	0.311	0.535	0.225	0.243	0.517	0.589	0.731
810	0.466	0.311	0.539	0.233	0.249	0.521	0.590	0.731
815	0.466	0.311	0.544	0.243	0.257	0.525	0.590	0.731
820	0.465	0.311	0.548	0.254	0.265	0.529	0.590	0.731
825	0.464	0.311	0.552	0.264	0.273	0.532	0.591	0.731
830	0.464	0.310	0.555	0.274	0.280	0.535	0.592	0.731

Appendix 27.2

Spectral reflectivity $R_i(\lambda)$ of the CIE 1974 Test-Color Samples (TCS) Nos. 9~14 (after CIE, 1995)

λ (nm)	R_9 (-)	R_{10} (-)	R_{11} (-)	R_{12} (-)	R_{13} (-)	R_{14} (-)
360	0.069	0.042	0.074	0.189	0.071	0.036
365	0.072	0.043	0.079	0.175	0.076	0.036
370	0.073	0.045	0.086	0.158	0.082	0.036
375	0.070	0.047	0.098	0.139	0.090	0.036
380	0.066	0.050	0.111	0.120	0.104	0.036
385	0.062	0.054	0.121	0.103	0.127	0.036
390	0.058	0.059	0.127	0.090	0.161	0.037
395	0.055	0.063	0.129	0.082	0.211	0.038
400	0.052	0.066	0.127	0.076	0.264	0.039
405	0.052	0.067	0.121	0.068	0.313	0.039
410	0.051	0.068	0.116	0.064	0.341	0.040
415	0.050	0.069	0.112	0.065	0.352	0.041
420	0.050	0.069	0.108	0.075	0.359	0.042
425	0.049	0.070	0.105	0.093	0.361	0.042
430	0.048	0.072	0.104	0.123	0.364	0.043
435	0.047	0.073	0.104	0.160	0.365	0.044
440	0.046	0.076	0.105	0.207	0.367	0.044
445	0.044	0.078	0.106	0.256	0.369	0.045
450	0.042	0.083	0.110	0.300	0.372	0.045
455	0.041	0.088	0.115	0.331	0.374	0.046
460	0.038	0.095	0.123	0.346	0.376	0.047
465	0.035	0.103	0.134	0.347	0.379	0.048
470	0.033	0.113	0.148	0.341	0.384	0.050
475	0.031	0.125	0.167	0.328	0.389	0.052
480	0.030	0.142	0.192	0.307	0.397	0.055
485	0.029	0.162	0.219	0.282	0.405	0.057
490	0.028	0.189	0.252	0.257	0.416	0.062
495	0.028	0.219	0.291	0.230	0.429	0.067
500	0.028	0.262	0.325	0.204	0.443	0.075
505	0.029	0.305	0.347	0.178	0.454	0.083
510	0.030	0.365	0.356	0.154	0.461	0.092
515	0.030	0.416	0.353	0.129	0.466	0.100
520	0.031	0.465	0.346	0.109	0.469	0.108
525	0.031	0.509	0.333	0.090	0.471	0.121
530	0.032	0.546	0.314	0.075	0.474	0.133
535	0.032	0.581	0.294	0.062	0.476	0.142
540	0.033	0.610	0.271	0.051	0.483	0.150
545	0.034	0.634	0.248	0.041	0.490	0.154
550	0.035	0.653	0.227	0.035	0.506	0.155
555	0.037	0.666	0.206	0.029	0.526	0.152
560	0.041	0.678	0.188	0.025	0.553	0.147
565	0.044	0.687	0.170	0.022	0.582	0.140
570	0.048	0.693	0.153	0.019	0.618	0.133
575	0.052	0.698	0.138	0.017	0.651	0.125
580	0.060	0.701	0.125	0.017	0.680	0.118
585	0.076	0.704	0.114	0.017	0.701	0.112
590	0.102	0.705	0.106	0.016	0.717	0.106

595	0.136	0.705	0.100	0.016	0.729	0.101
600	0.190	0.706	0.096	0.016	0.736	0.098
605	0.256	0.707	0.092	0.016	0.742	0.095
610	0.336	0.707	0.090	0.016	0.745	0.093
615	0.418	0.707	0.087	0.016	0.747	0.090
620	0.505	0.708	0.085	0.016	0.748	0.089
625	0.581	0.708	0.082	0.016	0.748	0.087
630	0.641	0.710	0.080	0.018	0.748	0.086
635	0.682	0.711	0.079	0.018	0.748	0.085
640	0.717	0.712	0.078	0.018	0.748	0.084
645	0.740	0.714	0.078	0.018	0.748	0.084
650	0.758	0.716	0.078	0.019	0.748	0.084
655	0.770	0.718	0.078	0.020	0.748	0.084
660	0.781	0.720	0.081	0.023	0.747	0.085
665	0.790	0.722	0.083	0.024	0.747	0.087
670	0.797	0.725	0.088	0.026	0.747	0.092
675	0.803	0.729	0.093	0.030	0.747	0.096
680	0.809	0.731	0.102	0.035	0.747	0.102
685	0.814	0.735	0.112	0.043	0.747	0.110
690	0.819	0.739	0.125	0.056	0.747	0.123
695	0.824	0.742	0.141	0.074	0.746	0.137
700	0.828	0.746	0.161	0.097	0.746	0.152
705	0.830	0.748	0.182	0.128	0.746	0.169
710	0.831	0.749	0.203	0.166	0.745	0.188
715	0.833	0.751	0.223	0.210	0.744	0.207
720	0.835	0.753	0.242	0.257	0.743	0.226
725	0.836	0.754	0.257	0.305	0.744	0.243
730	0.836	0.755	0.270	0.354	0.745	0.260
735	0.837	0.755	0.282	0.401	0.748	0.277
740	0.838	0.755	0.292	0.446	0.750	0.294
745	0.839	0.755	0.302	0.485	0.750	0.310
750	0.839	0.756	0.310	0.520	0.749	0.325
755	0.839	0.757	0.314	0.551	0.748	0.339
760	0.839	0.758	0.317	0.577	0.748	0.353
765	0.839	0.759	0.323	0.599	0.747	0.366
770	0.839	0.759	0.330	0.618	0.747	0.379
775	0.839	0.759	0.334	0.633	0.747	0.390
780	0.839	0.759	0.338	0.645	0.747	0.399
785	0.839	0.759	0.343	0.656	0.746	0.408
790	0.839	0.759	0.348	0.666	0.746	0.416
795	0.839	0.759	0.353	0.674	0.746	0.422
800	0.839	0.759	0.359	0.680	0.746	0.428
805	0.839	0.759	0.365	0.686	0.745	0.434
810	0.838	0.758	0.372	0.691	0.745	0.439
815	0.837	0.757	0.380	0.694	0.745	0.444
820	0.837	0.757	0.388	0.697	0.745	0.448
825	0.836	0.756	0.396	0.700	0.745	0.451
830	0.836	0.756	0.403	0.702	0.745	0.454

28

28 – White-light sources based on LED chips

As the trend of higher efficiencies in LEDs continues, the number of possible applications increases as well. A highly desirable application with a very large market size is *general lighting* in homes, offices, and public places. The field of ***solid-state lighting*** (SSL) is concerned with the development of solid-state sources for daylight illumination applications. LED chips are inherently monochromatic emitters. However, there are several ways to generate white light using LEDs. In the current chapter, approaches to white-light generation based on multiple LED chips (i.e. without a phosphor) will be discussed, whereas approaches based on LEDs and wavelength-converting materials (phosphors) will be discussed in subsequent chapters. A pivotal discussion of the promise of solid-state lighting was given by Bergh *et al.* (2001). An introduction to lighting technology using solid-state sources was given by Zukauskas *et al.* (2002a).

In the field of general daylight illumination, devices should have the following properties (i) high efficiency, (ii) high power capability, (iii) good color-rendering capabilities (iv) high reliability, (v) low-cost manufacturability, and (vi) environmental benignity. Such properties have allowed LEDs to replace conventional illumination sources, in particular incandescent and fluorescent lamps.

28.1 Generation of white light with LED chips

White-light sources generally have a broad emission spectrum that may encompass the entire visible spectral range. Accordingly, the three types of cones located in the retina of the human eye are excited. For the case of white light, the tristimulus values are such that the location of the chromaticity point is near the center of the chromaticity diagram.

The generation of white light can be accomplished with a multitude of possible emission spectra. The creation of white light out of monochromatic visible-light emitters can be based on dichromatic, trichromatic, or tetrachromatic approaches, as shown in ***Figure 28.1***. The optical sources can be classified in terms of their luminous efficacy and color-rendering properties.

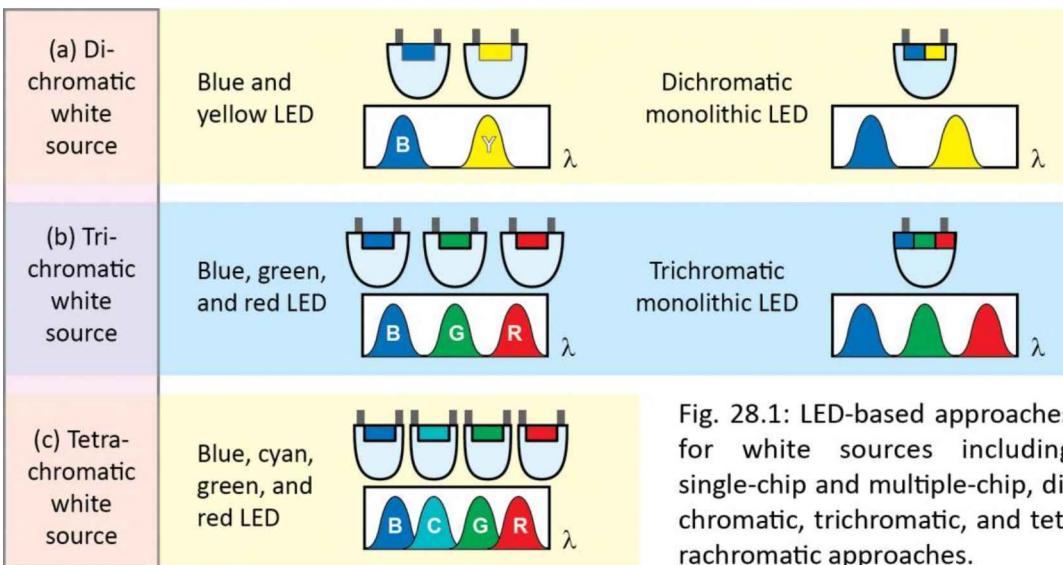


Fig. 28.1: LED-based approaches for white sources including single-chip and multiple-chip, di-chromatic, trichromatic, and tetrachromatic approaches.

Whereas high luminous efficacy is always desirable for high-power light sources, the desirability of a high color-rendering capability depends on the application. Generally, high-quality illumination applications, e.g. illumination in museums, homes, offices, and stores require a high color-rendering capability.

However, there are numerous applications where the color rendering capability is of lower priority, for example in the illumination of streets, parking garages, and stairwells. Finally, in signage applications, color rendering is irrelevant. Such signage applications include white pedestrian traffic lights, displays, and indicator lights.

There is a fundamental trade-off between the luminous efficacy of radiation and color rendering capability of a light source. Generally, dichromatic white light has the highest luminous efficacy and the poorest color-rendering capabilities. A trichromatic white source can have very acceptable color-rendering properties ($CRI > 80$) and luminous efficacies greater than 300 lm/W. Tetrachromatic sources can have color-rendering indices greater than 90.

Photographs of RGB LEDs (red-green-blue LEDs) are shown in *Figure 28.2* (Conrad, 2014). Three chips are mounted into the recess of a SMD package (surface-mount device package). The top of the SMD package may be covered with a black coating, which is common for devices used in information displays that are based on a matrix of RGB LEDs. The package may contain a supplementary element (e.g. a Zener diode) for protecting the LED chip from ESD (electro-static discharge). GaInN green and blue LED chips are more sensitive to ESD than AlGaInP red LED chips since the latter ones generally have a greater reverse leakage current.

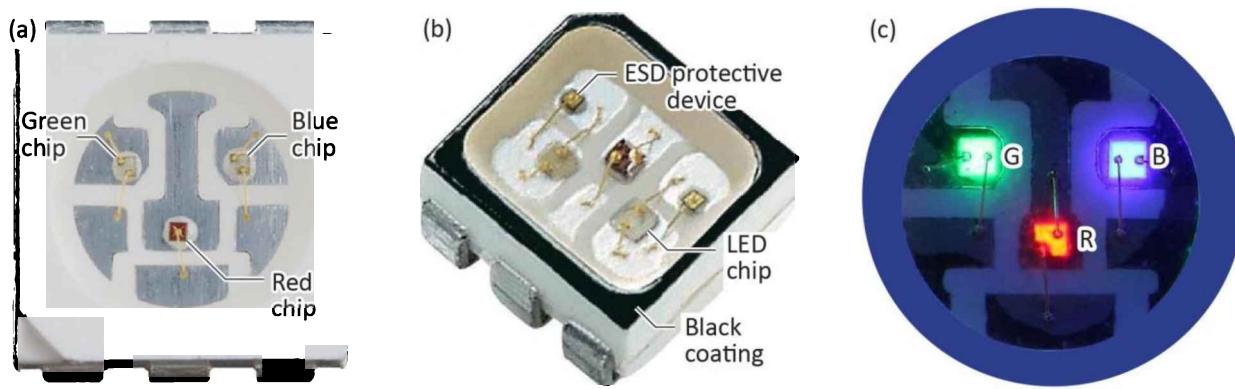


Fig. 28.2: Surface-mount device (SMD) LED packages containing a red (R), green (G), and blue (B) LED chip. (a) Top view; (b) perspective view with ESD protection and black coating; (c) Top view with chips lit (after Conrad, 2014).

28.2 Generation of white light by dichromatic sources

One method for generating white light is the use of two narrow emission bands, that is **complementary wavelengths** or **complementary colors**. Two complementary colors, at a certain power ratio, result in tristimulus values that appear as white. The wavelengths of complementary wavelength pairs are shown in **Figure 28.3**.

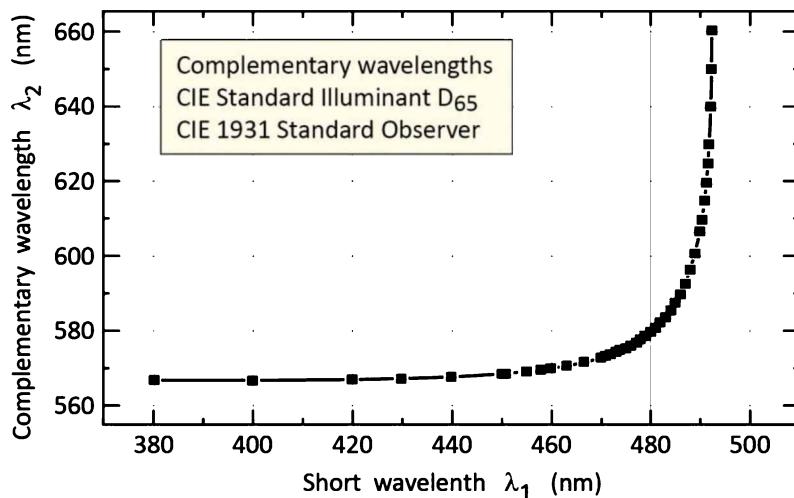


Fig. 28.3: Monochromatic complementary wavelengths resulting in the perception of white light at a certain power ratio (after Wyszecki and Stiles, 1982).

The numerical values for monochromatic complementary wavelengths are given in **Table 28.1**. The table also gives the power ratio required to attain the same chromaticity coordinate as Illuminant D₆₅.

Table 28.1: Wavelengths λ_1 and λ_2 of monochromatic complementary colors with respect to CIE Illuminant D₆₅ based on the CIE 1964 Standard Observer. Also given is the required power ratio. Illuminant D₆₅ has chromaticity coordinates $x_{D65} = 0.3138$ and $y_{D65} = 0.3310$ (after Wyszecki and Stiles, 1982).

Complementary wavelengths		Power ratio	Complementary wavelengths		Power ratio
λ_1 (nm)	λ_2 (nm)	$P(\lambda_2)/P(\lambda_1)$	λ_1 (nm)	λ_2 (nm)	$P(\lambda_2)/P(\lambda_1)$
380	560.9	0.000642	460	565.9	1.53
390	560.9	0.00955	470	570.4	1.09
400	561.1	0.0785	475	575.5	0.812
410	561.3	0.356	480	584.6	0.562
420	561.7	0.891	482	591.1	0.482
430	562.2	1.42	484	602.1	0.440
440	562.9	1.79	485	611.3	0.457
450	564.0	1.79	486	629.6	0.668

Next, we analyze the luminous efficacy of radiation of a source having two complementary emission lines. It is assumed that the two lines are thermally broadened to a full-width at half-maximum of ΔE . Emission lines of $\Delta E = 2kT$ to $10kT$ have been found experimentally for the GaInN system at room temperature ($kT = 25.9$ meV at 300 K). A **gaussian distribution** can be assumed for the two emission lines, so that the spectral power density is given by

$$P(\lambda) = P_1 \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\lambda-\lambda_1}{\sigma_1}\right)^2} + P_2 \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\lambda-\lambda_2}{\sigma_2}\right)^2} \quad (28.1)$$

where P_1 and P_2 are the optical powers of the two emission lines, and λ_1 and λ_2 are the peak wavelengths of the source. The gaussian standard deviation σ is related to the full-width at half-maximum of an emission spectrum, $\Delta\lambda$, by

$$\sigma = \Delta\lambda/(2\sqrt{2\ln 2}) = \Delta\lambda/2.355 . \quad (28.2)$$

The peak emission wavelengths λ_1 and λ_2 are chosen from **Table 28.1**. The table also gives the required power ratio of the two light sources. Although the table applies to strictly monochromatic sources ($\Delta\lambda \rightarrow 0$), the data can be used, with good approximation, for sources exhibiting moderate spectral broadening such as LEDs.

The luminous efficacy of radiation of a dichromatic source is shown in **Figure 28.4**. The figure reveals that the highest luminous efficacy of 440 lm/W occurs at a primary wavelength of $\lambda_1 = 445$ nm for $\Delta E = 2kT$. The very high value of the efficacy shows the great potential of dichromatic sources.

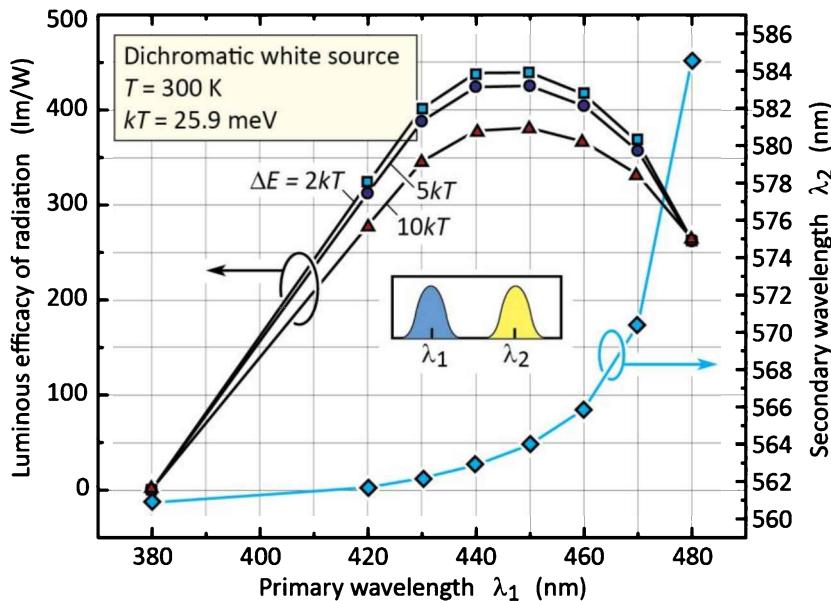


Fig. 28.4: Calculated luminous efficacy of dichromatic white-light source (with chromaticity point at D₆₅ standard illuminant) for different linewidths ΔE as a function of the primary wavelength. Also shown is the complementary secondary wavelength (after Li *et al.*, 2003).

Several approaches for the generation of white light by mixing two complementary colors have been demonstrated (Guo *et al.*, 1999; Sheu *et al.*, 2002; Dalmasso *et al.*, 2002; Li *et al.*, 2003). One possibility uses the mixing of light emitted by two LEDs, one emitting in the blue and the other one in the yellow spectral region. Another possibility, demonstrated by Guo *et al.* (1999), generates white light by using a GaN-based blue LED and a second semiconductor, AlGaN_P, as a wavelength converter. Sheu *et al.* (2002) demonstrated a codoped single active region quantum well white LED. The active region is doped with both Si and Zn. Blue light emission originates from quantum well band-to-band transitions, whereas a wide yellowish emission originates from donor-acceptor-pair (D-A) transitions. Because the D-A transition is spectrally wide, the codoped approach has the advantage of good color rendering. A dichromatic monolithic LED has been reported by Dalmasso *et al.* (2002), who employed two closely spaced GaInN active regions within the pn-junction region. A strong dependence of the emission spectrum on the injection current was found.

Li *et al.* (2003) reported a monolithic GaN-based LED with two active regions separated by a thin GaN layer. The device was designed for emission at 465 and 525 nm. The device structure is shown in **Figure 28.5**.

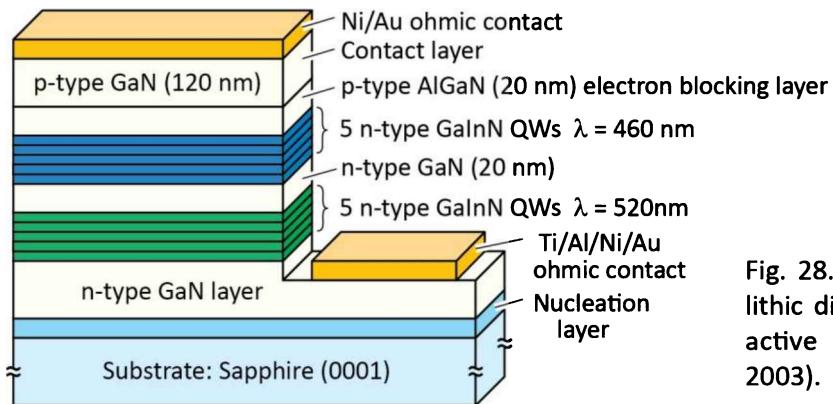


Fig. 28.5: Structure of a monolithic dichromatic LED with two active regions (after Li *et al.*, 2003).

Photoluminescence results are shown in **Figure 28.6 (a)**. The spectra exhibit two emission bands one centered at about 465 nm and one at 525 nm. As the excitation density is varied, the two peak positions do not change. However, the ratio of the two peak intensities changes with the excitation power density of the laser. The excitation-density-dependent emission ratio can be explained by the competition of different recombination paths (Li *et al.*, 2003). Neglecting non-radiative recombination, possible recombination paths of an electron in the blue quantum well (QW) are either direct radiative recombination in the blue QW or tunneling to the green QW with subsequent radiative recombination. Electrons and holes can tunnel from the blue QW to the green QW and recombine there radiatively. However, once in the green QW, carriers cannot tunnel back to the blue QW due to the higher energy of the blue QW.

Electroluminescence (EL) spectra of the device are shown in **Figure 28.6 (b)**. Two emission peaks are clearly observed with center wavelengths at about 450 nm and 520 nm. The blue peak is more intense than the green peak which can be attributed to the higher quantum efficiency of blue QWs compared with green QWs. In addition, holes are injected from the blue side (i.e. the side of the high-energy, blue-emitting QWs), whereas electrons are injected from the green side. As holes have a lower mobility and a higher effective mass, they are less likely to reach the green QWs, which can explain the higher intensity of the blue emission.

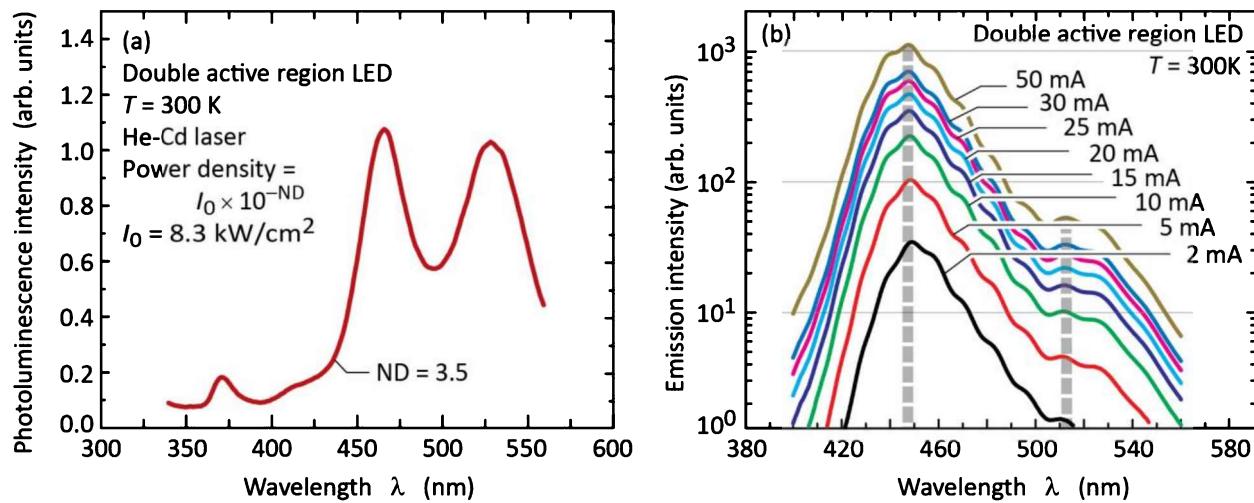


Fig. 28.6: Room temperature (a) photoluminescence and (b) electroluminescence spectra of monolithic dichromatic LED with two active regions (after Li *et al.*, 2003).

For current injection, holes are injected into the blue QWs, whereas electrons are injected into the green QWs. This is very different from optical excitation, where both types of carriers are injected from both sides of the active region. For the structure discussed here, the optical absorption length is longer than the distance between the active regions and the surface. This can explain the marked difference between the results for photoluminescence and electroluminescence (Li *et al.*, 2003).

Room temperature I - V curves of the double-active region LED exhibit excellent forward voltages < 3.0 V at small contact diameters of 100 μm indicating high-quality ohmic contacts. The increase in forward voltage for contacts with larger diameters was attributed to an increased voltage drop in the n-type buffer layer, because, at a given current density, the current in the buffer layer scales with the area of the contact A , but the access resistance through the n-type buffer layer scales with the circumference $A^{-1/2}$. In addition, the current crowding effect leads to non-uniform current injection particularly in large-diameter contacts and thus to an increased forward voltage.

28.3 Generation of white light by trichromatic sources

Whereas high-quality white light suitable for illumination applications cannot be generated by additive mixing of two narrow-band complementary colors, such high-quality white light can be generated by mixing of *three* or more primary colors. In a detailed analysis, Thornton (1971)

showed that mixing of discrete emission bands with peak wavelengths near 450, 540, and 610 nm resulted in a high-quality source. Thornton (1971) reported experiments with 60 human subjects who judged the quality of a trichromatic source in terms of its color rendition capability of meat, vegetables, flowers, and complexion, to be “very good, if not excellent”. Thornton (1971) taught to avoid the use of emitters near 500 and 580 nm when designing trichromatic sources with high-quality color rendition.

Although Thornton (1971) established that trichromatic sources can have high quality, the individual emission bands used in the experiments had a broad spectral width: The full-widths at half-maximum of the phosphor emitters employed in the study exceeded 50 nm. Semiconductors, with typical spectral widths < 50 nm, have narrower emission bands than phosphors.

The trichromatic emission spectrum of a white-light source made out of three types of LEDs emitting at 455, 525, and 605 nm is shown in *Figure 28.7 (a)*. The experimentally determined full-width at half-maximum of the spectra at room temperature (20°C) is $5.5kT$, $7.9kT$, and $2.5kT$ for the GaInN blue, GaInN green, and AlGaInP orange emitter, respectively, where $kT = 25.25$ meV.

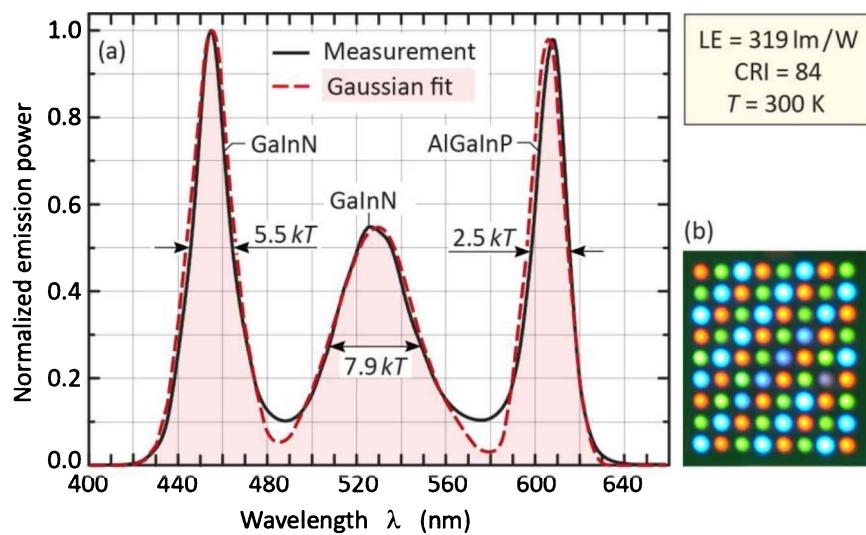


Fig. 28.7: (a) Emission spectrum of trichromatic white multi-LED source with color temperature of 6500 K (solid line) and gaussian fit (dashed line). The source has a luminous efficacy of radiation of 319 lm/W and a color rendering index of 84. (b) Photograph of source assembled from standard commercial devices (after Chhajed *et al.*, 2005).

The expression of the full-width at half-maximum in terms of kT is useful, as it can be related to the theoretical full-width at half-maximum of a thermally broadened emission band of a

semiconductor, which is $1.8kT$. Using the full-width at half-maximum given in the figure and the equation

$$\Delta\lambda = \frac{\lambda^2}{hc} \Delta E = \frac{(\lambda / \text{nm})^2}{1239.8} \frac{\Delta E}{\text{eV}}, \quad (28.3)$$

the full-widths at half-maximum of the blue, green, and orange sources are 23.2 nm, 44.3 nm, and 18.6 nm, respectively. Note that the green emission line is particularly broad, which can be attributed to alloy broadening and the formation of quantum-dot-like InN-rich regions within the high-In-content GaInN.

Also shown in the figure are gaussian fits to the experimental spectra. The gaussian fits match the experimental spectra well. Note that the gaussian curves (equations were given earlier in this chapter) are symmetric in terms of wavelength. *Asymmetric gaussian distributions*, which have a more pronounced long-wavelength tail, have been employed for phosphors (Ivey, 1963). The use of such asymmetric gaussian distributions does not appear to be warranted for semiconductors as their spectral power distribution is quite symmetric when plotted versus wavelength.

A photograph of the light source assembled from a large number of LEDs is shown in **Figure 28.7 (b)**. The power ratio of the orange, green, and blue emitters is adjusted to match the chromaticity of a planckian radiator with color temperature 6500 K. The LED source, assembled from standard commercial devices, has a luminous efficacy of radiation of 319 lm/W, a luminous source efficiency of 32 lm/W, and a color-rendering index of 84 (Chhajed *et al.*, 2005).

There are a large number of possible wavelength combinations for trichromatic sources. To attain a high efficacy of radiation, sources near the fringes of the visible spectrum (deep red and deep violet) should be avoided. Contour plots of the luminous efficacy of radiation and of the color rendering index of a trichromatic source with color temperature of 6500 K are shown in **Figure 28.8** for a full-width at half-maximum for each emission line of $5kT$. Inspection of the figure reveals that $\lambda_1 = 455$ nm, $\lambda_2 = 530$ nm, and $\lambda_3 = 605$ nm are particularly favorable in terms of the color-rendering index. The CIE general *CRI* is about 85 for this wavelength combination with the luminous efficacy of radiation being 320 lm/W.

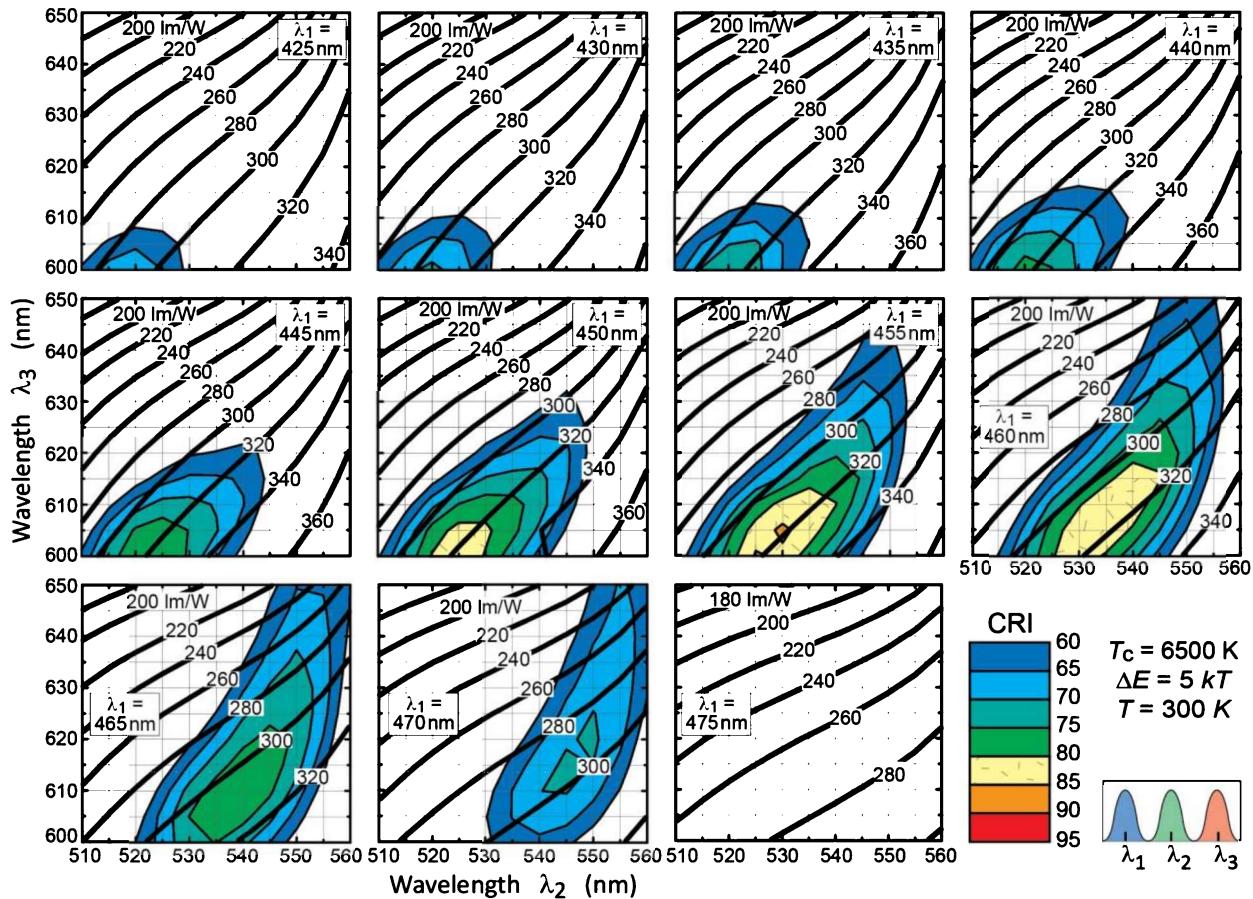


Fig. 28.8: Contour plot of luminous efficacy of radiation and CIE color-rendering index of white trichromatic LED source with color temperature 6500 K as a function of the three wavelengths for a linewidth (FWHM) of $5kT$ (after Chhajed *et al.*, 2005).

The figure also reveals that the *CRI* depends very sensitively on the exact peak positions. For example, changing the red peak wavelength from 605 to 620 nm decreases the *CRI* from 85 to 65. Similarly, changing the green wavelength from 530 to 550 nm decreases the *CRI* to values less than 60.

Contour plots of the luminous efficacy of radiation and of the color-rendering index of a trichromatic source with color temperature of 6500 K are shown in **Figure 28.9** for a full-width at half-maximum for each emission line of $8kT$. A higher *CRI* results from the broader emission lines. A very favorable combination in terms of a high *CRI* is obtained for $\lambda_1 = 450\text{--}455 \text{ nm}$, $\lambda_2 = 525\text{--}535 \text{ nm}$, and $\lambda_3 = 600\text{--}615 \text{ nm}$, where a *CRI* in the range 90~95 is obtained.

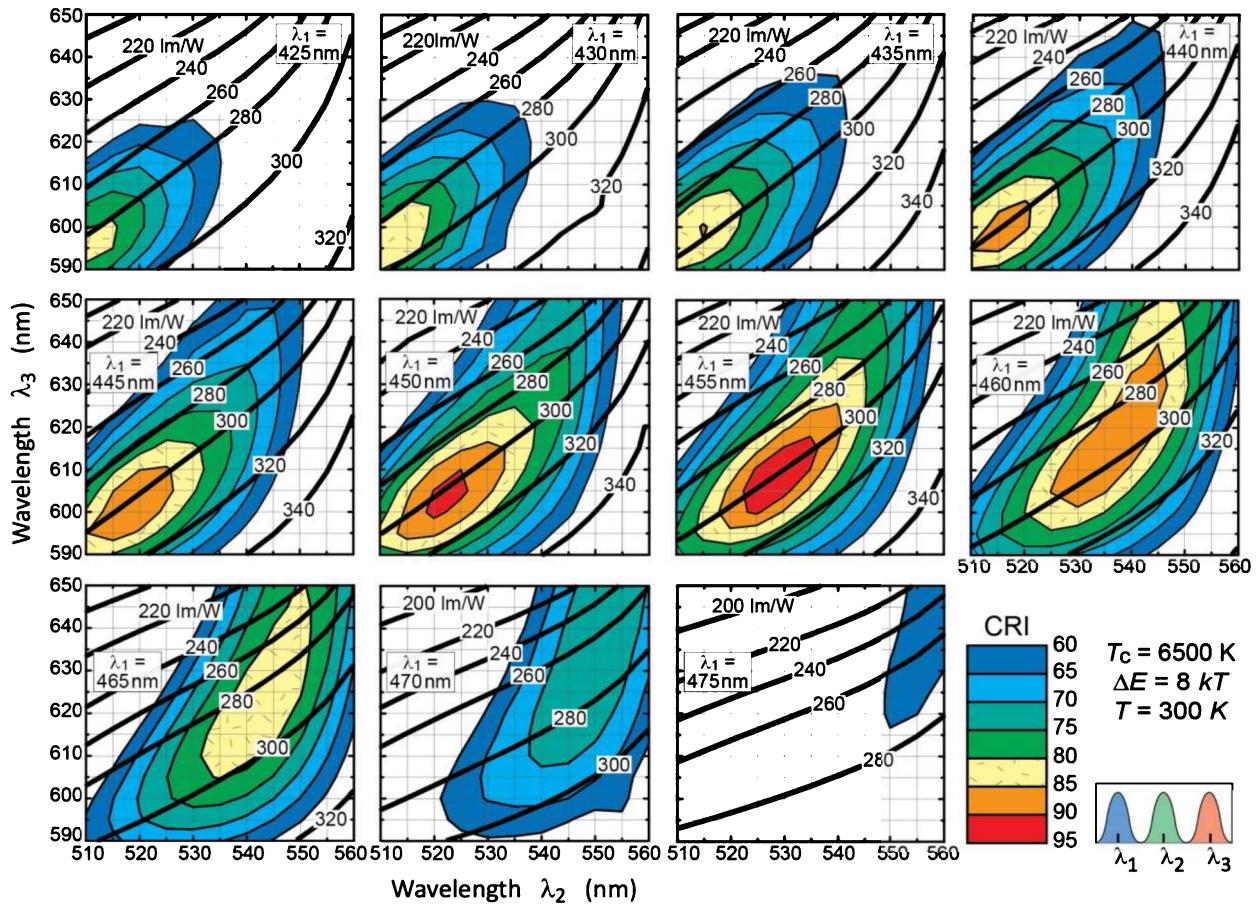


Fig. 28.9: Contour plot of luminous efficacy of radiation and CIE color-rendering index of white trichromatic LED source with color temperature 6500 K as a function of the three wavelengths for a linewidth (FWHM) of $8kT$ (after Chhajed *et al.*, 2005).

28.4 Temperature dependence of trichromatic LED-based white-light source

The relatively small range of wavelengths that enables a high color-rendering capability raises the question as to the stability of trichromatic sources with respect to junction and ambient temperature. The emission power (P), peak wavelength (λ_{peak}), and spectral width ($\Delta\lambda$) depend on temperature, with each of these quantities having a different temperature coefficient.

The optical output power of LEDs is temperature dependent in a manner that can be described by an exponential function and a characteristic temperature, $T_{\text{characteristic}}$. The light output power of an LED is then given by

$$P = P|_{300\text{ K}} \exp \frac{T - 300\text{ K}}{T_{\text{characteristic}}} . \quad (28.4)$$

As a result of these temperature dependences, the chromaticity point of a multi-LED white-light source changes with temperature. Consider a white-light source consisting of three types of

emitters emitting in the red, green, and blue. For such LEDs, the temperature coefficients of the peak emission wavelength, spectral width, and emission power have been measured and are given in **Table 28.2** (Chhajed *et al.*, 2005).

Table 28.2: Experimentally determined temperature coefficients for peak wavelength, spectral width, and emission power for blue, green, and red LEDs.

	Blue	Green	Red
$d\lambda_{peak}/dT$	0.0389 nm / °C	0.0308 nm / °C	0.156 nm / °C
$d\Delta\lambda/dT$	0.0466 nm / °C	0.0625 nm / °C	0.181 nm / °C
$T_{characteristic}$	493 K	379 K	209 K

Consider further that the three currents feeding the red, green, and blue LEDs are adjusted in such a way that the resulting chromaticity point equals that of Illuminant D₆₅ when the device temperature is 20°C. The optical spectrum of such a trichromatic white source is shown in **Figure 28.10**.

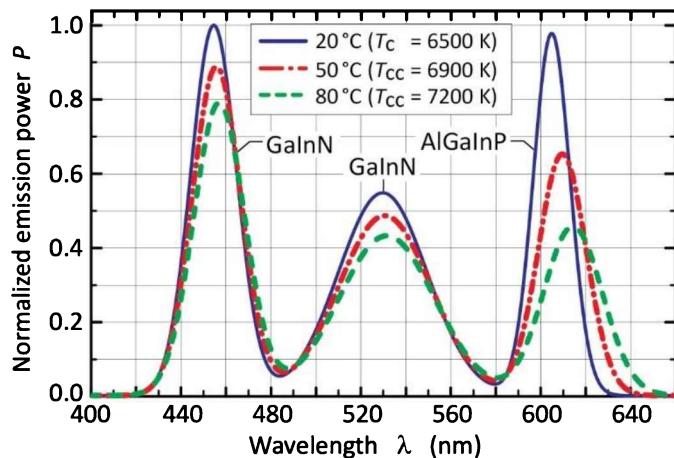


Fig. 28.10: Emission spectrum of trichromatic white LED source for different ambient temperatures (junction heating neglected). Optical power, linewidth, and peak wavelength change with temperature. As a result of these changes, the color temperature of the source increases (after Chhajed *et al.*, 2005).

However, as the device temperature increases, the chromaticity point of the trichromatic source changes due to the temperature dependences of the emission power, peak wavelength, and spectral width. This shift of the chromaticity point is shown in **Figure 28.11**. Inspection of the figure reveals that the chromaticity point shifts towards higher color temperatures. This can be explained by the stronger temperature dependence of the red LED's emission power. At high temperatures, the red component decreases more strongly (due to a low value of $T_{characteristic}$) than the green and blue component.

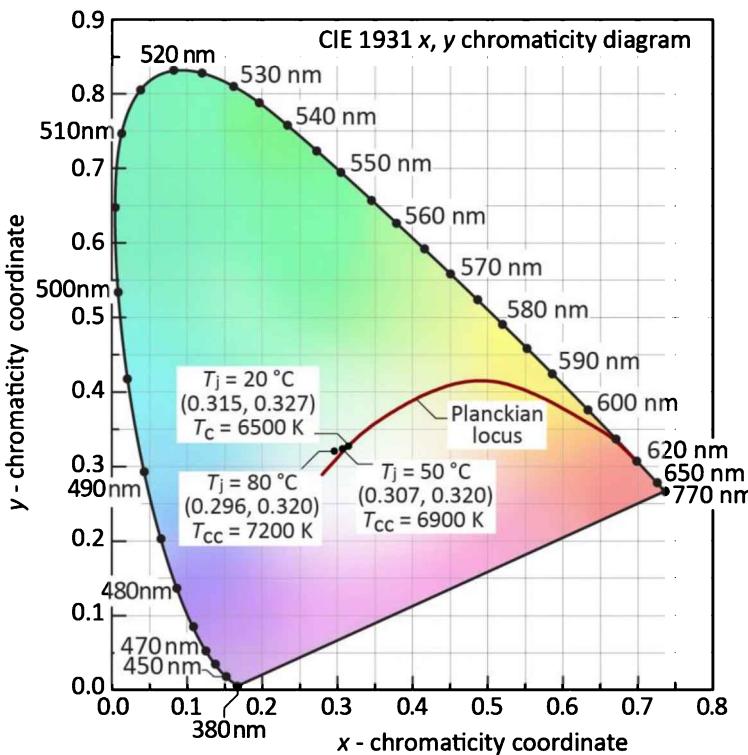


Fig. 28.11: Change in chromaticity of trichromatic white LED-based source. The source color temperature is 6500 K when devices are at room temperature. Due to the dependence of emission power, peak wavelength, and linewidth on temperature, the chromaticity point migrates off the planckian locus as the device temperature increases (after Chhajed *et al.*, 2005).

Figure 28.12 shows the chromaticity shift of the trichromatic source on a magnified scale in the CIE 1931 (x, y) chromaticity coordinate system as well as in the CIE 1976 (u', v') uniform chromaticity coordinate system along with the planckian locus. At $T_j = 50^\circ\text{C}$, the chromaticity point is 0.009 units away from the original point, and at 80°C , it is shifted 0.02 units from the original point. This shift causes a clearly noticeable change in color appearance and exceeds the limit of 0.01 units (“**0.01 rule**”) commonly used in the lighting industry as an acceptable chromaticity variation (Duggal, 2005).

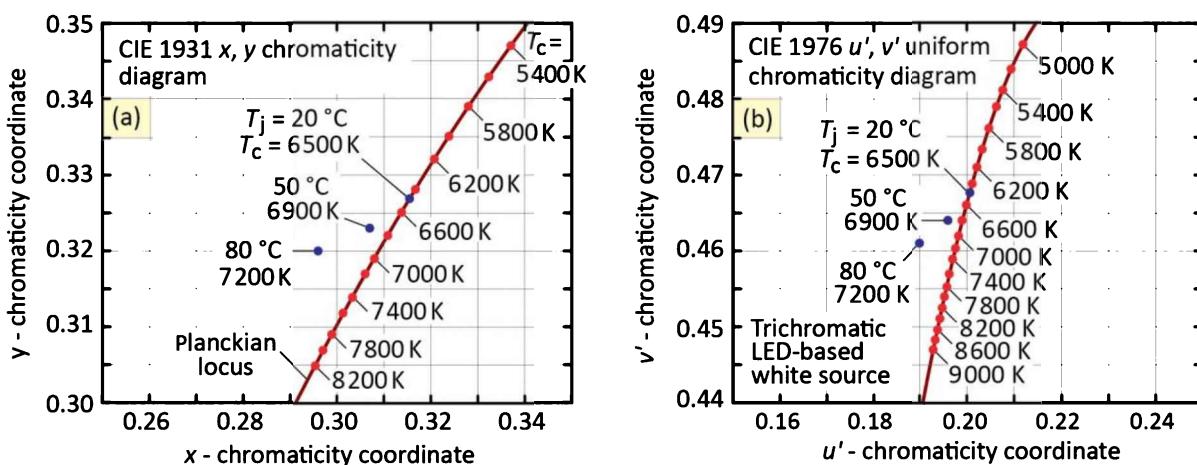


Fig. 28.12: Change in (a) x, y and (b) u', v' chromaticity of trichromatic white LED source. $T_c = 6500\text{ K}$ when p-n junctions are at room temperature (after Chhajed *et al.*, 2005).

The shift in chromaticity can be strongly reduced by adjusting the relative powers of the three LEDs. There are two possible implementations for adjusting the power ratio. In one implementation, the spectrum of the light source is constantly measured and a feedback control is used to adjust the optical power of the three components. In an alternative implementation, the device temperature is monitored and the optical power of the three components is adjusted using the known temperature dependence of the different types of emitters. The second method is generally easier to implement. However, the second method does not enable a compensation for device-aging effects.

28.5 Generation of white light by tetrachromatic and pentachromatic sources

Tetrachromatic and pentachromatic white sources use four and five types of LEDs, respectively (Zukauskas *et al.*, 2002a; Schubert and Kim, 2005). The color-rendering index of polychromatic sources generally increases with the number of sources. However, the luminous efficacy generally decreases with increasing number of sources. Thus, the color-rendering index and the luminous efficacy of tetrachromatic sources are generally higher and lower than those of trichromatic sources, respectively. However, the specifics depend on the exact choice of the emission wavelengths. Due to the great number of wavelength choices, the color temperatures of tetrachromatic and pentachromatic sources can be adjusted more liberally without compromising the color-rendering capability of the source.

28.6 References

- Bergh A., Crawford G., Duggal A., and Haitz R. "The promise and challenges of solid-state lighting" *Physics Today* p. 42 (December 2001)
- Chhajed S., Xi Y., Li Y.-L., Gessmann T., and Schubert E. F. "Influence of junction temperature on chromaticity and color rendering properties of trichromatic white light sources based on light-emitting diodes" *Journal of Applied Physics* **97**, 054506 (2005)
- Conrad "SMD LED RGB" < <http://www.conrad.de/ce/de/overview/0212266/LEDs-SMD> > accessed November (2014)
- Duggal A. R. "Organic electroluminescent devices for solid-state lighting" in *Organic Electroluminescence* edited by Z. H. Kafafi (Taylor & Francis Group, Boca Raton, Florida, 2005)
- Dalmasso S., Damilano B., Pernot C., Dussaigne A., Byrne D., Grandjean N., Leroux M., and Massies J. "Injection dependence of the electroluminescence spectra of phosphor free GaN-based white light emitting diodes" *Physica Status Solidi (a)* **192**, 139 (2002)
- Guo X., Graff J. W., and Schubert E. F. "Photon-recycling semiconductor light-emitting diodes" *IEDM Technical Digest, IEDM-99*, 600 (1999)
- Ivey H. F. "Color and efficiency of luminescent light sources" *Journal of the Optical Society of America* **53**, 1185 (1963)

- Li Y.-L., Gessmann Th., Schubert E. F., and Sheu J. K. "Carrier dynamics in nitride-based light-emitting p-n junction diodes with two active regions emitting at different wavelengths" *Journal of Applied Physics* **94**, 2167 (2003)
- Schubert E. F. and Kim J. K. "Solid-state light sources becoming smart" *Science* **308**, 1274 (2005)
- Sheu J. K., Pan C. J., Chi G. C., Kuo C. H., Wu L. W., Chen C. Chang H., S. J., and Su Y. K. "White-light emission from InGaN-GaN multiquantum-well light-emitting diodes with Si and Zn co-doped active well layer" *IEEE Photonics Technology Letters*, **14**, 450 (2002)
- Thornton W. A. "Luminosity and color rendering capability of white light" *Journal of the Optical Society of America* **61**, 1155 (1971)
- Wyszecki G. and Stiles W. S. *Color Science – Concepts and Methods, Quantitative Data and Formulae* 2nd edition (John Wiley and Sons, New York, 1982)
- Zukauskas A., Shur M. S., and Gaska R. *Introduction to Solid-State Lighting* (John Wiley and Sons, New York, 2002a)
- Zukauskas A., Vaicekauskas R., Ivanauskas F., Gaska R., and Shur M. S. "Optimization of white polychromatic semiconductor lamps" *Applied Physics Letters* **80**, 234 (2002b)

29

29 – Phosphors

The word “phosphor” originates from the Greek word “phosphoros” which means “light bearer” (“phosphor” should not to be confused with the chemical element P, phosphorus). A phosphor is a material that, when excited by primary optical radiation (λ_1), responsively emits secondary optical radiation at a different wavelength (λ_2). Although a phosphor, in the context of LEDs, usually is a solid inorganic material, a phosphor can be of any type (inorganic or organic) of any phase state (solid, liquid, or gaseous state).

We distinguish between ***down-converting*** phosphors, where $\lambda_1 < \lambda_2$, and ***up-converting*** phosphors, where $\lambda_1 > \lambda_2$. For down-converting phosphors, one photon of higher energy is converted to one photon of lower energy. For up-converting phosphors, to satisfy energy conservation, *two* primary photons need to be absorbed to emit *one* secondary photon. Accordingly, the efficiency of up-converting phosphors increases with the photon density. Nevertheless, the efficiency of up-converting phosphors is usually limited to a few percent. In contrast, down-converting phosphors can attain efficiencies close to 100%. This chapter concerns down-converting phosphors.

29.1 Introduction to phosphors

The phenomenon of “luminescence”, illustrated in ***Figure 29.1***, can be categorized by the type of excitation causing the responsively emitted luminescence. We distinguish photo-luminescence or PL (caused by an optical radiation), cathodo-luminescence or CL (caused by an electron-beam or cathode-ray), electro-luminescence or EL (caused by an electrical current), and chemi-luminescence (caused by a chemical reaction). Accordingly, a phosphor emits photo-luminescence or PL.

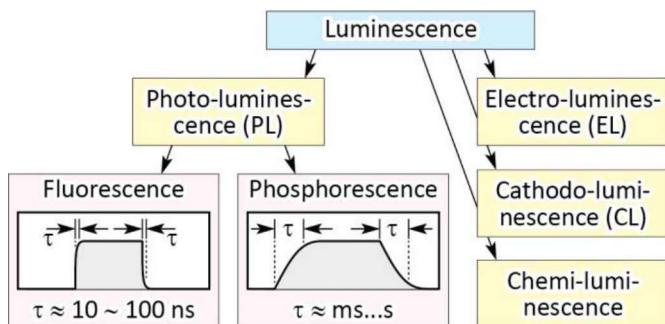


Fig. 29.1: Classification of luminescence into photo-luminescence (PL), electro-luminescence (EL), cathodo-luminescence (CL), and chemi-luminescence. PL is further categorized into fluorescence (short-lived, spin-allowed PL) and phosphorescence (long-lived, spin-forbidden PL). A phosphor material can emit either fluorescence or phosphorescence.

The radiative lifetime of an electron excited to a higher energy state in an atom can be short or long. The short-lifetime luminescence is called **fluorescence** with typical lifetimes of 10 ns to 100 ns. Fluorescence is due to spin-allowed electronic transitions. The long-lifetime luminescence is called **phosphorescence** with typical lifetimes of $> 100 \mu\text{s}$ to milliseconds to seconds. Phosphorescence is due to spin-forbidden electronic transitions resulting in a long lifetime. A phosphor material can emit either one, that is, a phosphor can emit either fluorescence or phosphorescence (Leverenz, 1949).

In the context of LEDs, which have high radiation intensities, phosphorescent materials would not be suitable since they tend to saturate at high radiation intensities. Accordingly, phosphors for LEDs preferably emit fluorescence. For example, the phosphor $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$, a fluorescent material, has a lifetime of 70~80 (Blasse and Bril, 1967).

Many inorganic phosphors have **activators**. An activator is an atom where the excited electron relaxes radiatively, i.e. by emission of a photon. Frequently, the activators are rare-earth atoms, specifically lanthanides, which are known to be optically active elements, such as cerium (Ce), europium (Eu), or erbium (Er). Furthermore, the radiative lifetime must be short (ns range) to enable high rate of transitions through the ground and excited states (a long radiative lifetime will lead to saturation at high excitation densities). In many cases, the activator itself absorbs the incoming primary radiation and responsively emits the secondary radiation.

Some phosphors employ **sensitizers**. They have the purpose of absorbing the primary radiation and shifting the absorption band to a desired wavelength range. Upon absorption, the excited electron transfers to the activator where it recombines. In some cases, a sensitizer can be a distinct atom such as manganese (Mn) (Kim *et al.*, 2014). In other cases, a sensitizer can be an oxygen (O) atom neighboring the activator atom. In yet other cases, the phosphor's host lattice acts as sensitizer. In yet other cases, the activator atom itself absorbs the incoming primary radiation so that there is no need for a sensitizer (e.g. in some organic, II-VI semiconducting, and

YAG:Ce phosphors). That is, phosphors can employ a range of activators and a range of sensitizers that are selected, based on the desired properties of the phosphor.

Activator and sensitizer are located in the **host lattice** of the phosphor. A phosphor host lattice has the following properties: First, the host lattice has a large bandgap energy so that it is optically transparent in the visible wavelength range. Second, the host lattice must possess a stable and stiff crystal structure; easily excited lattice vibrations in less stiff materials can lead to non-radiative relaxation, which decreases the phosphor efficiency. The requirement of large bandgap energy limits the host materials to compounds whose anion elements are selected from elements that are near the top of the columns of the periodic system of elements. In other words, phosphor host materials are typically selected from nitrides (top of column V), oxides and sulfides (top of column VI), fluorides (top of column VII), or combinations thereof.

The role of activators, sensitizers, and the host lattice is schematically illustrated in **Figure 29.2**. The general case is shown in **Figure 29.2 (a)**, whereas the more typical cases are shown in **Figure 29.2 (b)** and **(c)**.

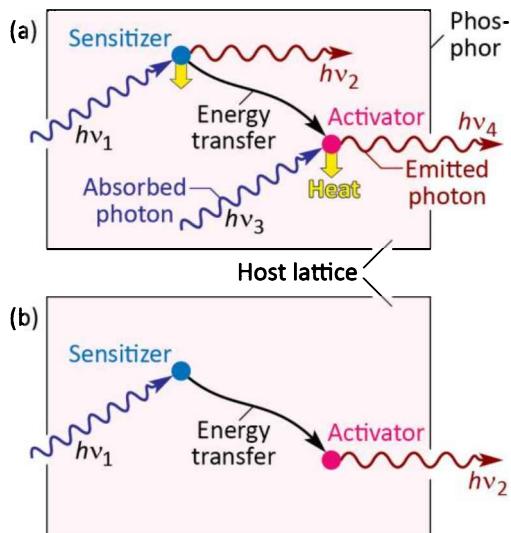


Fig. 29.2: (a) Phosphor having a sensitizer atom and an activator atom. Both, sensitizer and activator can absorb and emit radiation. The sensitizer also transfers energy to the activator. Some heat is generally generated as well. (b) Desired optical processes in a phosphor that uses a sensitizer and activator. Sensitizers allow one to shift the absorption wavelength to a desired value. (c) Desired optical processes in a phosphor that uses an activator but no sensitizer.

Common activator ions are shown in **Table 29.1** (OSU, 2003; Ronda and Srivastava, 2006; Cayless, 2012; Big Chemical Encyclopedia, 2016). Inspection of the table reveals that some activator elements (upper rows of table, e.g. Ce^{3+}) have broad emission bands whose peak wavelength depends on the host material, while other activator elements (lower rows of table, e.g. Er^{3+}) have narrow emission lines that do not depend on the host material. This difference can be explained by the configuration-coordinate model and crystal field theory, both to be discussed below.

Table 29.1: Common activator elements used in phosphors (photo-luminescent ions).

Ion and	Type	Excited State	Ground state	Emission color	λ_{peak}
Mn ²⁺ (3d ⁵)	Transition metal	t ₂ ⁴ e ¹ (⁴ T ₁)	t ₂ ³ e ² (⁶ A ₁)	Green-Orange-Red**	–
Sb ³⁺ (5s ²)	Group V	5s ¹ 5p ¹	5s ²	Blue**	–
Ce ³⁺ (4f ¹)	REE*	4f ⁰ 5d ¹	4f ¹ 5d ⁰	Near UV to red**	–
Eu ²⁺ (4f ⁷)	REE*	4f ⁶ 5d ¹	4f ⁷ 5d ⁰	Near UV to red**	–
Tm³⁺ (4f¹²)	REE*	¹G₄	³H₆	Blue	450 nm
Er³⁺ (4f¹¹)	REE*	⁴S_{3/2}	⁴I_{15/2}	Green	545 nm
Tb³⁺ (4f⁸)	REE*	⁵D₄	⁷F₅	Green	545 nm
Pr³⁺ (4f²)	REE*	³P₀	³H₅ (³F₂)	Red	605 (635) nm
Eu³⁺ (4f⁶)	REE*	⁵D₀	⁷F₂	Red	611 nm

* REE = Rare earth element

** The energy of transitions involving d, p, or s orbitals is very sensitive to the crystal field splitting induced by the lattice.

Cations are positively charged atoms whereas anions are negatively charged atoms. Lanthanide activators in common host materials (such as nitrides, oxides, sulfides, and fluorides) substitute for host-lattice cations (e.g. Y or Ca) and activators thus are also positively charged. The charge state of an atom is indicated right after the element's chemical symbol, so that Eu²⁺ is a doubly-ionized positive Eu ion and Eu³⁺ is a triply-ionized positive Eu ion.

The **configuration-coordinate model** explains several characteristics of phosphors (Leverenz, 1949). The model is based on the fact that the physical atomic configuration of the activator atom (e.g. the position of the activator atom) depends on the configuration and occupation of the electron shell. For example, if a valence electron (i.e. an electron forming a chemical bond with a neighboring atom) is excited to a higher level, then the atom will move to a new equilibrium position, that is, the physical configuration of the activator atom will change. The configuration-coordinate model is also based on the fact that changes on the electronic configuration (by excitation or de-excitation) are much faster than changes in the physical atomic configuration; this is due to the small mass of electrons when compared to the mass of the atom's nucleus.

As a result, we can visualize an ‘excitation – de-excitation sequence’ of an activator atom as follows: By means of optical excitation by a primary photon, an electron is excited from a “State 1” to a higher state, “State 2”. As a result of the change in electronic configuration, the activator atom will relax to a new, lower-energy equilibrium position (the atom assumes a new

physical atomic configuration) thereby converting some of the energy to heat. The relaxation of the atom occurs within about 1 ps (the period of thermal oscillation of the atom) while the average dwell time of the electron in State 2 is on the order of 10 to 100 ns (Leverenz, 1949). Finally, the electron radiatively transitions from State 2 to State 1. This sequence of events implies that the excitation energy (State 1 → State 2) necessarily is greater than the de-excitation energy (State 2 → State 1).¹³ That is, the secondary photon has a lower energy than the primary photon. The change in energy between excitation and de-excitation is known as the ***Stokes shift energy***. It is evident that the change in physical configuration is particularly strong if either State 1 or State 2 is a valence-electron state, since a change in valence electron certainly changes the atomic configuration of the activator atom. It is also evident that the change in physical configuration is weaker if State 1 and 2 are non-valence-electron states (e.g. orbitals lying closer to the nucleus of the atom than the valence-electron orbitals).

For example, if an optical emission transition of an activator atom is an intra-f-electron-shell transition with d-electrons forming the outer valences of the atom, then we expect a small effect of the host crystal on the energy of the transition. This is the case for, e.g., Nd³⁺ in Y₃Al₅O₁₂ or for Er³⁺ in Ca_{1-x}F_{2+x}:Er_x (Kramer and Boyd, 1981; Barriere *et al.*, 1994). If, on the other hand, an optical emission transition is a d-to-f-electron-shell transition with d-electrons forming the valences of the atom, then we expect a larger effect of the host crystal on the energy of the transition. This is the case for, e.g., Ce³⁺ in Y₃Al₅O₁₂.

Crystal Field Theory further explains the interaction between the activator atom and surrounding atoms of the phosphor's host lattice. The theory describes the breaking of degeneracies of electron orbital states, usually d or f orbitals, due to the static electric field produced by the surrounding charge distribution (electrons from anion neighbors). The coulombic interaction between an activator atom and host lattice atoms surrounding the activator atom includes attraction between the positively charged activator cation (activators usually are cations) and negative charge (electrons) of the nonbonding electrons of the anions of the host lattice (usually N, O, S, or F) as well as repulsion between electrons. Lanthanides, such as La, Ce, Nd, Eu, Gd, or Tb, which are cations in phosphors, have five degenerate d orbitals. As a cation moves towards one of the anions, bond-forming electrons in the intermediate region will

¹³ This holds even more true if a third state participates in the process, that is, if an electron is excited from a State 1 to a State 3 (1 → 3), then relaxes non-radiatively to a State 2 (3 → 2), and then relaxes radiatively to State 1 (2 → 1).

assume a higher energy due to coulombic repulsion. Similarly, as a cation moves away from one of the anions, bond-forming electrons in the intermediate region will assume a lower energy. As a result, the five degenerate d states split and form five non-degenerate d states having a range of energy. That is, the degeneracy of these orbital states is lifted so that these five levels split into five non-degenerate orbital states.

Electrons participating in an intra-f-electron-orbital transition of an activator atom are shielded from the crystal's electrostatic field by the closed outlying 5s and 5p shells; as a result, optical transitions due to intra-f transitions are spectrally narrow. On the other hand, the excited $4f^2\ 5d^1$ electron configuration produces broad spectral features, since the 5d electrons are not shielded from the crystal field (Kramer and Boyd, 1981).

29.2 Types of phosphor materials

There are different types of phosphors which may also be called “wavelength-conversion materials”. **Dyes** are organic phosphors. Many different dyes are commercially available.¹⁴ An example of a dye absorption and emission spectrum is shown in *Figure 29.3*.

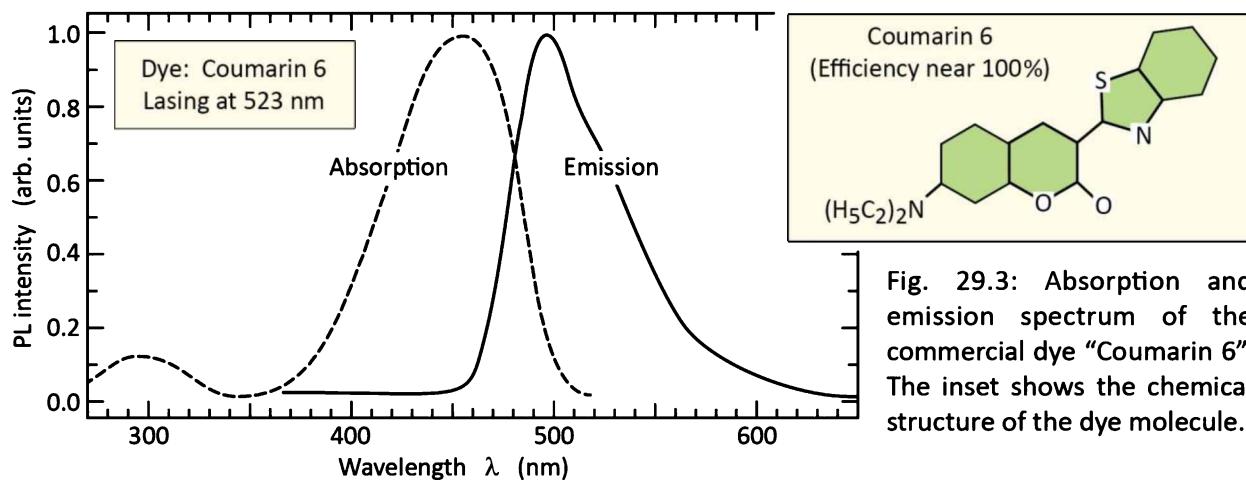


Fig. 29.3: Absorption and emission spectrum of the commercial dye “Coumarin 6”. The inset shows the chemical structure of the dye molecule.

Dyes can have quantum efficiencies > 90%. Dyes can be dissolved in transparent plastics such as acrylic glass, plexiglas, or PMMA and then are referred to as **pigments**. During the early development of white LEDs, organic dyes and pigments were considered by many research

¹⁴ Examples of commercially available dyes and pigments are the BASF Lumogen F Green 850, BASF Lumogen F Red 305, or BASF Lumogen F Yellow 083.

groups (for example Tokailin *et al.*, 1992). However, dyes, as organic molecules, lack the long-term stability afforded by inorganic phosphors. This is particularly true when the dye is subjected to the very high optical radiation intensity at the surface of an LED chip.¹⁵ As a result, dyes are rarely used in LED applications.

Semiconductors are another type of wavelength converter. Semiconductors are characterized by comparatively narrow emission lines with linewidths of the order of $2kT$. The spectral emission linewidth of semiconductors is narrower than the linewidth of many phosphors and dyes. Thus, semiconductors allow one to tailor the emission spectrum of a semiconductor wavelength converter with good precision.

A great variety of semiconductors is available. **Figure 29.4** shows elemental and binary compound semiconductors versus the semiconductor lattice constant. Using ternary or quaternary alloys, wavelength converters operating at virtually any visible wavelength can be fabricated. A number of semiconductor-based quantum-dot phosphors and nano-phosphors have been reported, including the II-VI semiconductor core/shell dots CdSe/ZnS and CdZnSe/CdZnS (Jang *et al.*, 2010).

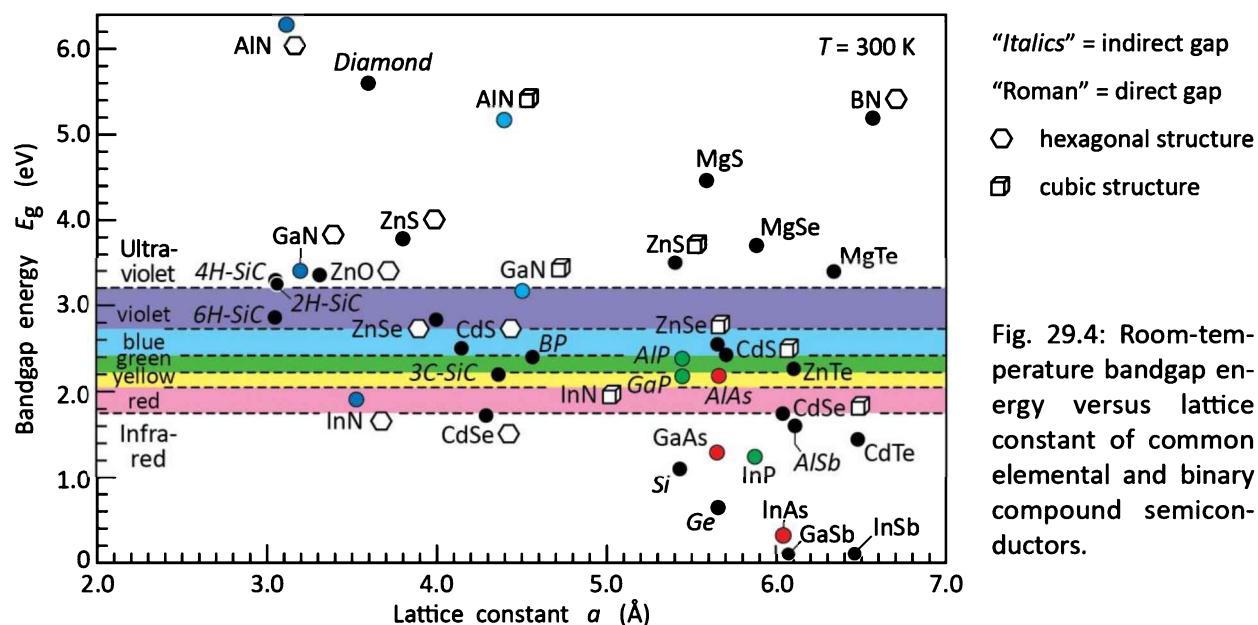
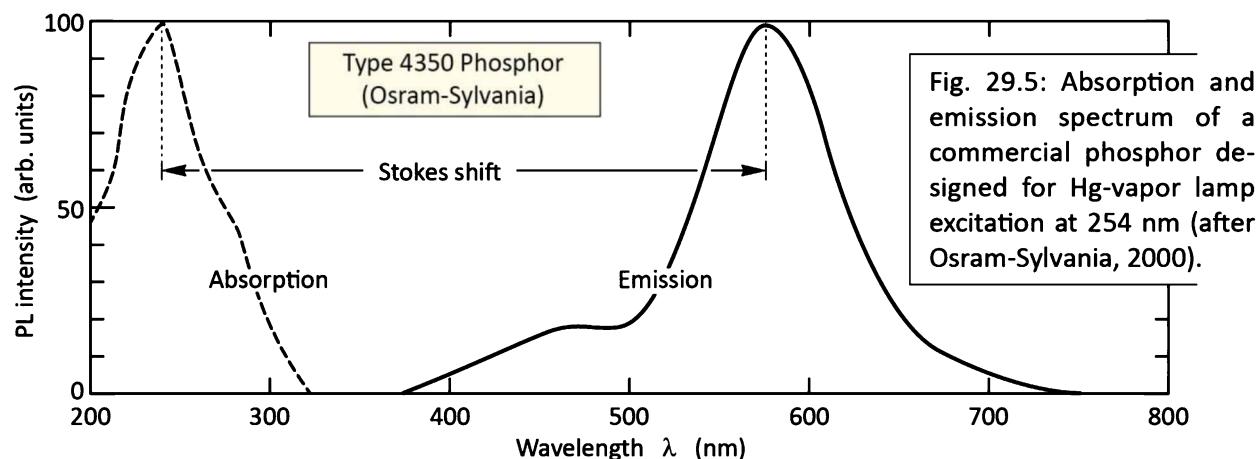


Fig. 29.4: Room-temperature bandgap energy versus lattice constant of common elemental and binary compound semiconductors.

¹⁵ The radiation intensity at the LED surface can exceed 500 kW/m². For comparison, the Sun's radiation intensity on Earth is about 1 kW/m².

Semiconductors can have internal quantum efficiencies > 90%. The light escape problem in semiconductor converters is less severe than it is in LEDs due to the fact that semiconductor converters do not need electrical contacts that could block the light.

Inorganic phosphors, in the form of powders, as used for LEDs, will be discussed in detail in the following section. Here, we will offer a short discussion of phosphors as used for the now-obsolete fluorescent lamps. The optical absorption and emission spectrum of a commercial phosphor, as had been used for a fluorescent lamp, is shown in *Figure 29.5*. The phosphor is excited by a mercury (Hg) vapor primary radiation source emitting in the deep UV range at 254 nm. The absorbed UV radiation is converted to white light by means of a phosphor that emits a broad band of wavelengths from 400 nm to 700 nm.



The key drawback of fluorescent lamps is apparent: The very large Stokes shift. By converting 254 nm photons to visible-range photons, about half of the primary photon's energy is lost. As a result, the efficiency of fluorescent lamps suffers from a significant setback, i.e. the enormous Stokes shift.

Advancements in the field of fluorescent lamps included the development of the Sb^{3+} and Mn^{2+} activated halo-phosphate phosphor with the chemical formula $\text{Ca}_5(\text{PO}_4)_3(\text{F}, \text{Cl}) : \text{Sb, Mn}$ in the 1940s (McKeag and Ranby, 1949; Johnson, 1961). The lighting industry underwent another revolution in the 1970s when a blend of three rare-earth ion containing phosphors activated by Eu^{2+} (blue), Tb^{3+} (green), and Eu^{3+} (red) was proposed for good color rendering, higher efficiency, and better color maintenance in fluorescent lamps (Ronda and Srivastava, 2006). However, color

quality of fluorescent lamps is inferior when compared directly to white LEDs.¹⁶

29.3 Phosphors used with LED chips

Most common phosphors used in LED applications are inorganic phosphors employing nitride, oxide, sulfide, and fluoride host materials that are doped with a rare-earth element (REE) specifically a lanthanide. The historically first type of phosphor used for white LEDs was $\text{Y}_3\text{Al}_5\text{O}_{12}$ doped with cerium, Ce (Bando *et al.*, 1996, 1998; Shimizu *et al.*, 1999; Nakamura and Fasol, 1997; Schlotter *et al.*, 1997, 1999). Since white LEDs fundamentally changed the field of lighting, they received widespread attention.

The host material of $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$ has a ***garnet crystal structure***, with the general chemical formula $\text{A}_3\text{B}_5\text{O}_{12}$ where A and B are cations and O is oxygen, the anion. Among the large group of garnets, yttrium aluminum garnet (YAG), $\text{Y}_3\text{Al}_5\text{O}_{12}$, is a particularly common host material. Phosphors having YAG as a host material are called ***YAG phosphors***. The optically active dopant is a rare-earth element, a rare-earth oxide, or another rare-earth compound. Most lanthanides are optically active. Light-emitting lanthanide elements include cerium (Ce) used in white LEDs, neodymium (Nd) used in lasers (Nd-doped YAG lasers), europium (Eu) used in LEDs, erbium (Er) used in optical amplifiers, and thorium (Th) oxide used in the mantle of gas lights.

Yttrium (Y) is a rare-earth element (REE) located in the same column as lanthanum (La). Lanthanides are REEs as well and thus naturally substitute for Y in $\text{Y}_3\text{Al}_5\text{O}_{12}$. Accordingly, the chemical formula of YAG:Ce can also be written as $\text{Y}_{3-x}\text{Ce}_x\text{Al}_5\text{O}_{12}$ where x is the molar fraction of Ce in YAG; x typically ranges from 1% to 10%,.

$\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$, whose emission spectrum is shown in ***Figure 29.6***, emits a broad band of wavelengths ranging from green to red so that, when properly combined with a blue LED chip, the resulting device emits white light. The $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$ phosphor is extraordinarily stable and can have efficiencies > 90%.

¹⁶ It has been realized that the CIE's 1995 color rendering index (*CRI*) is not suitable for comparing fluorescent lamps with white LEDs. Although respectable *CRIs* have been attained by fluorescent lamps, a side-by-side comparison with LED lamps reveals that the color quality of fluorescent lamps is significantly inferior to that of white LED lamps. See also "Colour rendering of white LED light sources" CIE 177:2007, ISBN-13: 978-3-901906-57-2 (2007).

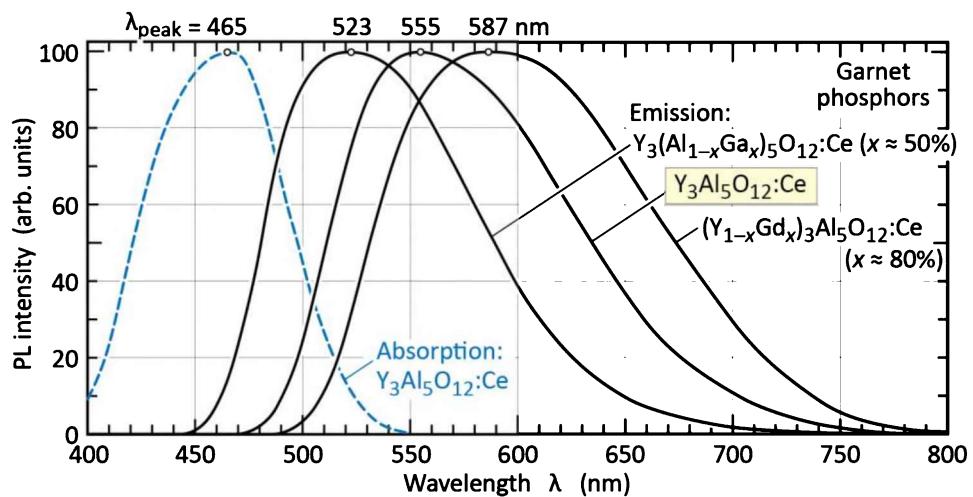


Fig. 29.6: Absorption for YAG:Ce and emission for YAG:Ce-based phosphors substituted with Ga and Gd (after Shimizu *et al.*, 1999; Mueller-Mach *et al.*, 2002; Bachmann, 2007).

The optical characteristics of YAG phosphors can be modified by partially substituting Gd for Y and Ga for Al so that the phosphor host has the composition $(Y_{1-x}Gd_x)_3(Al_{1-y}Ga_y)_5O_{12}$. The emission spectra for a Ce-doped $(Y_{1-x}Gd_x)_3(Al_{1-y}Ga_y)_5O_{12}$ phosphor with different compositions are shown in *Figure 29.6* (Shimizu *et al.*, 1999). The figure reveals that the addition of Gd shifts the emission spectrum to longer wavelengths, whereas the addition of Ga shifts the emission spectrum to shorter wavelengths.

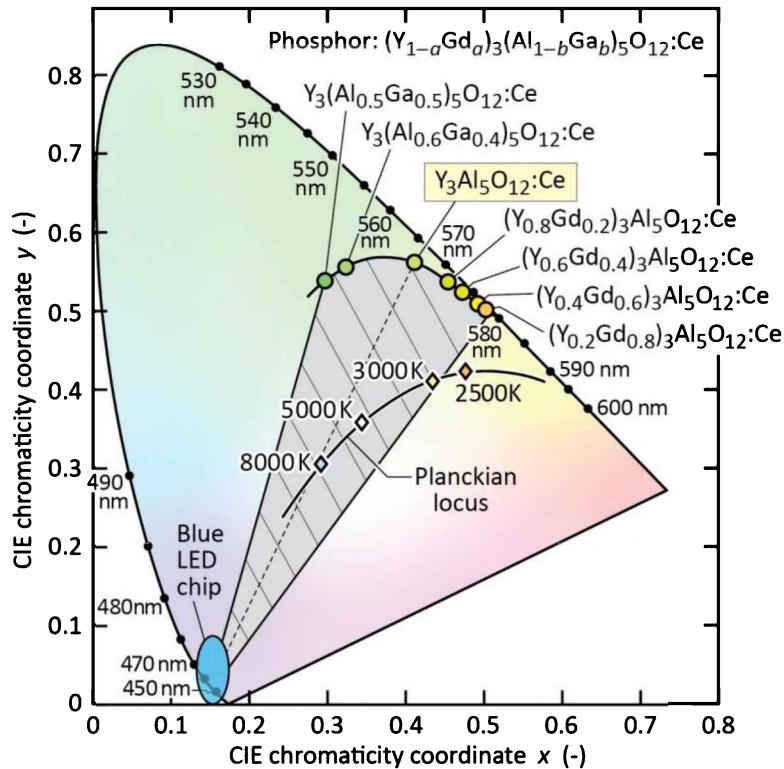


Fig. 29.7: Locations of YAG:Ce-based phosphors in the CIE chromaticity diagram. The base phosphor is pure yttrium aluminium garnet doped with cerium ($Y_3Al_5O_{12}:Ce$). The substitution of Ga for Al, and Gd for Y, shifts the phosphor's peak emission wavelength to shorter and longer wavelengths, respectively. By means of color mixing with a blue LED chip, a range of chromaticities can be created, as indicated by the hatched region, including white light on the Planckian locus with color temperatures ranging from about 2900 K to 8000 K (after Shimizu *et al.* 1999).

The chromaticity points of the YAG:Ce phosphors are shown in *Figure 29.7*. The shaded hatched region shows the chromaticities that can be attained by mixing light from a blue LED chip with

the light of Ga- and Gd-substituted YAG:Ce. The figure reveals that such emitters are suitable for white light with color temperatures ranging from about 3 000 K to 8 000 K.

An alternative to YAG phosphor is **TAG phosphor**, which is based on terbium aluminum garnet, $\text{Tb}_3\text{Al}_5\text{O}_{12}$ (Kummer *et al.*, 2003; Srivastava *et al.*, 2003). Emission spectra of TAG:Ce and YAG:Ce as well as of the alloy $(\text{Y}, \text{Tb})_3\text{Al}_5\text{O}_{12}:\text{Ce}$ are shown in **Figure 29.8** (Setlur *et al.*, 2004; Nazarov, 2005). The emission spectrum of TAG:Ce is similar but slightly red-shifted when compared to YAG:Ce. Both YAG and TAG crystallize in the garnet structure. Additionally, the Y^{3+} and Tb^{3+} ionic radii are very close ($r_{\text{Y}^{3+}} = 1.02 \text{ \AA}$ and $r_{\text{Tb}^{3+}} = 1.04 \text{ \AA}$). Consequently, the X-ray diffraction pattern of YAG does not strongly change as Y in YAG is substituted with Tb, even for Tb mole fractions of 30% (Potdevin *et al.*, 2005).

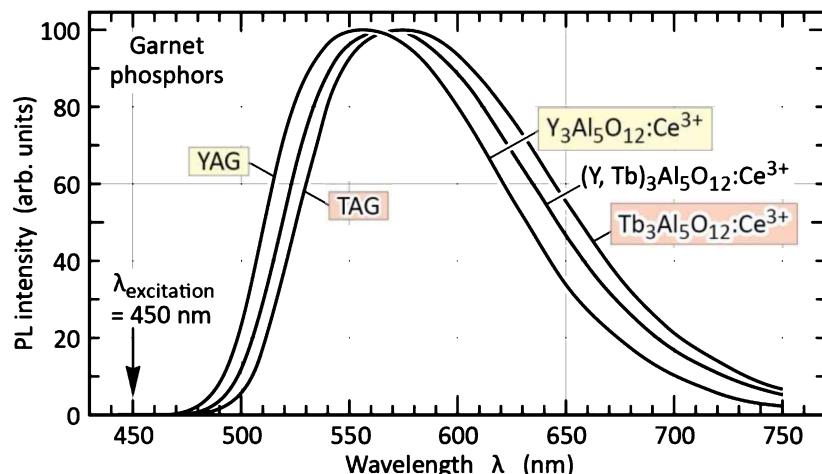


Fig. 29.8: Emission spectra of yttrium aluminium garnet (YAG), terbium aluminium garnet (TAG), and yttrium-terbium aluminium garnet, $(\text{Y}, \text{Te})_3\text{Al}_5\text{O}_{12}$, phosphors activated with Ce (after Setlur *et al.*, 2004, and Nazarov, 2005).

A systematic comparison between TAG:Ce and YAG:Ce was conducted by Chiang *et al.*, (2008). The authors found that the efficiency of TAG:Ce is somewhat lower than the YAG:Ce efficiency. The authors also found that TAG:Ce has a greater temperature dependence of the emission, i.e. a stronger thermal quenching behavior: At 150°C, YAG:Ce maintains about 90% of its room-temperature efficiency; in contrast, TAG:Ce maintains only about 60% of its room-temperature efficiency. Despite these drawbacks, TAG phosphors represent a viable alternative to YAG phosphors (Kim, 2005).

Examples of Eu-activator-doped sulfide-based phosphors are shown in **Figure 29.9** (Mueller-Mach *et al.*, 2002; Tian *et al.*, 2004). The emission of SrS activated with Eu is in the red spectral range. Sr (strontium) is a column-II metal (earth alkaline metal) and thus doubly positively

charged (Eu^{2+}). Accordingly, Eu^{2+} substitutes for Sr^{2+} which also is doubly positive charged. Ca (calcium), also an alkaline earth metal, located in the periodic system just above Sr, can substitute for Sr thereby forming the compound $\text{Sr}_{1-x}\text{Ca}_x\text{S}$. Due to the addition of Ca, the emission band shifts further into the red, as shown in *Figure 29.9*. By changing the host to SrGa_2S_4 (strontium thiogallate), the Eu activator emits in the green spectral range, as shown in the figure. Eu-activated red phosphors can be used in conjunction with YAG:Ce to enhance the red component of the phosphor.¹⁷

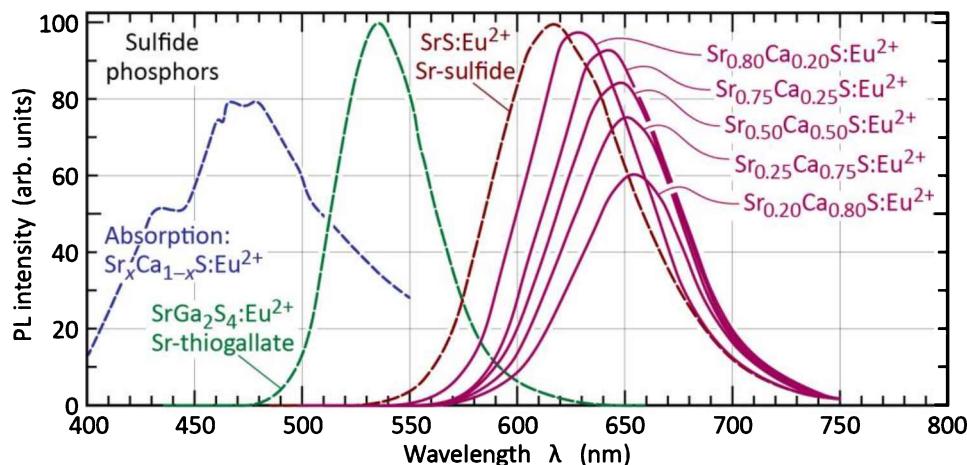


Fig. 29.9: Emission spectra of Eu-activated sulfide-based phosphors and excitation spectrum for $\text{Sr}_x\text{Ca}_{1-x}\text{S}: \text{Eu}^{2+}$ with $x \approx 1.0$ (after Mueller-Mach *et al.*, 2002, and Tian *et al.*, 2004).

Nitride host materials along with Eu or Ce activators can be efficient and stable phosphors. Particularly useful phosphor host materials are CaAlSiN_3 occasionally abbreviated as CASN (Uheda *et al.*, 2006; Piao *et al.*, 2007; Li *et al.*, 2008), SrCaAlSiN_3 occasionally abbreviated as SCASN (Kishimoto *et al.*, 2011), and SrCaSiN_3 occasionally abbreviated as SCSN, or as SCESN when doped with Eu (Yamada *et al.*, 2003). For a review of nitride phosphors for white LEDs, see Xie and Hirosaki (2007) and Setlur (2009). When doped with divalent Eu, the Eu is doubly positively charged (Eu^{2+}) and thus replaces the Ca^{2+} , so that the phosphor can be written as $\text{CaAlSiN}_3: \text{Eu}^{2+}$. The emission spectrum of this phosphor is shown in *Figure 29.10*. It shows a broad red emission band with peak wavelength of about 660 nm (Uheda *et al.*, 2006). Phosphor blends in white LEDs may contain $\text{CaAlSiN}_3: \text{Eu}^{2+}$, to enhance the red part of the emission spectrum.

¹⁷ YAG:Ce has a relatively weak emission in the red spectral range making it desirable to blend YAG:Ce with a red phosphor, e.g. SrS:Eu , especially for white LEDs with low color temperature ($\text{CT} < 3\,000\text{ K}$).

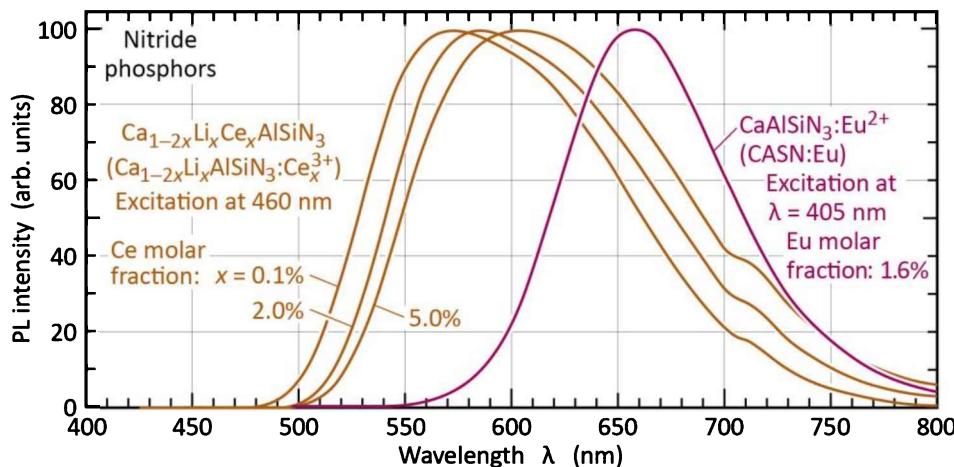


Fig. 29.10: Emission spectra of Eu-activated $\text{CaAlSiN}_3:\text{Eu}^{2+}$ red phosphor (CASN:Eu) and Ce-activated $\text{Ca}_{1-2x}\text{Li}_x\text{AlSiN}_3:\text{Ce}_x^{3+}$ yellow-orange phosphor (CLASN:Ce) (after Uheda *et al.*, 2006; Li *et al.*, 2008).

CaAlSiN_3 can also be doped with trivalent Ce, so that the Ce ion is triply positively charged (Ce^{3+}). Ca is an earth alkaline metal and thus a divalent cation (Ca^{2+}). For trivalent Ce^{3+} to substitute for the divalent Ca^{2+} , singly-valent Li^{1+} is added at the same amount as Ce^{3+} , so as to overall balance the valence electron count¹⁸ in the compound. It thus has the chemical formula $\text{Ca}_{1-2x}\text{Li}_x\text{AlSiN}_3:\text{Ce}_x^{3+}$, which can also be written as $\text{Ca}_{1-2x}\text{Li}_x\text{Ce}^{3+}\text{AlSiN}_3$ (Li *et al.*, 2008). Instead of adding Li, oxygen (O) can be added to substitute for the nitrogen (N) with the same motivation: Preservation of the valence electron count. The chemical formula of the phosphor then becomes $\text{Ca}_{1-x}\text{Ce}_x\text{AlSiN}_{3-2x/3}\text{O}_{3x/2}$ (Li *et al.*, 2008). The reasoning behind the addition of oxygen (O) is as follows: When subtracting 6 Ca^{2+} atoms and adding 6 Ce^{3+} atoms (i.e. when adding 6 valence electrons) then 4 N^{3-} are subtracted (minus 12 electrons) and 3 O^{2-} are added (plus 6 electrons) thereby making room for these 6 additional valence electrons. This electron-counting method ensures that the resulting compound can indeed form a stable chemical compound. Such compounds, i.e. compounds containing oxygen and nitrogen, are referred to as oxynitrides. Kameshima *et al.* (2010) disclosed various compositions of an efficient Eu-activated red-emitting nitride phosphor including $\text{Ca}_{0.99}\text{AlSiB}_z\text{N}_{3+z}:\text{Eu}_{0.001}^{2+}$ and $\text{Ca}_w\text{Al}_x\text{Si}_y\text{B}_z\text{N}_{((2/3)w+x+(4/3)y+z)}:\text{Eu}^{2+}$. These phosphors can emit light in the red spectral range with peak wavelengths of 650 nm without using high Eu activator concentrations (Eu is more expensive than the other elements of the phosphor, i.e. Ca, Al, Si, B, and N).

¹⁸ “Valence electrons” are those electrons that form the valences, i.e. form the chemical bonds.

Figure 29.10 also shows the emission spectrum of the $\text{Ca}_{1-2x}\text{Li}_x\text{AlSiN}_3:\text{Ce}_x^{3+}$ phosphors for Ce (and Li) molar concentrations of 0.1, 2.0 and 5.0%. Depending on the Ce (and Li) content, it is a yellow to orange emitting phosphor.

Silicate phosphors, when activated by Ce^{3+} or Eu^{2+} , are yellow phosphors emitting a broad band of wavelengths typically ranging from the green to the red spectral range. The phosphors absorb in the blue spectral range and are thus suited for white light generation in conjunction with blue LED chips. **Figure 29.11** shows $\text{Sr}_3\text{SiO}_5:\text{Eu}^{2+}$ (Eu-activated strontium silicate) with a peak wavelength of about 575 nm and emitting yellow light suited for white LEDs and as an alternative to YAG:Ce phosphors (Park *et al.*, 2004). The figure also shows the absorption and emission spectrum of $\text{Sr}_{3-2x}\text{SiO}_5:\text{Ce}_x^{3+}, \text{Li}_x^{1+}$ or simply $\text{Sr}_3\text{SiO}_5:\text{Ce}^{3+}, \text{Li}^{1+}$ (Ce-activated strontium silicate). The phosphor absorbs blue light, emits from green to red wavelengths, and has a peak wavelength of about 530 nm (Jang and Jeon, 2007). The figure finally shows the emission spectrum of $\text{Sr}_{2-x}\text{Ba}_x\text{SiO}_4:\text{Eu}$ (Eu-activated strontium barium silicate) emitting wavelengths from green to red and appearing yellow (Yoo *et al.*, 2005). By changing the Ba content of the phosphor, the peak wavelength can be tuned from 520 nm to 560 nm.

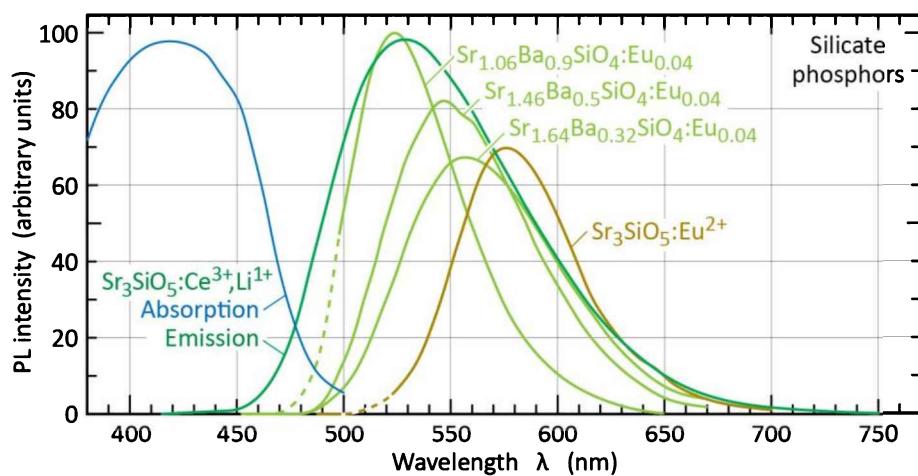


Fig. 29.11: Absorption and emission spectra of silicate phosphors: $\text{Sr}_3\text{SiO}_5:\text{Ce}^{3+}, \text{Li}^{1+}$ yellow phosphor (after Jang and Jeon, 2007), $\text{Sr}_3\text{SiO}_5:\text{Eu}^{2+}$ yellow phosphor (after Park *et al.*, 2004), and $(\text{Sr},\text{Ba})_2\text{SiO}_4:\text{Eu}^{2+}$ yellow phosphor (after Yoo *et al.*, 2005).

29.4 Phosphor body color

The body color of a phosphor is its color under daylight illumination. **Figure 29.12** shows the body colors of various phosphors (Intematix Company, 2014). Inspection of the figure reveals various body colors that can be described by, for example, the ordinary colors such as yellow, orange, or red.



Fig. 29.12: Phosphors having various body colors. The body color of rare-earth-doped phosphors generally tracks the emission color, although small differences between body color and emission color may exist. For example, YAG:Ce has a yellow body color and emits greenish-yellow light. In other cases, such as II-VI semiconductor phosphors, the body color and the emission color have no such close correlation (after Intematix Company, 2014).

The body color, or more specifically its chromaticity point, correlates with the chromaticity point of the emission spectrum. That is a yellow-light emitting phosphor generally appears yellowish. However, the correlation is not perfect. That is, a yellow-light-emitting phosphor can look orange, i.e. have an orange body color.

Inspection of LEDs or LED lamps allows one to derive some knowledge about a phosphor. For example, comparing *Figure 29.13 (a)* and *(b)* reveals that a phosphor with a yellow body color (see *(a)*) emits higher-color-temperature light than a phosphor with an orange body color (see *(b)*) (Cree Company, 2015). The phosphor shown in *(c)* is orange and the lamp's color temperature accordingly is low (2 700 K). The lamp shown in *(d)*, called the “L-Prize Lamp”, uses a (i) blue LED chip, (ii) a yellow phosphor, and (iii) a red LED chip. Accordingly, the lamp is capable of emitting a low-color-temperature white light (Philips Lumileds Company, 2013).

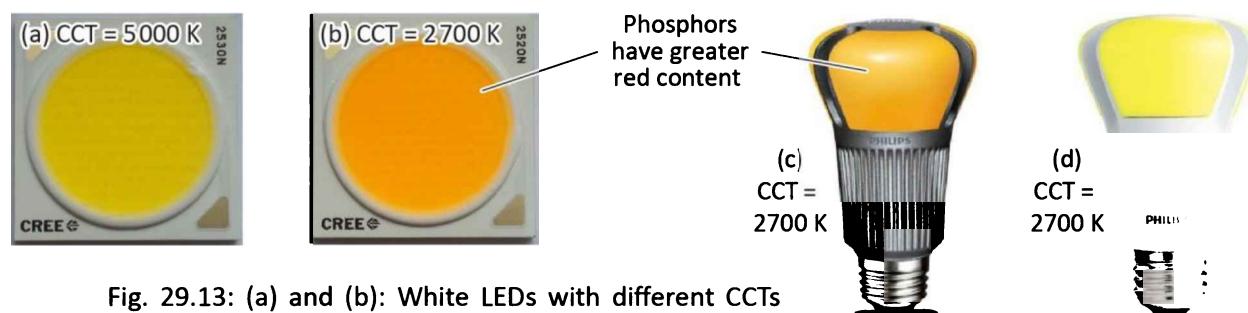


Fig. 29.13: (a) and (b): White LEDs with different CCTs and different phosphor blends. The lower-CCT LED uses a phosphor having a more reddish body color, indicating stronger red emission (after Cree Company, 2015). (c) and (d): Remote-phosphor white LED lamps with low CCT. To attain a low CCT, the lamps shown in (c) and (d) use a red-phosphor-containing phosphor blend and a red LED, respectively (after Philips Company, 2013).

29.5 Phosphor structure and synthesis

In this section, we restrict ourselves to phosphors that are activated by a rare earth element, specifically a lanthanide such as Ce or Eu. Host materials hosting activators must be optically transparent and possess good chemical, optical, and physical strength. The **cations** of such host structures should be elements for which the lanthanide activators can substitute. This includes elements that are close to lanthanum (La) in the periodic system of elements such as Ca (calcium), Sr (strontium), Sc (scandium), and Y (yttrium). A lanthanide (such as Ce or Eu) added to the host material will substitute for the cation. The host material of the phosphor also must be optically transparent, that is, have a large bandgap energy (E_g). Therefore, **anions** of the host structures are to be selected, due to the requirement of a large E_g , from the elements located near the top of the anion columns of the periodic system and therefore include N (nitrides), O (oxides), S (sulfides), and F (fluorides).

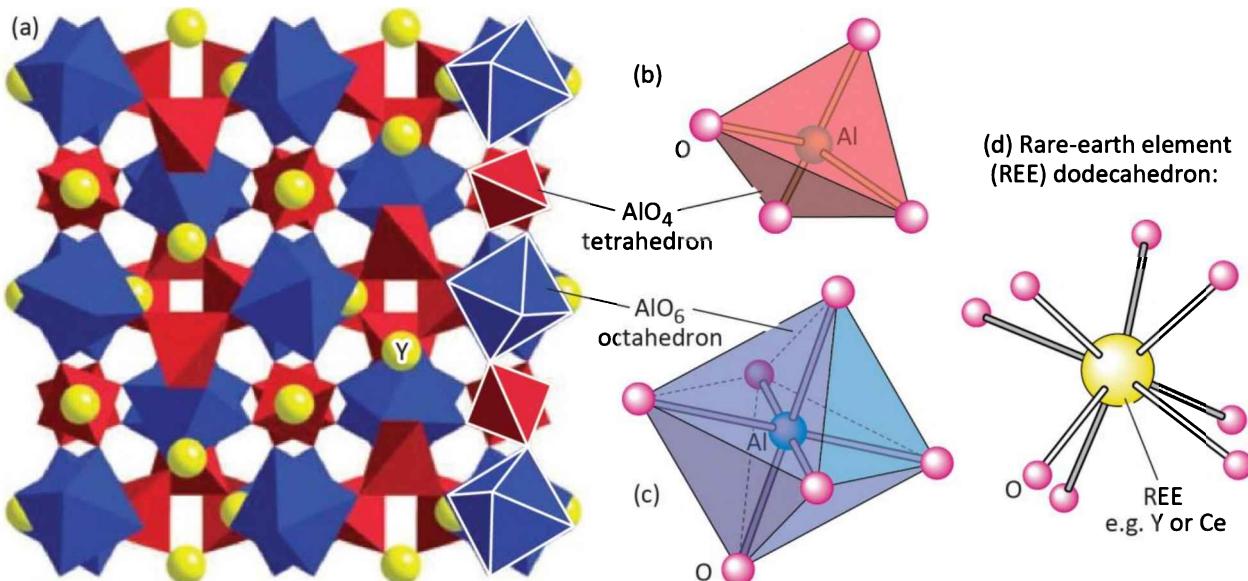
An additional requirement is the physical stability of the phosphor host material. *First*, the phosphor is subjected to the very high radiation intensities occurring at the surface of a blue LED chip.¹⁹ The high radiation intensities along with the heat and potential moisture illustrate the harshness of the conditions under which the phosphors must operate. *Second*, phosphor host materials must be stiff compounds. Easily excited lattice vibrations in less stiff materials can lead to non-radiative relaxation, which decreases the phosphor efficiency. Given these preliminary remarks, let us look at crystal structures of phosphors.

As an exemplary substance, we consider **aluminates**, specifically a **garnet phosphor**, yttrium aluminium garnet, $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$ or simply YAG:Ce. The garnet crystal is a cubic crystal structure with the unit cell consisting of 64 cations, 96 anions (oxygen), and thus a total of 160 atoms (!) (Shackelford and Doremus, 2008; Geiger, 2103). It can thus be concluded that garnets have a highly complex crystal structure. The $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$ garnet structure and its building blocks are illustrated in **Figure 29.14** and consist of the following:

- AlO_4 tetrahedrons consisting of a four-fold coordinated (four-fold bonded) central Al atom and four O atoms located at the four corners of the tetrahedron (see **Figure 29.14 (b)**). The tetrahedrons are highly stable, rigid, and stiff.

¹⁹ A 1 mm² surface-emitting LED chip consuming 1 W of electrical power and having an efficiency of 50% will have an optical radiation intensity of $500 \text{ mW/mm}^2 = 500 \text{ kW/m}^2$ which is 500 (!) times higher than the radiation intensity of the Sun (1 kW/m^2) on Earth. This comparison is a powerful illustration of the harsh conditions that exist at the surface of an LED chip.

- AlO_6 octahedrons consisting of a six-fold coordinated (six-fold bonded) central Al atom and six O atoms located at the six corners of the octahedron (see **Figure 29.14 (c)**). The octahedrons are highly stable, rigid, and stiff.
- A framework structure consisting of the tetra- and octahedrons where the O corner atoms are shared between adjacent tetra- and octahedrons (see **Figure 29.14 (a)**). The framework structure is a construct of corner-linked octahedra (center site being the B-site) and tetrahedra, with a larger cation in an 8-coordinated site (the A-site, or dodecahedral site).
- Voids formed by the framework structure. The voids are occupied by Y atoms. The Y atoms are eight-fold coordinated forming a dodecahedron (12-surface polyhedron) (see **Figure 29.14 (d)**). The Y atoms find themselves surrounded by rigid tetrahedrons and octahedrons, yet have a relatively spacious environment and are thus not rigidly bound.
- Ce atoms. Due to the chemical similarity with Y, Ce substitutes for Y. Accordingly, the Ce atoms find themselves surrounded by rigid tetrahedrons and octahedrons, yet have a relatively spacious environment and are thus not rigidly bound. The Ce atom is subject to thermal motion, is subject to the crystal field, and the configuration of states participating in the optical emission is strongly influenced by the local bonding environment. Consequently, the emission of Ce is a broad band and not a narrow line.



Similar principles apply to ***nitride phosphors***, e.g. the $\text{CaAlSiN}_3:\text{Eu}^{2+}$ phosphor. Individual building blocks of the nitride phosphor are AlN_4 and SiN_4 tetrahedrons, as shown in **Figure 29.15 (a)**

(Uheda *et al.*, 2006). These building blocks form a framework of tetrahedrons that share nitrogen (N) atoms at their corner points. The framework of tetrahedrons consists of sheets, a “Sheet A” and a “Sheet B” with one of the sheets being shown in *Figure 29.15 (b)* (Uheda *et al.*, 2006). The sheets are interconnected by N atoms that are shared between tetrahedrons belonging to two adjacent sheets.

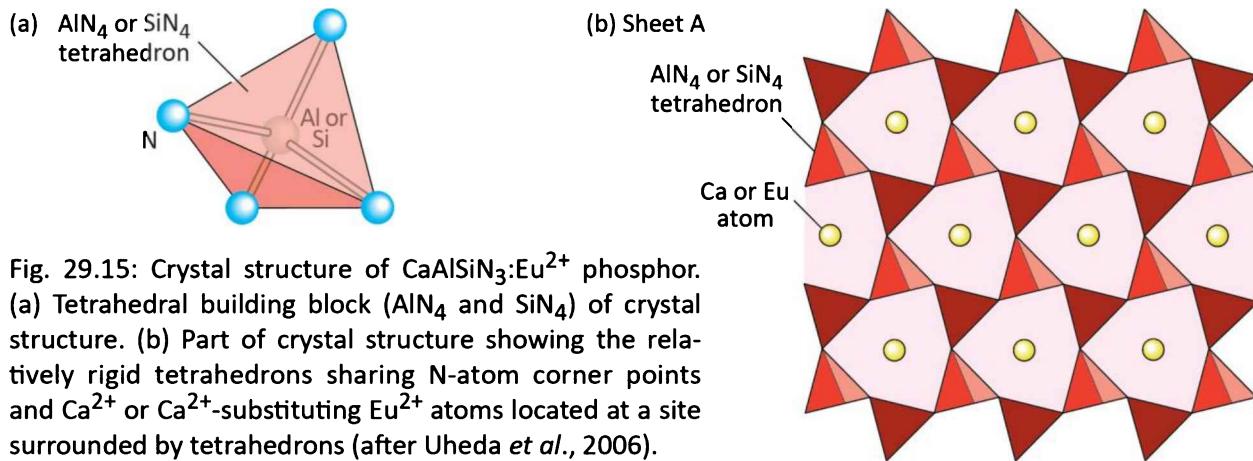


Fig. 29.15: Crystal structure of $\text{CaAlSiN}_3:\text{Eu}^{2+}$ phosphor. (a) Tetrahedral building block (AlN_4 and SiN_4) of crystal structure. (b) Part of crystal structure showing the relatively rigid tetrahedrons sharing N-atom corner points and Ca^{2+} or Eu^{2+} -substituting Eu^{2+} atoms located at a site surrounded by tetrahedrons (after Uheda *et al.*, 2006).

The framework of tetrahedrons forms relatively spacious voids that are occupied by Ca atoms. Ca is a doubly charged cation (Ca^{2+}). Eu, also being a doubly charged cation (Eu^{2+}), can substitute for Ca^{2+} and thus occupies the same site as Ca, i.e. a relatively spacious site with a relatively loose chemical bonding.

Silicate phosphors follow similar principles as discussed for garnet and nitride phosphors. The fundamental building block is the SiO_4 tetrahedron as shown in *Figure 29.16*. Multiple such tetrahedrons form a framework having voids for a larger anion such as Ca (calcium) and Sr (strontium). Ca and Sr have some chemical similarity to Y by being closely located to Y in the periodic system of elements. As a consequence, lanthanides such as Ce and Eu will substitute for the anion (Ca or Sr).



Fig. 29.16: The fundamental building block of silicate phosphors are SiO_4 tetrahedrons. Neighboring tetrahedrons share an O atom thereby forming a network of tetrahedrons. The lanthanide (e.g. Eu or Ce) are coordinated by multiple O atoms of the network of tetrahedrons.

Regarding the synthesis of phosphors and given that the above-discussed phosphors are highly stable materials, common synthesis methods employ high crystal growth temperatures > 1 000°C (Shionoya and Yen, 1999). As an example, we consider the synthesis of YAG:Ce as discussed by Shimizu *et al.* (1999). One synthesis method is ***solid-state synthesis***. In this method, powders of the oxides of elements forming the phosphor are the starting chemicals. For YAG:Ce phosphor, this includes yttrium oxide (Y_2O_3), aluminium oxide (Al_2O_3), and cerium oxide (CeO_2 and Ce_2O_3). The raw materials are mixed in stoichiometrical proportions using a crucible, mechanically compressed, and “fired” (crystallized) in an oven. The “firing” process may be carried out at high temperature ranging from 1 350 to 1 450°C and last for several hours (e.g. 2 to 5 hours). The ambient atmosphere during the “firing” process can be air or oxygen. The solid-state synthesis is relatively slow and may result in incomplete reactions since the starting materials are stable oxides, ceramics, which remain solid even at the high firing temperatures.

The disadvantages of solid-state synthesis can be addressed by using the ***flux-assisted solid-state synthesis***. The word “flux” implied the purpose of the flux, namely causing the “flow” of reagents. That is, the fundamental idea of flux-assisted synthesis is to locally liquefy the surfaces of the reaction’s starting materials so as to accelerate the reagents to form the equilibrium reaction product. It is well-known that chemical reactions proceed more readily in the liquid state than in the solid state, since the frequency of the various atoms or molecules bumping into each other is much higher in the liquid state than in the solid state.²⁰ The local liquefying of the starting materials is accomplished by employing a flux. The flux may be fluorides such as ammonium fluoride (NH_4F) (Shimizu *et al.*, 1999). Ammonia (NH_3), a large molecule, fluorine (F), and hydrogen fluoride (HF) do not readily incorporate into the garnet. Furthermore, fluorides and metal chlorides are usually water soluble so that the remnants of the flux can be washed away after the firing process.

The phosphor synthesis methods discussed above result in a “sintered cake” that forms in the crucible during the firing process. The “sintered cake” is first ground in a ball mill (or crushed by mortar and pestle) to form phosphor powder. The powder is then washed and dried. Additional steps can include the narrowing of the phosphor particle size distribution by sieving. The sieving

²⁰ Among some alchemists in the middle ages, it was believed that “chemical reactions occur in the liquid state only” (in latin “chemica reactionem in liquidus status solus”). Although this saying is of course not true in a strict sense, it is certainly true that the liquid state is much more amenable to chemical reactions than the solid state.

process allows one to eliminate phosphor particles with an undesirably large size (e.g. $> 30 \mu\text{m}$) or and undesirably small size (e.g. $< 5 \mu\text{m}$).

29.6 Quenching by temperature and activator concentration

There are two common phenomena in phosphors that relate to the **quenching** of the efficiency, i.e. the reduction in phosphor efficiency. The first phenomenon is **temperature quenching**, that is, the reduction of the phosphor efficiency at elevated temperatures.²¹ The second phenomenon is **concentration quenching**, that is, the reduction of the phosphor efficiency at high activator concentrations. Both phenomena will be discussed in this section.

Temperature quenching

The band diagram of a phosphor, YAG:Ce, shown in **Figure 29.17**, includes the conduction band and valence band of the host material, YAG, as well as relevant energy levels of the Ce activator (Ueda *et al.*, 2011). The excitation and de-excitation transitions of the Ce atom are also shown in the figure. At sufficiently high temperatures, an electron in an upper state of the Ce atom, the $5d_1$ and $5d_2$ state, may get excited to the conduction band. If this happens, the electron becomes delocalized and may diffuse to a defect site or the surface of the phosphor particle where non-radiative processes dominate. The process can be assumed to have a Boltzmann-type probability, i.e. $\exp{-(\Delta E/kT)}$, where ΔE is the energy it takes to elevate an electron from a Ce state to the bottom of the conduction band and T is the absolute temperature. As a consequence, the phosphor efficiency decreases with increasing temperature.

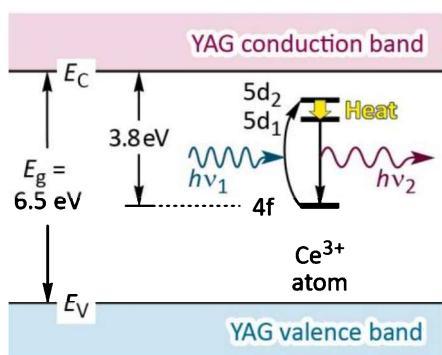


Fig. 29.17: Energy level structure of the Ce^{3+} atom in $\text{YAl}_5\text{O}_{12}$ (YAG). In equilibrium, the unexcited state, the 4f level of the Ce^{3+} atom, is occupied. Optical absorption transitions occur between the 4f and $5d_2$ level. Emission transitions occur between the $5d_1$ and 4f level. At higher temperatures, electrons may transition to the conduction band where they become delocalized and undergo non-radiative transitions at defect sites. That is, luminescence is “quenched” at high temperatures (after Ueda *et al.*, 2011).

²¹ For some phosphors, the quenching temperature may be below room temperature thereby making these phosphors unsuitable for room-temperature applications.

The thermal quenching characteristics of a phosphor will depend on multiple factors including the bandgap energy of the host material and the location of the Ce level within the bandgap of the host material. **Figure 29.18** shows the efficiency of various types of phosphors as a function of temperature (Setlur *et al.*, 2004; Bachmann *et al.*, 2009; Li, 2014). The efficiency is normalized to its room temperature (20°C) value. Inspection of the figure reveals that nitride phosphors (CASN:Eu) and garnet phosphors (YAG:Ce) show little quenching. At 100°C , the two types of phosphors maintain about 95% of their room-temperature efficiency; this is quite remarkable. On the other hand, yellow silicate phosphors exhibit a greater degree of quenching with the efficiency dropping to 87% and 32% at 100°C and 200°C , respectively. During operation, a phosphor temperature can exceed 100°C or even 150°C , particularly when phosphors are in close proximity to the LED chip. In order to keep a phosphor cool, it can be placed at a great distance from the LED chip (e.g. several mm or cm). Such remote phosphor distribution, commonly used for yellow silica phosphors, requires a greater quantity of phosphor but keeps the phosphor cooler (closer to room temperature).

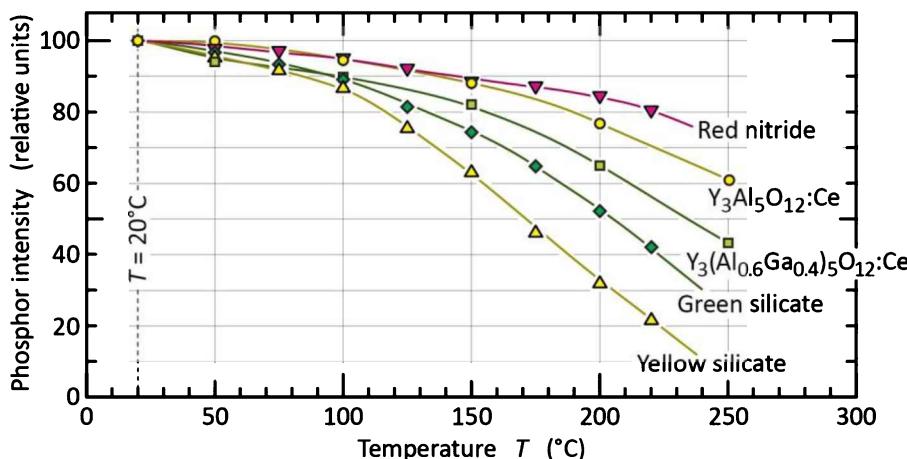


Fig. 29.18: Thermal quenching behavior of various phosphors. At $T = 20^{\circ}\text{C}$, the experimental data is normalized to 100% (after Setlur *et al.* 2004; Bachmann *et al.*, 2009; and Li, 2014).

Concentration quenching

Activator concentration quenching refers to the fact that at high activator concentrations, the quantum efficiency decreases ("quenches"). As the concentration of the activator increases, transport of carrier between activators is enabled. Carriers occupying activator states that are located in the vicinity of a defect will undergo non-radiative recombination. Therefore, the efficiency of a phosphor will decrease as the activator concentration increases. Typical optimum

concentrations of activators are in the range 1~5% (molar percentages). Concentration quenching typically starts at activator concentrations of about 5~10%.

An example of concentration quenching for CASN:Eu is shown in **Figure 19.19**. At zero Eu activator concentration, the efficiency naturally is zero. The efficiency reaches a maximum at a Eu molar concentration of 1.6%. At Eu concentrations > 5%, the efficiency drops substantially. At a Eu concentration of 20%, the efficiency has dropped to only about 11% of its maximum efficiency.

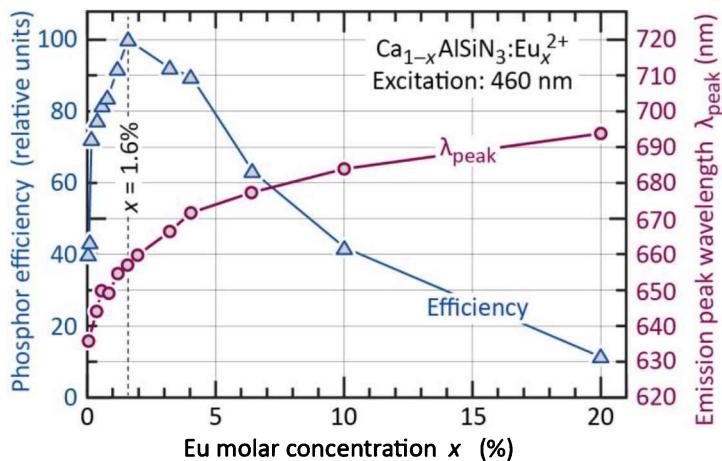


Fig. 29.19: Left ordinate: Relative efficiency of $\text{Ca}_{1-x}\text{AlSiN}_3:\text{Eu}_x^{2+}$ phosphor as a function of the Eu activator mole fraction. The decrease of the efficiency for high activator concentrations ($x > 1.6\%$) is known as *concentration quenching*. Right ordinate: Emission peak wavelength of the phosphor as a function of the Eu activator mole fraction (after Uheda *et al.* 2006).

One may note that the emission wavelength of a phosphor depends on the activator concentration. As a general principle, the emission wavelength shifts to longer wavelengths as the activator concentration increases. As shown in **Figure 29.19**, the emission peak wavelength of CASN:Eu shifts to longer wavelengths as the Eu activator concentration increases. That is, λ_{peak} shifts from about 540 nm to 694 nm as the Eu molar content increases from 0.2% to 20%. The about 50 nm shift is significant and will result in a noticeable color change of the phosphor. What is the origin of the wavelength shift? The question can be answered as follows: As the activator concentration increases, electron transport between activator atoms becomes possible. Some activators will be located in local *lower-energy* environments. Other activators will be located in local *higher-energy* environments. Electron transport will naturally proceed towards activators in local lower-energy environments so that a longer emission wavelength results.

Saturation at high excitation intensity

In principle, phosphor emission will saturate at high excitation intensity, specifically when the

majority of activators are in the excited state. The saturation can be avoided by using a high activator concentration and by selecting an activator that has a short radiative lifetime, e.g. 10~100 ns. In most applications that use fluorescent phosphors, saturation does not pose a problem.

29.7 Phosphor particle size

Phosphors consisting of nitride, oxide, sulfide, and fluoride host materials that are doped with a lanthanide activator, such as Ce or Eu, are almost always made in the form of a powder consisting of small phosphor particles with a typical diameter of 2~20 μm . A scanning electron micrograph (SEM) and a schematic illustration of a phosphor particle (with diameter d_{phosphor}) are shown in **Figure 29.20**.

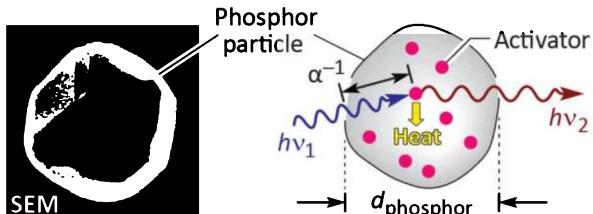


Fig. 29.20: Scanning electron micrograph (SEM) and illustration of phosphor particle with diameter d_{phosphor} absorbing a photon (energy $h\nu_1$) after absorption length α^{-1} , and emitting a photon (energy $h\nu_2$).

A key reason as to why phosphors are used in the form of powders consisting of small particles is the efficiency of the phosphor. Assuming that the optical absorption coefficient in a phosphor is α (at the wavelength of excitation), then the mean distance after which the exciting photon is absorbed is α^{-1} , i.e. the “mean absorption length”, or simply **absorption length**. Assuming that the same absorption coefficient applies to the emitted radiation, then the phosphor particle size should be on the order of a few absorption lengths (α^{-1}). If the phosphor-particle diameter would be much greater than the absorption length, e.g. $d_{\text{phosphor}} > 10 \alpha^{-1}$, then the light emitted from within the phosphor particle would likely be re-absorbed by the same phosphor particle. Multiple absorption processes within a phosphor particle must be avoided since the efficiency of the phosphor decreases with each absorption process. As a consequence, the phosphor particle size must not be too large.

On the other hand, as the phosphor particle size becomes smaller, the particle surface plays an increasingly large role.²² Since the surface structure, by its very nature, is non-periodic and

²² The surface-to-volume ratio of a particle, $S/V \approx 4\pi r^2 / ((4/3)\pi r^3) = 3/r$, increases with decreasing particle

thus defective, the efficiency decreases as the phosphor particle size decreases. As a consequence, the phosphor particle size must not be too small.

The optimum diameter of phosphor particles depends on a number of factors, including the chemistry of the host, activator element, and the activator concentration. Despite multiple factors playing a role, the optimum particle size typically ranges from 1 to 20 μm , frequently from 5 to 15 μm . For example, Höhn (2000) reported a garnet phosphor having particle sizes $< 20 \mu\text{m}$ with a median particle size $d_{50} < 5 \mu\text{m}$, e.g. $d_{50} = 2.5 \mu\text{m}$.

Besides efficiency, there are additional factors that are relevant with respect to phosphor particle size. These factors are:

- **Optical scattering:** As discussed elsewhere in this book, optical scattering by small particles includes geometric-optics scattering, Mie scattering, and Rayleigh scattering. Optical scattering causes a photon's propagation direction to be randomized and does not include wavelength conversion.²³ The **scattering length** may be defined as the mean distance after which light is scattered. Scattering increases with decreasing particle size. As a phosphor's particle size decreases, its tendency to scatter light (without wavelength conversion) increases.
- **Optical scattering versus wavelength conversion:** Depending on the concentration of the activator, the relative strengths of (i) optical scattering and (ii) wavelength conversion can be controlled. That is, a higher activator concentration results in more wavelength conversion whereas a lower activator concentration and smaller particle sizes result in more scattering.
- **Heat management:** As the phosphor particle size decreases, the surface-to-volume ratio changes. Heat is generated in the bulk of the particle whereas heat is dissipated through the surface area of the phosphor. Accordingly, small phosphor particles, with their higher surface-to-volume ratio, have more favorable heat-dissipation characteristics. This suggests that heating of phosphor particles is less of a problem with small phosphor particles. We note that by means of the phosphor operating temperature, the longevity (durability lifetime) of a phosphor will be influenced. That is, phosphors subject to high temperatures generally have a shorter durability lifetime.
- **Agglomeration:** Phosphor particles interact through their surfaces. Since the surface-to-volume ratio increases as the particle size decreases, agglomeration of phosphor particles becomes a problem. Agglomeration becomes more prevalent as the particle size decreases. Nano-particles usually have significant agglomeration problems. However, agglomeration can

size.

²³ Optical scattering is unrelated to wavelength conversion. During optical scattering, the photon's energy is preserved. However, the use of "scattering" is not always consistent in the technical literature. The word "scattering" is occasionally and incorrectly conflated with "wavelength conversion".

be reduced or eliminated by surface coating of the particles, for example with a thin-film shell of silicone or silica (SiO_2) surrounding the phosphor particle.²⁴

- **Sedimentation:** Large phosphor particles show a greater tendency towards sedimentation. That is, if the phosphor particles are dispersed in a transparent resin (such as silicone or epoxy) then large phosphor particles will be sinking in the resin more rapidly than smaller phosphor particles.²⁵

Scanning electron microscopy (SEM) images of phosphors are shown in **Figure 29.21**. Inspection of the SEM images reveals that the particle diameter varies from about 5 to 20 μm . If a phosphor is deposited in the form of a phosphor layer, then the layer must not be too thick so as to avoid self-absorption. An approximate practical rule is that the thickness of a densely packed phosphor layer should be about 1 to 3 times the particle size but not greater. For example, if particle size, d_{phosphor} , is 5 μm , the phosphor layer thickness may be 10 μm .

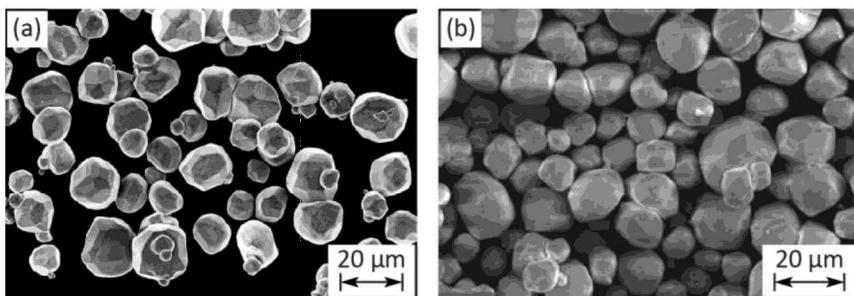


Fig. 29.21: Scanning-electron micrographs (SEMs) of phosphors: (a) $\text{Lu}_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$ (Solodovnyk *et al.*, 2015); (b) YAG:Ce LED phosphor (Intematix Company, 2014).

Phosphors consisting of small phosphor particles have a higher efficiency than a phosphor consisting of lumpy aggregations due to enhanced self-absorption in phosphors with a large particle size (Höhn *et al.*, 2000; Zachau *et al.*, 2008). It is therefore desirable to employ phosphors in the form of small particles or thin layers. It is also desirable to control the particle size so as to maintain a reasonably narrow particle size distribution. A particle size distribution (f) and cumulative size distribution (F) of a typical phosphor as used for white LED applications is shown in **Figure 29.22**. Inspection of the figure reveals that the distribution has a medium particle size of about 9 μm where $F = 50\%$. The distribution is relatively symmetric when drawn on a logarithmic abscissa scale but would become asymmetric when plotted on a linear abscissa scale. As a result, it would be difficult or impossible to describe the distribution by a Gaussian

²⁴ Phosphor particles may also be coated for surface passivation (passivation of surface states) so as to reduce the phosphor's degradation rate.

²⁵ For further discussion, see the Section on "Phosphor sedimentation" in this chapter.

distribution function. However, narrow phosphor particle distributions may be described by a Gaussian distribution function.

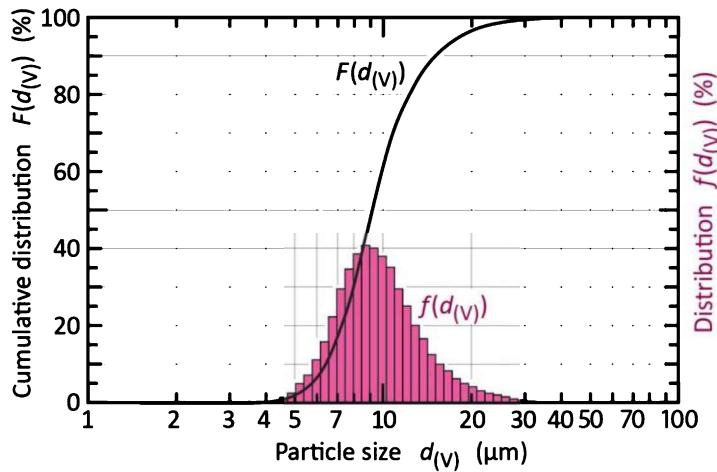


Fig. 29.22: Example of a phosphor particle size cumulative distribution, $F(d_{(V)})$, and distribution, $f(d_{(V)})$, as a function of $d_{(V)}$, where $d_{(V)}$ is diameter of a spherical particle that has the same volume as the phosphor particle. The median particle diameter, $d_{50(V)}$, is about 9 μm .

Phosphor particles are irregularly shaped particles and they are not spherical. Accordingly, a phosphor particle (just like a potato) does not have a unique diameter. To overcome this dilemma, a volume-based diameter, $d_{(V)}$, is defined: $d_{(V)}$ is the diameter of a spherical particle that has the same volume as the actual particle. **Figure 29.22** employs the volume-based particle diameter $d_{(V)}$. The **median particle diameter**, $d_{50(V)}$, is defined as follows: 50% of the particles have $d_{(V)} < d_{50(V)}$ and 50% of the particles have $d_{(V)} > d_{50(V)}$. That is, $d_{50(V)}$ is the size (diameter) in μm that splits the distribution in half, with half of the population residing below $d_{50(V)}$ and half of the population residing above $d_{50(V)}$ (Horiba, 2012). The phosphor distribution can be further defined by specifying the $d_{10(V)}$ and the $d_{90(V)}$ values; 10% of the particles have a diameter $< d_{10(V)}$; similarly, 10% of the particles have a diameter $> d_{90(V)}$.

How to control the phosphor particle size? The following factors influence and can be used to control the phosphor particle size:

- Phosphor synthesis conditions,
- Grinding conditions when powderizing the phosphor “sintered cake” after synthesis²⁶,
- Multiple sieving of the phosphor powder using sieves having different grid openings²⁷, and

²⁶ The grinding of the “sintered cake” converts it to a powder. The grinding can be accomplished by, e.g. using a ball mill. The size and mass of the balls as well as the milling time will determine the particle size of the powder.

²⁷ For example, a sieve having a grid opening of 10 μm will let particles with diameter $< 10 \mu\text{m}$ pass, while holding

- Sedimentation in a liquid so that light-weight particles float whereas heavy-weight particles sink.

How to measure the phosphor particle size distribution of a phosphor powder? The following techniques allow one to determine the particle size distribution of a phosphor (Horiba, 2012):

- Laser diffraction measurement,
- Dynamic light scattering (DLS), and
- Image analysis.

These measurement methods, when used with appropriate software support, can directly yield the particle size distribution including the $d_{50(V)}$ value.

29.8 Phosphor deposition

In the context of LEDs, phosphor can be deposited on an LED epitaxial wafer, on one or multiple LED chips, or on an LED package. In the interest of reducing manufacturing cost, the phosphor deposition process should be cost-efficient. One can distinguish between chip-level processes and wafer-level processes. A chip-level process is inherently serial whereas a wafer-level process is inherently parallel, implying a significant cost advantage for the wafer-level process. Therefore, we distinguish between chip-level and wafer-level deposition processes. For phosphor deposition on a single or on multiple chips, the following process can be used:

Dispensation: During the dispensation method, a suspension²⁸ of a phosphor powder in a polymeric resin (e.g. silicone or epoxy) is dispensed onto the surface of a chip or group of chips (Shimizu *et al.*, 1999). **Figure 19.23** shows examples of the dispensation process (Rahman and Jadwisienczak, 2014). A pipette with a nozzle opening is placed above the LED chip(s) and a quantity of the phosphor-resin suspension is dispensed. Sun *et al.* (2003) reported a volume of 500 nL (nano-liters) phosphor-resin suspension being dispensed with an accuracy of ± 68 nL onto an LED chip. The authors reported a 15% YAG:Ce phosphor volumetric loading, giving a net phosphor volume of 75 nL and a YAG:Ce mass of 342 μg .

back particles with diameter $> 10 \mu\text{m}$. Performing such sieving process repeatedly and with multiple grid sizes allows one to control, i.e. narrow, the phosphor particle distribution.

²⁸ A **suspension** usually is a liquid or viscous material that contains another material; an example is a phosphor-in-uncured-resin suspension. A **dispersion** usually is a solid material that contains another material; an example is a phosphor-in-cured-resin dispersion.

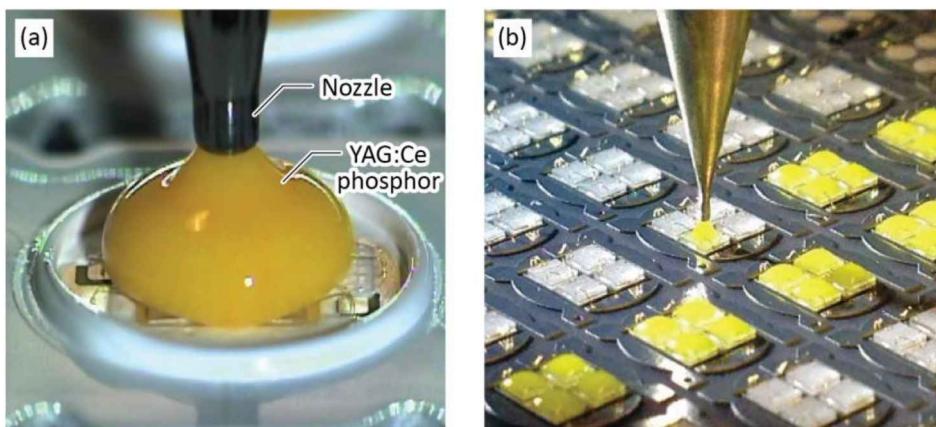


Fig. 29.23: Dispensation of resin containing yellow YAG:Ce phosphor by means of pipette having a dispensation nozzle for (a) chips-on-board (COB) LEDs and (b) surface-mount device (SMD) LEDs (after Rahman and Jadwiszak, 2014).

For phosphor deposition on an LED wafer, the following deposition processes have been practiced in the commercial manufacturing of LEDs:

Electrophoresis: In electrophoretic deposition, the phosphor is transported to the LED wafer surface by means of an electric field that is applied to a liquid phosphor-resin suspension that contains the phosphor powder (Grosso *et al.*, 1970; Collins *et al.*, 2003; Basin *et al.*, 2010; Wang *et al.*, 2010; Talbot, 2012). Electrophoretic deposition can provide uniform and densely packed phosphor layers. It can be applied repeatedly, thereby allowing for multiple layers of different types of phosphor layers.

Preform: Pre-forms of phosphors are layers of material containing the phosphor. The layer can be made of a phosphor ceramic plate (Bechtel *et al.*, 2008, 2010), or a solid film dispersion (e.g. cured silicone or epoxy containing the phosphor powder). The preform can be made to adhere to the LED wafer by means of an adhesive.

Spraying: In spray-deposition of phosphors, a suspension of a transparent resin (e.g. uncured silicone or epoxy) containing the phosphor is sprayed onto the wafer. The spraying process may be accompanied by a rotating wafer to improve uniformity of the phosphor.

Screen printing: Screen printing of an uncured phosphor-resin suspension is used. Screen printing allows for the masking of contact pads so that these pads are not coated by the phosphor-resin suspension. Subsequent to the printing process, the resin is cured to solidify (Yang *et al.*, 2013).

Spinning: Spinning can be used to deposit a phosphor-resin suspension onto a wafer. Subsequent to the spinning process, the resin is cured to solidify.

The processes listed above can be used with **phosphor blends** (a mixture of different types of phosphor). The processes can also be used sequentially for different types of phosphor, e.g. a yellow phosphor and a red phosphor. When using multiple phosphor layers, reabsorption processes should be minimized since each re-absorption process reduces the efficiency of the phosphor blend (Mueller and Mueller-Mach, 2007). To reduce or avoid reabsorption processes,

the different phosphors can be arranged in a manner shown in *Figure 29.24* where the red phosphor is deposited first, followed by the yellow phosphor. The configuration avoids re-absorption since the yellow phosphor will not absorb red light. Conversely, the red phosphor can absorb yellow light.

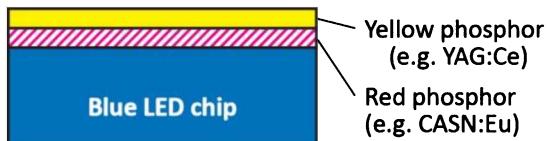


Fig. 29.24: Blue LED chip, red phosphor, and yellow phosphor arranged to avoid multiple absorption events.

With respect to spatial phosphor configurations, one generally distinguishes between ***proximate*** and ***remote phosphor configurations***. In proximate phosphor configurations, the phosphor is proximate to the LED (see, e.g. Shimizu *et al.*, 1999). In remote phosphor configurations, the phosphor is physically removed from the LED chip (Kim *et al.*, 2005). Remote phosphor configurations can reduce the operating temperature of the phosphor (thereby extending its durability lifetime) but require a greater quantity of phosphor. Remote phosphor configurations have been used in LED lightbulbs where the phosphor was applied to the lamp's bulb (see the images of LED lightbulbs discussed in a prior section of this chapter).

29.9 Phosphor sedimentation

An uncured phosphor-resin mixture is *liquid* or viscous and is referred to as phosphor-resin *suspension*. A cured phosphor-resin mixture is *solid* and is referred to as phosphor-resin *dispersion*. While in the liquid (or viscous) state, phosphor particles will, due to gravity, sediment (or sink) within the resin. The degree of sedimentation depends on multiple factors including the liquidity or viscosity of the uncured resin, time, temperature, the particle size of the phosphor,²⁹ and the specific gravity.³⁰

Figure 29.25 (a) and *(b)* shows a uniform and a sedimented phosphor distribution, respectively. It turns out that controlling the distribution of the phosphor can be quite advantageous. A controlled distribution is shown in *Figure 29.25 (c)*. Inspection of the figure reveals that many of the phosphor particles have become located directly on the LED chip by means of sedimentation (Wang *et al.*, 2014). Indeed, ***intentional sedimentation*** can be desirable:

²⁹ Generally speaking, large phosphor particles tend to sediment more rapidly than small phosphor particles.

³⁰ The specific gravity is defined as the ratio of the mass density (g/cm^3) of a substance to the mass density (g/cm^3) of a reference substance. For example, the specific gravity of YAG ($4.56 \text{ g}/\text{cm}^3$) in silicone ($1.25 \text{ g}/\text{cm}^3$) is 3.6.

First, by removing the phosphor from the resin-air surface, the phosphor is less prone to interact with moisture contained in air thereby enhancing the durability lifetime of the phosphor (Shimizu *et al.*, 1999). Second, by positioning the phosphor particles directly on the chip surface, heat generated by the phosphor particles can be conducted through the LED chip towards the heat sink (we note that the heat conductivity of a polymeric resin is a much lower than that of the chip). This allows the phosphor to efficiently dissipate its heat, operate at lower temperatures, and thereby enhancing its durability lifetime.

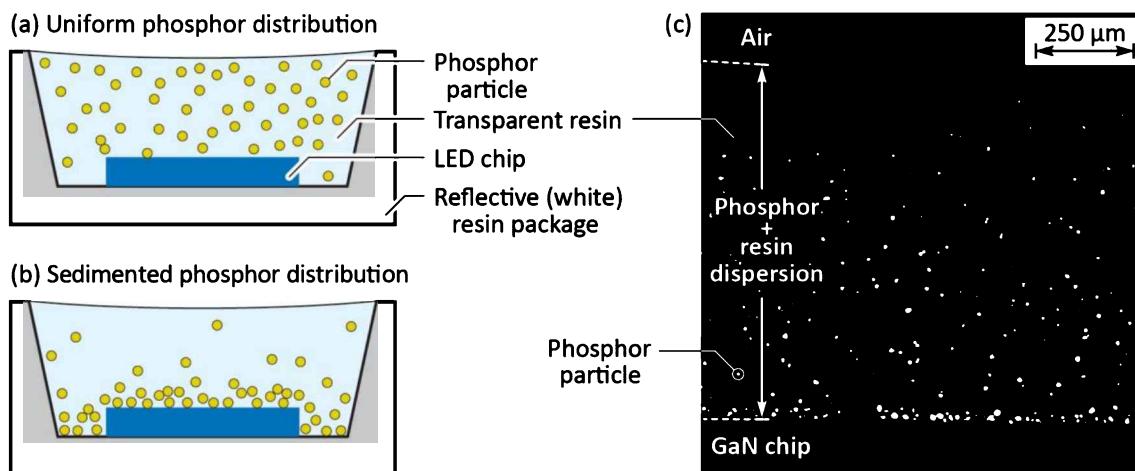


Fig. 29.25: (a) Uniform and (b) sedimented phosphor distribution in an optically transparent resin of an SMD LED. (c) Scanning electron micrograph of sedimented phosphor particles dispersed in a transparent resin (after Wang *et al.*, 2014).

Intentional sedimentation (also called “precipitation” or “settlement”) can be accomplished by different means, e.g. by using specific temperature profiles during the sedimentation and curing process (Shimizu *et al.*, 1999). For example, a phosphor-resin suspension may be heated up to make the suspension become more liquid so that sedimentation proceeds faster (higher temperatures will, however, also accelerate the curing process). Alternatively, a centrifuge can be used to force sedimentation.

29.10 Quantum-splitting phosphors

Quantum-splitting (or quantum-cutting) phosphors allow one to convert one short-wavelength photon into two longer-wavelength photons so that $h\nu_1 = h\nu_2 + h\nu_3$, where $h\nu_1$ is the energy of the photon absorbed by the phosphor and $h\nu_2$ and $h\nu_3$ are the energies of the photons emitted by the phosphors. Several quantum-splitting phosphors have been reported (Justel *et al.*, 1998; Wegh *et al.*, 1999; Srivastava and Ronda, 2003; Srivastava, 2004). The possibility of

quantum efficiencies approaching 200% for Eu^{3+} -doped LiGdF_4 has been proposed (Wegh *et al.*, 1999). The quantum-splitting phosphor $\text{YF}_3:\text{Pr}^{3+}$ at room temperature has a quantum efficiency of about 140% for 185 nm excitation (Justel *et al.*, 1998). However, viable quantum phosphors suitable for commercial applications have not yet been demonstrated.

29.11 Location of phosphors in the CIE chromaticity diagram

The light emitted by phosphors has a certain chromaticity and thus a location in the CIE chromaticity diagram. **Figure 29.26** shows the locations of various commercial phosphors in the CIE chromaticity diagram (Intematix Company, 2010). All phosphors are inorganic rare-earth-doped phosphors. The excitation range of the phosphors is in the short wavelength range (near UV, violet and blue) as is typically in phosphor-converted white LEDs. Inspection of the figure reveals that red phosphors are generally located *on the perimeter* of the chromaticity diagram.³¹

Green phosphors are generally located inside the perimeter of the chromaticity diagram.³²

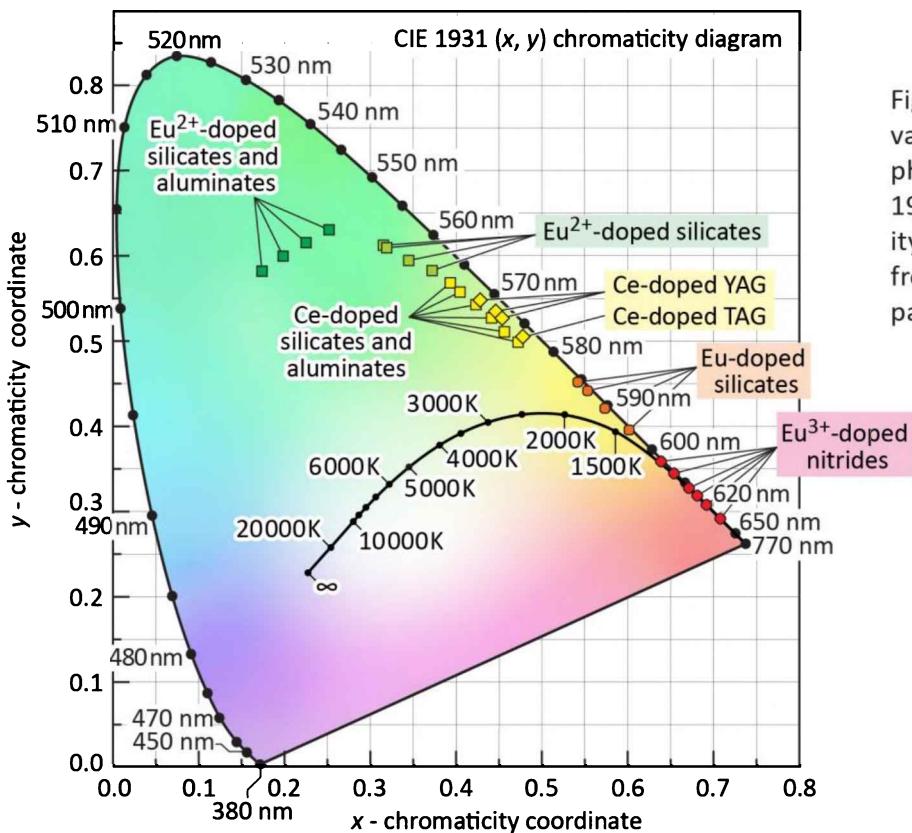


Fig. 29.26: Location of various commercial phosphors in the CIE 1931 (x, y) chromaticity diagram (adapted from Intematix Company, 2010).

³¹ Since the perimeter of the chromaticity diagram in the red region is a straight line, a broader spectral band will have a location on the perimeter.

³² Since the perimeter of the chromaticity diagram in the green region is strongly curved, a broader spectral band will have a location inside the perimeter.

29.12 References

- Bachmann V., Ronda C., and Meijerink A. "Temperature quenching of yellow Ce³⁺ luminescence in YAG:Ce" *Chemistry of Materials* **21**, 2077 (2009)
- Bando K., Noguchi Y., Sakano K., and Shimizu Y. "Development and application of high-brightness white LEDs" (in Japanese) *Technical Digest, Phosphor Research Society, 264th Meeting*, November 29 (1996)
- Bando K., Sakano K., Noguchi Y., and Shimizu Y. "Development of high-bright[ness] and pure-white LED lamps" The Illumination Engineering Institute of Japan – Special Issue "Lux Pacifica '97 3rd Pacific Basin Lighting Congress" *Journal of Light and Visual Environment* **22**, 2 (1998)
- Barriere A. S., Raoux S., Garcia A., L'Haridon H., Lambert B. and Moutonnet D. "Intra-4f-shell transitions at room temperature of Er³⁺ ions in Ca_{1-x}Er_xF_{2+x} thin films grown on Si (100)" *Journal of Applied Physics* **75**, 1133 (1994)
- Basin G., Martin P. S., West R. S., Morita Y., Heemstra T. "Blue LED with phosphor layer for producing white light and different phosphor in outer lens for reducing color temperature" US Patent 7,847,302 B2, issued December 7 (2010)
- Bechtel H., Schmidt P., Busselt W., Schreinemacher B. S. "Lumiramic™ – A new phosphor technology for high performance solid state light sources" *Proceedings of the SPIE* **7058**, 70580E-1 (2008)
- Bechtel H., Schmidt P. J., Tücks A., Heidemann M., Chamberlin D., Müller-Mach R., Müller G. O., Shchekin O. "Fully phosphor-converted LEDs with lumiramic™ phosphor technology" *Proceedings of the SPIE* **7784**, 77840W-1 (2010)
- Big Chemical Encyclopedia "Phosphors" Volume 15, Table 1: Common Activator Ions (Kirk Othmer Encyclopedia of Chemical Technology, 4th Edition), page 284 < <http://chempedia.info/page/042097093046027029110240084253243185203201099135/> > (2016)
- Blasse G. and Bril A. "A new phosphor for flying-spot cathode-ray tubes for color television: Yellow-emitting Y₃Al₅O₁₂-Ce³⁺" *Applied Physics Letters* **11**, 53 (1967)
- Cayless M.A. "Lamps and lighting" ISBN-13: 978-0340646182 (Routledge, London UK) page 138 (2012)
- Chiang C.-C., Tsai M.-S., and Hona M.-H. "Luminescent properties of cerium-activated garnet series phosphor: Structure and temperature effects" *Journal of the Electrochemical Society* **155**, B517 (2008)
- Collins III W. D., Krames M. R., Verhoeckx G. J., van Leth N. J. M. "Using electrophoresis to produce a conformally coated phosphor-converted light emitting semiconductor" US Patent 6,576,488 B2, issued June 10 (2003)
- Cree Company: Commercial chip-on-board LED products, such as the 19-mm CXA2530 LEDs < www.cree.com > (2015)
- Geiger C. A. "Garnet: A key phase in nature, the laboratory, and technology" *Elements* **9**, 447 (2013)
- Grosso P. F., Rutherford Jr. R. E., and Sargent D. E. "Electrophoretic deposition of luminescent materials" *Journal of the Electrochemical Society: Electrochemical Technology* **117**, 1456 (1970)
- Höhn K., Debray A., Schlotter P., Schmidt R., and Schneider J. "Wavelength-converting casting composition and its use" US patent 6,066,861, issued May 23 (2000)
- Horiba Instruments Company, Horiba Scientific "A guidebook to particle size analysis" < https://www.horiba.com/fileadmin/uploads/Scientific/Documents/PSA/PSA_Guidebook.pdf > (2012)
- Intematix Company: Phosphor Product Catalog (2010)
- Intematix Company: Commercial phosphors as shown on the Intematix Internet site < <http://www.intematix.com/products/led-phosphors> > (2014)
- Jang H. S. and Jeon D. Y. "Yellow-emitting Sr₃SiO₅:Ce³⁺, Li⁺ phosphor for white-light-emitting diodes and yellow-light-emitting diodes" *Applied Physics Letters* **90**, 041906 (2007)
- Jang E., Jun S., Jang H., Lim J., Kim B., and Kim Y. "White-light-emitting diodes with quantum dot color converters for display backlights" *Advanced Materials* **22**, 3076 (2010)
- Johnson P. D. "Some optical properties of powder and crystal halophosphate phosphors" *Journal of the Electrochemical Society* **108**, 159 (1961)
- Justel T., Nikol H., and Ronda C. R. "New developments in the field of luminescent materials for lighting

- and displays" *Angewandte Chemie* (International Edition) **37**, 3084 (1998)
- Kameshima M., Hosokawa S., Takashima S., Tamaki H., Shinohara T., Naitou T., and Kishimoto T. "Nitride phosphor, method for producing this nitride-phosphor, and light emitting device that uses this nitride phosphor" US Patent 7,854,859 B2, issued December 21 (2010)
- Kim S. J., Yun Y. J., Jung H.-K., and Choi S. "Color tuning and thermal quenching of different sensitizer ion, Mn²⁺ or Ce³⁺, doped Ba₂Mg(BO₃)₂:Eu²⁺ phosphor" *Optics Letters* **39**, 251 (2014)
- Kim J. K., Luo H., Schubert E. F., Cho J., Sone C., and Park Y. "Strongly enhanced phosphor efficiency in GaInN white light-emitting diodes using remote phosphor configuration and diffuse reflector cup" *Japanese Journal of Applied Physics* **44**, L649 (2005)
- Kishimoto K., Yokosawa Y., and Tomomura Y. "Illuminating device and vehicle headlamp" US patent application publication US 2011/0,248,624 (2011)
- Kramer M. A. and Boyd R. W. "Three-photon absorption in Nd-doped yttrium aluminum garnet" *Physical Review B* **23**, 986 (1981)
- Kummer F., Zwaschka F., Ellens A., Debray A., and Waitl G. "Luminous substance for a light source and light source associates therewith" US Patent 6,669,866 B1, issued December 30 (2003)
- Leverenz H. W. "Luminescent solids (phosphors)" *Science* **109**, 183 (1949)
- Li Y. Q., Hirosaki N., Xie R. J., Takeda T., and Mitomoet M. "Yellow-orange-emitting CaAlSiN₃:Ce³⁺ phosphor: Structure, photoluminescence, and application in white LEDs" *Chemistry of Materials* **20**, 6704 (2008)
- Li Y.-Q. "Phosphor manufacturing update from Intematix" *US Department of Energy 2014 Solid State Lighting Manufacturing R&D Workshop*, San Diego, California, USA, May 7~8 (2014)
- McKeag A. H. and Ranby P. W. "Alkaline earth halophosphate phosphors" US Patent 2,488,733; UK priority date June 17,1942; issued November 22 (1949)
- Mueller-Mach R., Mueller G. O., Krames M. R., and Trottier T. "High-power phosphor-converted light-emitting diodes based on III-nitrides" *IEEE Journal on Selected Topics in Quantum Electronics* **8**, 339 (2002)
- Mueller G. O. and Mueller-Mach R. B. "Wavelength converted semiconductor light emitting devices" US patent 7,250,715, issued July 31, 2007 (2007)
- Nakamura S., Fasol G. "The blue laser diode: GaN based light emitters and lasers" (Springer-Verlag, Berlin, Germany, 1997)
- Nazarov M. "Luminescence mechanism of highly efficient YAG and TAG phosphors" *Moldavian Journal of the Physical Sciences* **4**, 347 (2005)
- OSU, Ohio State Univ., "Optical Properties II: Emission of light, displays and transparent conductors" <https://chemistry.osu.edu/~woodward/ch754/lect2003/optical2_lect22.pdf> (2003)
- Park J. K., Kim C. H., Park S. H., Park H. D., and Choi S. Y. "Application of strontium silicate yellow phosphor for white light-emitting diodes" *Applied Physics Letters* **84**, 1647 (2004)
- Philips Lumileds Company: "Endura LED Lamp" and "L-Prize Lamp" commercial products <www.lumileds.com> (2013)
- Piao X., Machida K.-I., Horikawa T., Hanzawa H., Shimomura Y., and Kijima N. "Preparation of CaAlSiN₃:Eu²⁺ phosphors by the self-propagating high-temperature synthesis and their luminescent properties" *Chemical Materials* **19**, 4592 (2007)
- Potdevin A., Chadeyron G., Boyer D., Caillier B. and Mahiou R. "Sol-gel based YAG:Tb³⁺ or Eu³⁺ phosphors for application in lighting sources" (2005) *Journal of Physics D: Applied Physics* **38**, 3251 (2005)
- Rahman F. and Jadwisienczak W. "Phosphors: The driving force behind white LEDs" *Compound Semiconductor* <<http://www.compoundsemiconductor.net/article/94709phosphorsthedrivingforcebehindwhiteleds.html>> August 5 (2014)
- Ronda C. and Srivastava A. "Luminescence science and display materials" *The Electrochemical Society, Interface Spring 2006*, 55 (2006)
- Schlotter P., Schmidt R., Schneider J. "Luminescence conversion of blue light emitting diodes" *Applied Physics A* **64**, 417 (1997)
- Schlotter P., Baur J., Hielscher Ch., Kunzer M., Obloh H., Schmidt R., Schneider J. "Fabrication and

- characterization of GaN:InGaN:AlGaN double heterostructure LEDs and their application in luminescence conversion LEDs" *Materials Science and Engineering B* **59**, 390 (1999)
- Setlur A. A., Srivastava A. M., Comanzo H. A., Chandran G., Aiyer H., Shankar M. V., and Weaver S. E. "Ce³⁺ based phosphors for blue LED excitation" *Proceedings of the SPIE* **5187**, 142 (2004)
- Setlur A. A. "Phosphors for LED-based solid-state lighting" *The Electrochemical Society – Interface Winter 2009*, 32 (2009)
- Shackelford J. and Doremus R. H. (editors) "Ceramic and Glass Materials: Structure, Properties and Processing" (Springer, Berlin, 2008)
- Shimizu Y., Sakano K., Noguchi Y., Moriguchi T. "Light emitting device having a nitride compound semiconductor and a phosphor containing a garnet fluorescent material" (1999)
- Shionoya S. and Yen W. M. "Phosphor handbook", Chapter 4, page 317 "Methods of phosphor synthesis and related technology" (CRC Press, New York, NY, 1999)
- Srivastava A. M. "Phosphors" *Encyclopedia of Physical Science and Technology* (3rd edition) **11**, 855 (2004)
- Srivastava A. M., Setlur A. A., and Comanzo H. A. "Broad spectrum terbium-containing garnet phosphors and white-light sources incorporating the same" US Patent 6,596,195 B2, issued July 22 (2003)
- Srivastava A. M. and Ronda C. R. "Phosphors" *Interface (The Electrochemical Society)* **12** (2), p. 48 (2003)
- Sun X.-D., Wojnarowski R. J., Ayala R. E., McNulty T. F. "Fabrication of LED lamps by controlled deposition of a suspension media" US Patent 6,604,971 B1, issued August 12 (2003)
- Talbot J. B. "Electrophoretic Deposition of Phosphors for information displays and solid-state lighting" published as Chapter 7 in "Electrophoretic deposition of nanomaterials" edited by James H. Dickerson and Aldo R. Boccaccini (Springer, Berlin and New York, 2012)
- Tian Y., Zaremba D., and Hill I. G. "Red photoluminescent phosphors" US Patent 6,783,700 B2, issued August 31 (2004)
- Ueda J., Tanabe S., and Nakanishi T. "Analysis of Ce³⁺ luminescence quenching in solid solutions between Y₃Al₅O₁₂ and Y₃Ga₅O₁₂ by temperature dependence of photoconductivity measurement" *Journal of Applied Physics* **110**, 053102 (2011)
- Uheda K., Hirosaki N., Yamamoto Y., Naito A., Nakajima T., and Yamamoto H. "Luminescence properties of a red phosphor, CaAlSiN₃:Eu²⁺, for white light-emitting diodes" *Electrochemical and Solid-State Letters* **9**, H22 (2006)
- Wang K., Wu D., Chen F., Liu Z., Luo X., and Liu S. "Angular color uniformity enhancement of white light-emitting diodes integrated with freeform lenses" *Optics Letters* **35**, 1860 (2010)
- Wegh R. T., Donker H., Oskam K. D., and Meijerink A. "Visible quantum cutting in LiGdF₄:Eu³⁺ through downconversion" *Science* **283**, 664 (1999)
- Xie R.-J. and Hirosaki N. "Silicon-based oxynitride and nitride phosphors for white LEDs — A review" *Science and Technology of Advanced Materials* **8**, 588 (2007)
- Yamada M. Naitou T., Izuno K., Tamaki H., Murazaki Y., Kameshima M., and Mukai T. "Red-enhanced white-light-emitting diode using a new red phosphor" *Japanese Journal of Applied Physics* **42**, L20 (2003)
- Yang L., Chen M., Lv Z., Wang S., Liu X., and Liu S. "Preparation of a YAG:Ce phosphor glass by screen-printing technology and its application in LED packaging" *Optics Letters* **38**, 2240 (2013)
- Yoo J. S., Kim S. H., Yoo W. T., Hong G. Y., Kim K. P., Rowland J., and Holloway P. H. "Control of spectral properties of strontium-alkaline-earth-silicate-europium phosphors for LED applications" *Journal of the Electrochemical Society* **152**, G382 (2005)
- Zachau M., Becker D., Berben D., Fiedler T., Jermann F., and Zwaschka F. "Phosphors for solid state lighting" *Proceedings of the SPIE* **6910**, 691010-1 (2008)

30

30 – White LEDs based on phosphors

The generation of white light by an LED chip whose emitted light is partially or fully used to optically excite one or several phosphors, has become a viable and common method for generating white light for general illumination applications. There are several different methods for generating white light. These methods can be classified in dichromatic, trichromatic, and tetrachromatic approaches having two, three, or four distinct emission bands, respectively. **Figure 30.1** shows such multi-chromatic white LEDs based on one or more LED chips that optically excite one or more phosphors. Most commonly, these white LEDs include a blue LED chip. However, a violet or a UV LED chips can also be used.

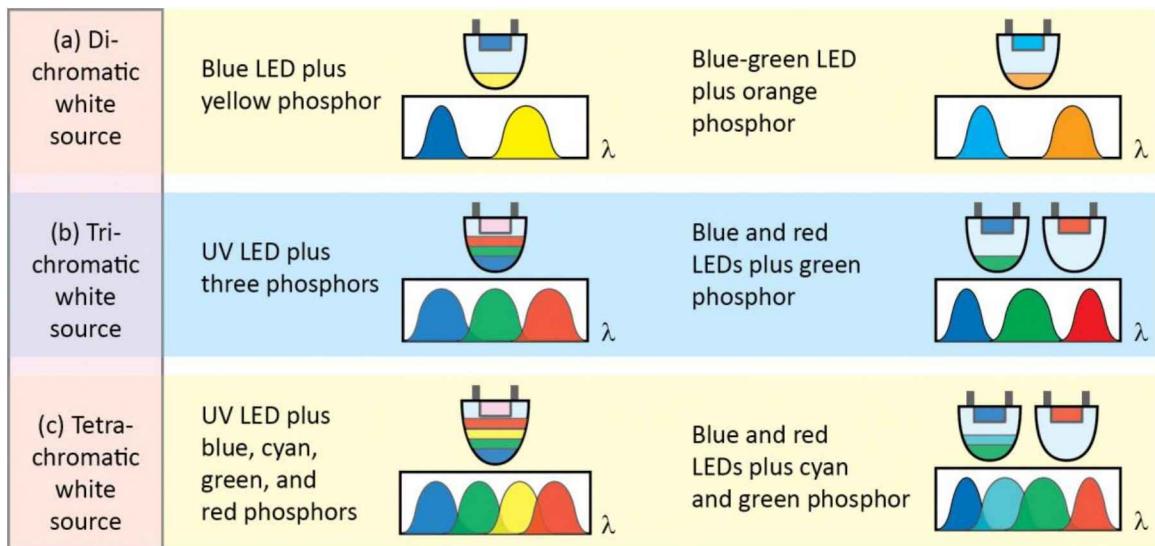


Fig. 30.1: White sources using a phosphor and excited by a blue, violet, or UV LED chip.

Figure 30.1 gives examples for each of the three categories (dichromatic, trichromatic, and tetrachromatic). Given the large range of choices, the figure cannot be comprehensive. However,

we can make two general statements: *Firstly*, the luminous source efficacy decreases with increasing multi-chromaticity of the source. That is, narrowband dichromatic sources have the highest luminous efficacy of radiation and also the highest potential luminous source efficacy. *Secondly*, the color-rendering capability is lowest for dichromatic sources and it increases with the multi-chromaticity of the source. The color-rendering index (general *CRI*) can reach values exceeding 95 for tetrachromatic sources.

30.1 Efficiency of wavelength-converter materials

The power-conversion efficiency of short-wavelength light to long-wavelength light by means of a **wavelength-converter** (λ -converter) or **phosphor** is determined by two factors: (i) the external quantum efficiency of the phosphor and (ii) the inherent quantum-mechanical-energy loss incurred during wavelength conversion.

The **phosphor quantum efficiency** (PQE) is given by

$$PQE = \frac{\text{number of photons emitted into free space by } \lambda\text{-converter per second}}{\text{number of photons absorbed by } \lambda\text{-converter per second}}. \quad (30.1)$$

The phosphor's external quantum efficiency originates in the *internal quantum efficiency* (IQE) and the *light-extraction efficiency* (LEE) of the phosphor according to $PQE = IQE \times LEE$. Note that the internal quantum efficiency depends on the inherent efficiency of the material whereas the extraction efficiency depends on the spatial structure and distribution of the phosphor, including its particle size and particle-size distribution. Generally, thin films have high extraction efficiencies whereas lumpy aggregations of phosphor materials have lower light-extraction efficiency due to reabsorption. It is therefore desirable to employ phosphor materials in the form of thin layers.

There is an inherent **wavelength-conversion loss** (sometimes called **quantum deficit** or **Stokes shift**) when a phosphor converts a photon with wavelength λ_1 to a photon with wavelength λ_2 ($\lambda_1 < \lambda_2$). The energy loss inherent to the conversion process is given by

$$\Delta E = h\nu_1 - h\nu_2 = \frac{hc}{\lambda_1} - \frac{hc}{\lambda_2}. \quad (30.2)$$

Thus, the wavelength-conversion efficiency (λCE) is given by

$$\lambda CE = \frac{h\nu_2}{h\nu_1} = \frac{\lambda_1}{\lambda_2} \quad (30.3)$$

where λ_1 is the wavelength of the photon absorbed by the phosphor and λ_2 is the wavelength of the photon emitted by the phosphor. Since the wavelength-conversion loss is fundamental in nature, it cannot be overcome with conventional λ -converter materials.

The phosphor power-conversion efficiency ($PPCE$) is the product of Eqns. (30.1) and (30.3)

$$PPCE = PQE \times \lambda CE. \quad (30.4)$$

The inherent wavelength-conversion loss is the reason that phosphor-based white LEDs have a fundamentally lower efficiency limit than white-light sources based on multiple LEDs.

The wavelength-conversion loss is highest for wavelength conversion from the UV to the red. For example, the conversion from UV (380 nm) to red (625 nm) can have a λ -conversion efficiency of only 61%. The low λ -conversion efficiency represents a driving force to employ red LEDs (rather than red phosphors) in highly efficient lighting systems.

Most white-light emitters use an LED chip emitting at short wavelength (i.e. blue) and a phosphor. Some of the light emitted by the blue LED is absorbed in the phosphor and then re-emitted as light with a longer wavelength. As a result, the lamp emits at least two different bands of wavelengths.

The possibility that white light can be generated in different ways raises the question as to which is the optimum way to generate white light? There are two parameters that need to be considered: First, the luminous efficiency; Second, the color-rendering index. For *signage applications*, the luminous efficiency is of primary importance and the color-rendering index is irrelevant. For *illumination applications*, both the luminous efficiency and the color-rendering index are important.

White-light sources employing two monochromatic complementary colors result in the highest possible luminous efficacy. However, the color-rendering index of such a dichromatic light source is low and unsuitable for general lighting applications.

The maximum luminous efficacy of radiation, attainable for white light created by two complementary monochromatic colors was calculated by MacAdam (1950). MacAdam showed that luminous efficacies exceeding 400 lm/W can be attained using a dichromatic source for white-light generation. The work of MacAdam (1950) was further refined by Ivey (1963) and

Thornton (1971). These authors showed that dichromatic white-light sources have high luminous efficacy but low color-rendering properties, making them perfectly suitable for signage applications but unsuitable for daylight illumination applications. In addition, Thornton (1971) showed that trichromatic white-light sources, i.e. sources creating white light by additive mixing of three discrete colors, have a color-rendering index suited for most applications. Thornton reported on an experiment in which 60 observers judged the color rendition of meat, vegetables, flowers, complexions, etc., when illuminated with a trichromatic light source with peak wavelengths at 450, 540, and 610 nm. The color rendition in this experiment was found to be “very good, if not excellent” illustrating the suitability of trichromatic white-light sources as potent daylight illumination sources.

A white-light source duplicating the sun’s spectrum would have good color-rendering capability. However, the radiation efficacy of such a light source is lower than what is possible with other spectral distributions, e.g. a trichromatic distribution. The sun’s spectrum has a high emission intensity near the boundaries of the visible spectrum (380 and 750 nm) where the eye sensitivity is very low. Thus, exact duplication of the sun’s spectrum is not a viable strategy for high-efficiency light sources.

30.2 Wavelength-converter materials

Converter materials have several parameters of interest, including the absorption wavelength, emission wavelength, quantum efficiency, radiative lifetime, chemical stability, stability with respect to high radiation intensities, and its temperature quenching characteristics. All wavelength-converter materials are commonly called **phosphors**. The term “phosphor” is not limited in any way. That is, a phosphor can...

- ... emit light by means of fluorescence or phosphorescence,
- ... be an inorganic or organic material, and
- ... be a solid, liquid, or gas.

Most common are inorganic solid phosphors. Generally, the chemical stability of inorganic phosphors is higher than that of organic phosphors. White LEDs based on inorganic phosphors will be discussed below.

However, white LEDs based on organic phosphors have also been demonstrated. Organic phosphors include dyes and pigments. A **dye** is an organic phosphor, frequently in a liquid form,

so that other materials can be soaked with the liquid dye. If a dye is dissolved in a transparent binder such as an epoxy, silicone, PMMA, acrylic glass, or polycarbonate the dye-based phosphor becomes a solid material. When crushed into small particles, such dye-containing binder is frequently referred to as ***pigment***.

White LEDs have been fabricated using organic dye molecules as wavelength converter materials. The dye can be incorporated in optically transparent polymers (also called binders or resins) such as the epoxy encapsulant of an LED (Schlotter *et al.*, 1997).

A drawback of organic dyes is their low chemical stability and thus their short reliability lifetime. Dye molecules “bleach out”, i.e. become optically inactive, after a certain number of photon absorption events. Typically, a dye molecule is stable for about 10^6 optical transitions (Jones, 2000). The lack of high molecular stability of dyes is a serious drawback. The reliability lifetime of dyes is considerably shorter than the lifetime of inorganic phosphors.

Dyes have a relatively small difference between the absorption and the emission band (Stokes shift). For example, the Stokes shift for the dye *Coumarin 6* is just 50 nm. This shift is smaller than the Stokes shift required for dichromatic white LEDs that need typical wavelength shifts of 100 nm or more (as inferred from the separation of two complementary wavelengths).

The most common phosphors used in conjunction with white LEDs consist of an inorganic host material that is doped with an optical activator as discussed in the chapter on phosphors in this book. Over the years it has become common practice to use more than one phosphor material in a white LED. Phosphors containing two or more chemically distinct phosphor materials are called ***phosphor blends***. For example, a common phosphor blend consists of a YAG:Ce phosphor and a red phosphor; such phosphor blend has been found suitable for white LEDs with a lower color temperature (e.g. 2 500~3 000 K).

A concern with white sources is ***spatial color uniformity***. The chromaticity of the white source should not depend on the emission direction. Color uniformity can be attained by a phosphor distribution that provides an about ***equal optical path length*** in the phosphor material independent of the emission direction (Reeh *et al.*, 2003).

Spatial uniformity can also be attained by adding ***mineral diffusers*** to the encapsulating resin (Reeh *et al.*, 2003). Such mineral diffusers are optically transparent substances, such as TiO_2 , CaF_2 , SiO_2 , CaCO_3 , and BaSO_4 , with a refractive index different from the surrounding resin. Mineral diffusers will cause light to reflect, refract, and scatter, thereby randomizing the propagation

direction and uniformizing the far-field distribution in terms of chromaticity and emission spectrum.

30.3 White LEDs based on YAG:Ce phosphor

The first white LED lamp using a phosphor and a blue GaInN/GaN optical excitation LED was reported by Bando *et al.* (1996), Shimizu *et al.* (1999), and subsequently reviewed by Nakamura and Fasol (1997).³³ The phosphor used as a wavelength converter was Ce-doped YAG with the chemical formula $\text{Y}_3\text{Al}_5\text{O}_{12} : \text{Ce}$.

The cross-sectional structure of the first white LED lamp is shown in *Figure 30.2 (a)*.³⁴ The figure shows the LED chip emitting in the blue and the YAG phosphor surrounding the chip. The YAG phosphor powder is dispersed in epoxy resin. During the manufacturing process, a droplet of the YAG phosphor dispersed in epoxy is deposited on the LED chip, so that the resin fills the cup-shaped depression in which the LED chip is located, as shown in *Figure 30.2 (b)*. As indicated in the figure, a fraction of the blue light is absorbed by the phosphor and re-emitted as longer-wavelength light.

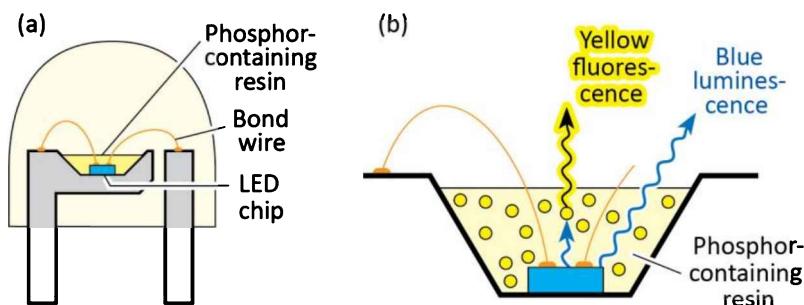


Fig. 30.2: (a) White LED lamp consisting of a GaInN blue LED chip and a phosphor. (b) Blue luminescence and wavelength-converted yellow fluorescence (after Nakamura and Fasol, 1997).

The emission spectrum of the phosphor-based white LED lamp thus consists of the blue emission band originating from the semiconductor LED chip and a longer-wavelength fluorescence emission, as shown in *Figure 30.3*. The thickness of the phosphor-containing epoxy and the concentration of the phosphor suspended in the epoxy determine the relative strengths of the two emission bands. The two bands can thus be adjusted to attain the desired balance between

³³ In these white LEDs, the GaInN/GaN LED chip used for optical excitation (“optical pumping”) was a device reported by Nakamura *et al.* (1995).

³⁴ The package of the first white LED was a traditional 5 mm radial LED package (occasionally referred to as a “bullet-shaped” LED).

the short-wavelength (blue) and long-wavelength (yellow) emission. One may note that the phosphor emission, although called yellow, has a broad spectral width ranging from green to red. This ensures that the resulting emission spectrum contains the three primary colors red, green, and blue.

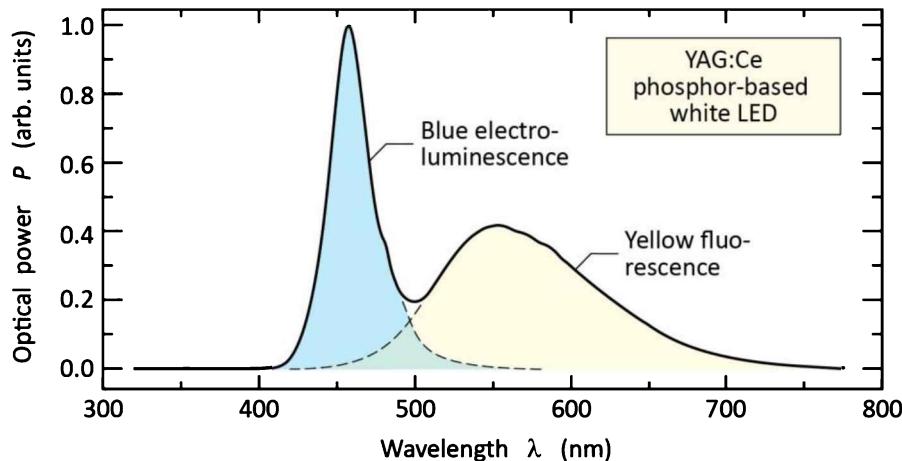


Fig. 30.3: Emission spectrum of a phosphor-based white LED manufactured by Nichia Company (Anan, Tokushima, Japan) in 2003.

The chemistry of YAG:Ce phosphor consists of the base host material $\text{Y}_3\text{Al}_5\text{O}_{12}$ and the activator element cerium (Ce). The base host material can be chemically modified by partial substitution of Y and Al with gadolinium (Gd) and gallium (Ga), respectively, so that the general chemical formula of the host material is $(\text{Y}_{1-x}\text{Gd}_x)_3(\text{Al}_{1-y}\text{Ga}_y)_5\text{O}_{12}$. The resulting change in the emission spectrum is shown in **Figure 30.4**: The addition of Gd shifts the peak emission wavelength to *longer* wavelengths (i.e. from about 550 nm to about 580 nm), whereas the addition of Ga shifts the peak emission wavelength to *shorter* wavelengths (i.e. from about 550 nm to about 520 nm). Such tuning of the peak emission wavelength is useful for producing white light with lower color temperatures (through the addition of Gd) or higher color temperatures (through the addition of Ga).

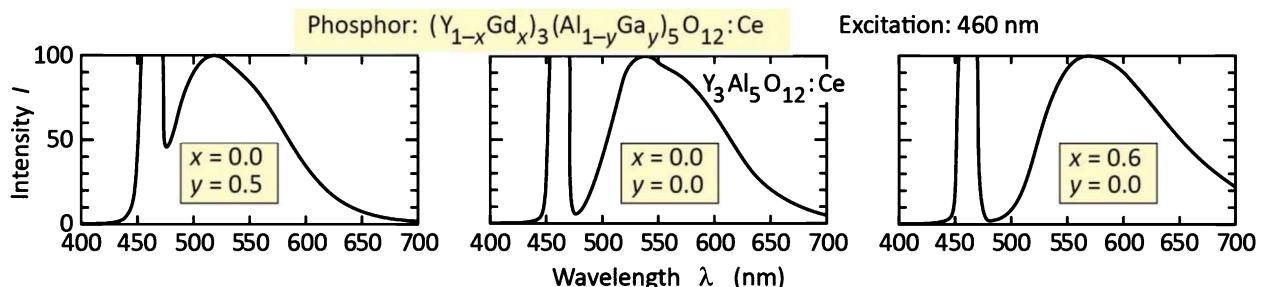


Fig. 30.4: Emission spectrum of Ce-doped yttrium aluminum garnet (YAG:Ce) for different chemical compositions. The excitation is at 460 nm (after Nakamura and Fasol, 1997).

The range of chromaticity points attainable with YAG phosphor is shown in *Figure 30.5*. It shows the chromaticity point of a blue LED chip and the range of chromaticity points attainable from YAG:Ce phosphor. By color mixing, a range of chromaticity points, indicated by the shaded area, can be attained. The black-body radiation locus (planckian locus) is also shown in the figure. Inspection of the figure yields that the blue-LED-chip-plus-YAG:Ce-phosphor combination is suitable for color temperatures from about 3 000 K to 8 000 K.

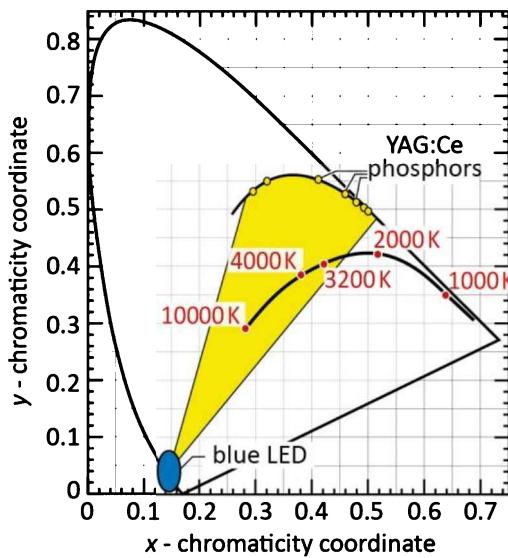


Fig. 30.5: Chromaticity points of YAG:Ce phosphor, and the general area (shaded) accessible to white emitters consisting of a blue LED and YAG:Ce phosphor (adapted from Nakamura and Fasol, 1997). Also shown is the planckian locus with color temperatures.

To attain even lower color temperatures, such as 2 700 K, the addition of an amber (i.e. yellow-orange), orange, or red phosphor is required. Note that a disadvantage of adding red phosphors is reduced luminous efficacy: The large Stokes shift of red phosphors reduces the efficiency (excitation at ~ 450 nm; emission at ~ 650 nm). Thus, although color rendering capabilities are improved, they are improved at the expense of luminous efficiency.

30.4 White LEDs based on phosphor blends

First-generation white LEDs from the Nichia Company had a relatively high color temperature (e.g. 4 000 to 8 000 K). The color temperature of white LEDs was reduced (e.g. $T_{CCT} < 3 000$ K) by including an additional phosphor that, when excited by 460 nm blue light, has a peak emission wavelength of 655 nm and a full-width at half-maximum of 110 nm (Narukawa, 2004; Narukawa *et al.* 2006, 2010; Niki *et al.*, 2004). As a result, the emission can be enhanced in the red range, as shown in *Figure 30.6*. Furthermore, by using an optimized phosphor mix, the pronounced depression at about 500 nm present in the first-generation white LED is reduced. The spectral

intensity enhancement near 500 nm and 675 nm enhance the light source's color rendering capability. Such white LEDs (Narukawa, 2004; Narukawa *et al.* 2006, 2010; Niki *et al.*, 2004) strongly render red colors and allow for a lower color temperature that can range between 2800 K (warm white) and 4700 K depending on the phosphor mix.

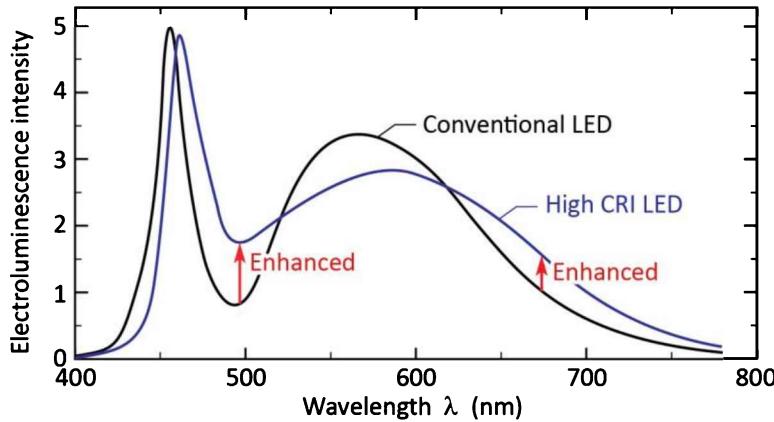


Fig. 30.6: Electroluminescence spectrum of conventional white LED and of high-color-rendering-index white LED. The high CRI results from the broader emission spectrum and the reduction of the notch in the spectrum (after Narukawa, 2004; Niki *et al.*, 2004).

By using an appropriate phosphor mix and a suitable balance between the primary blue emission (from the LED chip) and the phosphor emission, white light sources with color temperatures ranging from 2 000 K to 8 000 K can be fabricated. **Figure 30.7** shows representative emission spectra for high, medium, and low color temperature light sources (or correlated color temperature (CCT) light sources).

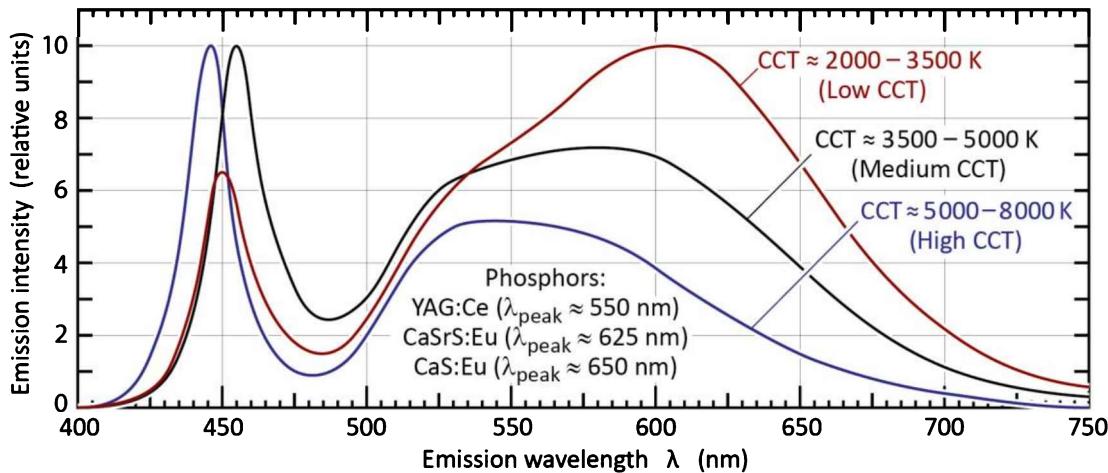


Fig. 30.7: White LEDs using various phosphors and conversion ratios to generate white light with correlated color temperatures (CCTs) ranging from 2 000 K to 8 000 K.

Inspection of the figure reveals that high-CCT light sources have a relatively strong blue emission line and weak red emission whereas low-CCT light sources have a relatively weak blue emission

line and strong red emission. YAG:Ce is the phosphor employed by all LEDs shown in the figure. For the low-CCT emission spectra shown in the figure, Eu-doped sulfide phosphors (CaSrS:Eu and CaS:Eu) are added to enhance the red emission. Eu-doped nitride and oxynitride phosphors are an alternative to the sulfide phosphors.

30.5 White LEDs using two LED chips and one phosphor

The relatively large Stokes shift losses incurred when a blue LED chip (~ 450 nm) excites a red phosphor (~ 650 nm) can be eliminated by replacing a red phosphor with red LED chip. Red LED chips based on the AlGaN/P/GaAs material system are relatively efficient with luminous efficacies of 100 lm/W being readily available.³⁵ Indeed, white LED lamps consisting of blue LED chips (450 nm), yellow phosphor (e.g. $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Ce}$), and red LED chips (625 nm) have been demonstrated (Shimizu *et al.*, 2003; Gielen *et al.*, 2014).

The emission spectrum of an LED lightbulb employing blue LEDs, a remotely located yellow phosphor, and red LEDs is shown in *Figure 30.8*. The inset shows a photograph of the lamp (Philips Lumileds, 2010; Gielen *et al.*, 2014).

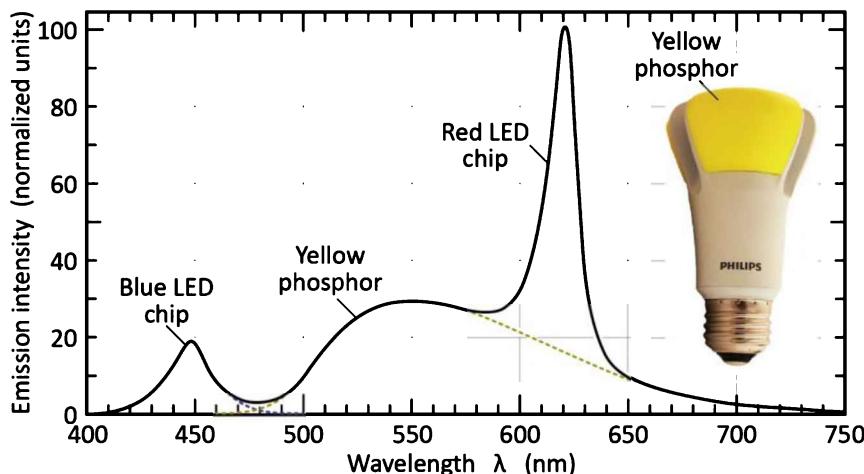


Fig. 30.8: Emission spectrum of white LED lamp with color temperature of 2700 K. The spectrum has three distinct components that are due to a blue LED chips, yellow phosphor, and red LED chips. The LED lamp ("L-Prize Lamp") is shown in the inset (Philips Lumileds, 2010).

In the OFF state, the lamp has yellow regions that are due to the yellow body color of the phosphor. The lamp emits a luminous flux of 940 lm and has an electrical input power of 10 W giving a luminous source efficacy of 94 lm/W. The correlated color temperature of the lamp is

³⁵ Typical high-power AlGaN/P / GaAs red LEDs (625 nm) emit a luminous flux in the range of 50~100 lm when injected with a current of 350 mA at a forward voltage of 2.1 V, giving a luminous efficacy of about 100 lm/W. See, for example <<http://www.osram-os.com/>>.

$\text{CTT} = 2\,700\text{ K}$ and its color rendering index is $\text{CRI} = 93$. The lamp was awarded a prize (“L-prize” of the US Department of Energy) and for this reason, is occasionally referred to as the “L-prize lamp”. It became available to consumers in about 2010.

30.6 White LEDs using violet excitation sources

The pursuit of higher quality light, i.e. light having a color rendering index (CRI) close to 100, has motivated the spectral broadening and smoothing of the emission band of white LEDs. On the short-wavelength side, this can be done by the employment of violet LEDs ($\sim 410\text{ nm}$) *instead of or in addition* to blue LEDs ($\sim 450\text{ nm}$). Furthermore, the use of multiple phosphors (e.g. a cyan, green, and red phosphor) allows for the smoothing of the emission spectrum so that all spectral regions have a significant emission power. This strategy has allowed for white LEDs with a CRI of 95. Such a high CRI value is virtually indistinguishable from an incandescent lamp ($\text{CRI} \approx 100$).

The emission spectrum of a white light source having a color temperature of 2 700 K is shown in **Figure 30.9**. The primary light source is a violet LED emitting a peak wavelength of 415 nm (Soraa Company, 2016).

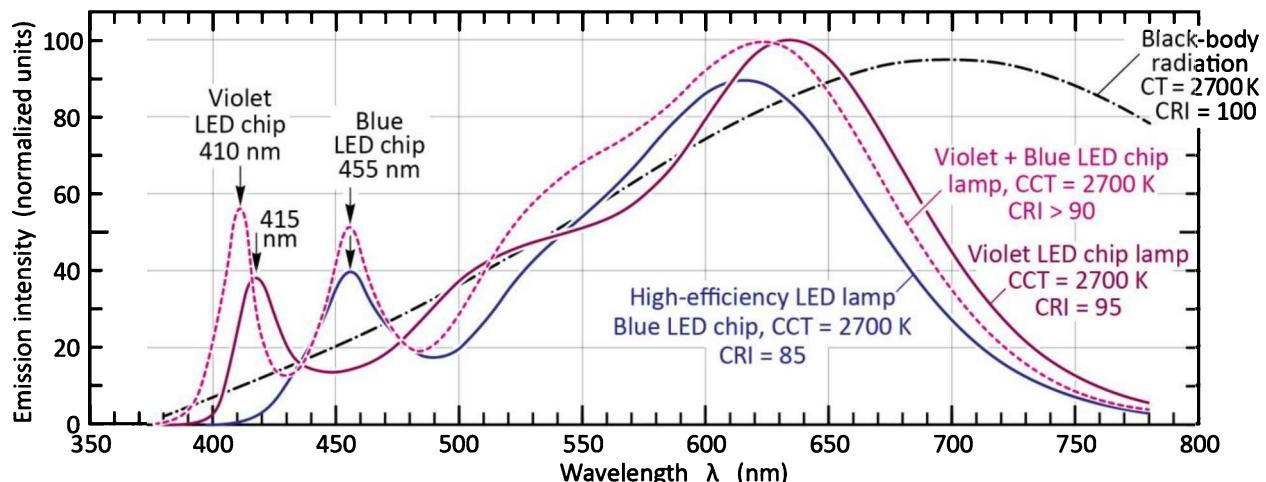


Fig. 30.9: Emission spectra of high-efficiency white LED, violet-LED-chip-based lamp, violet + blue-LED-chip-based lamp, and planckian black-body radiation, all having a color temperature of 2 700 K. Such lamps may be referred to as “high-CRI white”, “crisp white” or “vivid white” LED lamps (Soraa Company, 2016; Philips Lumileds Company, 2016).

Also shown in **Figure 30.9** is the emission spectrum of a blue-LED-based white light source. Comparison of the two spectra reveals that the high- CRI source has (i) greater emission intensity at the fringes of the spectrum and (ii) the spectrum more closely follows the black-body radiation

spectrum for a color temperature of 2 700 K (represented by the dashed line). These properties allow the high-CRI light source to have a CRI value that is 95, higher than the CRI value of 85 attained with the blue-LED-based white light source. A related approach has been pursued by the Philips Lumileds Company (2016): Adding a violet LED ($\lambda_{peak} = 410$ nm) to the blue-chip-plus-phosphor-blend approach. Philips' high-CRI lamps are available under the trade name “crisp white technology”.

The inevitable drawback of such high-CRI light sources is their lower source efficacy. Indeed, the luminous source efficacy of high-CRI light sources³⁶ can be 30% lower than the efficacy of blue-LED-based sources.

Another aspect of high-CRI white light sources is associated with using violet rather than blue excitation. Various materials (e.g. fabric, paper, and laundry detergent) contain fluorescent whitening agents.³⁷ Most of these fluorescent whitening agents are readily excited in the near-UV and violet but much less so in the blue part of the spectrum. As a result, white light sources based on violet (and near-UV) enable a whiter or “crisper white” appearance. The difference in appearance of sample objects when illuminated by a blue-based and violet-based white light source is shown in *Figure 30.10*. The sample objects shown in the figure (porcelain plate and napkin) can be assumed to contain fluorescent whitening agents (Philips Lumileds Company 2016; Soraa Company, 2016).

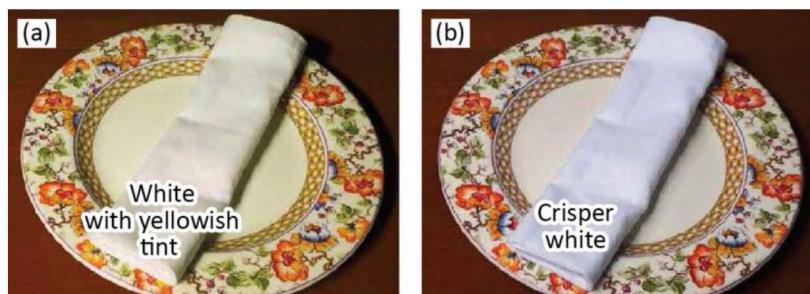


Fig. 30.10: Porcelain plate and napkin illuminated by (a) high-efficiency LED lamp, and (b) high-CRI LED lamp (“Crisp Color Technology”) both having a correlated color temperature of 2 700 K (Philips Lumileds, 2016; Soraa, 2016).

³⁶ High-CRI white LED light sources are associated with various trade names including “true white”, “crisp white” and “vivid white”.

³⁷ Fluorescent whitening agents (FWAs) are also called fluorescent brightening agents (FBAs) or optical brightening agents (OBAs). These chemical agents absorb UV and violet light (300~420 nm) and re-emit it in the blue range (typically 420~500 nm) by fluorescence. The blue fluorescence compensates for the yellowish tint of objects thereby making them appear truly white.

30.7 White LEDs using UV excitation sources

White LEDs can also be fabricated with optical excitation of phosphor in the ultraviolet (UV) wavelength range (Karlicek, 1999).³⁸ Semiconductor sources emitting in the near-UV (320~380 nm) and in the violet (close the edge of the visible spectrum 380~420 nm) can be used for such white sources. Indeed, GaInN / GaN LEDs emitting near 400 nm with remarkably high efficiencies have been reported (Morita *et al.*, 2004). For these LEDs, the primary emission has a very low or zero luminous efficacy of radiation. Accordingly, all light must be wavelength converted by means of phosphors. This *full-conversion* concept is in contrast to the *partial-conversion* concept of white LEDs that employ a blue LED chip as the primary light source.

For deep-UV semiconductor sources (200~320 nm), conventional phosphors, as used in fluorescent lighting, can be used for wavelength conversion. However, the large Stokes shift associated with deep-UV sources is a significant drawback for such sources. Furthermore the development of deep-UV LEDs is challenging due to the low p-type and n-type doping efficiency in AlGaN with high Al content and the difficulties encountered in epitaxially growing high-quality AlGaN with low dislocation and defect densities.

In deep- and near-UV-pumped white LEDs, the entire visible emission originates in the phosphor. Phosphors excited in the deep UV have been used since the 1950s in fluorescent light tubes and since the 1980s in compact fluorescent lamps (CFLs). Phosphors in fluorescent light sources are pumped by the UV emission coming from the low-pressure mercury-vapor discharge occurring inside the tube. The dominant emission of low-pressure mercury-vapor discharge lamps (**Hg lamps**) occurs in the UV at 254 nm. Phosphors with strong absorption in this wavelength range are readily available. Their color rendering properties are acceptable for many lighting applications.

A white LED using a near-UV AlGaN LED pump source and a tricolor phosphor blend was reported by Kaufmann *et al.* (2001). The LED pump source emitted at 380~405 nm, that is, near the boundary between the visible and UV spectrum.³⁹ The phosphor blend consisted of three

³⁸ The 1931 CIE chromaticity diagram shows that the visible spectrum ends to 380 nm with the eye sensitivity being very low at and near the 380-nm boundary. Accordingly, LEDs having a peak wavelength in the range 380~400 nm should be referred to as “visible LEDs”. However, in the technical literature, such LEDs are occasionally called “UV LEDs”. Consequently, the term “UV LED” will be used in a broader sense in the present section to include LEDs with peak wavelengths of 380~400 nm.

³⁹ Kaufmann *et al.* (2001) characterized the wavelength range of 380~405 nm as “UV”, even though, strictly speaking, the wavelength range is in the violet part of the visible spectrum.

phosphors emitting in the red, green, and blue parts of the spectrum. A color-rendering index of 78 was reported for the lamp.

The color-rendering index (*CRI*) of UV-excited phosphor mixes ranges between 60 and 100. Excellent *CRI*s as high as 97 were reported by Radkov *et al.* (2003) for phosphor blends excited near 400 nm. Furthermore, such UV-LED based sources exhibit independence of the phosphor-emission spectrum on the exact UV-LED excitation wavelength, because the visible emission is solely due to the phosphor. Consequently, UV-pumped white lamps have highly reproducible optical spectra so that “binning” may not be required. Monte Carlo simulations reported by Radkov *et al.* (2004) indeed showed a very low chromaticity point variation (entirely within the first MacAdam ellipse) for phosphor sources excited using LEDs coming from a variety chip bins with peak wavelengths ranging between 400 and 410 nm. The chromaticity variation was shown to be much lower for UV-LED/phosphor sources than for blue-LED/phosphor sources.

Sheu *et al.* (2003) reported a phosphor-converted LED fabricated by pre-coating red, green, and blue phosphors onto near-UV LED chips (400 nm) prior to packaging them into LED lamps.⁴⁰ At a 20 mA injection current, a correlated color temperature (CCT) of 5900 K and the color-rendering index (*CRI*) of 75 were found. The authors found that no changes in CCT and *CRI* when the injection current was increased from 20 to 60 mA. These results demonstrate that white LEDs based on a near-UV-excitation of a RGB phosphor blend are more optically stable (CCT stable, *CRI* stable, and more reproducible) than white LEDs based on a blue-LED-chip-plus-yellow-phosphor.

Doxsee and Jacob (2004) explained the difficulties with the blue-LED-chip-plus-yellow-phosphor approach:

The “color output of [blue-LED-chip-plus-yellow-phosphor] system varies greatly due to frequent, unavoidable routine deviations from desired parameters [...]. For example, the color output of the finished device is very sensitive to the thickness of the phosphor layer covering the LED. If the phosphor is too thin, then [...] the combined phosphor-LED output will appear bluish. In contrast, if the phosphor layer is too thick, then [...] [the] output will appear yellowish. Therefore, the thickness of the phosphor layer is an important variable affecting the color output of a blue LED based system. Unfortunately, the thickness of the phosphor layer is difficult to control during large scale production of LED-phosphor lamp systems, and the variations in phosphor thickness often result in relatively poor lamp-to-lamp color control. In addition, lamp-to-lamp variations occur due to the varying

⁴⁰ Sheu *et al.* (2003) characterized the wavelength 400 nm as “near-UV”, even though, strictly speaking, 400 nm is in the violet part of the visible spectrum.

of the optical power from chip-to-chip. The use of a UV LED chip to manufacture such a white-light system should give superior color performance compared to those based on blue LED chips since the UV chip is not appreciably contributing to the visible color of the LED.”

Chromaticity-point variation of LED lamps is a valid concern. One may ask: Is the chromaticity-point variation in blue-LED-chip-based white LEDs primarily caused by (i) the variation of the blue LED chip’s excitation wavelength or by (ii) the variation in the quantity of phosphor dispensed on the chip? Mueller-Mach *et al.* (2007) concluded that the primary cause for the chromaticity-point variation is the variation in the quantity of phosphor.

Even if color-control and color reproducibility would be superior in UV-pumped white LEDs, a fundamental drawback of UV-pumped white LEDs is the large energy loss (Stokes shift) incurred when converting UV light to white light. The potential luminous efficiency of UV-pumped white LED lamps is therefore markedly lower (e.g. by 30%) than for white sources based on a blue LED chip exciting a yellow phosphor. For this reason, UV-pumped white LEDs are not produced commercially at any significant scale.

30.8 White LEDs based on semiconductor converters (PRS-LED)

A white light-emitting diode using a semiconductor wavelength converter has been demonstrated by Guo *et al.* (1999, 2000). The schematic structure of the photon-recycling semiconductor LED (PRS-LED) is shown in *Figure 30.11 (a)*. The figure indicates that a fraction of the light emitted by the blue GaInN LED is absorbed by a AlGaN_xP secondary active region and re-emitted (or “recycled”) as lower-energy photons. In order to emit white light, the two light sources must emit at complementary wavelengths and the intensities of the two light sources must have a certain ratio. A detailed description and analysis of the PRS-LED can be found in the 2nd edition of this book.⁴¹

⁴¹ The 2nd edition of this book was published in 2006: E. F. Schubert “Light Emitting Diodes” (Cambridge University Press, Cambridge UK, 2006).

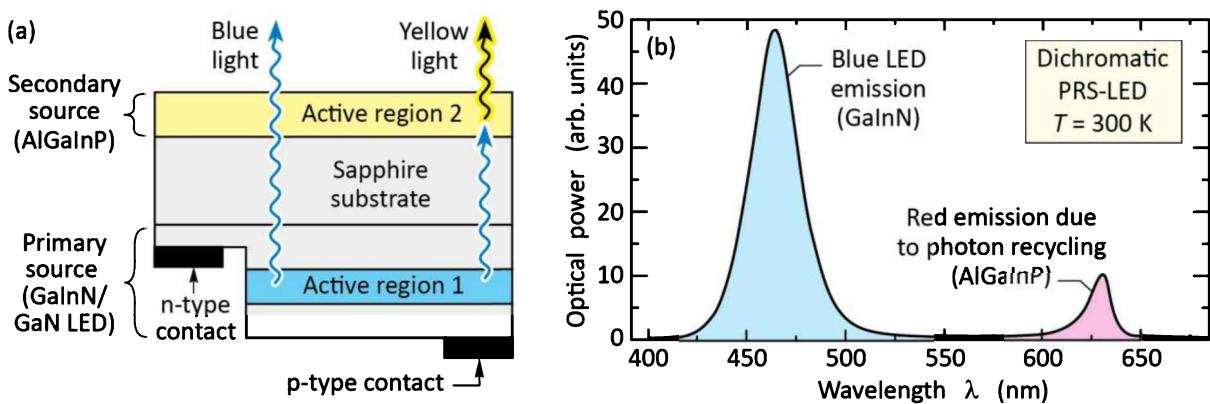


Fig. 30.11: (a) Schematic structure of a photon-recycling semiconductor LED (PRS-LED) with one current-injected active region (Active region 1) and one optically excited active region (Active region 2). (b) Emission spectrum of device (after Guo *et al.*, 1999, 2000).

The maximum theoretical efficiency of the PRS-LED has been calculated in the 2nd edition of this book assuming that both active regions have a quantum efficiency of 100%. The maximum luminous efficiency occurs if the primary source (i.e. the LED chip) emits at a peak wavelength of $\lambda_1 = 440$ nm. The secondary wavelength, λ_2 , must be complementary to $\lambda_1 = 440$ nm. This condition yields $\lambda_2 = 563$ nm. The required power ratio for attaining white light is given by $P(\lambda_2)/P(\lambda_1) = 1.79$.⁴² In this case, a theoretical luminous efficiency of 336 lm/W is obtained. Note that we assume that both light sources emit monochromatic light ($\Delta\lambda \rightarrow 0$). Taking into account the finite linewidth of the spontaneous emission from semiconductors (i.e. $\Delta E = 1.8 kT$), the calculated luminous efficiency is slightly lower.

PRS-LEDs have been demonstrated using a GaN/GaN LED emitting in the blue and an electrically passive AlGaNP photon-recycling semiconductor emitting in the red part of the spectrum (Guo *et al.*, 1999, 2000). The emission spectrum of the device, depicted in **Figure 30.11(b)**, shows the emission line of the primary LED at 470 nm and a second emission line at 630 nm due to absorption of the 470 nm light in the AlGaNP layer and re-emission of light at 630 nm. The recycling semiconductor used in this experiment is an AlGaNP/GaAs double heterostructure. The photon-recycling semiconductor is planar and no surface texturing was performed.

To avoid absorption of light in the GaAs substrate, the GaAs substrate of the AlGaNP epitaxial layer was removed. Firstly, the AlGaNP/GaAs recycling semiconductor was mounted on a glass

⁴² For the detailed calculation of the power ratio, the reader is referred to the 2nd edition of this book.

slide. Subsequently, the GaAs substrate was removed by polishing and selective wet chemical etching. Then the primary LED wafer and the photon-recycling wafer were bonded together.

Generally, dichromatic white LEDs have a higher luminous efficacy but lower color-rendering index (*CRI*) compared with trichromatic white LEDs. It can be shown that there is a fundamental trade-off between color rendering and the luminous efficacy of light-emitting devices (Walter, 1971). In order to improve the general *CRI* of dichromatic devices such as the PRS-LED, two possibilities can be considered. Firstly, the emission lines can be intentionally broadened, e.g. by compositional grading of the semiconductor active region. Secondly, a second photon-recycling semiconductor can be added thus creating a trichromatic PRS-LED. However, broadening of the two emission lines or the addition of a third emission line will decrease the luminous efficacy of the device.

30.9 White LEDs based on other methods

There are certainly additional strategies to generate white light by means of LEDs. Indeed, the number of specific ways by which white light can be generated is very large. It is not our goal to be comprehensive in this regard. Instead, it is the purpose of the present chapter to present strategies for white phosphor-converted LEDs that are viable, promising, or widely used.

To illustrate the variety of additional strategies for generation of white light, we point, as an example, to Oh *et al.* (2011) who presented an LED system consisting of four GaInN blue LED chips three of which were combined with one specific phosphor, i.e. a red, amber, and green phosphor. By means of full down-conversion, multiple monochrome light sources are fabricated and whose light is then combined and mixed to create spectrally tunable white light. Long-wave pass filters (LWPFs) on top of the LED packages blocked the blue light so as to promote full conversion. The principal advantage of this strategy, i.e. generation of white light by color-mixing the light of multiple monochrome fully phosphor-converted LEDs, is the dynamic control of the chromaticity including the attainability of color temperatures ranging from 2 700 to 6 500 K (Oh *et al.*, 2011). The strategy was shown to have better color rendering (higher *CRIs*) than the color-mixing approach based on the mixing of light from phosphor-free red, green, and blue (RGB) LEDs.

30.10 Spatial phosphor distributions

The ***spatial phosphor distribution*** in white LED lamps strongly influences the color uniformity and efficiency of the lamp. One can distinguish between ***proximate*** and ***remote phosphor distributions*** (Goetz, 2003; Holcomb *et al.*, 2003; Kim *et al.*, 2005; Luo *et al.*, 2005; Narendran *et al.*, 2005). In proximate phosphor distributions, the phosphor is located in close proximity to the semiconductor chip. Proximate phosphor distributions are shown in *Figure 30.12(a)* and (b). In remote phosphor distributions, the phosphor is spatially removed from the semiconductor chip. A remote phosphor distribution is shown in *Figure 30.12(c)*.

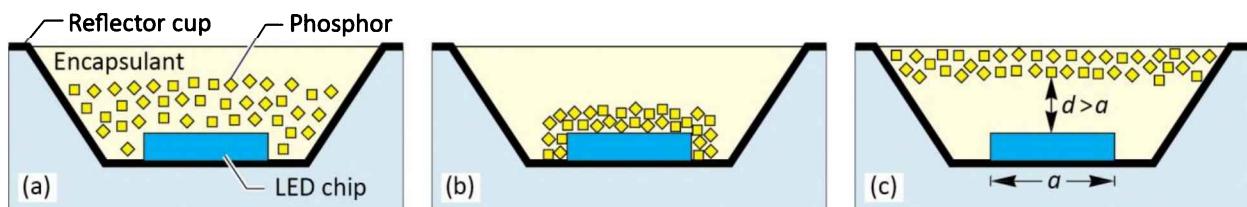


Fig. 30.12: (a) Proximate phosphor distribution, (b) proximate conformal phosphor distribution, and (c) remote phosphor distribution in which phosphor and chip are separated by at least one times the lateral dimension of the chip (after Kim *et al.*, 2005).

Photographs of the different phosphor distributions are shown in *Figure 30.13*. The proximate phosphor distribution shown in *Figure 30.13(a)* was introduced by Nichia Corporation in 1996. The phosphor particles are dissolved in the encapsulation material that is dispensed into the reflector cup. Gravity, buoyancy, and friction lead to a distribution of phosphor particles that favors larger phosphor particles to move downward, thereby bringing them closer to the chip surface.

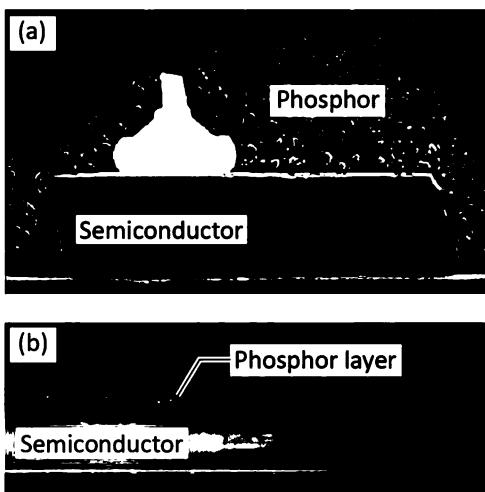
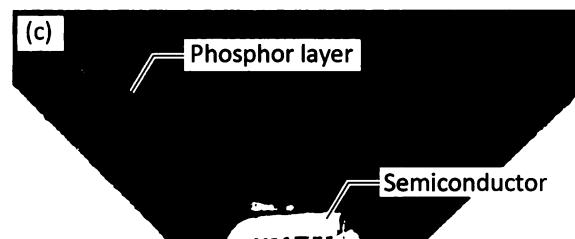


Fig. 30.13: Phosphor distributions in white LEDs:
 (a) Proximate phosphor distribution.
 (b) Proximate conformal phosphor distribution.
 (c) Remote phosphor distribution ((a) and (b) adapted from Goetz, 2003; (c) after Kim *et al.*, 2005).



Another proximate phosphor distribution, called the ***conformal phosphor distribution***, is shown in ***Figure 30.13 (b)***. Conformal phosphor distributions are accomplished by wafer-level phosphor dispensation thereby lowering the manufacturing cost as compared with a lamp-level phosphor dispensation. Conformal phosphor distributions provide a small emission area and high luminance, which is particularly relevant for imaging-optics applications. Imaging-optics applications (e.g. automotive headlights) frequently require the use of lenses. Optical design considerations show that point-like sources, i.e. sources with a small emission area, are desirable for these applications.

A general drawback of proximate phosphor distributions is the re-absorption of the phosphor's fluorescence by the semiconductor chip. Fluorescence emitted toward the semiconductor chip can be absorbed by the chip, e.g. by the metal contacts covering the chip. The reflectivity of the semiconductor chip and metal contacts can be limited.

This drawback can be avoided by remote phosphor distributions in which the phosphor is spatially distanced from the semiconductor chip (Kim *et al.*, 2005; Luo *et al.*, 2005; Narendran *et al.*, 2005). In such remote phosphor structures, it is less likely that phosphorescence impinges on the low-reflectance semiconductor chip due to the spatial separation between the primary emitter (semiconductor chip) and the secondary emitter (phosphor). The probability that fluorescence impinges on the semiconductor chip is greatly diminished if the distance between chip and phosphor is equal to or greater than the chip's lateral dimension, i.e. $d > a$, as shown in ***Figure 30.13 (c)***. As a result, higher phosphorescence efficiency is enabled. Ray-tracing simulations and experiments using a remote blue phosphor pumped by a GaInN emitter have indeed demonstrated phosphorescence efficiency improvements of 75% and 27%, respectively (Kim *et al.*, 2005; Luo *et al.*, 2005). Narendran *et al.* (2005) reported an average of 61% improvement in light output by using a remote phosphor configuration. At low currents, the remote phosphor configuration exceeded 80 lm/W, compared to 54 lm/W for a proximate phosphor configuration.

A phosphor can also be positioned at a very large distance from the LED chip. ***Figure 30.14*** shows an LED lightbulb using blue LED chips and a yellow phosphor arranged spherically around the chips (Tong *et al.*, 2011). The phosphor material is dispersed in a spherical binder or resin that surrounds the LED chips. The phosphor is thermally de-coupled from the LED chip thereby reducing the thermal load of both, phosphor and LED chips. However, a larger quantity of

phosphor material is required in such remote phosphor configuration than in a proximate phosphor configuration.

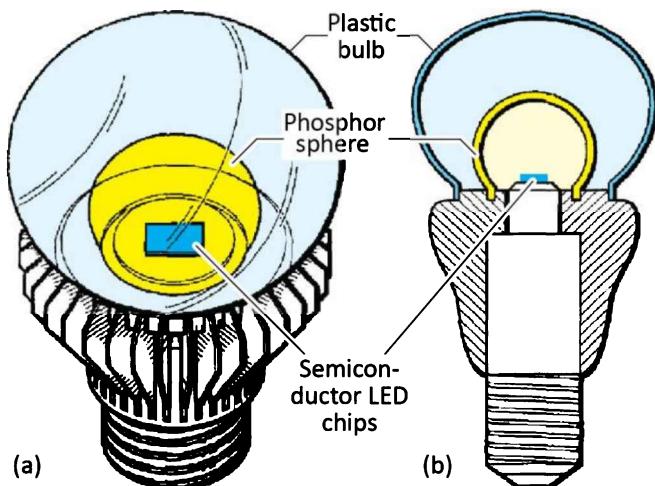


Fig. 30.14: (a) Perspective and (b) cross-sectional view of LED light bulb having a remote phosphor configuration. The phosphor is dispersed in a binder or resin that surrounds the LED chips. The phosphor is thermally de-coupled from the LED chip thereby reducing the thermal load of both, phosphor and LED chips. A larger amount of phosphor is required in such remote phosphor configuration than in a proximate phosphor configuration (after Tong *et al.*, 2011).

Considering the above-presented various phosphor spatial distributions, the following may be said: Proximate phosphor configurations require a smaller quantity of phosphor material thereby enabling lower phosphor costs. Furthermore, proximate phosphor configurations with the phosphor particles in physical contact with the chip enable the extraction of heat generated in the phosphor through the chip. Accordingly, proximate phosphor configurations (including the conformal phosphor configuration) are most common configurations. Additional considerations can be found in the chapter on phosphors in this book.

30.11 References

- Bando K., Noguchi Y., Sakano K., and Shimizu Y. (in Japanese) *Tech. Digest, Phosphor Res. Soc.*, 264th Meeting, November 29 (1996)
- Doxsee D. D. and Jacob C. "White light emitting diode device based on UV LED and phosphor blend" US patent 6,765,237, issued on July 20 (2004)
- Goetz W. "White lighting (illumination) with LEDs" *Fifth International Conference on Nitride Semiconductors*, Nara, Japan, May 25~30 (2003)
- Gielen V. S. D., Ansems J. P. M., Ter Weeme B. J. W., Bukkems P. J. M., and Xu J. "LED-based electric lamp" US patent 8,760,043, issued June 24 (2014)
- Guo X., Graff J. W., and Schubert E. F. "Photon-recycling semiconductor light-emitting diode" *IEDM Technical Digest, IEDM-99*, 600 (1999)
- Guo X., Graff J. W., and Schubert E. F. "Photon-recycling for high brightness LEDs" *Compound Semiconductors* 6, May/June (2000)
- Holcomb M. O., Mueller-Mach R., Mueller G. O., Collins D., Fletcher R. M., Steigerwald D. A., Eberle S., Lim Y. K., Martin P. S., and Krames M. "The LED light bulb: Are we there yet? Progress and challenges for solid-state illumination" *Conference on Lasers and Electro-Optics (CLEO)*, Baltimore, Maryland, June 1~6 (2003)

- Ivey H. F. "Color and efficiency of luminescent light sources" *J. Opt. Soc. Am.* **53**, 1185 (1963)
- Jones G. (Boston University), personal communication (2000)
- Karlicek Jr. R. F., personal communication (1999)
- Kaufmann U., Kunzer M., Köhler K., Obloh H., Pletschen W., Schlotter P., Schmidt R., Wagner J., Ellens A., Rossner W., and Kobusch M. "Ultraviolet pumped tricolor phosphor blend white emitting LEDs" *Phys. Stat. Sol. (a)* **188**, 143 (2001)
- Kim J. K., personal communication (2005)
- Kim J. K., Luo H., Schubert E. F., Cho J., Sone C., and Park Y. "Strongly enhanced phosphor efficiency in GaN white light-emitting diodes using remote phosphor configuration and diffuse reflector cup" *Jpn. J. Appl. Phys. – Express Letter* **44**, L649 (2005)
- Luo H., Kim J. K., Schubert E. F., Cho J., Sone C., and Park Y. "Analysis of high-power packages for phosphor-based white-light-emitting diodes" *Appl. Phys. Lett.* **86**, 243505 (2005)
- MacAdam D. L. "Maximum attainable luminous efficiency of various chromaticities" *J. Opt. Soc. Am.* **40**, 120 (1950)
- Morita D., Yamamoto M., Akaishi K., Matoba K., Yasutomo K., Kasai Y., Sano M., Nagahama S.-I., and Mukai T. "Watt-class high-output-power 365 nm ultraviolet light-emitting diodes" *Jpn. J. Appl. Phys.* **43**, 5945 (2004)
- Mueller-Mach R., Mueller G. O. and Krames M. R. "Phosphor-converted high power LEDs" *Proceedings of SPIE* **6797**, 67970G (2007)
- Nakamura S., Senoh M., Iwasa N., Nagahama S., Yamada T., and Mukai T. "Superbright green InGaN single-quantum-well-structure light-emitting diodes" *Japanese Journal of Applied Physics (Letters)* **34**, L1332 (1995)
- Nakamura S. and Fasol G. *The Blue Laser Diode* (Springer, Berlin, 1997)
- Narendran N., Gu Y., Freyssinier-Nova J. P., and Zhu Y. "Extracting phosphor-scattered photons to improve white LED efficiency" *Physica Status Solidi (a)* **202**, R60 (2005)
- Narukawa Y. "White light LEDs" *Optics & Photonics News* **15**, No. 4, p. 27 (2004)
- Narukawa Y., Narita J., Sakamoto T., Deguchi K., Yamada T., and Mukai T. "Ultra-high efficiency white light emitting diodes" *Japanese Journal of Applied Physics (Letters)* **45**, L1084 (2006)
- Narukawa Y., Ichikawa M., Sanga D., Sano M., and Mukai T. "White light emitting diodes with super-high luminous efficacy" *Journal of Physics D: Applied Physics* **43**, 354002 (2010)
- Niki I., Narukawa Y., Morita D., Sonobe S., Mitani T., Tamaki H., Murazaki Y., Yamada M., Mukai T. "White LEDs for solid state lighting" *Proceedings SPIE* **5187**, 1 (2004)
- Oh J. H., Oh J. R., Park H. K., Sung Y.-G., and Do Y. R. "New paradigm of multi-chip white LEDs: combination of an InGaN blue LED and full down-converted phosphor-converted LEDs" *Optics Express* **18**, A270 (2011)
- Philips Lumileds Company "Crisp White Technology" published on the internet: <<http://www.lumileds.com/technology/luxeon-technology/crispwhite-technology>> accessed on October 5 (2016)
- Radkov E., Setlur A., Brown Z., and Reginelli J. "High CRI phosphor blends for near UV LED lamps" *Proceedings SPIE* **5530**, 260 (2003)
- Radkov E., Bompiedi R., Srivastava A. M., Setlur A. A., and Becker C. "White light with UV LEDs" *Proceedings SPIE* **5187**, 171 (2004)
- Reeh U., Höhn K., Stath N., Waitl G., Schlotter P., Schneider J., and Schmidt R. "Light-radiating semiconductor component with luminescence conversion element" US Patent 6,576,930, issued June 10 (2003)
- Schlotter P., Schmidt R., and Schneider J. "Luminescence conversion of blue light emitting diodes" *Applied Physics A* **64**, 417 (1997)
- Sheu J. K., Chang S. J., Kuo C. H., Su Y. K., Wu L. W., Lin Y. C., Lai W. C., Tsai J. M., Chi G. C., and Wu R. K. "White-light emission from near UV InGaN–GaN LED chip pre-coated with blue/green/red phosphors" *IEEE Photonics Technology Letters* **15**, 18 (2003)
- Shimizu Y., Sakano K., Noguchi Y., and Moriguchi T. "Light emitting device having a nitride compound

- semiconductor and a phosphor containing a garnet fluorescent material" US patent 5,998,925 (priority dates of 1996) issued on December 7 (1999)
- Shimizu M., Shimomura Y., Tamura T., Nagai H., and Matsui N. "LED lamp" US patent 6,577,073, issued June 10 (2003)
- Soraa Company "High color-rendering, full-visible-spectrum LEDs by Soraa", published on the internet: <<https://www.led-professional.com/resources-1/articles>> June 30 (2016)
- Thornton W. A. "Luminosity and color-rendering capability of white light" *J. Opt. Soc. Am.* **61**, 1155 (1971)
- Tong T., Letoquin R., Keller B., Tarsa E., Youmans M., Lowes T., Medendorp Jr. N. W., Van De Ven A., Kung S., Negley G., Guschl P., Yuan Z. "Non-uniform diffuser to scatter light into uniform emission pattern" US patent application publication 2011/0,216,523 A1 (2011)
- Walter W. "Optimum phosphor blends for fluorescent lamps" *Appl. Opt.* **10**, 1108 (1971)

31

31 – Packaging: Types of packages for LEDs

LED packages provide a mechanically stable housing for an LED chip which facilitates the handling, mounting, soldering and further use of the LED. LED packages provide the following functions:

Electrical path that connects the external contacts (usually solderable contacts) with the electrodes of the LED chip.

Optical path that consists of a transparent or translucent encapsulating resin, that transmits the light emitted by the LED chip. Encapsulants also provide protection against unwanted mechanical shock, humidity, and chemicals.

Thermal path that allows the heat generated by the LED chip (and phosphor) to be extracted from the package.

In addition, a package may provide the following functions:

Wavelength conversion by means of a phosphor that is frequently dispersed in the above-mentioned transparent or translucent resin or deposited on the LED chip surface.

Electrostatic discharge (ESD) protection by means of e.g. a Zener diode that is integrated into the LED package.

Depending in the application, LEDs come in a variety of packages. Various packages, intended and optimized for specific applications, will be discussed in the present chapter.

Packaging of LEDs historically has been an inherently **serial** process. That is, LEDs are packaged in a serial way, one at a time. This is in contrast to LED chip manufacturing that proceeds in the form of wafers containing thousands of LED chips, thereby making LED chip manufacturing an inherently **parallel** process. The serial nature of LED package fabrication and the parallel nature of LED chip fabrication have cost implications. A parallel process inherently has cost advantages; a serial process is inherently expensive.

Given these considerations, it is desirable to “parallelize” the packaging of LEDs. As discussed below, in order to “parallelize” the manufacturing of LED packages, the packages are frequently processed in arrays so that many LEDs, e.g. hundreds of LEDs, are packaged at the same time.

31.1 Round 3 mm and 5 mm LED package

Virtually all LEDs are mounted in a package that provides two electrical leads, a transparent optical window for the light to escape, and, in power packages, a thermal path for heat dissipation. The chip-encapsulating material advantageously possesses high optical transparency, a high refractive index, chemical inertness, mechanical softness, high-temperature stability, and hermeticity against moisture. The refractive index contrast between the semiconductor and air is reduced by including an encapsulating resin thereby increasing the light-extraction efficiency. Virtually all encapsulants are polymers with a typical refractive index of 1.5 to 1.8. A reduced index change at the semiconductor surface increases the angle of total internal reflection thereby enlarging the light-escape cone and the light-extraction efficiency.

A traditional *low-power package* is shown in *Figure 31.1 (a)*. The LED chip is die-bonded to the bottom of a cup-like depression (“reflector cup”) thereby connecting the LED chip with one of the package’s lead wires (usually the cathode lead). A bond wire connects the LED top contact to the other lead wire (usually the anode lead). The LED package shown in the figure is referred to as a “3 mm” package, “5 mm” package, “T1-3/4” package, “radial LED” package, or “bullet-shaped” LED package.

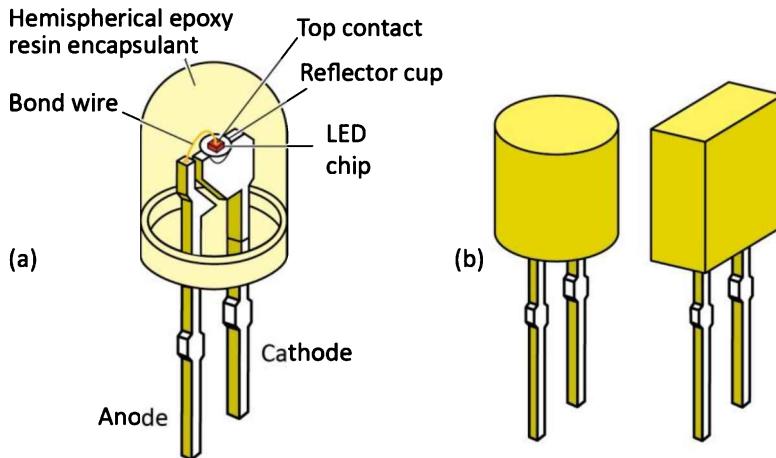


Fig. 31.1: Typical radial LED packages (also called 5 mm, T-1 3/4, or bullet-shaped LED packages); (a) LED with hemispherical encapsulant; (b) LEDs with cylindrical and rectangular encapsulant.

In low-power LEDs, the encapsulant has the shape of a hemisphere, as shown in *Figure 31.1 (a)*, so that the angle of incidence at the encapsulant-air interface is normal (90°). As a result, total internal reflection does not occur at the encapsulant-air interface. There are types of LEDs that do *not* have a hemispherical shape for the encapsulant. Some LEDs have a rectangular or cylindrical shape with a planar front surface. Examples of such shapes are shown in *Figure 31.1 (b)*. Planar-surface LEDs are frequently used under circumstances where the intended

viewing angle is close to normal incidence or where the LED's surface is intended to “blend in” (be co-planar) with another planar surface.

A photograph of an LED **leadframe** is shown in **Figure 31.2**. The individual leadframes are connected via a temporary stabilizing lead that is removed after die bonding, wire bonding, and the establishment of mechanical stability between anode and cathode lead by the encapsulant (typically epoxy encapsulant). The LED chip is die-bonded to the flat bottom of the **reflector cup** frequently using a silver-loaded electrically conductive epoxy. For high-power chips, the lower thermal resistance of die-bonding metal-based solders makes them preferable over Ag-loaded polymeric die-bonding materials.

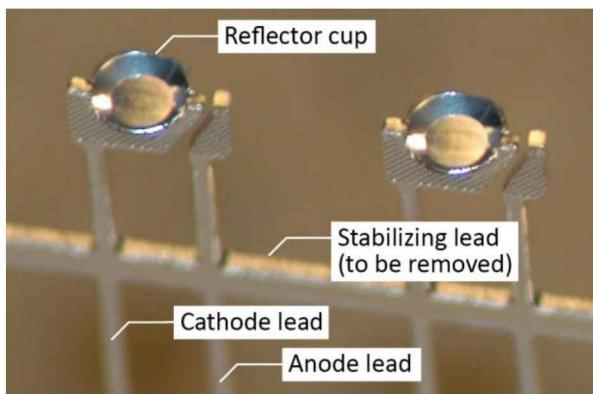


Fig. 31.2: Leadframe of a conventional radial LED package for mounting LED chips. The stabilizing lead (also called dam bar) is cut off once mechanical stability between the anode and cathode lead has been established by the epoxy encapsulant.

31.2 SMD LED injection molded package

Starting in the early 1990s, surface-mounted LEDs were introduced (Möllmer and Waitl, 1991). **Surface mount technology**, or **SMT**, is a method for which the components are mounted (or placed) directly onto the surface of a substrate such as a PCB (“printed circuit board” also called “printed wiring board”). SMT employs **surface-mount devices**, or **SMDs** that are mounted on the surface of the PCB. The mounting process does not rely on through-holes (“vias”) that penetrate the PCB.

A typical SMD LED package, as introduced in the 1990s, is shown in **Figure 31.3** (after Möllmer and Waitl, 1991). The SMD package consists of a reflective white resin body and contact leads that protrude the resin body at two sides and are folded around the resin body so as to provide solderable contact surfaces at the bottom surface of the package. The top surface of the white resin body has a recess in which the two contact leads are exposed. The LED chip is mounted on one of the leads and electrically connected to the other lead by means of a bond wire. The recess is filled with a transparent resin that encapsulates the chip.

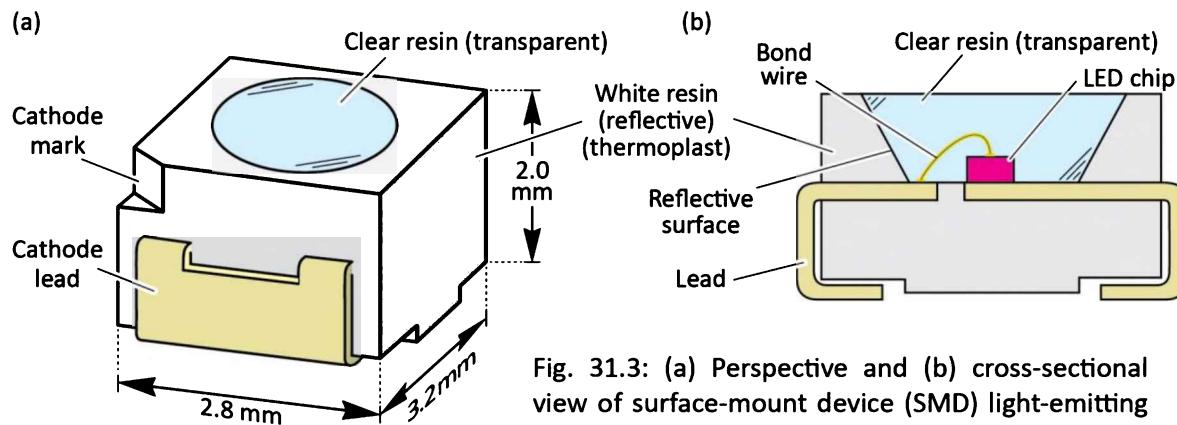


Fig. 31.3: (a) Perspective and (b) cross-sectional view of surface-mount device (SMD) light-emitting diode (LED) as introduced by the Siemens Company in 1991. The LED's anode and cathode leads are bent around the package body so that they fold under the bottom of the SMD LED's white resin part (after Möllmer and Waitl, 1991).

Various implementations of SMD LEDs based on an injection-molded package are shown in **Figure 31.4**. Inspection of the figure reveals that such SMD LED packages are suitable for housing one LED chip, as shown in **Figure 31.4 (a)** and **(b)**, as well as multiple LED chips, as shown in **Figure 31.4 (c)**. The latter shows a yellow phosphor dispersed in the encapsulating resin material.

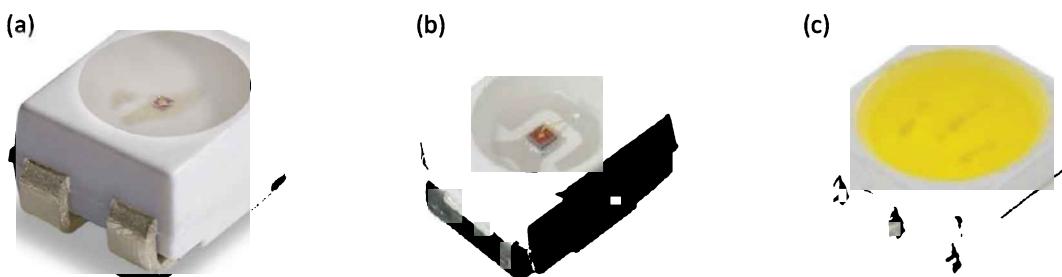


Fig. 31.4: Examples of various implementations of SMD LEDs based on thermoplastic white resin packages: (a) red LED, (b) orange LED, and (c) white LED.

In the 1990s and thereafter, the SMD LED packages were based in an ***injection molding process***⁴³. In the most common configuration of the process, a white ***thermoplastic material***⁴⁴ is heated and injected into a cavity formed by an upper mold and a lower mold with the lead frame being inserted between the upper and lower mold (“insert molding”). **Figure 31.5 (a)** and **(b)** shows the metal leadframe (before molding) and the molded leadframe (after molding), respectively.

⁴³ The term “injection-molding” is used in conjunction with the molding of a thermoplastic resin material. Heating of the thermoplastic resin is required during injection-molding to make the resin viscous and thus moldable.

⁴⁴ A thermoplastic material is a solid polymer material that becomes plastic (moldable, viscous) when heated to higher temperatures. The solid-to-plastic transition is a repeatable physical change that is not associated with a chemical change of the material.

Subsequent to the resin-molding process, the LED chips are die-bonded, wire-bonded, and encapsulated. Then the molded leadframe is cut and the contact leads of the package are folded around the white resin body of the package to attain the finished SMD LED.

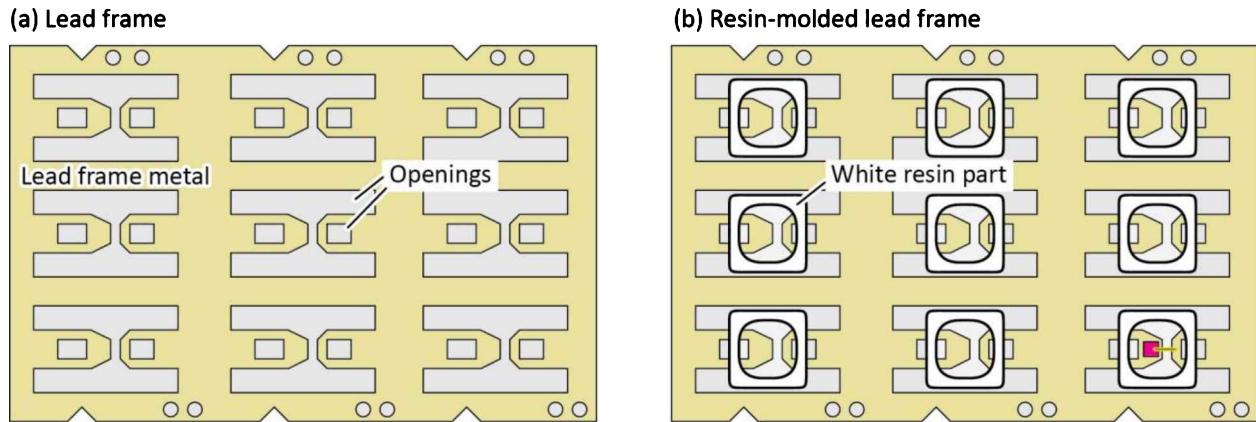


Fig. 31.5: (a) Schematic of an LED lead frame. The lead frame has openings that are punched or etched out. (b) Schematic of the lead frame after injection molding using a thermoplastic resin. The resin is light-reflective and thus has a white body color.

31.3 SMD LED package fabricated by using a resin-molded body

LED SMD packages have undergone significant improvements over traditional SMD LED packages. A newer generation of an SMD LED package is shown in *Figure 31.6* (Ichikawa *et al.*, 2013). Improvements include (i) better thermal coupling between the LED chip and the underlying PCB (enabled by metallic lead which is a good heat conductor), (ii) no need for bending and folding of the contact leads, (iii) lower profile (a package height of only 0.52 mm is shown in the figure), and (iv) improved manufacturing efficiency.

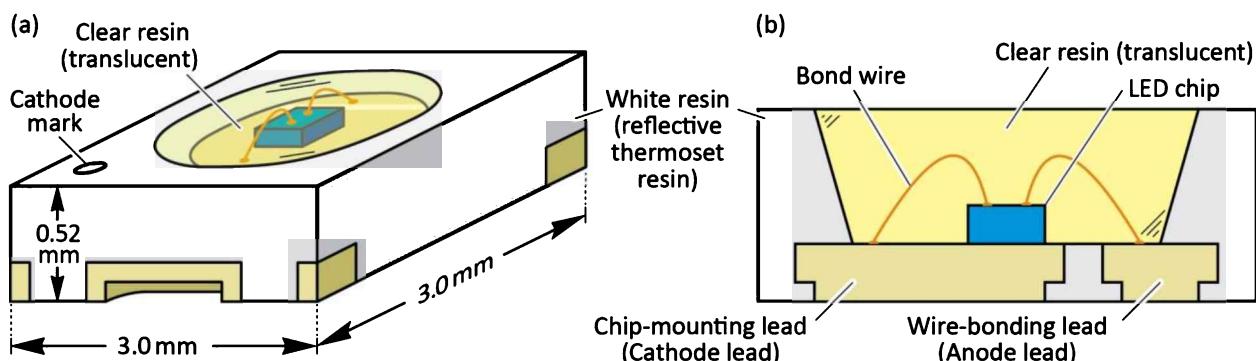


Fig. 31.6: (a) Perspective view and (b) cross-sectional view of surface-mount device (SMD) light-emitting diode (LED) that is fabricated by employing a resin-molded body. The LED has a low profile and the LED's leads are flat (after Ichikawa *et al.*, 2013).

Photographs of SMD LEDs based on a transfer-molded body are shown in **Figure 31.7**. Inspection of the figure reveals that such SMD LED packages can be made with a small package size. The white SMD LEDs contain a yellow or yellow-orange phosphor (after Nichia Company, 2016).

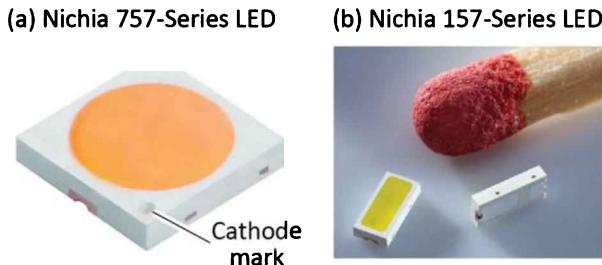


Fig. 31.7: Examples of SMD LEDs based on a molded array process using a thermosetting white resin packaging material: (a) Nichia 757 Series SMD LED ($3.0 \times 3.0 \times 0.52 \text{ mm}^3$). (b) Nichia 157 Series SMD LED ($3.0 \times 1.4 \times 0.52 \text{ mm}^3$).

The manufacturing process of the SMD LED is based on a ***transfer-molding process***⁴⁵ that employs a ***thermosetting material***.⁴⁶ Ichikawa *et al.* (2013) disclose that a metal lead frame can be over-molded using a resin so that an entire array of devices is fabricated without a space between individual devices as shown in **Figure 31.8**.

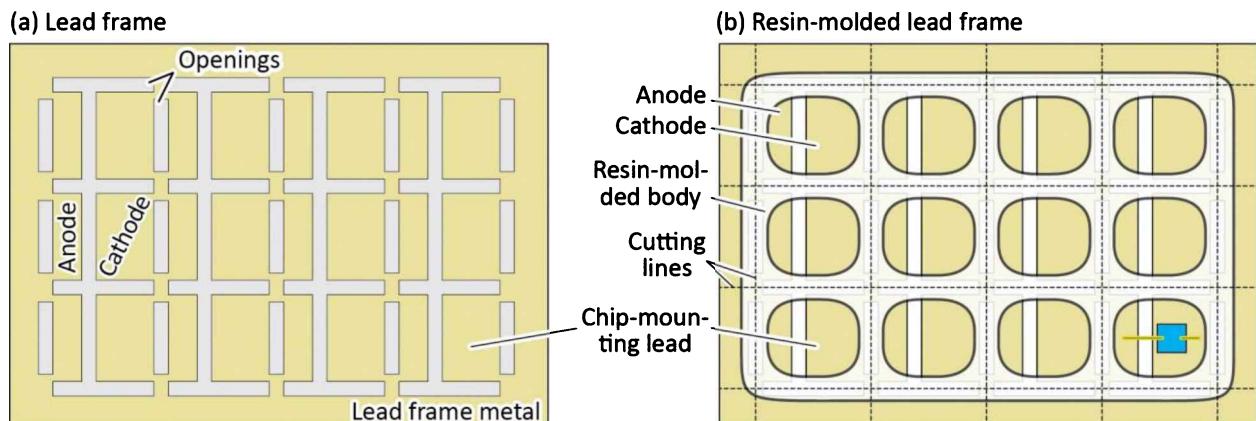


Fig. 31.8: (a) Lead frame for molded array SMD LEDs. Openings in lead frame are punched or etched out. (b) Lead frame after the transfer molding using a thermosetting resin. The resin is white and light-reflective. Die attach, wire bonding, and singulation are subsequent steps.

⁴⁵ The term “transfer-molding” is used in conjunction with thermosetting resin materials. In contrast to the injection-molding process, heating of the resin material is not required during the transfer-molding process.

⁴⁶ A thermosetting material is a liquid material that becomes solid (i.e. that “sets”) upon curing (e.g. by heating). Thermosetting materials can consist of a single or two chemical components. During the curing, the thermosetting material undergoes a chemical reaction that causes the material to solidify. A common thermosetting material is epoxy (EMC or epoxy molding compound).

That is, the resin is not supplied to individual devices but rather to an entire **resin-molded body**.⁴⁷ Subsequent to the resin-molding process, the LED chips are die-bonded, wire-bonded, and encapsulated. Then the molded leadframe (including the molded resin) is cut for singulation so as to attain a finished SMD device. Singulation can be accomplished by a precision circular saw that employs a thin diamond blade.

31.4 Ceramic-base LED package

High-power LEDs require an efficient thermal path that guides the heat from the LED chip to an underlying heat sink, such as a printed circuit board (PCB). Besides metals, ceramics are good heat conductors while being electrically insulating. The table below shows the thermal conductivity of various ceramic materials⁴⁸

<i>Material</i>	<i>Thermal conductivity</i>	<i>Comment</i>
Al_2O_3 (alumina)	30 W/(K m)	electrical insulator
Si_3N_4 (silicon nitride)	170~190 W/(K m)	electrical insulator
C (diamond)	1 000 W/(K m)	electrical insulator

For comparison, we give the thermal conductivity of a few other materials:

<i>Material</i>	<i>Thermal conductivity</i>	<i>Comment</i>
Plastic (acrylic glass or PMMA)	0.17~0.20 W/(K m)	electrical insulator
Al (aluminium)	237 W/(K m)	metal
Cu (copper)	401 W/(K m)	metal

That is, ceramics can have a thermal conductivity that is much higher than that of plastic materials. However, common ceramics (Al_2O_3 and Si_3N_4) have a lower thermal conductivity than common metals (Al and Cu).

An LED based on a ceramic substrate material is shown in *Figure 31.9* (Philips Lumileds Company, 2012 and Lumileds Company, 2015). The device package has two electrical contacts (anode and cathode) as well as a thermal pad located under the LED chip. The thermal pad is

⁴⁷ Note that the resin packages are *one body*, the resin-molded body, during the molding process. The individual LED resin packages emerge only during the singulation process.

⁴⁸ The thermal conductivities of various materials have been summarized by Wikipedia and can be found at <https://en.wikipedia.org/wiki/List_of_thermal_conductivities>

electrically insulated from anode and cathode. The LED device is also equipped with an ESD protection device that is a small black device located next to the LED chip.⁴⁹

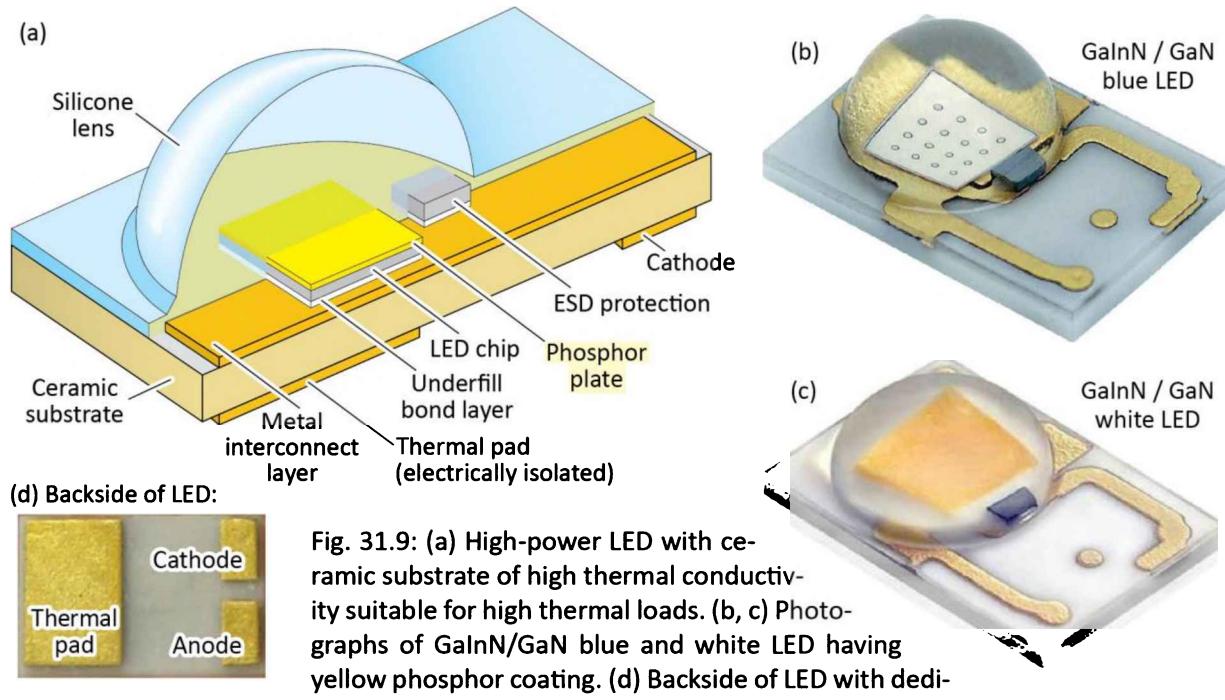


Fig. 31.9: (a) High-power LED with ceramic substrate of high thermal conductivity suitable for high thermal loads. (b, c) Photographs of GaInN/GaN blue and white LED having yellow phosphor coating. (d) Backside of LED with dedicated thermal pad electrically insulated from anode and cathode (Philips Lumileds Luxeon Rebel, 2012).

Such ceramic-substrate based devices are fabricated in arrays accommodating many devices, e.g. more than 200 devices, as evident from the array shown in **Figure 31.10**. Subsequent to die-bonding, ESD protection device bonding, wire bonding (if applicable), phosphor dispensation (if applicable), and encapsulation, the array is singulated by cutting, e.g. by a precision diamond circular saw.

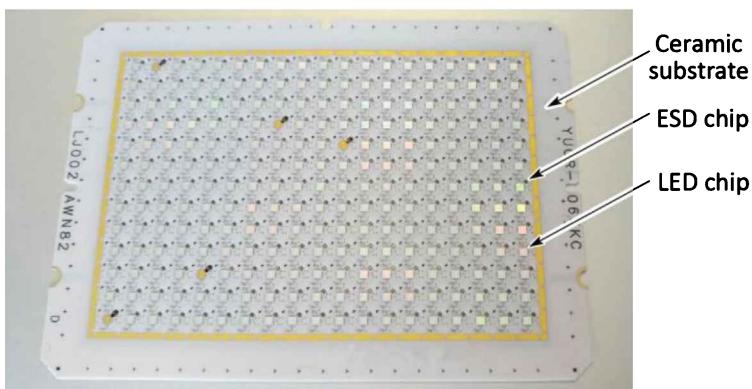


Fig. 31.10: Array of LED chips (19×13) and associated electrostatic discharge (ESD) protection devices (e.g. Zener diodes) on a ceramic substrate.

⁴⁹ Typical ESD (Electro-Static Discharge) protection devices are Zener diodes that are connected in an anti-parallel manner to the LED chip.

31.5 Side-view LED package

Side-view LEDs emit light in the lateral direction (i.e. along the plane of a PCB or FCB⁵⁰).⁵¹ Side-view LEDs generally have a low profile with an elongated light-emission window so that the emitted light can be efficiently coupled into a light guide. **Figure 31.11** shows a side-view LED (Ishida, 2004). Inspection of the figure shows that the lead contacts protrude the sides of the resin package and are bent so as to be solderable to a PCB or FCB. In order to accommodate the low profile of the package, rectangular rather than square-shaped chips are used. The primary application of side-view LEDs is backlighting of LCDs (Liquid Crystal Displays). Given the abundance of mobile display devices, side-view LEDs are manufactured in large quantities.

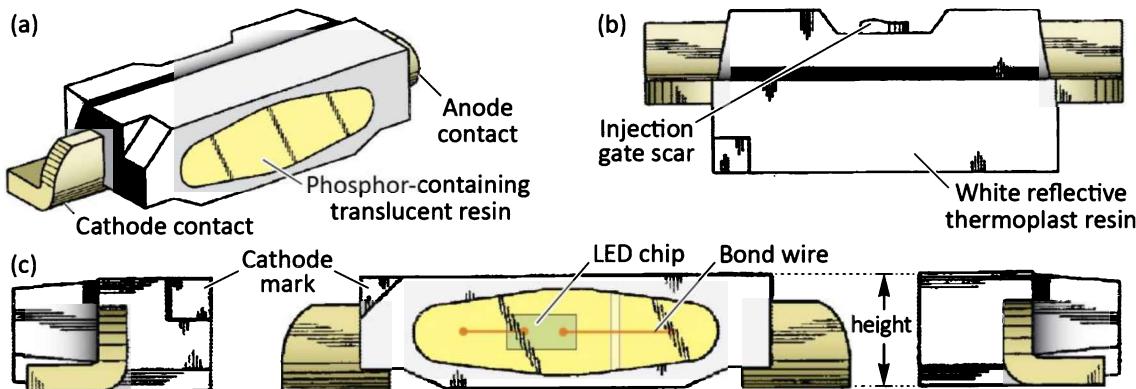


Fig. 31.11: (a) Perspective view, (b) top view, and (c) side views of an injection-molded side-view LED (SV-LED) package as used for the back-lighting of displays with an edge-lit light guide. Heights of SV-LEDs typically range from 0.5 to 1.0 mm (after Ishida, 2004).

Photographs of several side-view LEDs are shown in **Figure 31.12**. In order to reduce the width of the device, the contact leads are bent multiple times so that they can be folded under the white resin package. The package height of side-view LEDs is generally small with typical heights below 1.0 mm and as low as 500 µm.

⁵⁰ PCB = Printed Circuit Board; FCB = Flexible Circuit Board

⁵¹ The devices may more appropriately be called “side-emitting SMD LEDs”. However, the name “side-view LEDs” is widely established so that we will use it in the present chapter.



Fig. 31.12: (a), (b), and (c) Photographs of various side-view white LEDs as used for back-lighting in display applications, such as liquid-crystal display (LCD) applications.

As mentioned above, the primary application of side-view LEDs is the back-lighting of liquid-crystal displays (LCDs). A back-lighting unit is positioned behind the LCD and provides the LCD with a uniform intensity of light. The side-view LEDs couple light into a light guide, typically a plastic sheet, that provides the backlight for the LCD. A schematic setup is shown in *Figure 31.13* where two side-view LEDs inject light into the edges of the light guide. Side-view LEDs can illuminate one, two, three, or four of edges of the light guide.

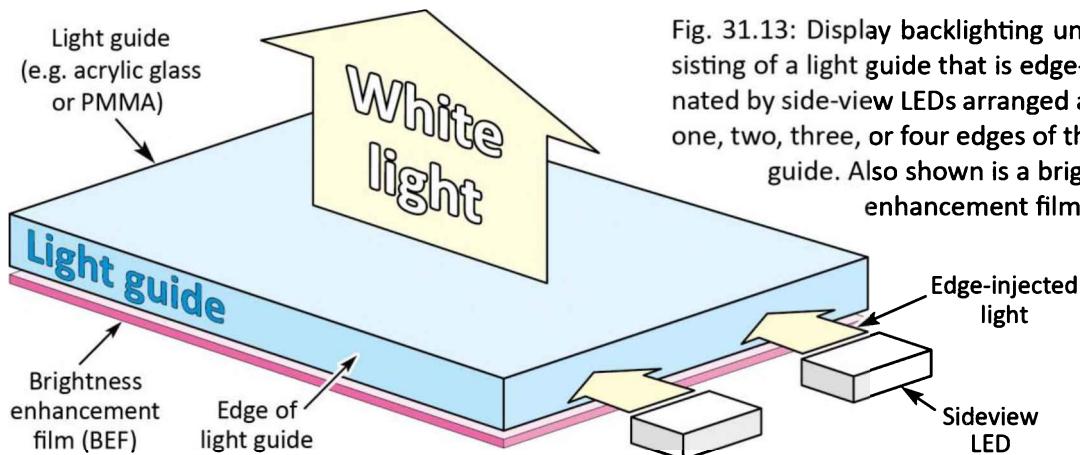


Fig. 31.13: Display backlighting unit consisting of a light guide that is edge-illuminated by side-view LEDs arranged around one, two, three, or four edges of the light guide. Also shown is a brightness enhancement film (BEF).

A cross section of the light guide is shown in *Figure 31.14*. As indicated in the figure, the light guide is not perfectly planar but rather equipped with optical scattering centers that scatter or re-direct the light guided by the light guide towards the exit surface (top surface). In order to reduce the undesirable optical losses on the back side, the back side is equipped with optical films that reflect the light. A common reflector film is the brightness enhancement film (BEF) and the dual brightness enhancement film (DBEF).

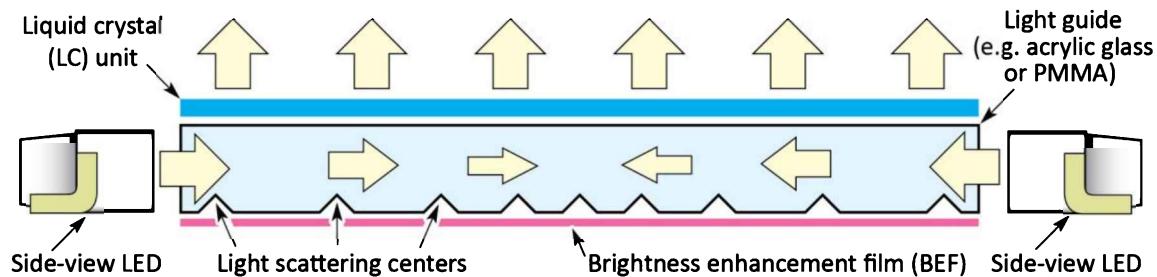


Fig. 31.14: Cross-sectional view of a liquid-crystal display (LCD) that includes a light guide having light-scattering centers to uniformly out-couple the light, a brightness enhancement film (BEF), and a liquid crystal (LC) unit.

For completeness, a back-illuminated light guide is shown in **Figure 31.15**. Back-illuminated light guides do not employ side-view LEDs but rather regular SMD LEDs. Back-illuminated light guides are typically used for larger display sizes.

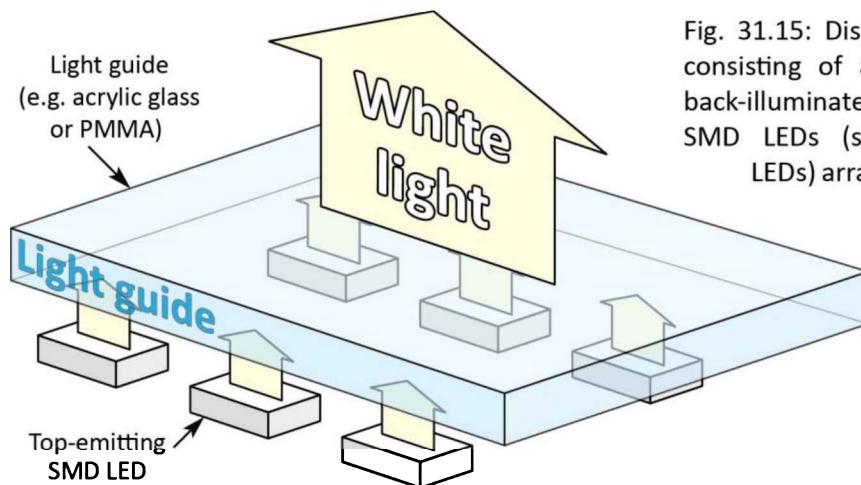


Fig. 31.15: Display backlighting unit consisting of a light guide that is back-illuminated by top-emitting SMD LEDs (surface-mount device LEDs) arranged on the back side of the light guide.

31.6 Other examples of LED packages

A high-power LED package that employs a dedicated large metallic heat sink is shown in **Figure 31.16** (after Krames, 2003). Power packages have a *direct, thermally conductive path* from the LED chip, through the package, to a heat sink, e.g. a printed circuit board. The power package shown in the figure has several features. Firstly, the package contains an Al or Cu heatsink slug with low thermal resistivity to which the LED submount is soldered by a metal-based solder. Secondly, the chip is encapsulated with silicone. Because standard silicone retains mechanical softness in its cured state, the silicone encapsulant is covered with a plastic cover that also serves as lens. Thirdly, the chip is directly mounted on a Si submount that includes electrostatic discharge protection (ESD).

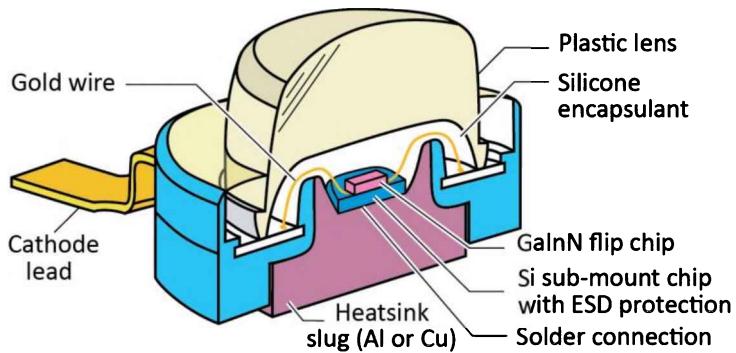


Fig. 31.16: Cross section through high-power package. The heatsink slug can be soldered to a printed circuit board for efficient heat removal. This package, called the *Barracuda package*, was introduced by Philips Lumileds Company (adapted from Krames, 2003).

LED packages may also include a lens that is shaped according to the application. Generally, convex lenses collimate or focus light whereas concave lenses further disperse light. A secondary optics lens that is intended to disperse the emitted light and enhance the light intensity along the lateral direction is shown in **Figure 31.17**. Such device is intended for use in a backlighting unit that employs a back-illuminated light guide.

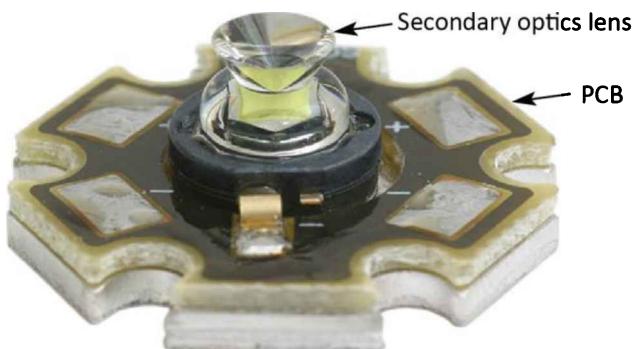


Fig. 31.17: Example of packaged LED having secondary optics that disperses light along the lateral direction. This is accomplished by a lobe-shaped lens that has a concave top surface. The LED is mounted on a PCB (printed circuit board).

31.7 Chip-on-board LED packages (COBs)

In **chip-on-board** (COB) LED packages, multiple LED chips are mounted on a common board (or substrate). Such devices are known as COB LEDs. The number of LED chips can range from 5 to 50. Due to the large number of LED chips, COB devices are suitable for high electrical input powers such as 10 to 100 W. The multiple chips are subjected to a single encapsulation process. **Figure 31.18** shows examples of COB devices. Due to their high-power capability, COB devices are used in applications that require a high luminous flux such as high-bay lighting, street lighting, high-output track lighting, and downlights.

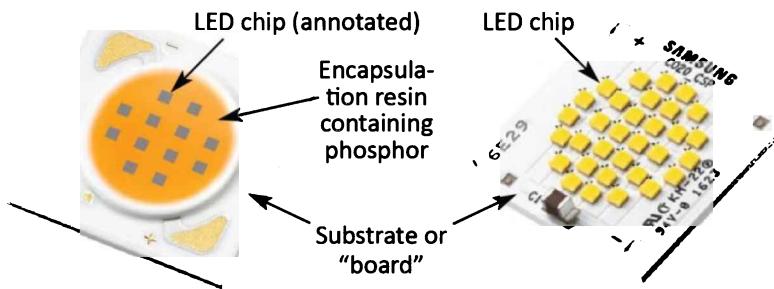


Fig. 31.18: Example of a “chip-on-board” LED module or COB LED that has multiple LED chips placed on one board thereby making COB LEDs suitable for high electrical input powers such as 10 to 100 W.

The LED chips in a COB device are frequently wired in a series circuit so that the operating voltage of a COB device exceeds the drive voltage of a single diode. Drive voltages of COB devices are, for example 12 V, 24 V, or 48 V.

31.8 Chip-scale LED packages (CSPs)

There is a general desire to reduce the packaging costs of LEDs. Packaging, although not as sophisticated as LED chips, nevertheless is highly advanced. Importantly, because packaging is more serial in nature, packaging is inherently expensive. Accordingly, efforts have been made to replace conventional packaging by **chip-scale packaging**. This trend is found in the field of silicon integrated circuits (Garrou, 2000; Hotchkiss *et al.*, 2001; Kim *et al.*, 2001) as well as well as in the field of LEDs (Kim *et al.*, 2009; Chen *et al.*, 2010; Lee *et al.*, 2012).

The basic idea behind chip-scale packaging is to perform all operations, including the packaging process on the wafer scale. As a result, chip-scale packaged LEDs resemble bare chips to the casual observer. However, closer inspection, shows that an additional substrate, frequently Si or glass, is used to provide the electrical contact pads on the bottom side of the CSP LED. An example of a CSP LED is shown in *Figure 31.19* (after Samsung Company, 2016).

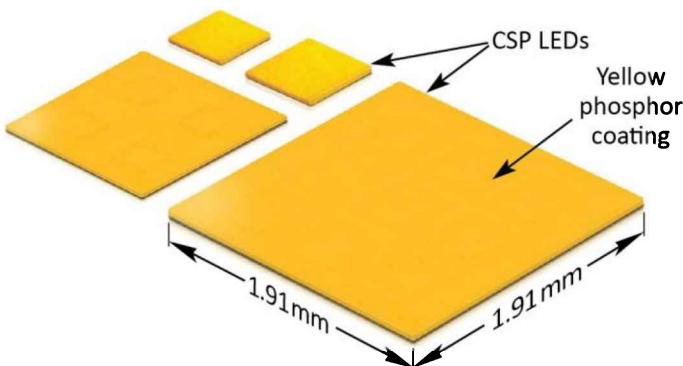


Fig. 31.19: Example of a “chip-scale-package” LED or CSP LED. The LED does not have a package in the conventional sense (after Samsung Company, 2016).

31.9 References

- Chen K., Zhang R., and Lee S. W. R. "Integration of phosphor printing and encapsulant dispensing processes for wafer level LED array packaging" *11th International Conference on Electronic Packaging Technology & High-Density Packaging*, ISBN 978-1-4244-8142-2, page 1386 (2010)
- Ichikawa H., Hayashi M., Sasaoka S., and Miki T. "Light emitting device, resin package, resin-molded body, and methods for manufacturing light emitting device, resin package and resin-molded body" US Patent 8,530,250 B2, issued September 10 (2013)
- Lee S. W. R., Zhang R., Chen K., Lo J. C. C. "Emerging trend for LED wafer level packaging" *Frontiers in Optoelectronics* **5**, 119 (2012)
- Lumileds Company "DS58 LUXEON Rebel White & PC Amber Product Datasheet" San Jose, California, USA (2015)
- Garrou P. "Wafer level chip scale packaging (WL-CSP): An overview" *IEEE Transactions on Advanced Packaging* **23**, 198 (2000)
- Hotchkiss G., Amador G., Edwards D., Hundt P., Stark L., Stierman R., and Heinen G. "Wafer level packaging of a tape flip-chip chip scale packages" *Microelectronics Reliability* **41**, 705 (2001)
- Ishida M. "Light emitting diode" US Patent D 491,538 S, issued June 15 (2004)
- Kim N. S., Jang D. H., Kang S. Y., Kwon H. K. "Chip scale packages and methods for manufacturing the chip scale packages at wafer level" US Patent 6,187,615 B1, issued February 13 (2001)
- Kim S.-M., Lee K.-C., Yu Y. M., Baek J. H., and Jung G. Y. "Wafer-level packaged light-emitting diodes using photo-dielectric resin" *IEEE Electron Device Letters* **30**, 638 (2009)
- Krames M. R. "Overview of current status and recent progress of LED technology" US Department of Energy Workshop "Solid State Lighting – Illuminating the Challenges" Crystal City, VA, Nov. 13~14 (2003)
- Möllmer F. and Waitl G. "Siemens-SMD-TOP-LEDs – LEDs for Surface Mounting" *Siemens Components* **26**, 147 (1991)
- Nichia Company "2016 LED catalogue" <http://www.nichia.co.jp/en/product/led_catalogue.html> (2016)
- Philips Lumileds Company "LUXEON Rebel and LUXEON Rebel ES Product Datasheet DS68" San Jose, California, USA (2012)
- Samsung Company "Samsung introduces full line-up of LED components based on chip-scale packaging technology based on chip-scale packaging technology" *Samsung Press Release*, March 14 (2016)

32

32 – Packaging: Materials used in packaging

This chapter is intended to convey some basic properties of materials used in LED packaging. The chapter is not intended to be comprehensive but rather give an overview of materials issues encountered in the field of LEDs.

32.1 Resin materials

Resin materials play a prominent role in microelectronic and optoelectronic packaging. Encapsulating resins in the context of LEDs have several requirements including high transparency, high refractive index, chemical stability, high-temperature stability, stability under high optical radiation intensities, and hermeticity (barrier against moisture). The word “resin” was historically used for the viscous transparent substance that is extruded from trees, specifically coniferous trees⁵² such as the pine tree. An image of a tree resin is shown in **Figure 32.1**.



(a)



Fig. 32.1: (a) Tree resin, a viscous optically transparent (or translucent) material. The tree resin hardens over time and becomes amber. (b) Amber may have inclusions such as bark, dust, or the remnants of an insect. A synthetic resin is a man-made material that bears some similarities with tree resins.

Accordingly, a **synthetic resin** is a man-made substance that has properties that, to some degree, resemble the properties of tree resins. Synthetic resins⁵³ are generally transparent polymeric organic materials that can be mechanically soft. Indeed, the degree of softness varies among the

⁵² Coniferous trees are trees with needle-shaped leaves.

⁵³ For convenience, we may refer to “synthetic resins” simply as “resins” as is customary within the technical community.

synthetic resins and can range from very soft (gel like), to intermediate (resilient), to hard as, for example, plexiglass.

The chemical structure of synthetic resins, shown in *Figure 32.2*, is a polymer that has an atomic “**backbone**”, e.g. a sequence of C atoms. Chemically bonded to the backbone are atoms, such as H atoms, or molecular groups such as the $-\text{CH}_3$ group (methyl group). The atomic or molecular groups bonded to the atomic backbone are referred to as **functional groups**.⁵⁴ Inspection of the figure reveals that the polymeric chains can be branched and cross-linked. The density of branching points and crosslinking points determined the mechanical viscosity and hardness of a polymer. A polymer completely lacking branches or cross-links tend to be liquid.⁵⁵ A polymer becomes viscous and ultimately solid as the density of branches and cross-links increases.

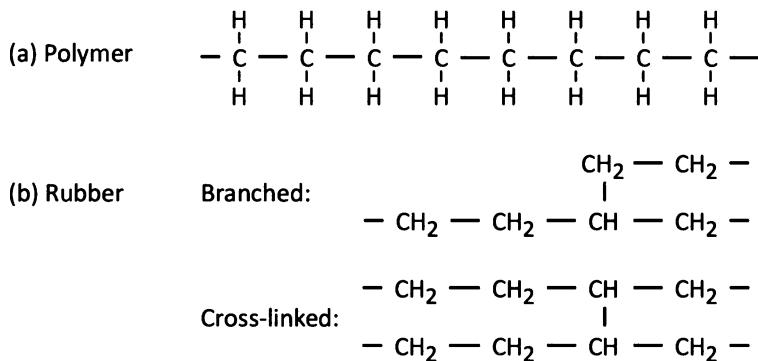


Fig. 32.2: Chemical structure of a polymer including the atomic “backbone” e.g. a carbon-atom backbone. Polymers can be branched and cross-linked which determines the mechanical hardness of the polymer.

Rubber compounds lack transparency and cannot be used as LED encapsulants. However, it is well known that oxides are frequently transparent. In fact, encapsulants used for LEDs frequently contain oxygen.

32.2 Types of resins

There are two major types of synthetic resins: The first type is the **thermoplastic resin**. It is a solid resin that becomes plastic (and moldable) when heated. The solid-to-plastic transition is a physical change that is repeatable. In contrast, the second type of resin, the **thermosetting resin**,

⁵⁴ “Functional groups” is used in conjunction with polymers. Coordination chemistry concerns compounds having a central metal atom. A functional group attached to a central metal atom is referred to as a **ligand**.

⁵⁵ Polymers consisting of very long atomic backbones, and not having any branches or cross-links, can be solid as well.

undergoes a chemical change during the curing process upon which the resin irreversibly hardens.

Thermoplastic resins

A thermoplastic resin is a solid polymer material that becomes plastic (moldable, viscous) when heated to higher temperatures. The solid-to-plastic transition is a repeatable physical change that is not associated with a chemical change of the material. Because the solid-to-plastic transition of thermoplastic polymers is reminiscent to the solid-glass-to-viscous-glass transition of glasses (such as SiO_2), the temperature at which the solid-to-plastic transition (or solid-to-viscous-transition) occurs is called the **glass transition temperature**. The glass-transition temperature of thermoplastic polymers can vary over a wide range with typical values being between 70°C to 200°C. The melting temperature is always higher than the glass-transition temperature. Thermoplastic resins, as used for optoelectronic applications, are typically transparent in their pure but can be modified to become reflective, as discussed below.

An exemplary thermoplastic material is PMMA or poly-(methyl methacrylate). It is also known as acrylic glass and under the trade name plexiglass. PMMA is a transparent resin with a glass transition temperature of 105°C. That is, for temperatures $> 105^\circ\text{C}$, PMMA becomes moldable. PMMA has been employed as LED encapsulation material. The chemical structure of one of PMMA's molecular unit cells is shown in **Figure 32.3**. The relatively low refractive index of PMMA ($\bar{n} = 1.49$ in the wavelength range 500~650 nm) results in a limited light-extraction efficiency when used with high-index semiconductors.

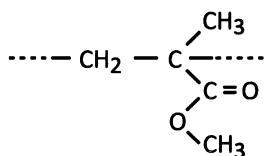


Fig. 32.3: Chemical structure of poly methyl methacrylate (PMMA) or acrylic glass.

Another exemplary thermoplastic material is nylon. Nylon-6,6 in the dry state has a glass transition temperature of about 70°C.

Molding thermoplastic resins is commonly accomplished by means of **injection molding**. The thermoplastic material is provided in the form of resin pellets. The molding usually employs a lower mold and an upper mold with a hollow molding cavity between the two molds. While being injected into the mold's cavity, the resin must have a temperature above the glass transition

temperature. This can be accomplished through either heating the resin pellets or application of high pressure in an auger-type (screw-type) injector; pressure causes the temperature to increase so that the solid pellets become viscous and moldable resin. In addition, the mold may be heated to temperatures exceeding the resin's glass-transition temperature. During injection molding, the pressure is kept sufficiently low so as to avoid damage (deformation) to inserts (e.g. a metal lead frame insert) by the viscous resin. Examples of ***thermoplastic*** polymeric resins are:

Liquid-crystal polymer (LCPs) is a class of aromatic polymers. Below the glass transition temperature (T_{glass}), thermoplastic polymers possess molecular orders and thus have a semi-crystalline arrangement of the polymer molecule. At temperatures above T_{glass} , the semi-crystalline order is usually lost so that polymers become a disordered, viscous liquid. However, in LCPs the molecular order is *retained* in their *viscous liquid* state (above T_{glass}). Hence, the name “liquid crystal polymer”. Like traditional semi-crystalline thermoplastic polymers, the glass transition temperature in LCPs is associated with the co-operative movements of the polymer molecule in the amorphous phase.

Nylon is a common thermoplastic polymeric material having a C backbone, and containing N (replacing some of the C atoms), O (double-bonded to C), and H.

Polycarbonate contains aromatic rings, is highly transparent, and mechanically strong.

Polyphenylene sulfide (PPS) is a polymer consisting of aromatic rings linked by sulfides (sulfur atoms). The material has good thermal and chemical resistance.

Poly (methyl methacrylate)⁵⁶ (PMMA) is a common thermoplastic material frequently used for LEDs.

Polyphthalamide (PPA) is a semi-crystalline, aromatic high-performance polyamide. PPA is strong, stiff, insensitive to moisture, and resistive to chemical fatigue.

Thermosetting resins

A thermosetting resin is a liquid (or viscous) material that becomes solid (i.e. that “sets”) upon curing (e.g. by heating). Thermosetting materials can consist of a single or two chemical components. The uncured compounds of the resin can be highly liquid and thus easily flow into a mold. During the curing, the thermosetting material undergoes a chemical reaction that causes the material to solidify. Pure thermosetting resins typically are optically transparent. Exemplary thermosetting resin materials are epoxy (EMC or epoxy molding compound) and silicone. Both will be discussed in greater detail below.

⁵⁶ Note that some acrylates are thermosetting materials.

Molding thermosetting resins is accomplished by ***transfer molding***. The molding usually employs a lower mold and an upper mold with a hollow molding cavity formed between the two molds. Due to the high liquidity (low viscosity) of some uncured resins (e.g. a two-component uncured epoxy resin not employing a filler), these resins can easily fill into the mold cavity, even at room temperature.⁵⁷ If a mold cavity is rich in features, a high pressure is desirable to minimize molding defects.

In contrast to thermoplastic resins, thermosetting resins (in the uncured and cured state) can have **–OH** groups (hydroxyl groups) at various (end-) points of the polymer molecule (Colas, 1997; Kuramoto and Kishimoto, 2006). The hydroxyl groups can form oxygen bonds as well as hydrogen bridge bonds⁵⁸ with adjoining substrate materials (such as a metal lead frame), thereby enabling good adhesion⁵⁹ of thermosetting resins (Colas, 1997; Kuramoto and Kishimoto, 2006). Examples of common ***thermosetting*** polymeric resins are:

Acrylic resins⁶⁰ can be a thermosetting polymer resin that is derived from acrylic acid, methacrylic acid or other related compounds.

Epoxy is frequently used in the LED and microelectronics industry. Epoxy usually is made from two components, component A and B.⁶¹ Upon mixing and curing, the resin is formed. The curing reaction frequently is exothermic, thereby heating the resin and accelerating the curing process. Epoxy is mechanically strong, is chemically inert, has high optical transparency, and provides a good barrier against moisture.

Polyurethane⁶² (PUR) can be thermosetting polymer resins composed of organic units joined by carbamate (urethane) links. While most polyurethanes are thermosetting polymers that do not melt when heated, thermoplastic polyurethanes are also available.

⁵⁷ The liquidity (and viscosity) of the resin depend on the quantity of fillers used. Fillers generally decrease the liquidity (i.e. increase the viscosity) of a resin.

⁵⁸ H atoms, when covalently bonded to a highly electronegative atom (such as N, O, or F), are polarized and carry a partial positive charge. H-bridge bonds are due to the electrostatic attraction between two polar groups. H-bridge bonds provide good adhesion to other materials and are due to Van der Waals forces. “Van der Waals force” is a general term used to describe attractive electrostatic intermolecular forces.

⁵⁹ Adhesion is the result of the physical and chemical bonding between two materials. Highly reactive groups are the hydroxyl group (**–OH**) or the carbonyl group (**=C=O**). These groups enable chemical attraction through van der Waals forces as well as the formation of oxygen bonds with the adjoining substrate material. (Unrelated to microelectronic applications, sugar molecules have multiple hydroxyl groups (**–OH**); as a result, sugars, are generally sticky and can even be used for making glues.)

⁶⁰ Note that some acrylic resins are thermoplastic materials.)

⁶¹ Chemical component A and B are sometimes called “resin” and “hardener”. Upon curing, the resulting substance is called “resin” as well.

⁶² Note that some polyurethanes are thermoplastic materials.

Silicone is frequently used in the LED and microelectronics industry. Silicone can be made by curing a single chemical component (more common) or two components (less common). Uncured silicone may be cooled for shipping. Silicone is mechanically soft, chemically inert, has high optical transparency, and is resistant to photo-degradation.

32.3 Modification of resins by fillers

Resins, in their pure state, are optically transparent. In order to change the properties of resins, they are modified by the addition of **fillers**.⁶³ These fillers include light-scattering particles that modify the resin from being transparent (transparent-glass-like), to translucent (frosted-glass-like), and ultimately to reflective (white-colored porcelain-like). Suitable light-diffusing particles, are inorganic, non-absorbing at the wavelength of interest, and include TiO₂ (titania), Al₂O₃ (alumina), SiO₂ (silica), and BaTiO₃ (barium titanate). Since the resin and the light-scattering particles are non-absorbing, the modified resin is non-absorbing as well. An example of a clear resin (no fillers) and white resin (filled with light-scattering particles) is shown in **Figure 32.4** (a) and (b), respectively.



Fig. 32.4: (a) Clear or transparent resin. (b) White or reflective resin.
(c) Black or absorbing resin.

On the other hand, if it is desired that the resin absorbs light, a light-absorbing pigment can be added to the resin. A primary example of light absorbing pigment is the “carbon black” pigment.⁶⁴ Addition of such light-absorbing pigment makes the resin dark-colored or black as shown in **Figure 32.4** (c). Light-absorbing resins are common in the packaging of electronic devices such as Si integrated circuits (ICs) in order to avoid illumination of the IC chip. Light incident on the IC chip would undesirably modify its properties by means of photo-currents.

⁶³ “Fillers” includes any types of diffusing fillers (diffusants) and absorbing fillers.

⁶⁴ The “carbon black” pigment is a strongly light-absorbing powder that consists of mostly carbon (C) in a configuration that has a high-surface-to-volume ratio. The “carbon black” pigment can be made by incomplete combustion of heavy petroleum products such as tar.

For a white light-reflecting resin that is filled with light-scattering particles such as TiO₂, reflectivities greater than 90% can be achieved. For a light-absorbing resin (black resin) that is filled with a light-absorbing pigment such as carbon-black pigment, transmittances less than 0.1% can be achieved.

Encapsulants containing **mineral diffusers** cause light to reflect, refract, and scatter, thereby randomizing the propagation direction and making the far-field distribution isotropic. For multi-color devices (e.g. multi-chip white LEDs), mineral diffusers make the color distribution more uniform. Mineral diffusers are optically transparent substances, such as TiO₂, CaF₂, SiO₂, CaCO₃, and BaSO₄, with a refractive index different from that of the encapsulant (Reeh *et al.*, 2003).

Fillers may also be used to reduce the resin's coefficient of thermal expansion (CTE). Pure resin may have a relatively high CTE, e.g. 50 ppm/°C. Inorganic mineral fillers may have a relatively low CTE, e.g. 5 ppm/°C. Consequently, the addition of a filler can reduce the CTE of the resin, e.g. to values less than 50 ppm/°C.

The maximum volume fill factor of a filler consisting of equally-sized spherical particles that are arranged in the form of a simple cubic lattice with lattice constant a is $\frac{4}{3}\pi\left(\frac{1}{2}a\right)^3/a^3 = 52.4\%$. If the equally-sized spherical particles are arranged in a densely-packed hexagonal lattice, the maximum volume fill factor is given by $\pi/(3 \times 2^{1/2}) = 74\%$.⁶⁵ If there are additional smaller spheres that are packed into the gaps between the larger spheres, the maximum fill factor can exceed 74%.

A translucent resin can have an advantage over a transparent resin. In particular, some scattering of an LED-chip-encapsulating resin can be advantageous to enhance light extraction. It is known that light propagating in a clear (transparent) encapsulating resin of an LED package is subject to some trapped optical modes that cannot be outcoupled from the package. Such trapped modes include “whispering gallery modes” that propagate along the periphery of a spherical or hemispherical resin package (Luo *et al.*, 2006). In the absence of scattering, the light rays are deterministic paths that correspond to trapped optical modes. Accordingly, some scattering may be desirable. This is confirmed by ray-tracing simulations, shown in **Figure 32.5**, which were performed on LED chips covered by a thin sheet of an encapsulating resin (Shiang *et al.*, 2004; Shiang and Duggal, 2004; Luo *et al.*, 2007). Inspection of the figure reveals that there is an optimum amount of scattering that maximizes the light-extraction efficiency.

⁶⁵ The maximum volume fill factor for the dense arrangement of equally-sized spheres was derived by Carl Friedrich Gauss. Detailed information can be found in Wikipedia under the entry “Close-packing of equal spheres” (2017).

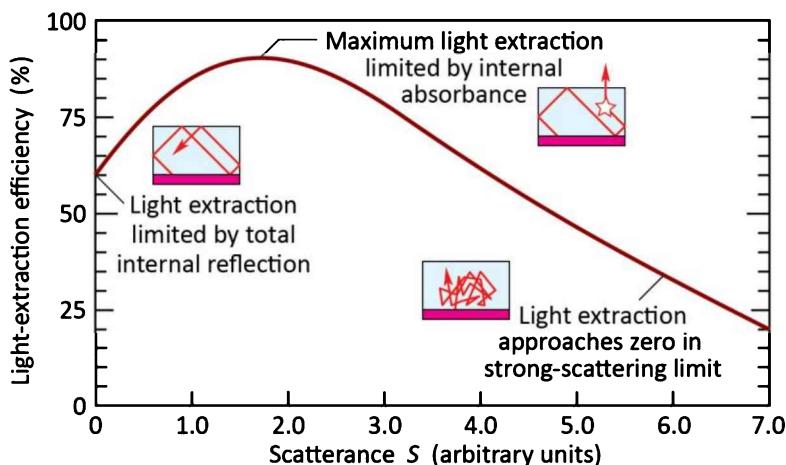


Fig. 32.5: Light-extraction efficiency of an LED coated by a thin-film encapsulating resin having a constant thickness. The light extraction has a maximum for a moderate scattering strength in the encapsulating resin (after Shiang *et al.*, 2004; Shiang and Duggal, 2004; and Luo *et al.*, 2006).

32.4 Epoxy resins

A common encapsulant is ***epoxy resin*** (also called epoxy), which has low shrinkage upon curing, remains transparent, and does not show degradation over many years for long-wavelength visible-spectrum and IR LEDs. However, it has been reported that epoxy resins lose transparency in LEDs emitting at shorter wavelengths, i.e. in the blue, violet, and UV (Barton *et al.*, 1998). Epoxy resins are chemically stable up to temperatures of about 120°C (Denk and Reeh, 1979). Prolonged exposure to temperatures greater than 120°C leads to yellowing (loss of transparency).

The ***epoxy group***, shown in ***Figure 32.6***, contains an O atom attached to two C atoms already bonded to each other. Such a three-membered ring consisting of one oxygen and two carbon atoms is shown in ***Figure 32.6*** and the epoxy group occurs particularly at the ends of the epoxy resin molecule (Denk and Reeh, 1979). The epoxy molecule shown in the figure is “diglycidyl ether of bisphenol-A” and includes two benzene rings that ensure mechanical flexibility (steric flexibility) of the molecule allowing it to form a large number of co-polymerization cross links.

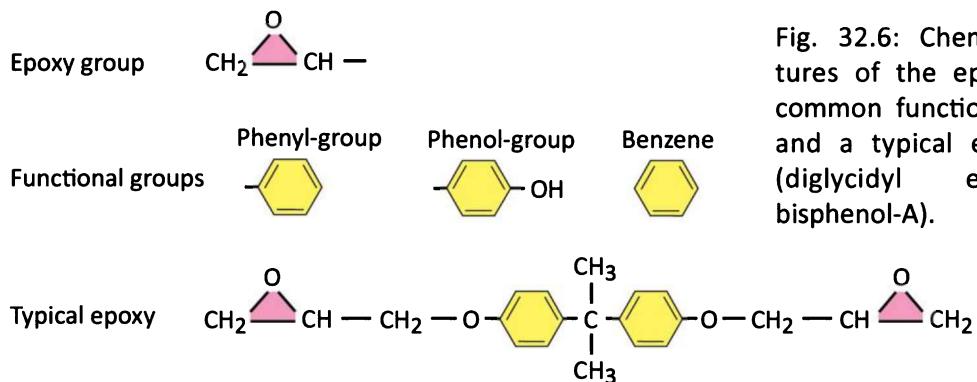


Fig. 32.6: Chemical structures of the epoxy group, common functional groups, and a typical epoxy resin (diglycidyl ether of bisphenol-A).

An epoxy molecule can also have various functional groups chemically bonded to the molecule including the phenyl- and phenol-group, both of which derive themselves from the benzene molecule (Carmody and Kelly, 1943).

Solid epoxy resins can be formed by stoichiometrically mixing two liquid compounds, the epoxy compound with another compound, a hardener or curing agent. The curing agent has at least two sites that tend to chemically bond with epoxy groups. Frequently, the atomic sites bonding with the epoxy group are part of amino groups⁶⁶, hydroxyl groups⁶⁷, or organic acid anhydride groups⁶⁸ (Denk and Reeh, 1979). Upon curing the epoxy group of the epoxy resin molecule reacts with the amino group (or hydroxyl or anhydride group) of the hardener. The reaction between the epoxy group and an amine curing agent is shown in **Figure 32.7**. Inspection of the figure reveals that during the thermal-setting process, the epoxy groups co-polymerize with amino groups (or hydroxyl groups or anhydride groups) of the curing agent.

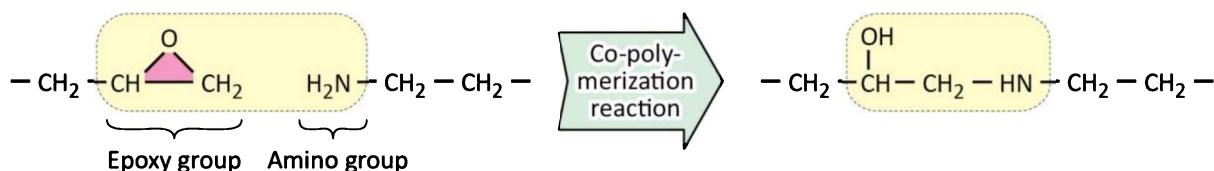


Fig. 32.7: Co-polymerization reaction of the epoxy group of a polymer with the amino group of another polymer.

As a result of the curing process (i.e. co-polymerization reactions), a highly cross-linked polymer is obtained, as shown in **Figure 32.8**, where each one of the curing agent's amino groups has two bonding sites. The tetrahedral bonding configuration of the carbon atom affords the epoxy polymer molecule significant steric flexibility so that a large number of crosslinks is enabled.

⁶⁶ Amino groups are $-\text{NH}_2$ groups that are typically connected to a carbon atom, i.e. $\equiv\text{C}-\text{NH}_2$. The group is also called an “amino radical” thus indicating its relatively high reactivity. An organic compound that contains an amino group is called an amine.

⁶⁷ Hydroxyl groups are $-\text{O}-\text{H}$ groups that are typically connected to a carbon atom, i.e. $\equiv\text{C}-\text{O}-\text{H}$. Related to hydroxyl groups are organic acid anhydrides with $-\text{COOH}$ being the organic acid and $-\text{COO}$ being the organic acid anhydride.

⁶⁸ Organic acids are characterized by the $-\text{COOH}$ functional group (carboxyl group). Two $-\text{COOH}$ groups can react to form an anhydride (i.e. a compound “lacking water”), i.e. $-\text{COOH} + -\text{COOH} \rightarrow -\text{CO}-\text{O}-\text{CO}- + \text{H}_2\text{O}$.

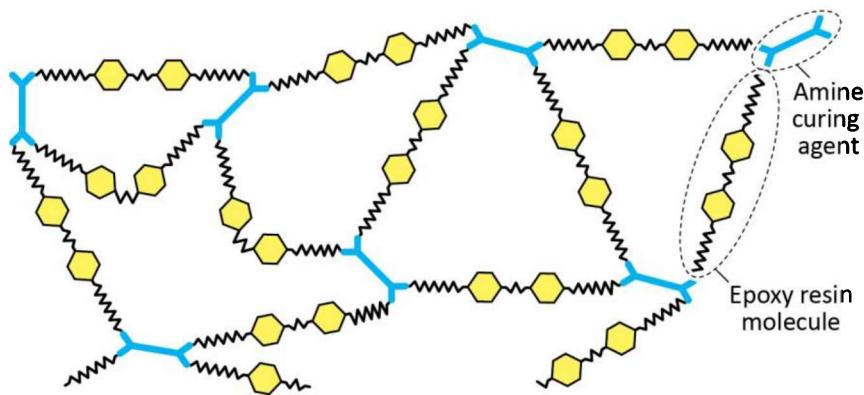


Fig. 32.8: Chemical structure of cured epoxy resin. The resin is strongly cross-linked (co-polymerized) by means of an amine curing agent. The epoxy molecule with its C atoms and benzene rings offers steric flexibility, thereby enabling many cross-links.

There is a variety of epoxy-based resins available. During the curing process, the epoxy resin molecule may be made to react (co-polymerize) with either (i) itself⁶⁹ or (ii) a curing agent (hardener). In both cases, the resulting epoxy resin is a highly cross-linked polymer. If a curing agent is used, the epoxy resin molecule's epoxy group reacts (co-polymerizes) with a functional group of a curing agent. Suitable functional groups of the curing agent include amines⁷⁰, organic acids⁷¹ (and acid anhydrides), phenols, alcohols and thiols⁷². Epoxy resins include thermally curable, UV-curable, and microwave-curable epoxy resins (Flick, 1993; Gorczyk, 2001; Kumar *et al.* 2006).

According to Kumar *et al.* (2006), typical epoxy resins for encapsulation are thermally curable two-part liquid systems consisting of bisphenol-A based or cycloaliphatic epoxide and an anhydride. The formation of such epoxy resin requires a short high-temperature cure (120°C). The two-part system has to be in stoichiometric proportions. Resin-rich compositions lead to lower glass transition temperatures while hardener-rich compositions may lead to discoloration of the encapsulant. The refractive index of epoxy resin is near $\bar{n} = 1.6$. (Kumar *et al.*, 2006).

In conclusion, in the context of LEDs, pure epoxy resins are known for having high transparency, good mechanical characteristics (low shrinkage upon curing), good thermal stability, and excellent hermeticity.⁷³ However, prolonged exposure of the epoxy to

⁶⁹ The epoxy group of one epoxy molecule reacts with the functional group of another epoxy molecule thereby cross-linking the two molecules.

⁷⁰ Amines are organic molecules having an amino group ($-\text{NH}_2$ group).

⁷¹ Organic acids are characterized by the $-\text{COOH}$ functional group (carboxyl group). Two $-\text{COOH}$ groups can react to form an anhydride (i.e. a compound “lacking water”), i.e. $-\text{COOH} + -\text{COOH} \rightarrow -\text{CO}-\text{O}-\text{CO}- + \text{H}_2\text{O}$.

⁷² Thiols are organic molecules containing an $-\text{SH}$ group. “Thiol” is a compound word that derives itself from “theion” (Greek: Sulfur) and the “ol” ending (in analogy alcohol, i.e. the $-\text{OH}$ group, or here, thiol, the $-\text{SH}$ group).

⁷³ Hermeticity is of importance for all electronic and optoelectronic components including LEDs. Hermeticity ensures that water, moisture, and humidity is kept away from the semiconductor component. Water, as a liquid, can be

temperatures exceeding 120°C will lead to discoloration and loss of transparency (Denk and Reeh, 1979). Additionally, the photo-degradation of epoxy (loss of transparency) by high intensity of blue and UV light has been reported (Barton *et al.*, 1998).

32.5 Silicone resins

The transition of LEDs from low-power to high-power devices created the demand for synthetic resins that are more stable than epoxy resin. The weaknesses of epoxy resin include limited thermal stability and degradation under high-intensity blue light and UV radiation. **Silicone resin** (also called **silicone rubber**, **poly-siloxane resin** or simply **siloxane resin**⁷⁴), a solution to this demand, has been used in the LED industry since the early 2000s (Norris *et al.*, 2005; Bahadur *et al.*, 2006). Silicone consists of an inorganic $-\text{Si}-\text{O}-\text{Si}-\text{O}-$ backbone to which organic functional groups are attached (Colas, 1997). Silicone is thermally stable up to temperatures of about 170~190°C, significantly higher than epoxies (Crivello, 2004; Bahadur *et al.*, 2006; Lin *et al.*, 2010). Furthermore, silicone has a higher threshold for photo-degradation than epoxy resin (Lin *et al.*, 2010). Silicone is flexible (and remains flexible for decades) thereby reducing the mechanical stress on the semiconductor chip and other elements such as bond wires. Various curing methods are available including thermal curing and UV curing. For LED packaging, the dominant cure system is “addition cure” (also called “neutral cure”) that releases no cure by-products and shows minimal cure shrinkage (Norris *et al.*, 2005).⁷⁵ Silicones can be formulated as one-part or two-part products (Norris *et al.*, 2005).

The basic structure of Silicone is shown in **Figure 32.9**. The polymeric “backbone” of silicone is an alternating chain of Si and O atoms. The Si atom has four bonds so that the two not forming the atomic chain, connect to organic functional groups. Possible functional groups include the methyl group ($-\text{CH}_3$), ethyl group ($-\text{C}_2\text{H}_5$ or $-\text{CH}_2-\text{CH}_3$), phenyl group ($-\text{C}_6\text{H}_5$), and phenol group ($-\text{C}_6\text{H}_4\text{OH}$). The Si atom connecting to two methyl groups yields **poly-dimethyl siloxane** (PDMS),

considered as an accelerator of most degradation reactions because chemical reactions in the liquid state proceed in a highly accelerated manner. Indeed, alchemists in the middle ages believed that “chemical reactions occur in the liquid state only” (in latin “chemica reactionem in liquidus status solus”). Although this is not true in a strict sense, it elucidates the need of keeping water, moisture, and humidity away from electronic components.

⁷⁴ “Siloxane” refers to the $-\text{Si}-\text{O}-\text{Si}-$ linkage.

⁷⁵ “Addition cure” and “neutral cure” refer to co-polymerization that is based on the pure addition of branches and cross-links without producing reaction-by-products such as volatile gases. Curing reactions that involve minimal shrinkage are desirable and available.

a common silicone oil.⁷⁶ If the C atom connects to two phenyl groups yields ***poly-diphenyl siloxane*** (PDPS), a common silicone for LED encapsulation (Kim *et al.*, 2010). The characteristics of silicone are determined by the choice of the functional groups and abundance of cross-links: Larger functional groups such as the propyl group ($-C_3H_7$ or $-CH_2-CH_2-CH_3$) or the phenyl group as well as a larger abundance of cross-links will introduce steric hindrance and thus make the silicone less flexible. Furthermore, the addition of larger-molecule functional groups such as the phenyl or phenol-group will increase the mass density of the polymer and thus the optical density of the material, i.e. increase the refractive index. The use of all methyl substituents on the silicone backbone results in a refractive index of $\bar{n} = 1.40$ (at $\lambda = 633$ nm); the use of phenyl groups on the silicone backbone allows for a refractive index of $\bar{n} = 1.55$ (Norris *et al.*, 2005). A high refractive index is desirable for LED-encapsulating resins for enhancing light extraction.⁷⁷

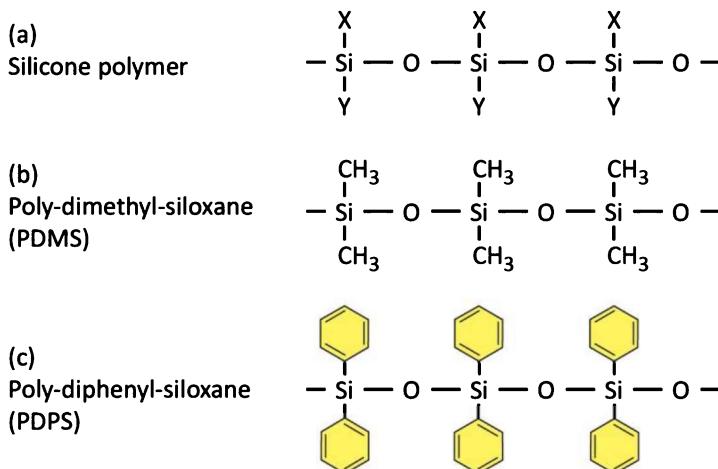


Fig. 32.9: (a) Chemical structure of silicone polymers. X and Y represent atoms or molecules (ligands) such as H (hydrogen), $-CH_3$ (methyl group), or $-C_6H_5$ (phenyl group). (b), (c) Common silicones are poly-dimethyl-siloxane (PDMS) and poly-diphenyl-siloxane (PDPS).

By choice of the functional groups and by the abundance of branches and cross-links, the mechanical consistency of silicone can be controlled. Indeed, silicones can be liquid (usually not of interest for LEDs), gel-like (very soft and deformable), elastomer-like (flexible and resilient such as rubber), to hard-plastic-like (solid and not easily indentable) (Norris *et al.*, 2005).⁷⁸ An

⁷⁶ PDMS has a low abundance of cross links and thus is a liquid.

⁷⁷ Norris *et al.* (2005) reported refractive indices of silicones to range between 1.38 and 1.58.

⁷⁸ The mechanical hardness of materials can be quantified by the Durometer Shore Hardness Scale, specifically the Hardness Scales "Shore 00" (very soft gels with Shore 00 ≤ 45), "Shore A" (elastomers with $0 \leq$ Shore A ≤ 100), and "Shore D" (hard resins with Shore D ≥ 60). The Durometer Shore Hardness Scale measures the hardness of polymers that range in hardness from very soft and flexible (soft gels), to medium and somewhat flexible (rubber like), to hard with almost no flexibility at all (hard plastic). Each Shore Hardness Scale ranges from 0 to 100 with 0

example of a highly cross-linked silicone resin structure is shown in *Figure 32.10*. The mechanical characteristics of the resin are of interest when considering thermal expansion of packaged devices.

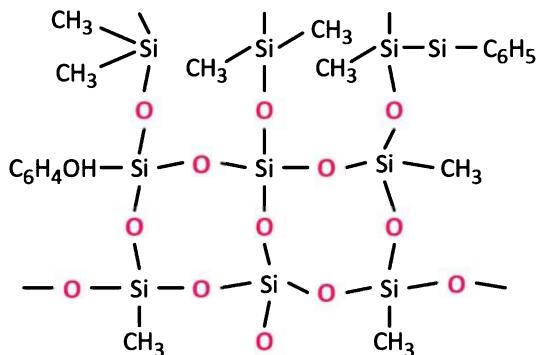


Fig. 32.10: Example of a strongly cross-linked silicone resin. The abundance of cross-links between $-\text{Si}-\text{O}-\text{Si}-\text{O}-\text{Si}-\text{O}-$ backbones determines the flexibility of silicone resins which can range from soft (gel-like) to flexible (rubber-like) to hard (hard-plastic-like).

Some insight can be gained by comparing silicone resin to glass (SiO_2). The atomic “backbone” of silicone is a sequence of an alternating series of Si and O atoms, i.e. $-\text{Si}-\text{O}-\text{Si}-\text{O}-$. As such, silicone has some resemblance to the amorphous state of SiO_2 (silica or glass). Indeed, the resemblance of silicone to silica is greater than that of epoxy resins.⁷⁹ The atomic structure of silica is shown in *Figure 32.11*. Inspection of the figure allows one to identify the same $-\text{Si}-\text{O}-\text{Si}-\text{O}-$ sequence. Given the excellent thermal and chemical stability and very high transparency of silica (Crivello, 2004), the resemblance suggests that silicone encapsulants are chemically and thermally stable and do not lose transparency as easily as epoxy resins (Crivello, 2004). This indeed is the case. We note that although silica has excellent chemical stability, thermal stability, and high transparency, its rigidity, brittleness, and lack the flexibility make silica (SiO_2) unsuitable for use in conjunction with LEDs (Crivello, 2004).

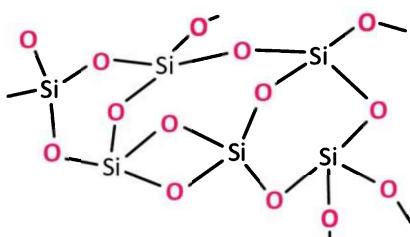


Fig. 32.11: Structure of amorphous SiO_2 (glass). The structure may be thought of as having a $-\text{Si}-\text{O}-\text{Si}-\text{O}-$ backbone with Si atoms connecting to additional O atoms (rather than functional groups as in the case of silicone). The SiO_2 structure is strongly interconnected, hard, and brittle. Silicone on the other hand is less interconnected, softer, and more flexible.

being relatively soft and 100 being relatively hard (see Wikipedia entry on “Shore Durometer”, accessed 2017).

⁷⁹ In the glass state, SiO_2 also has long chains of $\text{Si}-\text{O}-\text{Si}-\text{O}-\text{Si}-\text{O}-\text{Si}-\text{O}$. Recall that the Si atom with its four bonds. In silicone, two of the Si bonds connect with functional groups, whereas in SiO_2 , these two bonds connect with two additional $-\text{Si}-\text{O}-\text{Si}-$ chains. In SiO_x (oxygen-deficient silica) with $x < 2$, double bonds may be formed between some Si and O atoms ($\text{Si}=\text{O}$).

Silicone resins exhibit excellent adhesion. It may be concluded that the adhesive properties of silicone are due to the steric flexibility of the **–Si–O–Si–O–** backbone, which enables good contact with a mating surface, the polarity of the **Si–O** bond, as well as the abundance of oxygen atoms (and the occurrence of terminal **–OH** hydroxyl groups in uncured silicone), thereby allowing for the formation of oxygen bonds between the silicone and the adjoining substrate material (Colas, 1997). As a consequence, excellent adhesion properties can be attained, e.g. between silicone and a lead-frame metal of an LED.

Synthetic silicone resins that also contain an epoxy group are a combination of pure silicones with pure epoxies. Wang *et al.* (2000) and Rubinsztajn (2003) disclosed such silicone epoxy resin. The resin includes a silicone “backbone” with epoxy groups included in the polymer molecule, particularly at its ends. The motivation for this combination is to demonstrate a resin that has good UV resistance and mechanical strength.

32.6 Advanced resin structures

Whereas white resins (light-reflective resins) are used for the lower part of an SMD LED package, clear resins (transparent or translucent resins) are used for encapsulating the LED chip.⁸⁰ The clear encapsulating resin desirably has a high refractive index because a relatively high refractive index enlarges the light-escape cone and thus light extraction. Whereas the refractive index of polymers typically is in the range $\bar{n} = 1.4$ to 1.6 , semiconductors have refractive indices much greater, for example, $\bar{n} = 2.5$ (GaN) or 3.4 (GaAs). Consequently, it is desirable to increase the refractive index of encapsulating polymers (Zhmakin, 2011).

One common strategy in increasing the refractive index of polymers is to increase their mass density. Increasing the mass density will increase the number of polarizable atoms and thus the polarizability of the material, i.e. the refractive index. Increasing the mass density can be accomplished by employing functional groups that contain more atoms. For example, a silicone polymer having phenyl functional groups will generally have a higher mass density and refractive index than a silicone polymer having methyl functional groups. Liu and Ueda (2011) reported that aromatic rings, sulfur-containing groups, halogens (except fluorine), and organometallic groups are often utilized to increase the refractive index of polymers. This strategy allows one to increase

⁸⁰ SMD LEDs (surface-mount device LED) typically consist of a bottom white resin housing having a recess and a transparent resin encapsulation filling the recess.

the refractive index of silicone-based polymers to about $\bar{n} = 1.65$ and in some cases even as high as 1.80 (Liu and Ueda, 2011).

Another strategy is filling the resin with particles having a high refractive index. Encapsulants containing ***nanoparticles*** with a high refractive index (e.g. titania, magnesia, yttria, zirconia, alumina, GaN, AlN, ZnO, ZnSe) have been proposed by Lester *et al.* (1998). The refractive index of the nanoparticle-loaded encapsulant is given by

$$\bar{n} = \frac{\bar{n}_{host} V_{host} + \bar{n}_{nano} V_{nano}}{V_{host} + V_{nano}} \quad (32.1)$$

where V_{host} and V_{nano} refer to the volume of the host and nanoparticles, respectively. For high loading factors, the refractive index of the encapsulant can significantly exceed that of the host, thereby enlarging the semiconductor escape cone and increasing light extraction.

Given the relatively high refractive index of TiO_2 (titania), $\bar{n}_{\text{TiO}_2} = 2.488$ (anatase phase), 2.583 (brookite phase), and 2.609 (rutile phase), titania is a particularly promising filler for enhancing the refractive index (Mont *et al.*, 2008). It is well known that fillers cause optical scattering (Mie and Rayleigh scattering).⁸¹ The relative strength of the scattering decreases as the particle size decreases. In order to minimize optical scattering, TiO_2 nano-particles are employed. Nano-particle-loaded polymer resins with a refractive index of 1.67 (Mont *et al.*, 2008) and even greater than 1.80 have been reported (Liu and Ueda, 2009).

The maximum volume fill factor of a filler consisting of equally-sized spherical particles that are arranged in the form of a simple cubic lattice with lattice constant a is $\frac{4}{3}\pi (\frac{1}{2}a)^3/a^3 = 52.4\%$. If the equally-sized spherical particles are arranged in a densely-packed hexagonal lattice, the maximum volume fill factor is given by $\pi/(3 \times 2^{1/2}) = 74\%$.⁸² If there are additional smaller spheres that are packed into the gaps between the larger spheres, the maximum fill factor can exceed 74%.

Graded-index encapsulants consisting of several layers with different refractive indices were demonstrated by Lee *et al.* (2004). The layer of highest refractive index is in contact with the

⁸¹ Optical scattering needs to be kept low so that light can traverse the encapsulating resin without being excessively scattered. One strategy to reduce optical scattering is to use ***nano***-particles. Another possible strategy to reduce scattering is to reduce the randomness in the distribution of the nano-particles, e.g. by arranging them into an ordered pattern.

⁸² The maximum volume fill factor for the dense arrangement of equally-sized spheres was derived by Carl Friedrich Gauss. Detailed information can be found in Wikipedia under the entry “Close-packing of equal spheres” (2017).

semiconductor chip. The outer layers of the encapsulant have lower refractive indices. Such refractive-index graded encapsulants should have extraction efficiencies exceeding those of non-graded refractive index encapsulants.

32.7 Degradation of resins under harsh operating conditions

Reflective resins and transmissive resins (i.e. white resins and clear resins) can be subject to harsh operating conditions, in particular to high temperatures, high radiation intensities, and moisture or humidity (H_2O).

With respect to temperature, LED chips and phosphors convert some of the electrical input power to heat so that the temperature of chip and phosphor rises to potentially exceed 100°C or even 200°C. Epoxies are generally stable up to temperatures of about 120°C (Denk and Reeh, 1979). Silicones are generally stable up to temperatures of about 170~190°C (Crivello, 2004; Bahadur *et al.*, 2006; Lin *et al.*, 2010). Exceeding these temperatures will result in a loss of optical transparency which manifests itself by yellowing or browning of the resin.

With respect to high radiation intensities, it is worthwhile considering the optical intensities at the surface of an LED chip: A $1 \times 1 \text{ mm}^2$ GaInN / GaN surface-emitting LED chip with an input power of about 1 W (350 mA at 3.0 V) having a power-conversion efficiency of 50% emits an optical power of 0.5 W and thus the radiation intensity at the chip surface is $0.5 \text{ W} / \text{mm}^2 = 500 \text{ kW/m}^2$. This is a very high radiation intensity. Indeed, the intensity is 500 times greater than that of solar radiation (1 kW/m^2).⁸³ It is reasonable to expect that degradation will occur under such high radiation intensities. This is particularly true for blue, violet, and UV LEDs due to the relatively high energy of short-wavelength photons.⁸⁴ Accordingly, the degradation of resins under high radiation intensities is a concern. Indeed, Barton *et al.* (1998) reported the yellowing of epoxy resins when used for encapsulation in blue LEDs. Furthermore, a study by Lin *et al.* (2010) showed that epoxy resins degrade more rapidly than silicone resins when exposed to UV radiation.

The combination of high temperature with high radiation intensities constitutes harsh operating conditions that are prone to degrade resins. If, in addition, moisture and humidity

⁸³ Solar radiation (sunlight) on Earth has an intensity of about 1 kW/m^2 .

⁸⁴ The violet wavelength of 405 nm carries a photon energy of about 3 eV and thus is sufficiently high to break a chemical bond.

(generally H₂O) reach the LED chip or phosphor, significant degradation is to be expected. Therefore, it is desirable that encapsulating resins hermetically seal the LED chips and phosphors.⁸⁵ It is generally found that epoxy is a better barrier against moisture than silicone (Crivello, 2004).⁸⁶

The degradation of resins under harsh operating conditions can be reduced by using chemically stable inorganic fillers. Such inorganic fillers include, for example, Al₂O₃ and SiO₂. This strategy works particularly well for white resins, for which optical scattering is desired.

32.8 References

- Bahadur M., Norris A. W., Zarifsi A., Alger J. S., and Windiate C. C. "Silicone materials for LED packaging" *Proceedings of SPIE* **6337**, 63370F-1 (2006)
- Barton D. L., Osinski M., Perlin P., Helms C. J., and Berg N. H. "Life tests and failure mechanisms of GaN / AlGaN / InGaN light-emitting diodes" *Proceedings of the SPIE* **3279**, 17 (1998)
- Carmody W. H. and Kelly H. E. "Method of making phenolated resins" US Patent 2,319,386, issued May 18 (1943)
- Colas A. "Silicone chemistry overview" Publication of the Dow Corning Company of Midland, Michigan, USA <https://www.dowcorning.com/content/publishedlit/51-960A_silicone-chemistry-overview.pdf> (1997)
- Crivello J. V., Rensselaer Polytechnic Institute (RPI), personal communication (2004)
- Denk H. and Reeh U. "Casting resin compositions for sealing opto-electronics components" US Patent 4,178,274 issued December 11 (1979)
- Flick E. W. *Epoxy Resins, Curing Agents, Compounds, and Modifiers: An Industrial Guide* (Noyes Data Corporation/Noyes Publications, Park Ridge NJ, 1993)
- Gorczyk J., Bogdal D., Pielichowski J., and Penczek P. "Synthesis of high molecular weight epoxy resins under microwave irradiation" *Fifth International Electronic Conference on Synthetic Organic Chemistry* (ECSOC-5), <http://www.mdpi.org/ecsoc-5.htm>, 1–30 (September 2001)
- Kim J.-S., Yang S. C., and Bae B.-S. "Thermally stable transparent sol-gel based siloxane hybrid material with high refractive index for light emitting diode (LED) encapsulation" *Chemistry of Materials* **22**, 3549 (2010)
- Kumar R. N., Keem L. Y., Mang N. C., and Abubakar A. "Ultraviolet radiation curable epoxy resin encapsulant for light-emitting diodes" *Journal of Applied Polymer Science* **100**, 1048 (2006)
- Kuramoto M. and Kishimoto T. "Resin molding, surface mounted light emitting device, and manufacturing method therefor" Japanese Unexamined Patent Application Publication Number JP 2006-156704, published June 15 (2006)
- Lester S. D., Miller J. N., Roitman D. B. "High refractive index package material and a light emitting device encapsulated with such material" US Patent 5,777,433 issued July 7 (1998)
- Lin Y.-H., You J. P., Lin Y.-C., Tran N. T., and Shi F. G. "Development of high-performance optical silicone for the packaging of high-power LEDs" *IEEE Transactions on Components and Packaging Technologies* **33**, 761 (2010)
- Liu J.-G. and Ueda M. "High refractive index polymers: Fundamental research and practical applications" *Journal of Material Chemistry* **19**, 8907 (2009)
- Luo H., Kim J. K., Xi Y. A., Schubert E. F., Cho J., Sone C., and Park Y. "Trapped whispering-gallery optical

⁸⁵ The permeability of resins is characterized by "Water Vapor Transmission Rates" frequently abbreviated as WVTR.

⁸⁶ However, the permeability of silicone for water vapor can be increased by choosing a highly cross-linked silicone.

- modes in white light-emitting diode lamps with remote phosphor" *Applied Physics Letters* **89**, 041125 (2006)
- Luo H. "Highly efficient package configurations for white-light-emitting diode lamps with novel remote phosphor distribution and electrostatic discharge protection" *Dissertation*, Rensselaer Polytechnic Institute, Troy NY (2006)
- Mont F. W., Kim J. K., Schubert M. F., Schubert E. F., Siegel R. W. "High-refractive-index TiO₂-nanoparticle-loaded encapsulants for light-emitting diodes" *Journal of Applied Physics* **103**, 083120 (2008)
- Norris A. W., Bahadur M., Yoshitake M. "Novel silicone materials for LED packaging" *Proceedings of SPIE* **5941**, 594115-1 (2005)
- Reeh U., Höhn K., Stath N., Waitl G., Schlotter P., Schneider J., and Schmidt R. "Light-radiating semiconductor component with luminescence conversion element" US Patent 6,576,930 B2, issued June 10 (2003)
- Rubinsztajn M. I. and Rubinsztajn S. "Composition comprising silicone epoxy resin, hydroxyl compound, and anhydride and curing catalyst" US Patent 6,632,892 B2, issued October 14 (2003)
- Shiang J. J. and Duggal A. R. "Application of radiative transport theory to light extraction from organic light emitting diodes" *Journal of Applied Physics* **95**, 2880 (2004)
- Shiang J. J., Faircloth T. J., and Duggal A. R. "Experimental demonstration of increased organic light emitting device output via volumetric light scattering" *Journal of Applied Physics* **95**, 2889 (2004)
- Wang W. J., Perng L. H., Hsiue G. H., and Chang F. C. "Characterization and properties of new silicone-containing epoxy resin" *Polymer* **41**, 6113 (2000)
- Zhmakin A. I. "Enhancement of light extraction from light emitting diodes" *Physics Reports* **498**, 189 (2011)

33

33 – Packaging: Protection against electrostatic discharge (ESD)

33.1 The human body model and the IEC 61000-4-2 standard

Electrostatic discharge (ESD) is a common catastrophic failure mechanism for electronic and optoelectronic components caused by the sudden discharge of an electrical charge through the component (Voldman, 2004). Frequently, the electrostatic discharge is caused by a human being touching the component. ESD can be understood in terms of the **human body model** (HBM), illustrated in *Figure 33.1* (Voldman, 2004). The human body model represents a typical human body by means of a capacitance $C = 100 \text{ pF}$ and a series resistance $R_{\text{human}} = R_{\text{series}} = 1500 \Omega$. When walking across a carpet or a plastic floor, a person may charge up the capacitor. The associated voltages can be as high as 10 kV. If the person touches an electronic component, the capacitor discharges instantaneously through the electronic component and may permanently damage the component. It is desirable that electronic components withstand an ESD event unharmed. Accordingly, in ESD tests that simulate ESD events, a capacitor is charged to certain ESD test voltages, typically a few hundred to a few thousand volts (e.g. 500 V to 15 kV). Subsequently, the charge on the capacitor is discharged through the **device under test** (DUT) as illustrated in *Figure 33.1*.

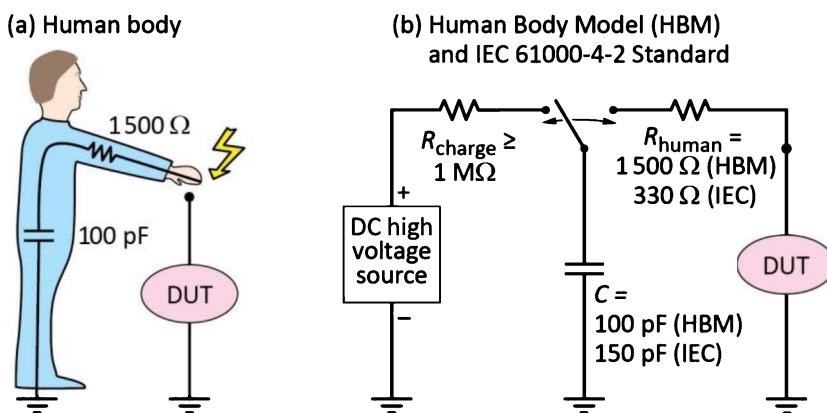


Fig. 33.1: (a) Human body causing an electrostatic discharge (ESD) on a “device under test” (DUT). (b) Equivalent circuit used for ESD testing based on the Human Body Model and the IEC 61000-4-2 standard.

There are multiple standards for testing the stability of an electronic component against ESD. One common model is the Human Body Model (HBM); another common standard is the IEC 61000-4-2 Standard (IEC, 2017).⁸⁷ Both standards, HBM and IEC, are similar in that both use a capacitor and a series resistor. However, the numerical values of the capacitor and resistor are different:

Human Body Model:

Discharge capacitor = 100 pF; Discharge resistor = 1 500 Ω

International Electrotechnical Commission Standard IEC 61000-4-2:

Discharge capacitor = 150 pF; Discharge resistor = 330 Ω

Standardized discharge voltages include 2 kV, 4 kV, 6 kV and 8 kV. As will become evident from the subsequent discussion that the IEC test conditions (due to its higher capacitance and lower series resistance) are more challenging than the HBM test conditions.

33.2 The electrostatic discharge process

Consider a capacitor that carries a charge $+Q$ and thus has a voltage of $V = Q/C$. The energy stored in the capacitor is given by $E = \frac{1}{2}CV^2$. During the discharging event, an upper limit for the energy to be dissipated in the device under test (DUT) is $E = \frac{1}{2}CV^2$. If the discharge event occurs during the time τ , then the power dissipated in the DUT is given by E/τ .

We first assume that the DUT is a resistor with resistance R_{DUT} as shown in **Figure 33.2 (b)**. Assuming IEC testing conditions ($C = 150$ pF and $R_{series} = 330$ Ω) and if $R_{DUT} \gg R_{series}$, then most of the energy will be dissipated in R_{DUT} , i.e. the DUT. On the other hand, if $R_{DUT} \ll R_{series}$, then most of the energy will be dissipated in R_{series} . Accordingly, a low-resistance DUT will be beneficial for minimizing the impact of the ESD event. As an example, let us consider $R_{DUT} = 1$ kΩ so that $R_{DUT} \gg R_{series}$. An ESD event with a voltage of 1 kV has a duration of the RC time constant, i.e. $\tau = RC = 1330\ \Omega \times 150\text{ pF} = 200\text{ ns}$, so that the instantaneous power during the ESD event is $P_{DUT} \approx E/\tau = \frac{1}{2}C V^2 / \tau \approx 375\text{ W (!)}$. This is a respectable power for a small electronic component, so that the component could get damaged. This example illustrates that reducing the DUT resistance will be critical in reducing the power dissipated in the DUT.

⁸⁷ IEC = International Electrotechnical Commission based in Geneva, Switzerland

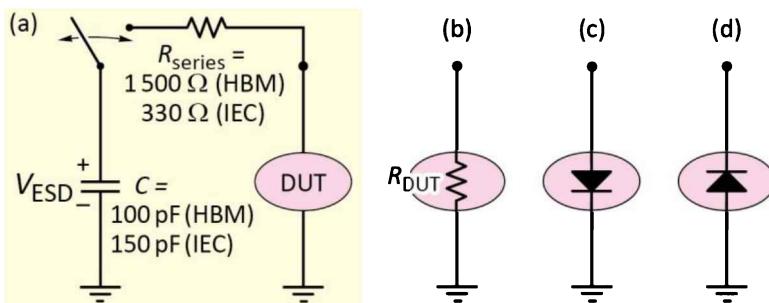


Fig. 33.2: (a) Discharge of capacitor through a device under test (DUT). The DUT can be, for example, a (b) resistor, (c) forward-biased diode, or (d) reverse-biased diode.

Let us next consider that the DUT is a diode and that the discharge current flows through the diode in the *forward* direction as shown in *Figure 33.2 (c)*. Taking the equivalent circuit of the forward-biased diode as a voltage source with voltage V_{th} with a small series resistor with resistance $R_{DUT} \ll R_{series}$, then most of the energy carried by the capacitor will be dissipated in R_{series} (i.e. the human body), i.e. *not* in the diode. Accordingly, damage to the diode is unlikely.

Let us next consider that the DUT is a diode and that the discharge current flows through the diode in the *reverse* direction as shown in *Figure 33.2 (d)*. Taking the equivalent circuit of the forward-biased diode as a resistor R_{DUT} in the $k\Omega$ range, then $R_{DUT} \gg R_{series}$, so that most of the energy carried by the capacitor will be dissipated in R_{DUT} (i.e. the diode). Accordingly, damage to the diode is more likely.

Thus, it is evident that the energy dissipated in the diode during a *reverse* discharge event is much greater than the energy dissipated in the diode during a *forward* discharge event, suggesting that *reverse* discharges are more damaging than *forward* discharges. This has indeed been confirmed by experiments (Wen *et al.*, 2004).

Accordingly, an ESD-protection circuit for an LED will function to (i) not influence the LED under normal operating conditions and to (ii) offer a low-resistance by-pass if a transient ESD voltage pulse occurs at the LED. The current caused by an electrostatic discharge event is intended to bypass the LED and flow through the ESD-protection circuit. The protection is particularly important when the ESD current pulse would flow through the LED in the reverse direction. As will be seen in the subsequent section, this functionality of an ESD-protection device is well fulfilled by a Si Zener diode.

33.3 ESD-protection circuits

A variety of ESD-protection circuits are shown in *Figure 33.3*. The simplest protection circuit, shown in *Figure 33.3 (a)*, has one Zener diode that limits the reverse voltage at the LED to the

Zener voltage. The circuit has the disadvantage of current flow through the Zener diode, if the voltage at the LED is accidentally reversed. This disadvantage is avoided for the circuit shown in **Figure 33.3 (b)** where the two Zener diodes are used. A low-capacitance configuration of the ESD-protection circuit is shown in **Figure 33.3 (c)**.⁸⁸ A bi-directional low-capacitance ESD-protection circuit is shown in **Figure 33.3 (d)**.

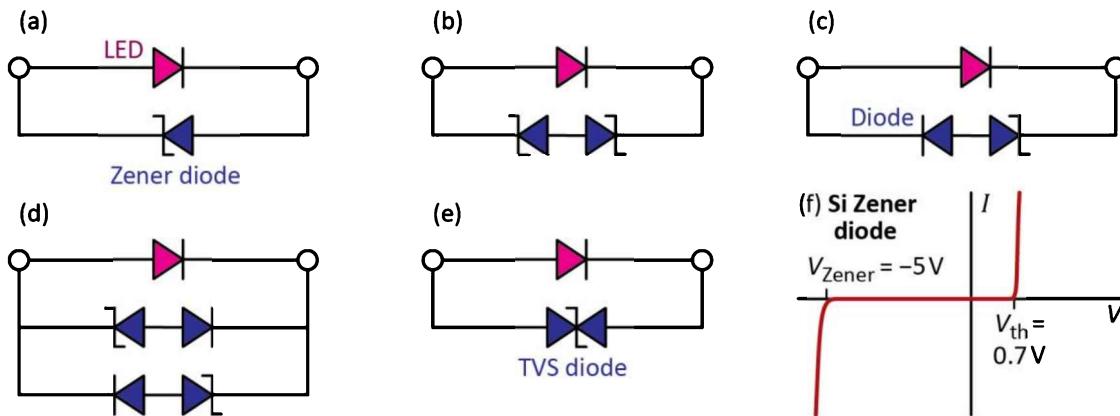


Fig. 33.3: (a)–(e) Electrostatic discharge (ESD) protection circuits using Si Zener diodes, Si pn junction diodes, and Si transient voltage suppression (TVS) diodes. (f) Current–voltage characteristic of a Si Zener diode with a Zener voltage of $V_{Zener} = -5\text{ V}$.

Common Zener diodes are intended for voltage stabilization in power supplies. Zener diodes that are optimized for ESD applications are referred to as **transient voltage suppression diodes** (TVS diodes). Although the TVS diode come in multiple varieties, its operating principle is based on the Zener diode. Unlike a Zener diode, TVS diodes can have a symmetric current–voltage characteristic. That is, a TVS diode can become highly conductive for $|V| > V_{Zener}$ irrespective of the polarity of the TVS diode. Symmetric current–voltage characteristics are attained by using a bi-directional TVS diode, i.e. two opposing Zener diodes. An ESD-protection circuit using a bi-directional TVS diode is shown in **Figure 33.3 (e)**.⁸⁹

Wide-bandgap diodes (such as GaN-based diodes) are particularly prone to ESD failures, due to inherently high values of R_{DUT} under reverse-bias conditions (wide-bandgap diodes have low

⁸⁸ The ESD-protection circuit has a low capacitance since the capacitances of the ESD-protection diodes are in a series circuit.

⁸⁹ While Zener diodes and TVS diodes are very common ESD-protection devices in the field of LEDs, ceramic high-voltage capacitors are another type of ESD-protection device. Capacitors naturally resist abrupt voltage changes since the energy stored in a capacitor ($\frac{1}{2}CV^2$) cannot “jump” (change discontinuously). However, the larger size of ceramic high-voltage capacitors is undesirable in the context of LEDs.

reverse saturation currents and high breakdown voltages). This has motivated the development of ESD-protection circuits particularly for III–V nitride diodes (Steigerwald *et al.*, 2002; Sheu, 2003).

Electrostatic discharge-protection circuits (such as a Zener or TVS diode) can be added to the LED package. An example of a ESD-protection device that is included in an LED package is shown in **Figure 33.4** (Philips Lumileds Company, 2012). The EDS-protection device has a dark grey color indicating a black resin package or a bare Si diode chip.

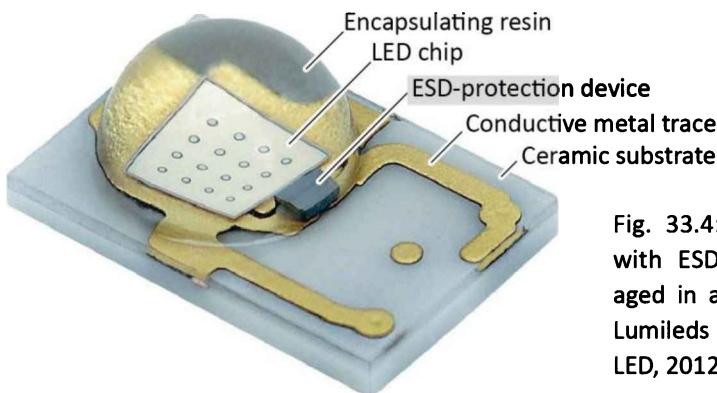


Fig. 33.4: GaInN / GaN blue LED with ESD-protection device packaged in a ceramic package (Philips Lumileds Company "Luxeon Rebel" LED, 2012).

Furthermore, ESD-protection devices have been integrated in Si submounts for LED chips, as shown in **Figure 33.5**. That is, an ESD-protection device is included in the Si submount that carries the LED chip (Steigerwald *et al.*, 2002; Philips Lumileds, 2004). Using the Zener diodes increases the threshold voltage of the ESD circuit to values beyond the turn-on voltage of the LED. Thus, under normal operating conditions, the current through the ESD-protection circuit is negligibly small.

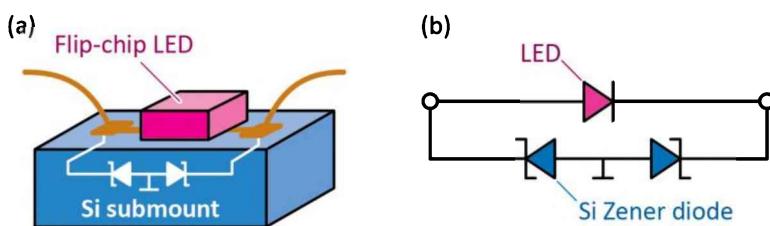


Fig. 33.5: (a) Electrostatic discharge (ESD) protection integrated into a Si submount. (b) ESD protection circuit (after Lumileds Company, 2004).

33.4 ESD-protection circuits integrated with the LED

Monolithic integration of the ESD-protection device with the LED chip is desirable for miniaturization and cost reasons. Sheu (2003) proposed a Schottky diode integrated with the LED chip. The structure, shown in **Figure 33.6** (a), consists of a large-area p-n junction diode and,

separated by a deep trench, a small-area Schottky diode. The Schottky diode, fabricated on the n-type buffer layer of a GaInN LED, is forward biased when the LED is biased in the reverse polarity. For reverse electrostatic discharges, the current flows mostly through the Schottky diode, thereby bypassing the p-n junction and preventing damage to the p-n junction. For forward electrostatic discharges, the current flows through the p-n junction. In an alternative structure, shown in *Figure 33.6 (b)*, the Schottky diode is replaced by a p-n junction diode (Cho, 2005). Furthermore, it has been reported that the ability of LEDs to withstand ESD can be enhanced by including a superlattice structure in the LED epitaxial layer stack (Wen *et al.*, 2004).

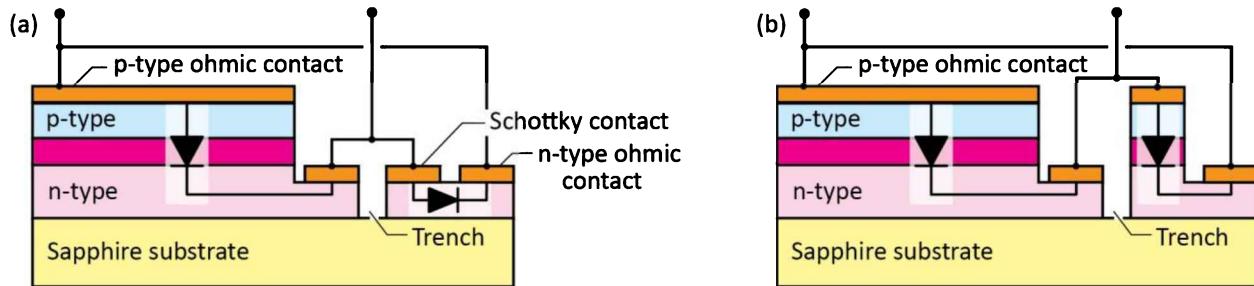


Fig. 33.6: On-chip ESD protection using (a) a small-area Schottky diode on the n-type buffer layer of a GaInN LED (Sheu, 2003) and (b) a small-area p-n junction diode.

33.5 References

- Cho J., Samsung Advanced Institute of Technology, Suwon, Korea, personal communication (2005)
- IEC 61000-4-2 Standard, International Electrotechnical Commission, Geneva, Switzerland; Internet site: <<http://www.iec.ch/>> (accessed 2017)
- Lumileds Company “Luxeon reliability” Application Brief **AB25**, 11 (2004)
- Philips Lumileds Company “LUXEON Rebel and LUXEON Rebel ES Product Datasheet DS68” San Jose, California, USA (2012)
- Sheu J.-K. “Group III–V element-based LED having ESD protection capacity” US Patent 6,593,597 B2 (2003)
- Steigerwald D. A., Bhat J. C., Collins D., Fletcher R. M., Holcomb M. O., Ludowise M. J., Martin P. S., and Rudaz S. L. “Illumination with solid state lighting technology” *IEEE Journal of Selected Topics in Quantum Electronics* **8**, 310 (2002)
- Voldman S. H. *ESD: Physics and Devices* (John Wiley and Sons, New York, 2004)
- Wen T. C., Chang S. J., Lee C. T., Lai W. C., and Sheu J. K. “Nitride-based LEDs with modulation-doped AlGaN-GaN superlattice structures” *IEEE Transactions on Electron Devices* **51**, 1743 (2004)

34

34 – Packaging: Junction and carrier temperatures

The temperature of the active region crystal lattice, frequently referred to as the *junction temperature*, is a critical parameter. The junction temperature is relevant for several reasons. Firstly, the internal quantum efficiency depends on the junction temperature. Secondly, high-temperature operation shortens the device lifetime. Thirdly, a high device temperature can lead to degradation of the encapsulant. It is therefore desirable to know the junction temperature as a function of the drive current.

Heat can be generated in the contacts, cladding layers, and the active region. At low current levels, heat generation in the parasitic resistances of contacts and cladding layers is small due to the I^2R dependence of Joule heating. The dominant heat source at low current levels is the active region, where heat is created by non-radiative recombination. At high current levels, the contribution of parasitics becomes increasingly important and can even dominate.

There are several different ways to measure the junction temperature, which include micro-Raman spectroscopy (Todoroki *et al.*, 1985), threshold voltage (Abdelkader *et al.*, 1992), thermal resistance (Murata and Nakada, 1992), photothermal reflectance microscopy (Epperlein, 1990), electroluminescence (Epperlein and Bona, 1993), photoluminescence (Hall *et al.*, 1992) and a non-contact method based on the peak ratio of a dichromatic source (Gu and Narendran, 2003). Most methods are *indirect* methods that infer the junction temperature from an easily measurable parameter. In this chapter, we discuss a method based on the shift of the peak emission wavelength with the temperature and a method based on the shift of the diode forward voltage with temperature. We also discuss the carrier temperature as inferred from the high-energy slope of the emission spectrum.

34.1 Carrier temperature and high-energy slope of spectrum

The Boltzmann distribution of carriers, applicable to the high-energy part of the emission spectrum, results in an exponential dependence of the emission intensity on energy, i.e.

$$I \propto \exp [-h\nu / (kT_c)] \quad (34.1)$$

where T_c is the **carrier temperature**. The high-energy slope of the spectrum is given by

$$\frac{d(\ln I)}{d(h\nu)} \propto -\frac{1}{kT_c}. \quad (34.2)$$

Thus, the carrier temperature can be directly inferred from the slope. Because the carrier temperature is generally higher than the junction temperature, e.g. due to high-energy injection of carriers into the active region, this method gives an *upper limit* for the actual junction temperature.

Figure 34.1 shows the evaluation of the carrier temperature from the emission spectrum of a GaInN and an AlGalnP LED (Chhajed *et al.*, 2005; Gessmann *et al.*, 2003). Inspection of the figure reveals that the carrier temperature increases along with the current level. At low current levels, the GaInN device has a carrier temperature of 221°C and the AlGalnP device has a carrier temperature of 212°C. At high current levels, the carrier temperature increases to 415°C and 235°C for the GaInN and AlGalnP LED, respectively. Due to the alloy-broadening effect occurring in ternary and quaternary semiconductor alloys, these temperatures overestimate the true carrier temperature.

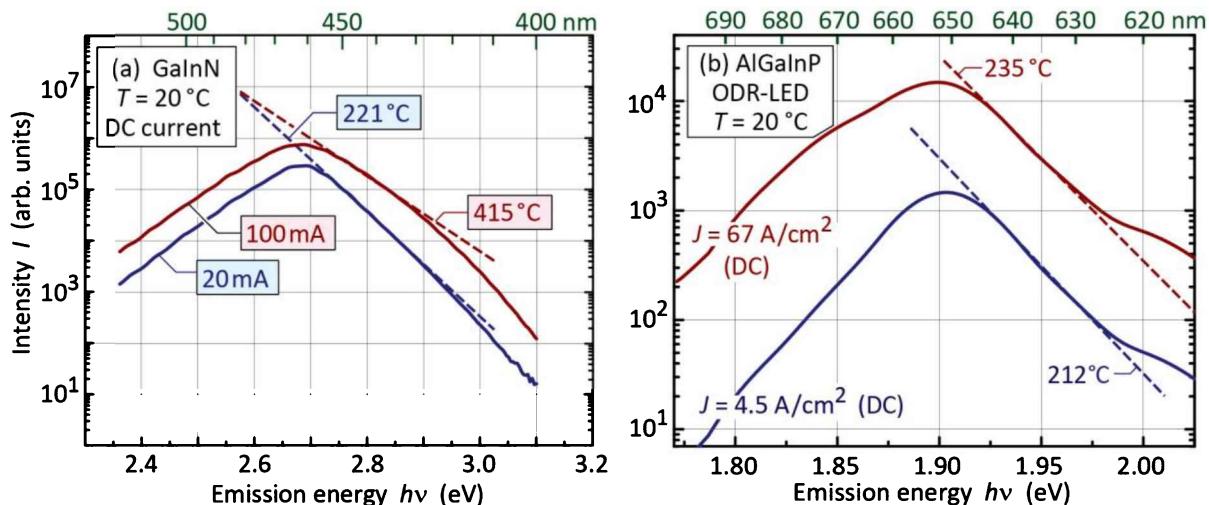


Fig. 34.1: Carrier temperatures in (a) GaInN blue and (b) AlGalnP red LEDs inferred from the high-energy slope of emission spectrum. Due to the alloy-broadening effect, the measured carrier temperatures overestimate the true carrier temperature (after Chhajed *et al.*, 2005; Gessmann *et al.*, 2003).

Semiconductor alloys exhibit substantial broadening of the emission spectrum (and its high-energy slope) due to alloy broadening, i.e. the statistical fluctuation of the chemical composition occurring in ternary and quaternary semiconductors (Schubert *et al.*, 1984). De-convolution of the alloy-broadening effect and the kT -broadening effect allows for a more accurate estimate of the carrier temperature.

The determination of the carrier temperature using the high-energy slope works best for binary compounds such as GaAs or InP. Such semiconductors do not exhibit alloy broadening and thus the high-energy slope is more representative of the true carrier temperature.

34.2 Junction temperature and peak emission wavelength

This method makes use of the dependence of the bandgap energy (and thus the peak emission wavelength) on temperature. The method consists of a calibration measurement and a junction-temperature measurement. In the calibration measurement, the peak energy is measured at different ambient temperatures, typically in the range 20°C to 120°C, by placing the device in a temperature-controlled oven. The device is injected with a range of pulsed currents with a duty cycle $\ll 1$ to minimize additional heating. As a consequence, the ambient temperature in the oven and the junction temperature can be assumed to be identical. The calibration measurement establishes the junction-temperature versus emission-peak-energy relation for a range of currents. Calibration data for a deep UV LED are shown in **Figure 34.2 (a)** (Xi *et al.*, 2004; 2005).

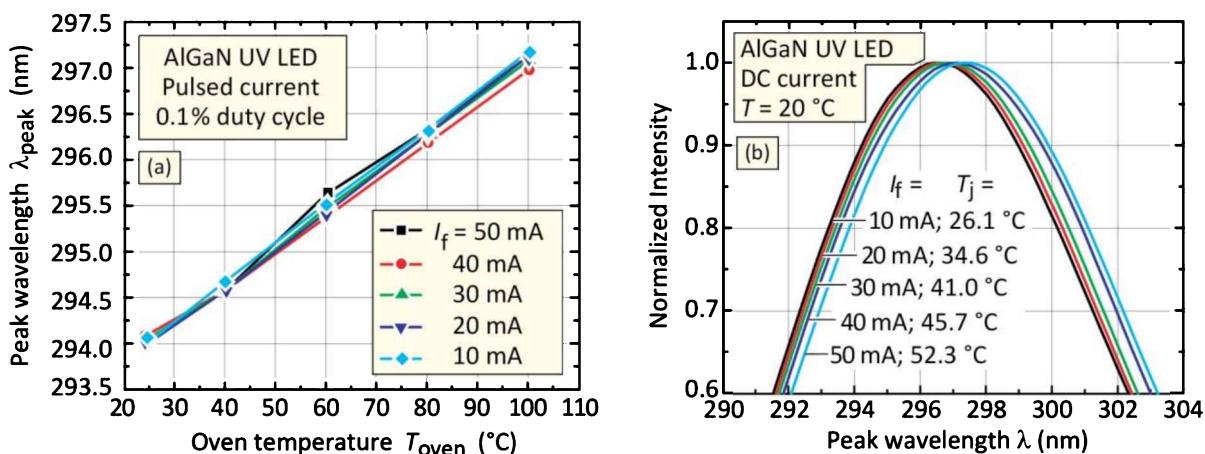


Fig. 34.2: (a) Peak emission wavelength versus oven temperature of an AlGaN UV LED for pulsed current injection with 0.1% duty cycle. (b) Emission spectra and junction temperatures for different DC currents with device at 20 °C (after Xi *et al.*, 2005).

Subsequent to the calibration, the peak emission energy is measured as a function of the DC injection current with the device in a room-temperature ambient. The junction temperature for each current level can then be determined by using the calibration data. *Figure 34.2 (b)* shows the emission spectra of the UV LED for different injection currents. The junction temperatures inferred from the calibration measurement are shown in *Figure 34.3* (Xi *et al.*, 2005).

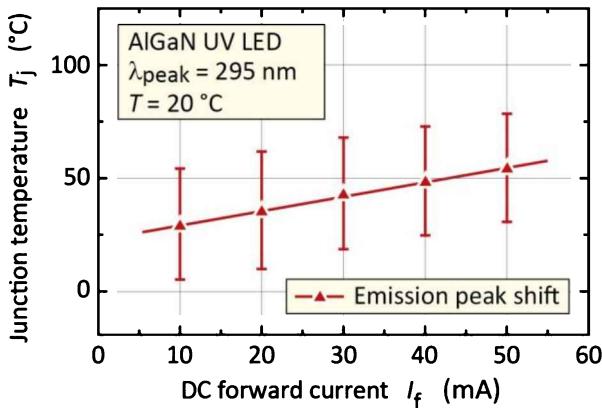


Fig. 34.3: Junction temperature inferred from emission peak energy as a function of DC injection current for a 300 $\mu\text{m} \times 300 \mu\text{m}$ deep UV LED emitting at 295 nm. The error bar stems from an uncertainty in the peak energy (after Xi *et al.*, 2005).

The accuracy of the method is limited by the ability to determine the peak wavelength. As a rule of thumb, the error bar of the peak wavelength is about 5~10% of the full-width at half-maximum of the luminescence line. Alloy-broadening effects and kT broadening impose a limitation on the accuracy of the method.

The shift of the emission energy with respect to temperature (dE_g/dT) is due to the temperature dependence of the energy gap. The change of the energy gap can be calculated from the Varshni parameters to be discussed in the subsequent section of this chapter.

Note that **band-filling effects** should not influence the results as they also enter the calibration measurement. The peak emission energy shifts to *higher* energies due to band filling occurring at high injection current densities. In contrast, the peak emission energy shifts to *lower* energies due to bandgap shrinkage. Although it is difficult to clearly separate the junction-temperature-induced shift from the band-filling-induced shift, the former effect dominates under typical experimental conditions.

34.3 Theory of temperature dependence of diode forward voltage

The derivation of the temperature dependence of the forward voltage presented here follows the analysis first given by Xi *et al.* (2004; 2005). The I – V characteristic of an ideal p-n junction

diode is given by the Shockley equation

$$J = J_s (e^{eV_f / (n_{ideal} kT)} - 1) \quad (34.3)$$

where J_s is the saturation current density. For non-degenerate semiconductors and under forward bias conditions, $V_f \gg kT/e$, one obtains

$$\frac{dV_f}{dT} = \frac{d}{dT} \left(\frac{n_{ideal} kT}{e} \ln \frac{J_f}{J_s} \right). \quad (34.4)$$

The saturation current density depends on the diffusion constants of electrons and holes, the lifetimes of electrons and holes, the effective density of states at the conduction band and valence band edges, and the bandgap energy, all of which depend on the junction temperature. The temperature dependence of the effective density of states is given by $N_{c,v} \propto T^{3/2}$. Assuming phonon scattering, the temperature dependence of the carrier mobility is $\mu \propto T^{-3/2}$. Using the Einstein relation, the diffusion constant depends on temperature according to $D \propto T^{-1/2}$. The minority carrier lifetime can either decrease (non-radiative recombination) or increase (radiative recombination) with temperature. Due to this uncertainty, the minority carrier lifetime is assumed to be independent of temperature. Using these temperature dependences in Eqn. (34.4) and executing the derivative yields

$$\frac{dV_f}{dT} = \frac{eV_f - E_g}{eT} + \frac{1}{e} \frac{dE_g}{dT} - \frac{3k}{e}. \quad (34.5)$$

This equation gives the fundamental temperature dependence of the forward voltage. The first, second, and third summands on the right-hand side of the equation are due to the temperature dependence of the intrinsic carrier concentration, bandgap energy, and effective densities of states, respectively. The equation includes the temperature dependence of the bandgap energy, which had not been taken into account in earlier derivations of dV_f/dT (Millman and Halkias, 1972).

LEDs are typically operated at forward voltages close to the built-in voltage, i.e. $V_f \approx V_{bi}$. Thus, for non-degenerate doping concentrations, we can write

$$eV_f - E_g \approx kT \ln \left(\frac{N_D N_A}{n_i^2} \right) - kT \ln \left(\frac{N_c N_v}{n_i^2} \right) = kT \ln \left(\frac{N_D N_A}{N_c N_v} \right). \quad (34.6)$$

The second summand on the right-hand side of Eqn. (34.5) is due to the changes in bandgap energy. As the temperature increases, the energy gap of semiconductors generally decreases. The temperature dependence of the energy gap of a semiconductor can be expressed by the **Varshni formula** (Varshni, 1967)

$$E_g = E_g|_{T=0\text{ K}} - \frac{\alpha T^2}{T + \beta} \quad (34.7)$$

where α and β are fitting parameters, frequently called the **Varshni parameters**. The bandgap energy versus temperature for several semiconductors is shown in **Figure 34.4** along with the values of the Varshni parameters (α and β).

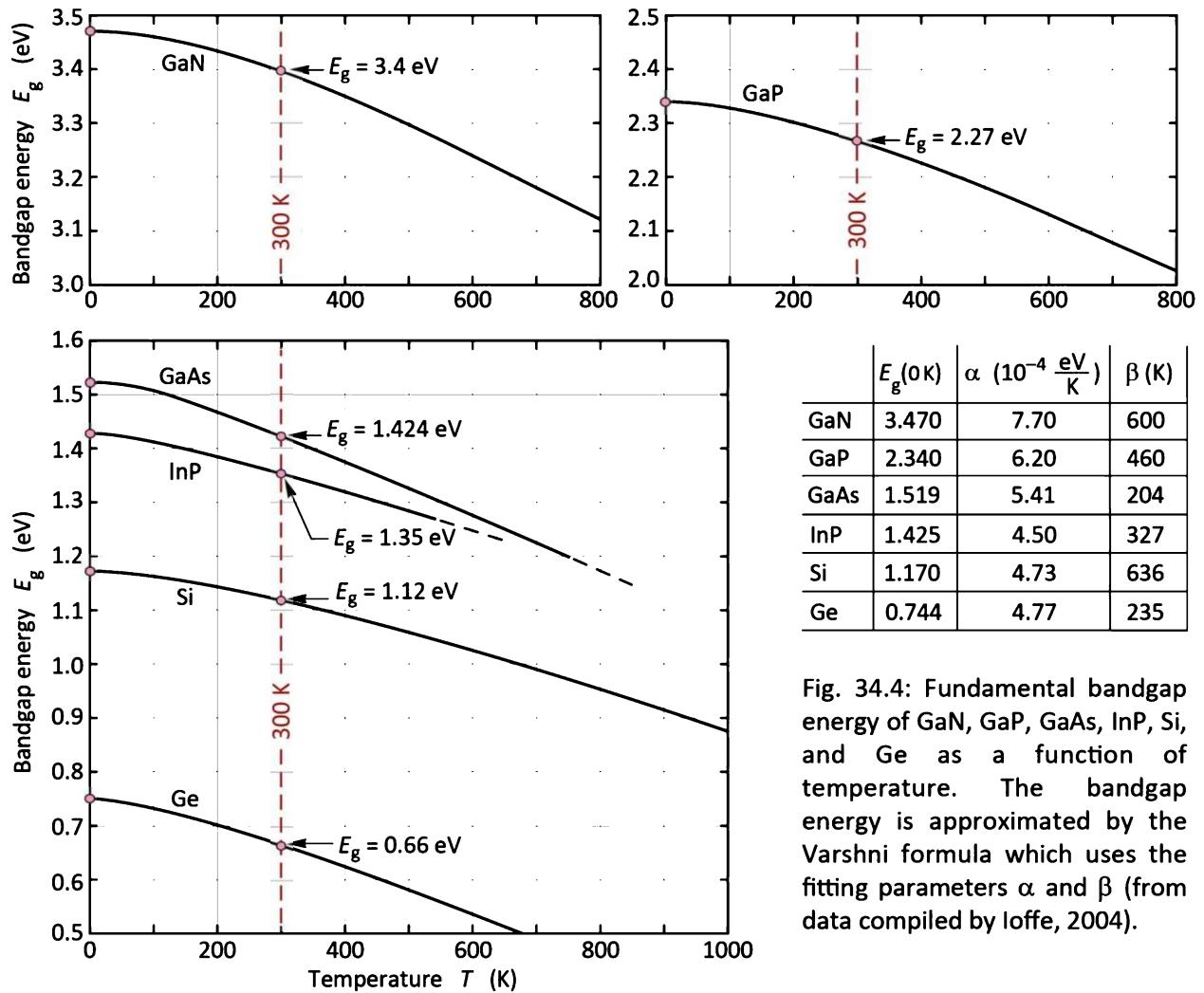


Fig. 34.4: Fundamental bandgap energy of GaN, GaP, GaAs, InP, Si, and Ge as a function of temperature. The bandgap energy is approximated by the Varshni formula which uses the fitting parameters α and β (from data compiled by Ioffe, 2004).

The Varshni parameters for several semiconductors, compiled by the Ioffe (2004), are given in **Table 34.1**. Substituting Eqns. (34.6) and (34.7) into Eqn. (6.5) yields

$$\frac{dV_f}{dT} \approx \underbrace{\frac{k}{e} \ln\left(\frac{N_D N_A}{N_c N_v}\right)}_{\text{due to } T \text{ dependence of } n_i} - \underbrace{\frac{\alpha T (T + 2\beta)}{e (T + \beta)^2}}_{\frac{1}{e} \frac{dE_g}{dT}} - \underbrace{\frac{3k}{e}}_{\text{due to } T \text{ dependence of DOS}} . \quad (34.8)$$

This equation is a very useful expression for the temperature coefficient of the forward voltage.

Table 34.1: Varshni parameters of common semiconductors (from data compiled by Ioffe, 2004).

Semiconductor	E_g at 0 K (eV)	α (10^{-4} eV/K)	β (K)	Validity range
AlN (wurtzite)	6.026	18.0	1462	$T \leq 300$ K
GaN (wurtzite)	3.47	7.7	600	$T \leq 600$ K
GaP	2.34	6.2	460	$T \leq 1\,200$ K
GaAs	1.519	5.41	204	$T \leq 1\,000$ K
GaSb	0.813	3.78	94	$T \leq 300$ K
InN (wurtzite)	1.994	2.45	624	$T \leq 300$ K
InP	1.425	4.50	327	$T \leq 800$ K
InAs	0.415	2.76	83	$T \leq 300$ K
InSb	0.24	6.0	500	$T \leq 300$ K
Si	1.170	4.73	636	$T \leq 1\,000$ K
Ge	0.744	4.77	235	$T \leq 700$ K

For GaN diodes, Xi *et al.* (2004; 2005) reported a calculated dV_f/dT of -1.76 mV/K, which is in good agreement with the experimental value of -2.3 mV/K. Deviations between theory and experiment were attributed to the temperature coefficient of the resistivity of the neutral regions, which decreases with increasing temperature due to a higher doping activation (Xi *et al.*, 2005).

The temperature dependence of a GaPAs/GaAs LED is illustrated in **Figure 34.5**, which shows the I – V characteristic at 77 K and at room temperature. Inspection of the figure reveals that the threshold voltage as well as the series resistance of the diode increases as the diode is cooled. If the device were driven at a constant voltage, e.g. 1.9 V, a large current change would result from the change in temperature.

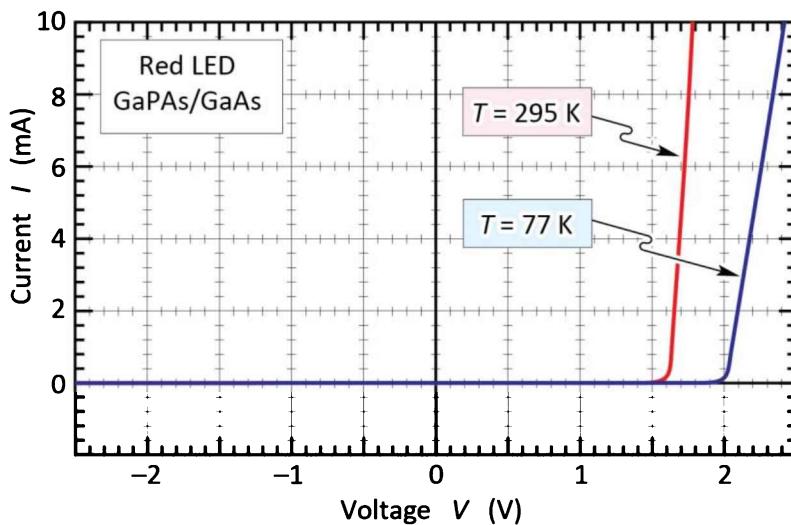


Fig. 34.5: Current–voltage characteristic of GaAs/GaAs LED emitting in the red part of the visible spectrum, measured at 77 and 295 K. The threshold voltages are 2.0 and 1.6 V, at 77 and 300 K, respectively.

Exercise: Temperature dependence of diode forward voltage. Experimentally determined linear temperature coefficients (dV_f/dT) for GaAs diodes range from 1.2 to 1.4 mV/K. Calculate the linear temperature coefficient of the forward voltage of a GaAs diode with $N_A = N_D = 2 \times 10^{17} \text{ cm}^{-3}$ at room temperature. What is the decrease in forward voltage if the ambient temperature is increased from 20 to 40°C and the internal heating in the diode can be neglected?

Solution: For GaAs with $\alpha = 5.41 \times 10^{-4} \text{ eV/K}$ and $\beta = 204 \text{ K}$, one obtains, at room temperature $dV_f/dT = -1.09 \text{ mV/K}$. The decrease in diode voltage for the 20°C temperature increase is $\Delta V_f = 21.9 \text{ mV}$.

34.4 Measurement of junction temperature using forward voltage

This method consists of a V_f calibration measurement under pulsed-current injection, and a V_f measurement under DC-current injection. The two measurements are illustrated in *Figure 34.6*. In the calibration measurement, the device under test is located in a temperature-controlled oven, so that the temperature of the device and junction is known. The temperature is varied from typically 20°C to 120°C. The calibration measurement is performed in a pulsed mode with a very small duty cycle (e.g. 0.1%), so that the heat generated by the injection current becomes negligibly small. The forward voltage is measured at each temperature for the current levels of interest. The calibration measurement establishes the relation between forward voltage and junction temperature for the I_f levels of interest.

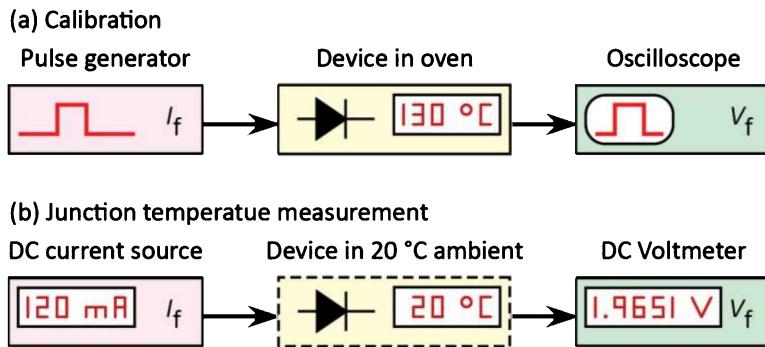


Fig. 34.6: (a) Pulsed calibration procedure establishing the forward voltage versus junction temperature (V_f versus T_j) relation and (b) determination of junction temperature for different DC forward currents.

Subsequently the device is exposed to room-temperature ambient and subjected to a series of DC currents. Forward voltages are measured once thermal steady state has been reached. The measured DC forward voltages and the calibration measurement data are used to establish the junction temperature for different current levels. A calibration measurement and a junction temperature measurement for an AlGaN deep UV LED is shown in *Figure 34.7* (Xi *et al.*, 2005).

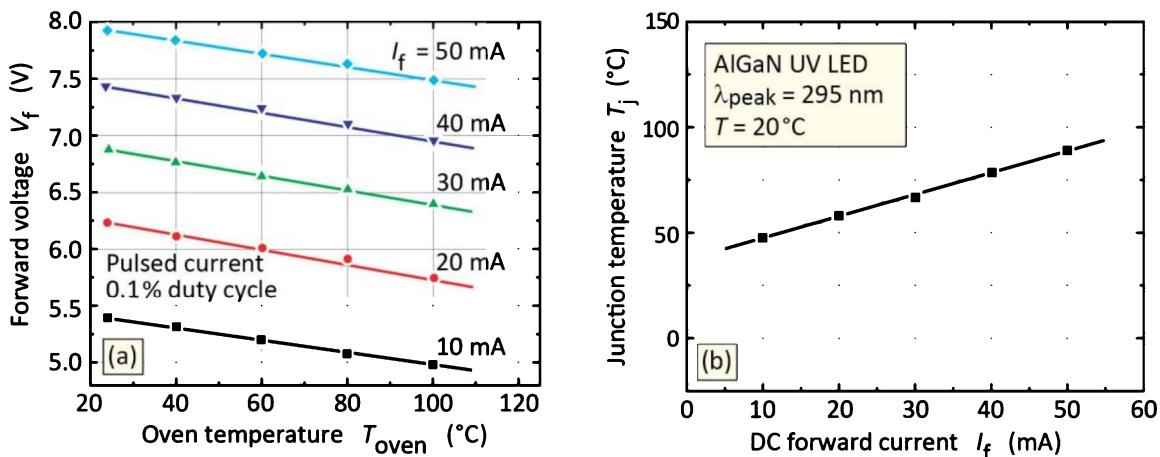


Fig. 34.7: (a) Pulsed calibration measurement (duty cycle 0.1%) and (b) junction temperature (T_j) versus DC current of an AlGaN UV LED (after Xi *et al.*, 2005)

Junction temperatures of several different devices are shown in *Figure 34.8* including red (AlGaInP, $\lambda = 625 \text{ nm}$), green (GaInN, $\lambda = 525 \text{ nm}$), blue (GaInN, $\lambda = 460 \text{ nm}$), and UV (GaInN, $\lambda = 370 \text{ nm}$) devices packaged in conventional 5 mm packages (Chhajed *et al.*, 2005). The forward-voltage method is accurate to within a few degrees. The V_f method is more accurate than the peak-wavelength method. The latter method is limited by the uncertainty in the peak wavelength, which is difficult to determine accurately for broadened emission bands. Also shown in the figure is the carrier temperature derived from the high-energy slope of the spectrum. The

accuracy of the carrier temperature suffers from alloy broadening, which decreases the high-energy slope (and thus increases the apparent carrier temperature).

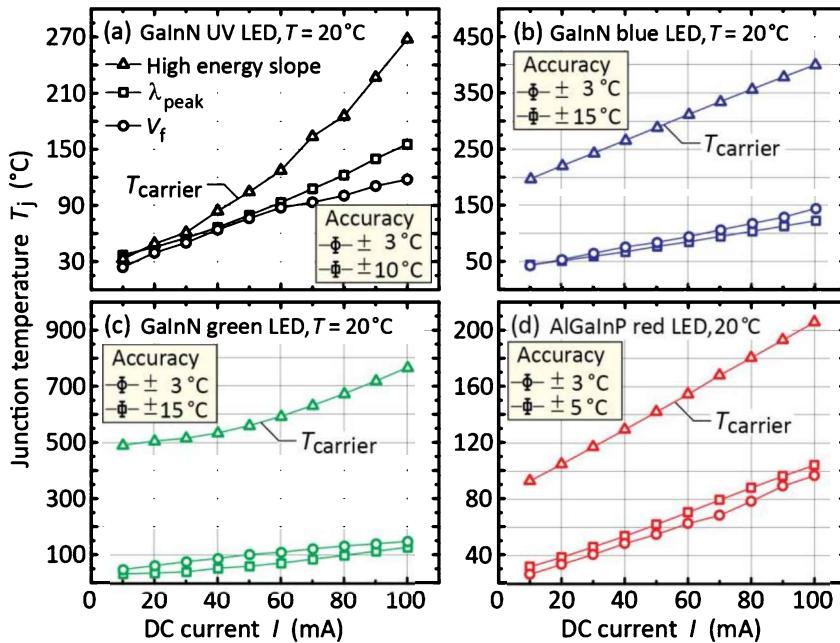


Fig. 34.8: Junction and carrier temperature of devices in conventional 5 mm packages as a function of DC injection current. The measured carrier temperature over-estimates the true carrier temperature due to alloy broadening (after Chhajed *et al.*, 2005).

Next, we describe an *alternative, very useful method* to measure the junction temperature of an LED (Li, 2007). The method consists of (i) a *calibration measurement* and (ii) the actual *junction temperature measurement*. Since the method *requires only one calibration measurement*, it consumes less time. Another advantage of the method is that the *calibration measurement is done under DC* (rather than pulsed) conditions, which generally is more stable and easier to perform.

In the *calibration measurement*, the device is subjected to a small DC forward current (probe current) that does not cause significant heating, and placed in an oven where it is subjected to a series of different temperatures. At each temperature, the DC forward voltage is measured so that a calibrated V_f -versus- T relationship is established. The probe current is selected to have a small value (e.g. 1 mA for small-area LEDs and 10 mA for large-area LEDs), so that no significant heating results from this current. The calibration measurement is schematically shown in **Figure 34.9 (a)**.

For the *junction temperature measurement*, the device is placed in a room-temperature ambient and then subjected to the quasi-DC forward currents of interest, e.g. $I_f = 100$ mA. A “negative” pulse is applied to the device, during which the forward current is reduced to the

probe current value (e.g. 1 mA) and the forward voltage is measured during the 1 mA pulse, as shown in **Figure 34.9 (b)**. Using the previously calibrated V_f -versus- T relationship, the corresponding junction temperature is obtained. Note that the measurement pulse must be very short, that is, shorter than the thermal time constant of the LED (< 1 ms or even < 0.1 ms); in addition, the duty cycle should be very small (e.g. 1 %), so that the probe pulse does not change the junction temperature.

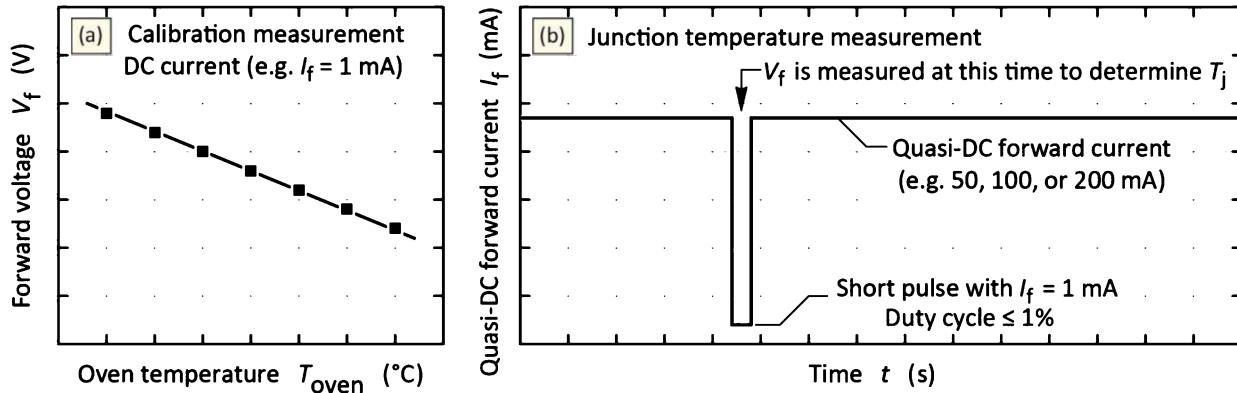


Fig. 34.9: (a) DC calibration measurement at very small current, for example 1 mA, so that it does not significantly heat the junction. (b) Quasi-DC-forward-current-versus-time trace used to measure the junction temperature.

34.5 Thermal considerations for pulsed measurements

The temperature depends on the power consumed in a device according to the following equation (Mena *et al.*, 1999)

$$T_j = T_{\text{ambient}} + (IV - P_{\text{optical}}) R_{\text{th}} - \tau_{\text{th}} \frac{dT_j}{dt} \quad (34.9)$$

where IV is the electrical input power, P_{optical} is the optical power emitted by the device, R_{th} is the thermal resistance of the device, and τ_{th} is the **thermal time constant of the device**.

Let us assume that the input electrical power is instantaneously switched “on” or “off”, so that the input power is changed in a step-function-like fashion. In this case, the temperature changes according to the exponential function, $\exp(-t/\tau_{\text{th}})$, which can be verified by solving the above differential equation. If the power is a constant for $t < 0$ and switched “ON” or “OFF” at $t = 0$, then the temperature change is given by

$$\text{Power "ON": } T_j(t) = T_{ambient} + \left(T_j|_{t \rightarrow \infty} - T_{ambient} \right) (1 - e^{-t/\tau}) \quad (34.10)$$

$$\text{Power "OFF": } T_j(t) = T_{ambient} + \left(T_j|_{t=0} - T_{ambient} \right) e^{-t/\tau}. \quad (34.11)$$

Let us further assume that the thermal time constant is 1 ms. In this case, the temperature versus time changes as shown in *Figure 34.10*. The exponential heat-up and cool-down function is shown on a linear scale and logarithmic time scale in *Figure 34.10* (a) and (b), respectively. Inspection of the figure reveals that the duration of a pulse must be much shorter than the thermal time constant in order to ensure that the temperature does not significantly change during the pulse. This condition, $\tau_{pulse} \ll \tau_{th}$, must be met for all pulsed measurements in which heating or cooling effects are to be neglected.

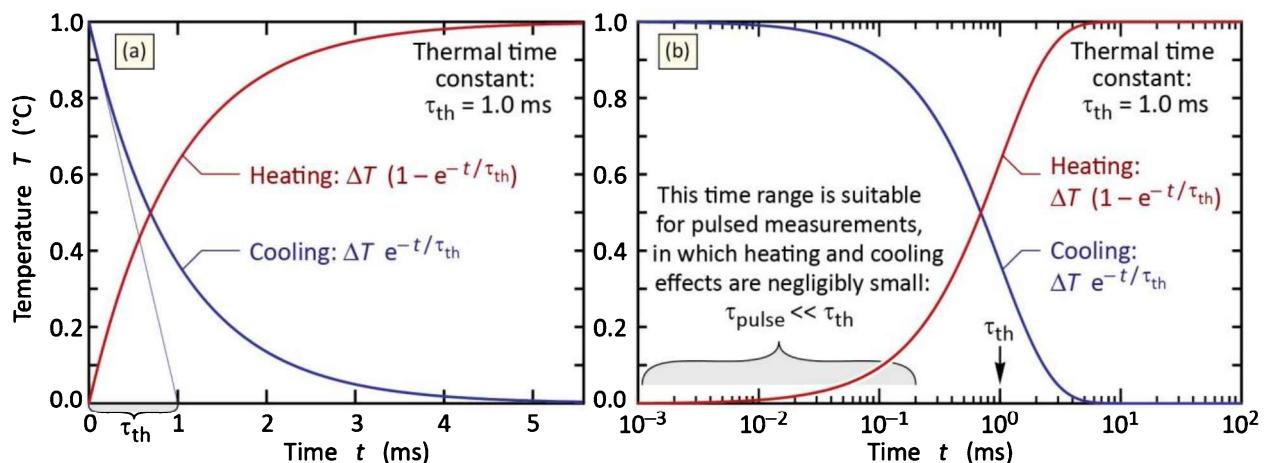


Fig. 34.10: Exponential temperature decay and rise on (a) linear and (b) logarithmic time axis. If the pulse duration (τ_{pulse}) is much smaller than the thermal time constant (τ_{th}), then the measurement is not subject to the evolving temperature change.

For LEDs, the thermal time constant, τ_{th} , has been estimated to be on the order of 1 ms (Shan *et al.*, 2010). Thus, the pulse duration for a LED should be $\tau_{pulse} < 1$ ms. A typical value for the pulse duration in pulsed measurements is $\tau_{pulse} = 50$ μ s along with a duty cycle of $\leq 5\%$.

For vertical-cavity surface-emitting lasers, the thermal time constant, τ_{th} , has been estimated to be on the order of 1 μ s (Hasnain *et al.*, 1991; Mena *et al.*, 1999). Thus, the pulse duration for a laser diode should be $\tau_{pulse} < 1$ μ s. A typical value for the pulse duration in pulsed measurements is $\tau_{pulse} = 500$ ns along with a duty cycle of $\leq 5\%$.

34.6 Constant-current and constant-voltage DC drive circuits

Different considerations play a role in designing the DC drive circuit of an LED operated under steady-state conditions. These considerations include the simplicity and cost of the drive circuit, the power-conversion efficiency, and the compensation of the temperature dependence of the light intensity emitted from the LED.

A simple drive circuit is a **constant-voltage supply** such as a battery or the rectified AC output of a transformer. There are, however, two drawbacks to constant-voltage drives of LEDs. Firstly, the diode current depends exponentially on the voltage, so a small variation in the drive voltage results in a large change in the current. Secondly, the threshold voltage of a diode depends on temperature, so any temperature change results in a significant change in current. A constant-voltage operating characteristic of a diode is shown in **Figure 34.11**.

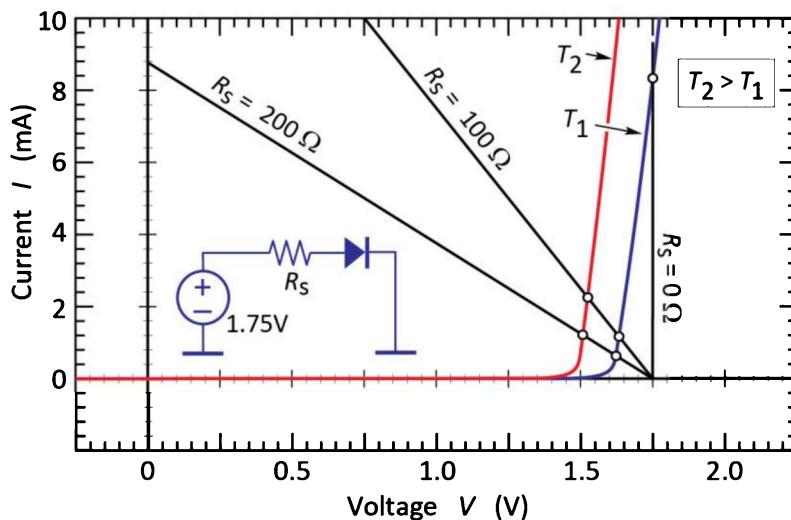


Fig. 34.11: LED drive circuit with series resistance R_s . The intersection between the diode I - V characteristics and the load lines are the points of operation. Small series resistances result in an increased diode current at high temperatures, thus allowing for compensation of a lower LED radiative efficiency.

If a resistor is connected in series with the diode, the strong temperature dependence of the diode current is reduced. The series resistance together with the temperature dependence of the diode determines the temperature coefficient of the diode current.

When a diode is driven with a constant current, then the emission intensity decreases with increasing temperature. A constant-voltage power supply with a series resistance can be used to reduce the temperature dependence of the emission intensity. The emission intensity of LEDs generally decreases with increasing temperature due to non-radiative recombination. In addition, the threshold voltage decreases with increasing temperature. However, for a constant-voltage supply, the diode current increases when the temperature increases, as shown in

Figure 34.11. Thus, a series resistor can be used to compensate for the emission intensity decrease at elevated temperatures. It should be noted that the electrical-to-optical power-conversion efficiency drops due to the power consumed in the series resistor.

The temperature dependence of LED intensity is an important factor for LEDs used in outdoor applications. On hot summer days, the temperature and the ambient light intensity are high. Owing to the high temperature, the LED intensity drops. In addition, high brightness is required due to high ambient light levels. This effect can be compensated for by driving the LEDs with a higher current as the temperature increases.

A constant-current drive circuit can consist of a transistor with the LED as a load. A constant-current drive circuit allows one to drive the LED irrespective of the diode threshold voltage and the diode temperature. However, a constant-current drive circuit does not compensate for the decrease of LED emission at elevated temperatures.

Exercise: Compensation of the temperature dependence of an LED with a drive circuit. Consider an LED with a characteristic temperature $T_1 = 100$ K, a turn-on voltage of 1.4 V at 20°C, a temperature coefficient of the turn-on voltage of -2.1 mV/K, and a linear I – V characteristic with a differential resistance of $5\ \Omega$ for forward voltages larger than the turn-on voltage. Assume that the temperature dependence of the emission intensity is given by $I = I|_{300\text{ K}} \exp[-(T - 300\text{ K})/T_1]$.

Design a drive circuit consisting of a constant-voltage source and a resistor, which compensates for the temperature dependence of the emission intensity of the LED so that the LED emission intensity is the same at the water freezing-point temperature (0°C) and 60°C. The LED should draw 20 mA at the freezing-point temperature.

Solution: At 60°C, the current needs to be 36.4 mA in order to keep the emission intensity independent of temperature. Constructing a load line that intersects the 0°C and 60°C diode I – V characteristic at 20 mA and 36.4 mA, respectively, yields the following values for the drive circuit: Constant-voltage source with $V = 1.6$ V and series resistance of $2.7\ \Omega$.

34.7 References

- Abdelkader H. I., Hausien H. H., and Martin J. D. “Temperature rise and thermal rise-time measurements of a semiconductor laser diode” *Review of Scientific Instrumentation* **63**, 2004 (1992)
- Chhajed S., Xi Y., Li Y.-L., Gessmann Th., and Schubert E. F. “Influence of junction temperature on chromaticity and color rendering properties of trichromatic white light sources based on light emitting diodes” *Journal of Applied Physics* **97**, 054506 (2005)
- Epperlein P. W. “Reflectance modulation – a novel approach to laser mirror characterization” in *Proceedings of 17th International Symposium of Gallium Arsenide and Related Compounds*, IOP Conference Series, IOP, London, **112**, 633 (1990)
- Epperlein P. W. and Bona G. L. “Influence of the vertical structure on the mirror facet temperatures of visible GaInP quantum well lasers” *Applied Physics Letters* **62**, 3074 (1993)

- Gessmann Th., Schubert E. F., Graff J. W., Streubel K., and Karnutsch C. "Omni-directionally reflective contacts for light-emitting diodes" *IEEE Electron Devices Letters* **24**, 683 (2003)
- Gu Y. and Narendran N. "A non-contact method for determining junction temperature of phosphor-converted white LEDs" *Third International Conference on Solid State Lighting, Proceedings of SPIE*, San Diego, Calif., 2003 (to be published) see also J. Taylor "Non-intrusive techniques help to predict the lifetime of LED lighting systems" *Compound Semiconductors* October (2003)
- Hall D. C., Goldberg L., and Mehuys D. "Technique for lateral temperature profiling in optoelectronic devices using a photoluminescence microprobe" *Applied Physics Letters* **61**, 384 (1992)
- Hasnain G., Tai K., Yang L., Wang Y. H., Fischer R. J., Wynn J. D., Weir B., Dutta N. K., and Cho A. Y. "Performance of gain-guided surface emitting lasers with semiconductor distributed Bragg reflectors" *IEEE Journal of Quantum Electronics* **27**, 1377 (1991)
- Ioffe Physico-Technical Institute (Saint Petersburg, Russia) "Physical properties of semiconductors" < www.ioffe.ru/SVA/NSM/Semicond > (2004)
- Li Y. L. "Charles", personal communication (2007)
- Mena P. V., Morikuni J. J., Kang S.-M., Harton A. V., and Wyatt K. W. "A simple rate-equation-based thermal VCSEL model" *IEEE Journal of Lightwave Technology* **17**, 865 (1999)
- Millman J. and Halkias C. *Integrated Electronics: Analog and Digital Circuits and Systems* (McGraw-Hill, New York, 1972)
- Murata S. and Nakada H. "Adding a heat bypass improves the thermal characteristics of a 50 µm spaced 8-beam laser diode array" *Journal of Applied Physics* **72**, 2514 (1992)
- Rommel J. M., Gavrilovic P. and Dabkowski F. P. "Photoluminescence measurement of the facet temperature of 1 W gain-guided AlGaAs/GaAs laser diodes" *Journal of Applied Physics* **80**, 6547 (1996)
- Schubert E. F., Göbel E. O., Horikoshi Y., Ploog K., and Queisser H. J. "Alloy broadening in photoluminescence spectra of AlGaAs" *Physical Review B* **30**, 813 (1984)
- Shan Q., Dai Q., Chhajed S., Cho J., and Schubert E. F. "Analysis of thermal properties of GaInN light-emitting diodes and laser diodes" *Journal of Applied Physics* **108**, 084504 (2010)
- Todoroki S., Sawai M., and Aiki K. "Temperature distribution along the striped active region in high-power GaAlAs visible lasers" *Journal of Applied Physics* **58**, 1124 (1985)
- Varshni Y. P. "Temperature dependence of the energy gap in semiconductors" *Physica* **34**, 149 (1967)
- Xi Y. and Schubert E. F. "Junction-temperature measurement in GaN ultraviolet light-emitting diodes using diode forward-voltage method" *Applied Physics Letters* **85**, 2163 (2004)
- Xi Y., Xi J.-Q., Gessmann T., Shah J. M., Kim J. K., Schubert E. F., Fischer A. J., Crawford M. H., Bogart K. H. A., and Allerman A. A. "Junction temperature measurements in deep-UV light-emitting diodes" *Applied Physics Letters* **86**, 031907 (2005)

35

35 – Communication: Optical communication

LEDs are used in communication systems transmitting low and medium data rates (< 1 Gbit/s) over short and medium distances (< 10 km). These communication systems are based on either *guided light waves* (Keiser, 1999; Neyer *et al.*, 1999; Hecht, 2001; Mynbaev and Scheiner, 2001; Kibler *et al.*, 2004) or *free-space waves* (Carruthers, 2002; Heatley *et al.*, 1998; Kahn and Barry, 2001). In guided-wave communication, individual optical fibers or fiber bundles are used as the transmission medium and LED-based optical communication links are limited to distances of a few kilometers. Optical fiber systems include *silica* and *plastic* optical fibers. Free-space communication is usually limited to a room, even though longer distances are possible. In this chapter we discuss the characteristics of transmission media used for LED communication.

35.1 Types of optical fibers

The cross section of optical fibers consists of a circular core region surrounded by a cladding region. The core region has a higher refractive index than the cladding region. Typically, the core refractive index is about 1% higher than the cladding refractive index. Light propagating in the core is guided inside the core by means of *total internal reflection*. The condition of total internal reflection can be inferred from Snell's law. A light ray is *totally internally reflected* whenever it is incident at the core–cladding boundary. In a ray-optics picture, light rays propagating inside the core follow a *zigzag path*.

There are three types of optical fibers used in communication systems. These types are the (*i*) *step-index multimode fiber*, (*ii*) *graded-index multimode fiber*, and (*iii*) *single-mode fiber*. The three types of fibers are shown in **Figure 35.1** along with the refractive index profiles.

Step-index multimode fibers have a relatively large core diameter. Typical core diameters are 50, 62.5, and 100 μm for silica fibers used in communication systems. Plastic optical fibers have larger core diameters, typically 1 mm. An important advantage of multimode fibers is *easy coupling* of the light source to the fiber. Usually, a $\pm 5 \mu\text{m}$ accuracy of alignment is sufficient for multimode fibers with a core diameter of 50 μm . The main disadvantage of multimode fibers is the occurrence of *modal dispersion*.

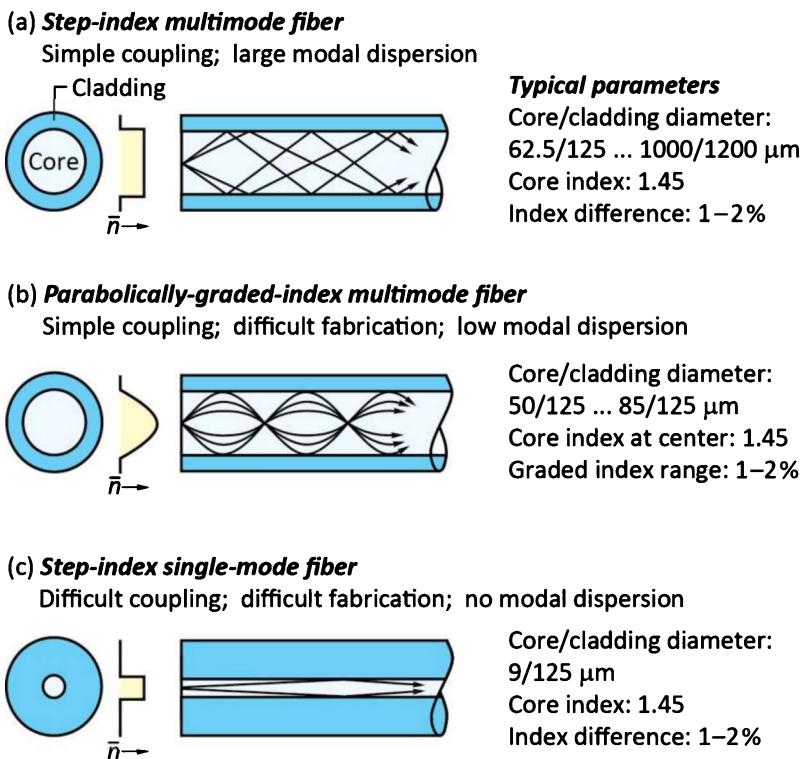


Fig. 35.1: (a) Step-index multimode fibers allow for the propagation of several optical modes. (b) Parabolically graded-index multimode fibers allow for the propagation of several modes with similar propagation constant. Graded-index multimode fibers have a lower modal dispersion than step-index multimode fibers. (c) Step-index single-mode fibers have a small core diameter and no modal dispersion.

Because the core diameter in multi-mode fibers is much larger than the operating wavelength, several optical modes can propagate in the waveguide. These optical modes have different propagation constants so that different modes arrive at the end of the fiber at different times, even if they were launched at the same time. This leads to a broadening of the optical pulse and limits the maximum bit rate that can be transmitted over a fiber of a given length.

Modal dispersion is reduced by *graded-index* multimode fibers. Graded-index multimode fibers have a parabolically graded core index leading to a reduction in modal dispersion.

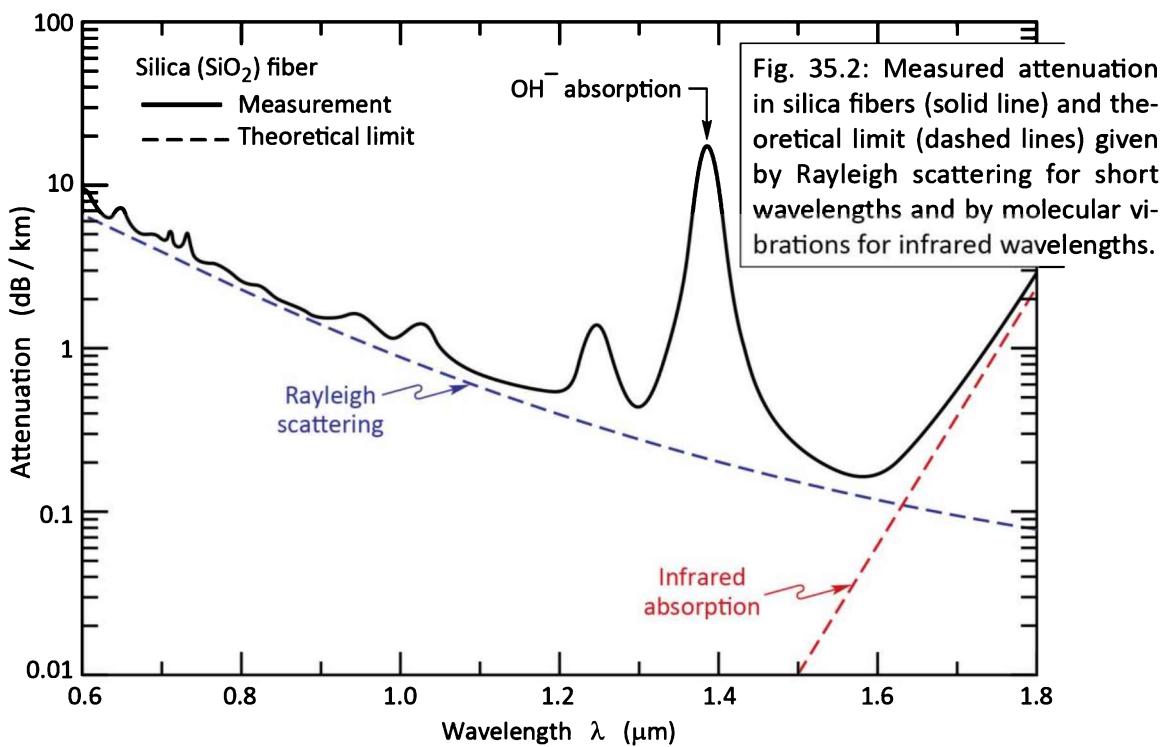
Single-mode fibers have such a small core diameter that only a single optical mode can propagate in the fiber. Typical single-mode core diameters are 5~10 µm. The main advantage of single-mode fibers is the *lack of modal dispersion*. The main disadvantage of single-mode fibers is *difficult coupling* due to the small core diameter. A small core diameter requires light sources with high brightness such as lasers. However, LEDs, in particular edge-emitting LEDs and superluminescent LEDs are also occasionally used with single-mode fibers. Coupling of light into single-mode fibers requires precise alignment with tolerances of a few micrometers.

If the *optical power* to be transmitted over an optical fiber is of prime interest, the core diameter should be as large as possible and the core–cladding index difference should also be as

large as possible. Specialty fibers with core diameters of > 1 mm are available. Such fibers are not suitable for communication applications due to the large modal dispersion.

35.2 Attenuation in silica and plastic optical fibers

Silica (SiO_2) has excellent optical properties including great long-term stability. A large variety of glasses and fibers are available. The attenuation of silica fibers is shown in *Figure 35.2*. Inspection of the figure reveals that a minimum loss of 0.2 dB occurs at a wavelength of 1.55 μm .



There are several optical “windows” for communication over silica fibers. These communication windows are at 0.85, 1.3, and 1.55 μm . The 0.85 μm communication window is suitable for communication with GaAs-based LEDs and lasers. However, this window is limited to short distances due to the material dispersion and to the high attenuation of silica fibers at that wavelength. The 1.3 μm communication window is also suited for communication with LEDs and lasers. This window has relatively low loss and zero dispersion, allowing for high-bit-rate transmission, in particular in graded-index and single-mode fibers. The 1.55 μm communication window is characterized by the lowest loss of all three windows. Consequently, this window is used for long-distance high-bit-rate communication. To allow for high bit rates, single-mode

fibers must be used. Since it is difficult to efficiently couple light emerging from an LED into a single-mode fiber, lasers are preferred over LEDs at $1.55\text{ }\mu\text{m}$.

Plastic optical fibers are becoming increasingly popular for short-distance communication (Neyer *et al.*, 1999; Kibler *et al.*, 2004). However, plastic fibers have losses that are about 1 000 times greater than the losses in silica fibers. Therefore, the transmission distances are limited to just a few meters to a few hundred meters, e.g. for communication within an automobile (Kibler *et al.*, 2004) or airplane.

The attenuation in plastic fibers is shown in *Figure 35.3*. The preferred communication window of plastic fibers is at 650 nm, where the loss is of the order of $0.1\sim0.2\text{ dB per meter}$. At even shorter wavelengths, the attenuation in plastic fibers decreases. However, the material dispersion increases, thus making the 650 nm wavelength the preferred communication wavelength in plastic optical fibers.

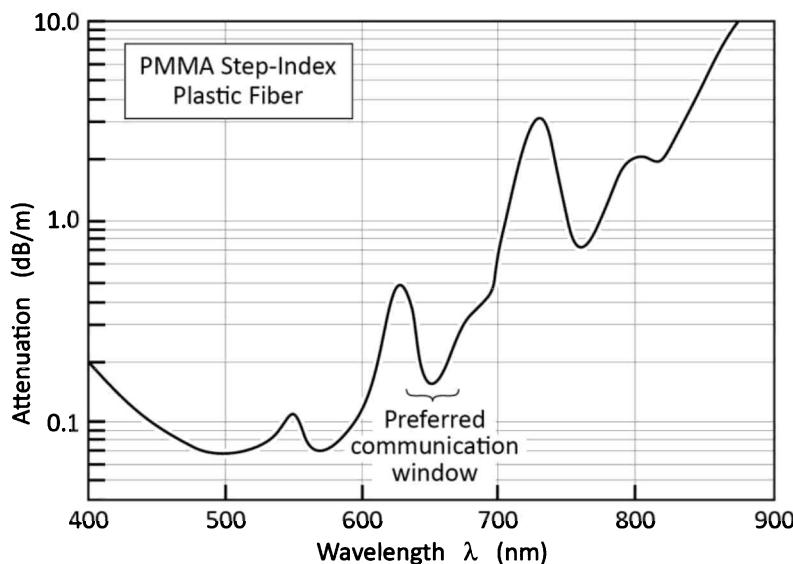


Fig. 35.3: Attenuation of a PMMA step-index plastic optical fiber. At 650 nm, the preferred communication wavelength, the attenuation is about 150 dB/km (after data sheet of Toray Industries Ltd., 2002).

35.3 Modal dispersion in fibers

Modal dispersion occurs in multimode fibers that have a larger core diameter or a larger index difference between the core and the cladding than single-mode fibers. Typical core diameters range from 50 to 1 000 μm for multimode fibers and 5 to 10 μm for single-mode fibers. In the ray optics model, *different optical modes* correspond to light rays propagating at *different angles* in the core of the waveguide. The derivation of propagation angles in multimode fibers would go beyond the scope of this chapter. Here an approximate calculation will be performed to obtain the modal dispersion.

Consider a fiber waveguide with refractive indices of the core and cladding of \bar{n}_1 and \bar{n}_2 , respectively. Assume that the waveguide supports the propagation of more than one optical mode. Two of these modes are shown schematically in **Figure 35.4**. Owing to the difference in optical path length, the mode with the smaller propagation angle θ will arrive earlier at the end of the multimode fiber. The **modal dispersion** is the time delay between the fastest and the slowest optical mode normalized to the length L of the waveguide.

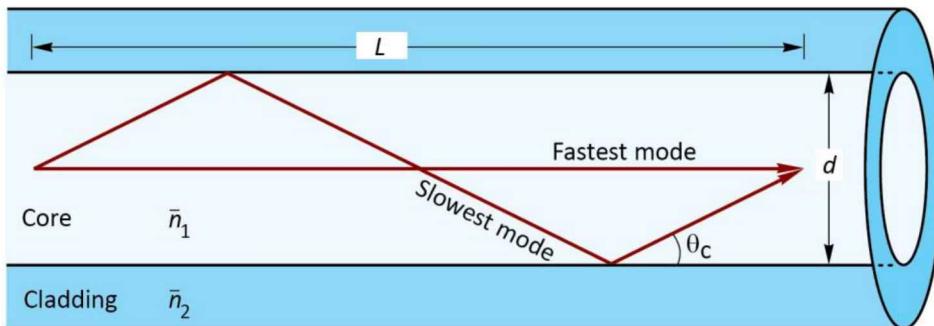


Fig. 35.4: Geometry used for calculation of the modal dispersion in a multimode fiber waveguide.

In the calculation, assume that the phase and group velocity are given by $v_{ph} = c/\bar{n}_1 \approx v_{gr}$. The fastest mode has the smallest propagation angle and we approximate the smallest angle by $\theta_{m=0} \approx 0^\circ$. The slowest mode has the largest propagation angle and we approximate the largest angle by $\theta_m \approx \theta_c$, where θ_c is the critical angle of total internal reflection. This approximation can be made without loss of accuracy for multimode fibers which carry many modes.

The propagation times for the fastest and slowest modes per unit length of the fiber are given by

$$\tau_{fast} = \frac{L}{c/\bar{n}_1} \quad \text{and} \quad \tau_{slow} = \frac{L/\cos \theta_c}{c/\bar{n}_1} \quad (35.1)$$

where the critical angle for total internal reflection can be derived from Snell's law and is given by

$$\theta_c = \arccos (\bar{n}_2 / \bar{n}_1) . \quad (35.2)$$

The time delay per unit length, or modal dispersion, is then given by

$$\frac{\Delta\tau}{L} = \frac{\tau_{slow} - \tau_{fast}}{L} = \frac{\bar{n}_1}{c} \left(\frac{1}{\cos \theta_c} - 1 \right). \quad (35.3)$$

A waveguide supporting *many* modes has a large time delay between the fastest and slowest modes. Thus, modal dispersion increases with the number of optical modes supported by the waveguide.

Exercise: Modal dispersion in waveguides. Calculate the time delay between the slowest and the fastest modes, and the maximum possible bit rate for a 1 km long multimode fiber waveguide with core refractive index $\bar{n}_1 = 1.45$ and cladding refractive index $\bar{n}_2 = 1.4$.

Solution: Using Snell's law (Eqn. 35.2), one obtains $\theta_c \approx 15^\circ$. The time delay calculated from Eqn. (35.3) for a 1 km long fiber amounts to $\Delta\tau = 170$ ns. The minimum time required to transmit one bit of information is given by $\Delta\tau$. This yields an approximate maximum bit rate of $f_{max} = 1/170$ ns = 5.8 Mbit/s. The calculation shows that modal dispersion can be a significant limitation in optical communication. Graded-index multimode fibers or single-mode fibers are therefore required for high-speed communication systems.

35.4 Material dispersion in fibers

Material dispersion is another mechanism limiting the capacity of optical fibers. Material dispersion is due to the dependence of the refractive index on the wavelength. **Figure** 35.5 shows, as a function of wavelength, the phase refractive index and the group refractive index of silica. The indices are defined as

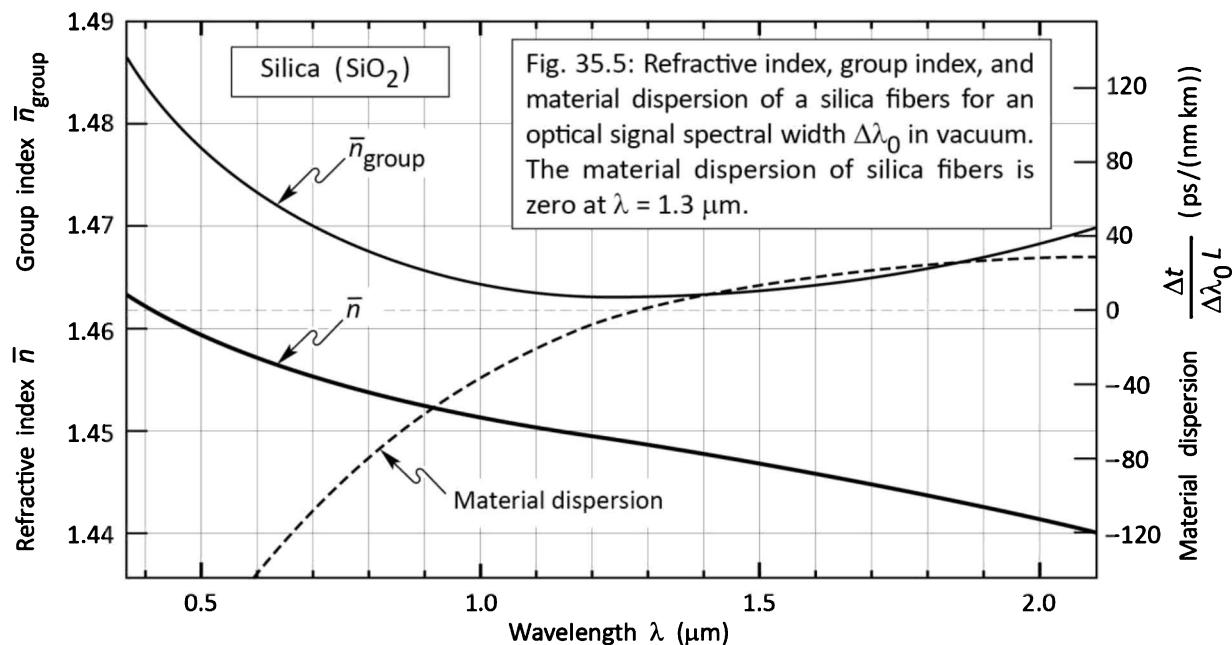
$$\bar{n} = \frac{c}{v_{ph}} \quad (\text{phase refractive index}) \quad (35.4)$$

and

$$\bar{n}_{gr} = \frac{c}{v_{gr}} \quad (\text{group refractive index}) \quad (35.5)$$

where v_{ph} and v_{gr} are the phase and group velocity in silica, respectively. The **phase refractive index** and the **group refractive index** are related by

$$\bar{n}_{gr} = \bar{n} - \lambda \frac{d\bar{n}}{d\lambda} = \bar{n} - \lambda_0 \frac{d\bar{n}}{d\lambda_0}. \quad (35.6)$$



If the fiber is dispersive, the difference in group velocity between the “slowest color” and the “fastest color” of the optical signal is given by

$$|\Delta v_{gr}| = \frac{c}{\bar{n}_{gr}^2} \frac{d\bar{n}_{gr}}{d\lambda} \Delta\lambda \quad (35.7)$$

where $\Delta\lambda$ is the spectral width of the optical signal.

The time delay between the leading edge and the trailing edge of an optical signal after traveling in the fiber for a length L , called the **material dispersion**, is given by

$$\Delta\tau = \frac{L}{v_{gr}^2} \Delta v_{gr} = \frac{L}{c} \frac{d\bar{n}_{gr}}{d\lambda} \Delta\lambda = \frac{L}{c} \frac{d\bar{n}_{gr}}{d\lambda_0} \Delta\lambda_0. \quad (35.8)$$

The material dispersion is measured in $\text{ps}/(\text{nm km})$ and it is illustrated for silica fibers in **Figure 35.5**. LEDs have a broad emission linewidth. Therefore, material dispersion is, along with modal dispersion, the bandwidth-limiting factor in optical fiber communication systems operated with LEDs.

Exercise: Material dispersion in waveguides. Derive Eqns. (35.6) and (35.7). Why does material dispersion have a much smaller significance for semiconductor lasers than for LEDs?

Solution:

Derivation of Eqn. (35.6): The group refractive index is defined as

$$\bar{n}_{gr} = \frac{c}{v_{gr}} = c \frac{dk}{d\omega} = c \frac{d}{d\omega} \frac{\omega \bar{n}}{c} .$$

In this equation we have used

$$k = \frac{2\pi}{\lambda} = \bar{n} k_0 = \bar{n} \frac{2\pi}{\lambda_0} = \bar{n} \frac{2\pi}{c/v} = \bar{n} \frac{\omega}{c} .$$

Performing the derivative yields

$$\bar{n}_{gr} = c \left(\frac{\bar{n}}{c} + \frac{\omega}{c} \frac{d\bar{n}}{d\omega} \right) = \bar{n} + \omega \frac{d\bar{n}}{d\omega} .$$

Using $\omega = 2\pi v = 2\pi c/\lambda_0$, and using $d(1/\lambda_0) = -\lambda_0^2 d\lambda_0$, one obtains

$$\bar{n}_{gr} = \bar{n} + 2\pi \frac{c}{\lambda_0} \frac{1}{2\pi c} \frac{d\bar{n}}{d\frac{1}{\lambda_0}} = \bar{n} - \lambda_0 \frac{d\bar{n}}{d\lambda_0} ,$$

what was to be shown.

Derivation of Eqn. (35.7): It is $v_{gr} = c/\bar{n}_{gr}$. Forming the derivative with respect to λ yields $dv_{gr}/d\lambda = c (d/d\lambda) \bar{n}_{gr}^{-1} = c (-1/\bar{n}_{gr}^2) (d\bar{n}_{gr}/d\lambda)$. Therefore, $\Delta v_{gr} = c (-1/\bar{n}_{gr}^2) (d\bar{n}_{gr}/d\lambda) \Delta\lambda$. Taking the absolute value of both sides of the equation yields the equation that was to be shown.

Why is material dispersion much less relevant for lasers than for LEDs? According to Eqn. (35.8), material dispersion is proportional to the spectral linewidth of the source, $\Delta\lambda_0$. Because lasers (longitudinal multi-mode as well as single-mode lasers) have a much narrower linewidth than LEDs, material dispersion is generally much less relevant for laser-based communication systems than it is for LED-based communication systems.

Substantial material dispersion exists in plastic fibers at all wavelengths of interest. These wavelengths are the local loss minimum at 650 nm and the low-loss region of 500~600 nm. The material dispersion is given in **Table 35.1**. The data indicates that 650 nm is the wavelength of least dispersion, making 650 nm the preferred communication wavelength in plastic optical fibers.

Table 35.1: Material dispersion in PMMA plastic optical fibers (courtesy of R. Marcks von Wurtemberg, *Mitel* Corporation, 2000).

Wavelength	525 nm	560 nm	650 nm
Material dispersion	700 ps/(nm km)	500 ps/(nm km)	320 ps/(nm km)

Exercise: Comparison of material and modal dispersion. Consider a 62.5 μm core diameter multimode step-index fiber of 3 km length with a core index of $\bar{n}_1 = 1.45$ and a cladding index of $\bar{n}_2 = 1.4$. Assume that the fiber inputs come from either an LED or a laser emitting at 850 nm. Assume that the LED and the laser have a linewidth of 50 and 5 nm, respectively. Calculate the modal and the material dispersion for each case and explain the result.

Solution: Modal dispersion: The calculation of the modal dispersion depends only on the fiber and is independent of the source. Using Snell's law, one obtains $\theta_c \approx 15^\circ$. The time delay due to modal dispersion calculated from Eqn. (35.3) for a 3 km long optical fiber amounts to $\Delta\tau_{\text{modal dispersion}} = 510 \text{ ns}$. **Material dispersion:** Figure 35.5 shows that at 850 nm, the material dispersion in silica is $\Delta\tau_{\text{material dispersion}}/(\Delta\lambda_0 L) \approx -65 \text{ ps / (nm km)}$. For a 3 km long optical fiber that is fed by a source with a 50 nm line width (LED) and 5 nm line width (laser), the dispersion amounts to 9.75 ns and 0.975 ns, respectively. This exercise shows that modal dispersion typically dominates over material dispersion when multimode fibers are used.

35.5 Numerical aperture of fibers

Owing to the requirement of total internal reflection, the only light rays that can propagate losslessly in the core of an optical fiber are those that have a propagation angle smaller than the critical angle for total internal reflection. Light rays for which the propagation angle is too large will consequently not couple into the fiber. Here we consider the coupling of light from an LED light source into an optical fiber. We assume that the fiber end has a polished planar surface normal to the optical axis of the fiber, as shown in **Figure 35.6**.

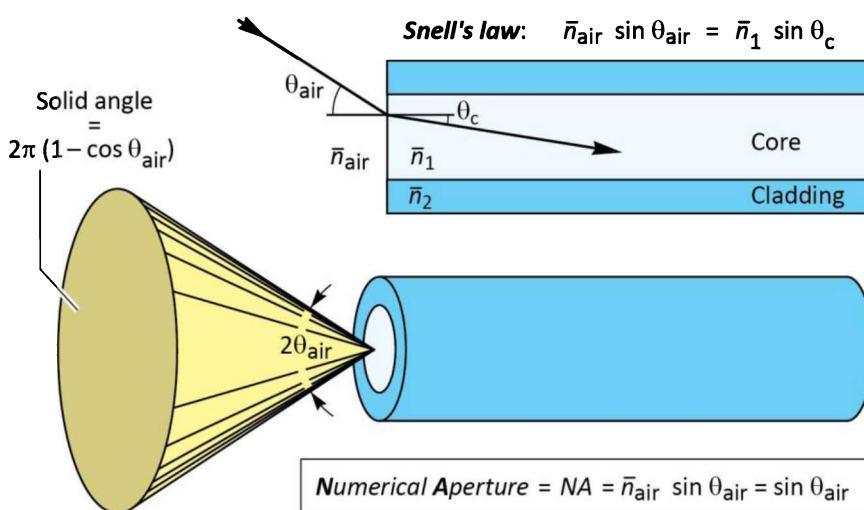


Fig. 35.6: Illustration of the *numerical aperture* (*NA*) of a fiber. For example, the light acceptance angle in air is $\theta_{\text{air}} = 11.5^\circ$ for a numerical aperture of $NA = 0.2$.

As a consequence of the requirement of total internal reflection, only a range of angles will be “accepted” by the fiber for lossless propagation. Outside the **acceptance angle** range, light rays

will be refracted into the cladding layer where they will incur losses.

The range of allowed angles can be inferred from Snell's law. As illustrated in **Figure 35.6**, the maximum angle for acceptance in the fiber is given by

$$\bar{n}_{air} \sin \theta_{air} = \bar{n}_1 \sin \theta_c . \quad (35.9)$$

Since the refractive index of air is approximately unity, the maximum acceptance angle in air is given by

$$\theta_{air} = \arcsin (\bar{n}_1 \sin \theta_c) . \quad (35.10)$$

The maximum acceptance angle defines a cone of allowed angles, as shown in **Figure 35.6**. Light rays incident on the core of the optical fiber with propagation angles within the cone can propagate without loss.

Another way to express the acceptance cone is the **numerical aperture** of the fiber. The numerical aperture (*NA*) is defined as

$$NA = \bar{n}_1 \sin \theta_c = \bar{n}_{air} \sin \theta_{air} = \sin \theta_{air} \approx \theta_{air} \quad (35.11)$$

where the approximation $\sin \theta_{air} \approx \theta_{air}$ is valid for *small* numerical apertures. Typical *NAs* for silica single-mode fibers are 0.1 and typical *NAs* for silica multimode fibers are 0.15~0.25. Plastic optical fibers can have higher *NAs*, typically 0.2~0.4.

The solid angle corresponding to a certain *NA* is given by

$$\text{Solid angle } \Omega = 2\pi(1 - \cos \theta_{air}) = 2\pi[1 - \cos(\arcsin NA)] \approx \pi NA^2 \quad (35.12)$$

where the small-angle approximations $\sin \theta_{air} \approx \theta_{air}$ and $\cos \theta_{air} \approx 1 - \frac{1}{2} \theta_{air}^2$ have been used. The power emitted by an LED is proportional to the solid angle. Thus, the power coupled to an LED is proportional to the *NA squared* of the fiber in the small-angle approximation.

Exercise: Coupling efficiency of a fiber butt-coupled to an LED. Consider an LED with a point-like emission region that emits an optical power of 1 mW into the hemisphere. For simplicity, assume that the intensity emitted by the LED is independent of the emission angle. What is the maximum acceptance angle of a single-mode fiber with *NA* = 0.1 and multimode fiber with *NA* = 0.25? What is the power that can be coupled into the two fibers?

Solution: The maximum acceptance angles of the single-mode and multimode fibers in air are $\theta_{air} = 5.7^\circ$ and 14.5° , respectively. The solid angle defined by an acceptance angle θ_{air} is given by $\Omega = 0.031$

and 0.20 for the single-mode and multimode fiber, respectively. Since the entire hemisphere has a solid angle of 2π , the power coupled into the single-mode and multimode fibers is given by 0.0049 mW and 0.032 mW, respectively.

35.6 Coupling with lenses

The low coupling efficiency of LEDs to optical fibers can be improved with convex lenses, if the light-emitting region of the LED is smaller than the optical fiber core. In this case, the light-emitting region can be imaged on the fiber core, thereby reducing the angle of incidence. The light source is *adapted* to the *NA* of the fiber (“*NA-matching*”).

A convex lens can produce an image with height I of a light-emitting object with height O . If the image is larger than the object, the angles of the light incident from the lens on the image are less divergent than the light emanating from the object towards the lens. The *smaller divergence* obtained for *magnified* images allows one to increase the coupling efficiency to fibers. The principle of coupling with a convex lens is shown in **Figure 35.7**.

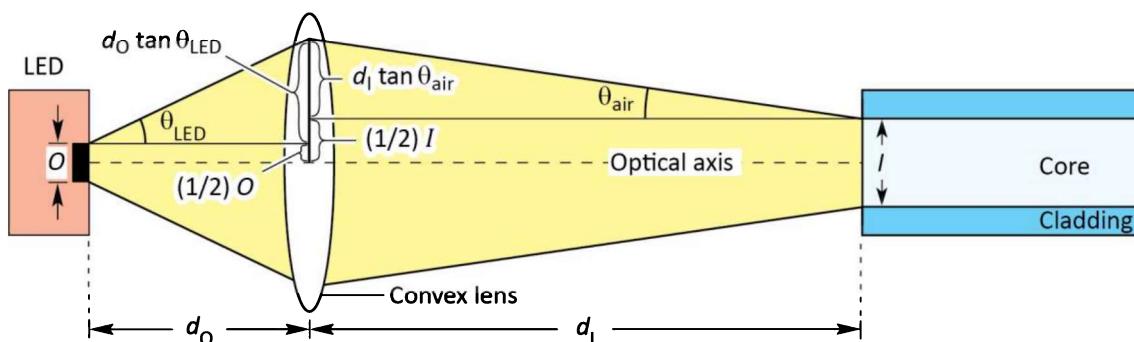


Fig. 35.7: Coupling by imaging the light-emitting region of an LED onto the core of an optical fiber. The LED's emission region has diameter O (Object). The emission region is imaged onto the fiber core with diameter I (Image) using a convex lens with focal length f .

The condition for a focused image (minimum image size) is given by the lens equation

$$\frac{1}{d_o} + \frac{1}{d_I} = \frac{1}{f} \quad (35.13)$$

where d_o and d_I are the distances of the object and the image from the lens, respectively, and f is the focal length of the lens. The magnification of the image of the LED light source on the core of the fiber is given by

$$M = \frac{I}{O} = \frac{d_I}{d_O} . \quad (35.14)$$

As shown in **Figure 35.7**, it is

$$\frac{1}{2}O + d_O \tan \theta_{LED} = \frac{1}{2}I + d_I \tan \theta_{air} . \quad (35.15)$$

If the LED and the core of the optical fiber are much smaller than the diameter of the lens and if the angles are relatively small, then Eqn. (35.15) can be approximated by

$$\theta_{LED} = \frac{d_I}{d_O} \theta_{air} = \frac{I}{O} \theta_{air} . \quad (35.16)$$

Since d_I is larger than d_O , the acceptance angle for light emanating from the LED is *larger* than that of the fiber, implying *increased* coupling efficiency. Consequently, we can define the numerical aperture of the LED, NA_{LED} , which is the maximum angle of light emanating from the LED that is accepted by the fiber. Using Eqn. (35.14) and the small-angle approximation, NA_{LED} is given by

$$NA_{LED} = \frac{I}{O} NA . \quad (35.17)$$

Since the coupling efficiency is proportional to NA^2 (see Eqn. 35.12), the coupling efficiency is increased to

$$\text{Coupling efficiency} \propto NA_{LED}^2 = \left(\frac{I}{O} NA \right)^2 . \quad (35.18)$$

The result shows that high coupling efficiencies are obtained for LEDs with small-diameter light-emitting regions, large fiber-core diameters, and large-NA fibers.

Lensed LEDs are frequently used in communication applications. A micrograph of an LED with a monolithically integrated lens is shown in **Figure 35.8**. The light-emitting region of the LED is 20 μm and the lens shown in the figure has a diameter of about 80 μm .

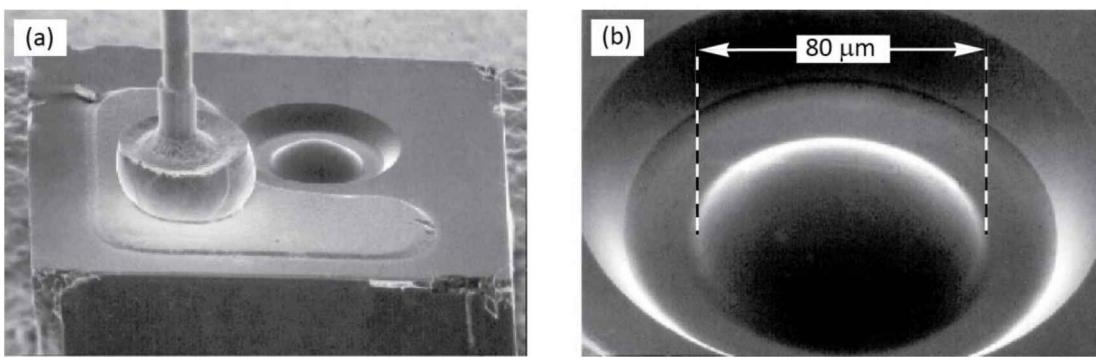


Fig. 35.8: (a) Communication LED chip with integrated lens. (b) Micrograph of the lens etched by a photochemical process into the GaAs substrate (AT&T ODL product, 1995).

Exercise: Coupling efficiency of a fiber coupled to an LED with a lens. Consider an LED circular emission region with diameter 20 μm coupled to a silica multimode fiber with $NA = 0.2$ and a core diameter of 62.5 μm. The LED emits a power of 1 mW into the hemisphere lying above the planar LED surface. For simplicity, assume that the LED emission intensity is independent of the emission angle. What is the maximum power that can be coupled into the multimode fiber?

Solution: Improved coupling can be obtained by imaging the LED emission region on to the core of the optical fiber. For maximum coupled power, a convex lens with magnification $M = 62.5 \mu\text{m} / 20 \mu\text{m} = 3.125$ can be used. Using the lens, the acceptance angle of the fiber is increased from $\theta_{air} = 11.5^\circ$ to $\theta_{LED} = 35.9^\circ$. The solid angle defined by the LED acceptance angle θ_{LED} is given by $\Omega = 1.19$. Since the LED emits 1 mW into the entire hemisphere (with solid angle $\Omega = 2\pi$), the power coupled into the fiber is given by 0.189 mW.

35.7 Free-space optical communication

Free-space optical communication (Carruthers, 2002; Heatley *et al.*, 1998; Kahn and Barry, 2001) is suitable for low to medium bit rates. The most common application of free-space optical communication is the remote control of consumer appliances such as stereos and television sets. Other applications are the remote control of automobile door locks and the cordless interface between computers and peripheral devices such as a mouse, keyboard, and printer.

Free-space optical communication is limited to line-of-sight applications since obstacles such as walls and floors will block the path of light. Furniture may also block the path of light. However, a light beam may be reflected from the ceiling so that communication may still be possible even if there is no direct line of sight connection between the optical transmitter and the receiver.

The wavelength of choice for free-space optical communication is the near infrared. GaAs LEDs emitting with good efficiency are readily available. Infrared light is preferred over visible light sources because the former does not provide a distraction to anyone near the optical

transmitter.

Eye safety considerations limit the maximum power of optical transmitters. At a wavelength of 870 nm, the optical power is limited to typically a few mW. Other wavelengths, such as 1500 nm, allow for higher optical powers. The 1500 nm wavelength range is termed “eye safe”, since the cornea absorbs 1500 nm light, thus preventing light from reaching the sensitive retina. The wavelength 1500 nm thus allows for higher optical powers than 870 nm sources.

If we restrict our considerations to small distances, the transmission medium air can be considered to be totally lossless. However, the optical signal strength decreases for uncollimated light beams due to spatial divergence. For *isotropic emitters*, the intensity decreases with the square of the radius, i.e.

$$I = P / (4\pi r^2) \quad (35.19)$$

where P is the optical power emitted by the source and r is the distance from the source. The decrease in intensity thus has a very different dependence compared with the intensity in fiber communication.

The rapidly decreasing intensity limits the maximum range of optical communication. Collimated light beams can overcome this problem. Transmission distances of several km are possible without significant loss provided that atmospheric conditions are good, i.e. in the absence of fog or precipitation. Semiconductor lasers are used for such collimated transmission systems due to the ability to form collimated beams with very little spatial dispersion.

Multipath distortion or **multipath time delay** severely limits the data rate in free-space optical communication systems. A schematic illustration of multipath distortion is shown in **Figure 35.9**. A light beam emanating from the optical transmitter may take several different paths from the transmitter to the receiver. This is especially true for rooms with high-reflectivity surfaces such as white ceilings, walls, or mirrors.

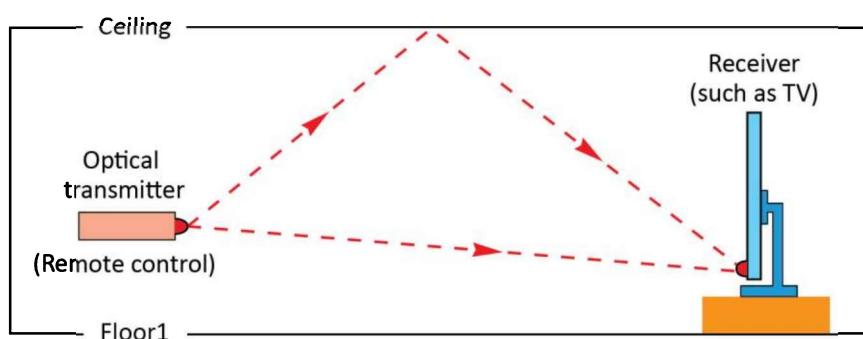


Fig. 35.9: Illustration of multipath distortion of a free-space optical signal, which limits the maximum data rate of the signal.

As an approximate rule, the longest path is assumed to be twice as long as the shortest path between the transmitter and the receiver. This approximate rule leads to a multipath distortion time delay of

$$\Delta\tau = L/c \quad (35.20)$$

where L is the transmitter–receiver distance and c is the speed of light in vacuum. The maximum data rate is then limited to

$$f_{max} \approx 1/\Delta\tau. \quad (35.21)$$

For a room size of 5 m, the multipath delay is about $\Delta\tau = 17$ ns. Thus, the data rate will be limited to about 60 MHz.

Another limitation of free-space optical communication is the detector noise. Sunlight and incandescent light sources have strong emission in the infrared. Thus, a large DC photocurrent is generated in the detector, especially under direct sunlight conditions. The detector noise can be reduced by limiting the bandwidth of the receiver system. By reducing the bandwidth of the receiver system, and thereby also the system data rate, the detector noise is reduced, since the noise spectrum is much wider than the system bandwidth.

The detector noise due to ambient light sources can also be reduced by using optical band-pass filters, long-wavelength-pass filters, or a combination of both filters. Such filters prevent unwanted ambient light from reaching the detector.

35.8 References

- Carruthers J. B. "Wireless infrared communications" in *Wiley Encyclopedia of Telecommunications* edited by J. G. Proakis (John Wiley and Sons, New York, 2002)
- Heatley D. J. T., Wisely D. R., Neild I., and Cochrane P. "Optical wireless: the story so far" *IEEE Communications Magazine* **36**, 72 (1998)
- Hecht J. *Understanding Fiber Optics* (Prentice Hall, Upper Saddle River, New Jersey, 2001)
- Kahn J. M. and Barry J. R. "Wireless infrared communications" *Proceedings of the IEEE* **85**, 265 (2001)
- Keiser G. *Optical Fiber Communications* 3rd edition (McGraw-Hill, New York, 1999)
- Kibler T., Poferl S., Böck G., Huber H.-P., and Zeeb E. "Optical data buses for automotive applications" *IEEE Journal of Lightwave Technology* **22**, 2184 (2004)
- Mynbaev D. K. and Scheiner L. L. *Fiber-Optic Communications Technology* (Prentice Hall, Upper Saddle River, New Jersey, 2001)
- Neyer A., Wittmann B., and Jöhnck M. "Plastic-optical-fiber-based parallel optical interconnects" *IEEE Journal of Selected Topics in Quantum Electronics* **5**, 193 (1999)

36

36 – Communication: Communication LEDs

LEDs can be used for either free-space communication or for fiber communication applications. ***Free-space communication applications*** include the remote control of appliances such as television sets and stereos, and data communication between a computer and peripheral devices. LEDs used in optical ***fiber communication applications*** are suited for distances of a few km and bit rates up to about 1 Gbit/s. Most fibers used with LEDs are multimode (step-index and graded-index) fibers. However, some applications employ single-mode fibers.

36.1 LEDs for free-space communication

Free-space communication LEDs are commonly made with GaAs or GaInAs active regions and are grown on GaAs substrates. The GaInAs layer is pseudomorphic, i.e. sufficiently thin that it is coherently strained, and no dislocations are generated. The emission wavelength of GaAs and coherently strained GaInAs LEDs is limited to wavelengths in the IR ranging from 870 nm (for GaAs active regions) to about 950 nm (for GaInAs active regions).

The wavelength of free-space communication LEDs is in the infrared so that the light emitted is invisible to the human eye and does not distract. Since free-space communication usually involves transmission distances of less than 100 m, the transmission medium (air) can be considered, to a good approximation, to be lossless and dispersionless.

The ***total light power*** is an important figure of merit in free-space communication LEDs, so that the internal efficiency and the extraction efficiency need to be maximized. The emission pattern (far field) is another important parameter. The emission pattern should be wide to reduce the requirement of aiming the emitter towards the receiver.

36.2 LEDs for fiber-optic communication

Light-emitting diodes are the light source of choice for local area low and medium bit rate optical communication. Owing to the spontaneous emission lifetime of about 1 ns in highly excited semiconductors, the maximum bit rates attainable with LEDs are limited to rates ≤ 1 Gbit/s.

Thus, multi-Gbit/s transmission rates are not feasible with LED sources. Transmission rates of several hundred Mbit/s are fully sufficient for many local-area communication applications.

LEDs used for fiber communication applications are very different from LEDs used in lamp applications. In communication LEDs, high coupling efficiency of the light emanating from the LED to the fiber is essential. Only the light emanating from *one* surface, namely the surface of the LED abutting the optical fiber, can be coupled into the fiber. Therefore, it is essential to maximize the light emission from *one* surface of the LED. In LEDs used for fiber communication, the *power emitted per unit area* is a useful figure of merit. This is in contrast to free-space communication LEDs where the *total power* emitted by the LED is the appropriate figure of merit.

In order to maximize the LED-to-fiber coupling efficiency, the light-emitting spot should be smaller than the core diameter of the optical fiber. Typical are circular emission regions with diameters of 20~50 µm for devices used with multimode fibers. Silica multimode fibers have typical core diameters of 50~100 µm.

Plastic optical fibers, on the other hand, can have core diameters as large as 1 mm. Accordingly, LEDs with larger light-emitting areas can be used with plastic fibers.

36.3 Surface-emitting Burrus-type communication LEDs emitting at 870 nm

One of the first LED structures suitable for optical fiber communication was developed by Charles Burrus of AT&T Bell Laboratories (Burrus and Miller, 1971; Saul *et al.*, 1985). The Burrus-type LED, shown in *Figure 36.1* (a), consists of a double heterostructure with a GaAs active region that is grown lattice-matched on a GaAs substrate. The original structure proposed and demonstrated by Burrus was just a homojunction. However, such homojunction LEDs are no longer in use due to unwanted light reabsorption in the layers adjoining the active region.

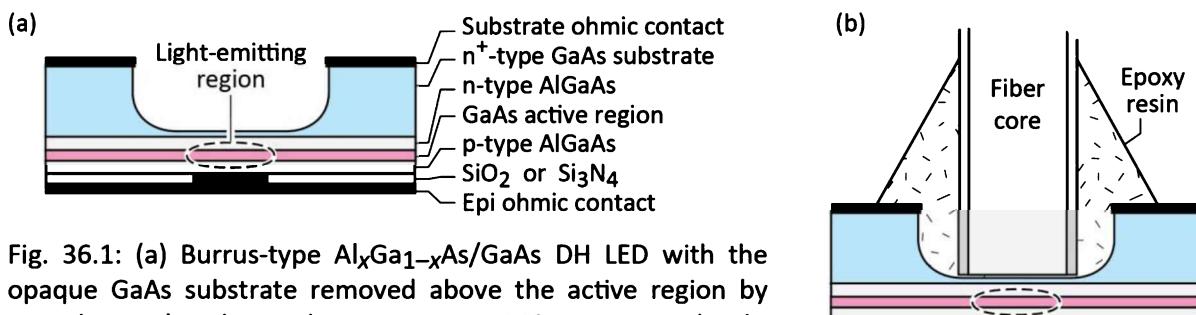


Fig. 36.1: (a) Burrus-type Al_xGa_{1-x}As/GaAs DH LED with the opaque GaAs substrate removed above the active region by wet chemical etching. The Burrus-type LED is mounted substrate-side up. (b) Optical fiber coupled to a Burrus-type LED.

The Burrus-type structure has several features making it suitable for communication applications. Firstly, the light is generated in an active region of small lateral extent. The lateral size of the active region is determined by the p-type ohmic contact size of the LED. If the p-type confinement layer is sufficiently thin, no current spreading occurs, and the lateral extent of the light-emitting region is the same as the contact size. By design, the lateral extent of the active region is smaller than the core diameter of the optical fiber to maximize coupling efficiency. Secondly, the opaque GaAs substrate is partially removed, as shown in **Figure 36.1**, to reduce absorption of light in the substrate. The substrate can be thinned by chemically assisted mechanical polishing to about 150 μm thickness followed by a wet chemical etch. If the GaAs membrane created by etching is too thin, it tends to break easily during fiber coupling. On the other hand, a thick membrane reduces radiative efficiency due to absorption.

Figure 36.1(b) shows a typical coupling arrangement of the Burrus-type LED to an optical fiber. Epoxy is used to permanently attach the fiber to the LED. Note that the p-type contact also serves as a sink for the heat generated in the active region. The heat sink is particularly efficient if the p-type contact includes a thick layer of electroplated gold.

36.4 Surface-emitting communication LEDs emitting at 1300 nm

Communication LEDs emitting at 1300 nm, when used with graded-index silica fibers, are suited for high-speed data transmission. A communication LED structure emitting at 1300 nm is shown in **Figure 36.2** (Saul *et al.*, 1985). The light is emitted through the InP substrate, which is transparent at 1300 nm. Accordingly, the device is mounted epi-side down in the LED package.

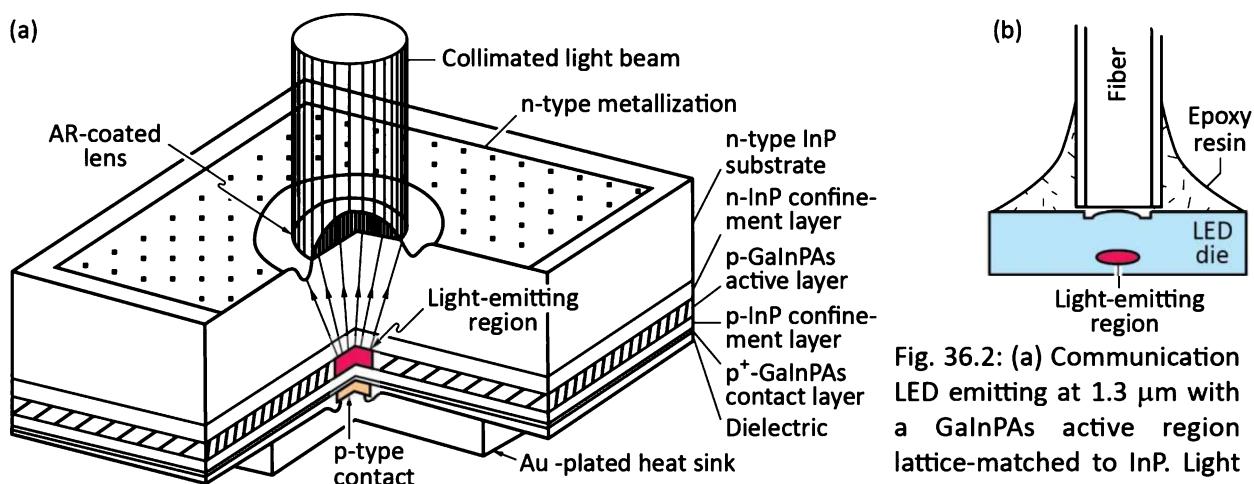


Fig. 36.2: (a) Communication LED emitting at 1.3 μm with a GaInPAs active region lattice-matched to InP. Light generated in the active re-

gion is transmitted through the InP substrate. The lateral dimension of the light-emitting region is defined by current injection through the p-type contact with diameter of 20 μm . An AR-coated lens, etched into the substrate, collimates the light beam. (b) LED-to-fiber coupling using resin.

The device has a GaInPAs active region lattice matched to the InP substrate. No current-spreading layer is used so that the light-emitting region is located directly above the p-type contact. At the light-exit point, an optical lens collimates the light to improve the LED-to-fiber coupling efficiency. The lens is integral to the LED chip and is etched into the InP substrate by a photo-electrochemical process (Ostermayer *et al.*, 1983).

The emission spectra measured from the surface and from the edge of the device are shown in **Figure 36.3**. The two emission spectra are markedly different. The spectrum emitted towards the edge has a smaller spectral width due to self-absorption. During self-absorption, predominantly photons of the high-energy part of the spectrum are reabsorbed.

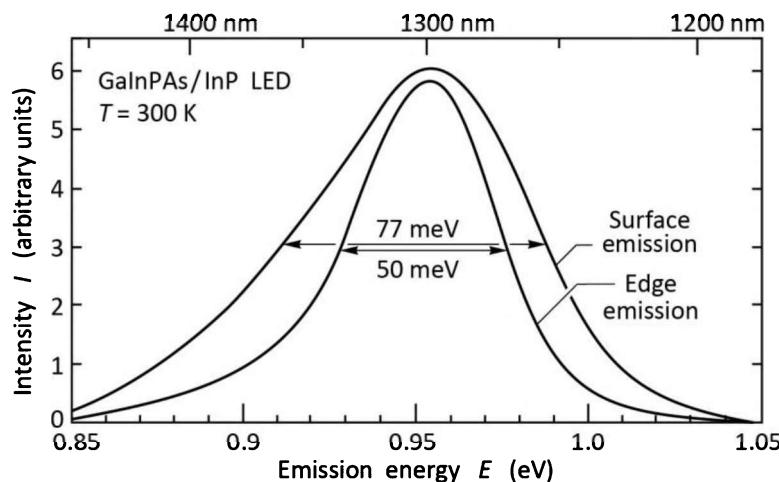


Fig. 36.3: Emission spectra along the edge and surface of a GaInPAs/InP communication LED emitting at 1300 nm. The spectrum emitted along the edge of the LED is narrower due to self-absorption.

A scanning electron micrograph of a GaInPAs/InP LED wafer is shown in **Figure 36.4**. The surface of contacts displays roughness due to the annealing process that follows contact deposition. To reduce Fresnel reflection losses at the semiconductor-air boundary, the lens is anti-reflection coated.

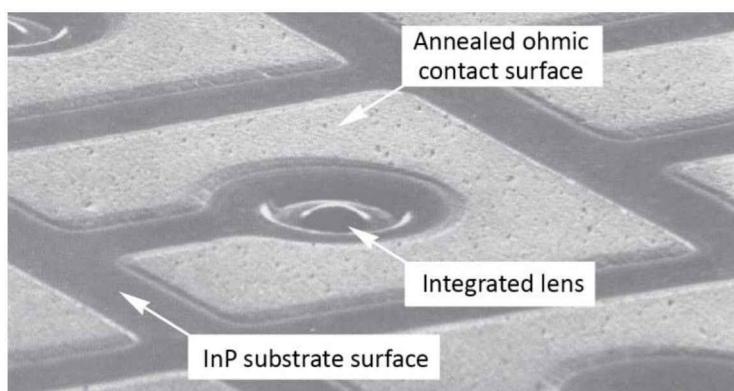


Fig. 36.4: GaInPAs communication LED grown on an InP wafer. The LED has an integrated semiconductor lens. The ohmic contact metal surface has a textured appearance due to the annealing process (after Ostermeyer *et al.*, 1983).

Note that the wavelength 1500 nm is of interest for long-distance silica fiber communication. Long-distance communication fibers must have a small core diameter to be single mode to avoid modal dispersion. Due to low LED-to-single-mode-fiber coupling efficiency and due to the broad emission band of LEDs, LEDs are not used in long-distance silica fiber communication.

36.5 Communication LEDs emitting at 650 nm

LEDs emitting at 650 nm are useful for plastic optical fiber communication. These fibers have a loss minimum and relatively low dispersion at 650 nm. Communication LEDs emitting in the 600~650 nm range are based on the $(\text{Al}_x\text{Ga}_{1-x})_{0.5}\text{In}_{0.5}\text{P}$ material system just as for 650 nm visible-spectrum LED lamps.

Typical LED structures used for 650 nm plastic fiber communication applications are shown in **Figures 36.5 (a) and (b)**. Both LEDs are top-emitting devices due to the opaqueness of the GaAs substrate. The LEDs employ current-blocking layers that guide the current to the active region. Light emitted from this region is not obstructed by the metallic ring contact.

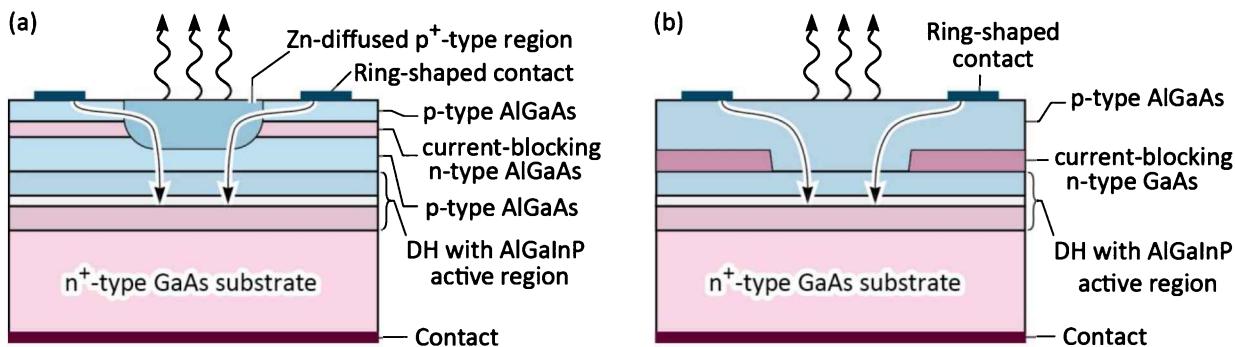


Fig. 36.5: 650 nm AlGaNp/GaAs LEDs for plastic fiber communications using a current-blocking layer fabricated by (a) p⁺-type diffusion and (b) epitaxial re-growth. Current is funneled to center of the active region where emitted light is not obstructed by top contact.

The figure shows two types of **current-blocking regions**. **Figure 36.5 (a)** shows a current-blocking p-n-p structure in the transparent top layer. A Zn-diffused region in the center region of the LED overcompensates donors in the current-blocking layer making it p-type. As a result, current flows only through the Zn-diffused layer. **Figure 36.5 (b)** shows a current-blocking layer fabricated by epitaxial regrowth. After growth of an n-type blocking layer on top of the p-type confinement layer, the wafer is taken out of the growth system for patterning and etching. Subsequently, the

wafer is re-introduced into the growth system for resumption of epitaxial growth. Epitaxial regrowth is a more expensive process and usually the device yield decreases for structures employing regrowth. The transparent window layers can consist of AlGaAs, as shown in *Figure 36.5*, GaP, or another transparent semiconductor.

Resonant-cavity LEDs (RCLEDs) have several advantages over conventional LEDs including high brightness and a narrow spectral width (Schubert *et al.*, 1992, 1994; Schubert and Hunt, 1999). An RCLED emitting at 650 nm is shown in *Figure 36.6* (Streubel and Stevens, 1998; Streubel *et al.*, 1998; Whitaker, 1999; Mitel, 2000). As in the two previous structures, the top contact is ring-shaped and the current is guided to the center opening of the ring. An ion-implanted region is used for current blocking. Hydrogen and oxygen have been employed to render the implanted region insulating. Oxygen implants are more stable than hydrogen implants since small hydrogen atoms tend to easily diffuse out of the semiconductor at moderate annealing temperatures.

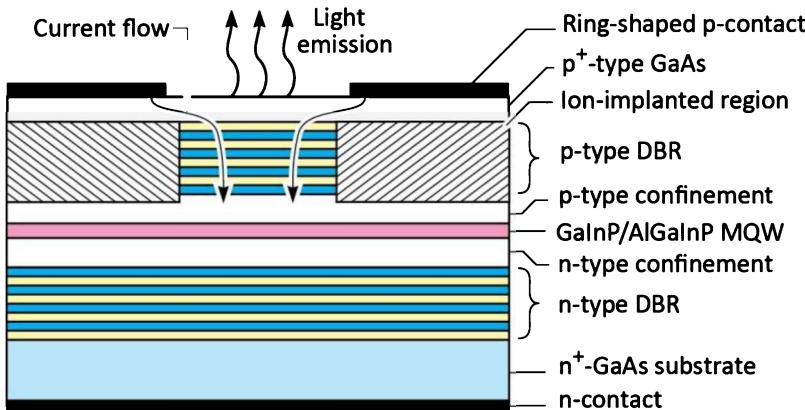


Fig. 36.6: RCLED emitting at 650 nm. Two distributed Bragg reflectors (DBRs) form the cavity. The active region is a GaInP/AlGaInP multiple-quantum well (MQW) structure (after Whitaker, 1999).

The emission spectra of an RCLED and a conventional LED are shown in *Figure 36.7*. The spectra shown are the spectra of the fiber-coupled light. Inspection of the spectra reveals two features. Firstly, the RCLED exhibits a higher spectral purity thereby reducing chromatic dispersion. Secondly, the fiber-coupled intensity of the RCLED is higher due to the more directed emission pattern. High-speed transmission of 250 Mbit/s over plastic optical fibers has been demonstrated with 650 nm RCLEDs (Streubel and Stevens, 1998).

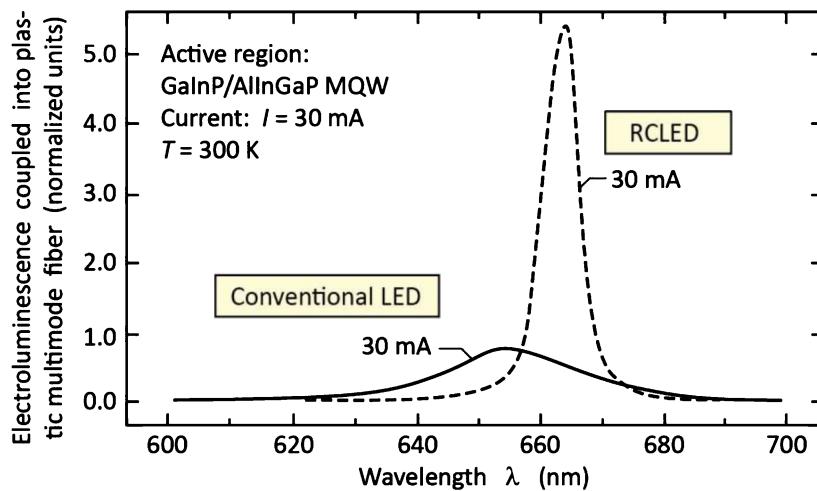


Fig. 36.7: Fiber-coupled ($NA = 0.275$) emission spectrum of RCLED and conventional LED at injection current of 30 mA. The microcavity effect of the RCLED enhances the emission intensity and reduces the emission linewidth, especially for low NA fibers (after Whitaker, 1999).

36.6 Edge-emitting superluminescent diodes (SLDs)

Edge-emitting LEDs are motivated by the need for high-brightness LEDs that allow for high-efficiency coupling to optical fibers. Edge-emitting LEDs comprise an optical waveguide region that guides light emitted along the waveguide by total internal reflection.

Superluminescent light-emitting diodes or **superluminescent diodes** (SLDs) are broad-band high-intensity emission sources that emit incoherent light. Incoherent light does not result in “speckle patterns” obtained from coherent light sources such as lasers. SLDs are suitable as communication devices used with single-mode fibers and also as high-intensity light sources for the analysis of optical components (Liu, 2000).

Light is guided in the **core region** of the waveguide. Total internal reflection occurs at the boundaries between the core region and the upper and lower **cladding layers** as shown in **Figure 36.8**. In order to make waveguiding possible, the core layer must have a higher refractive index than the cladding layers. Photons emitted with a sufficiently small angle of incidence at the core–cladding interface will be guided by the waveguide, as indicated in **Figure 36.8**.

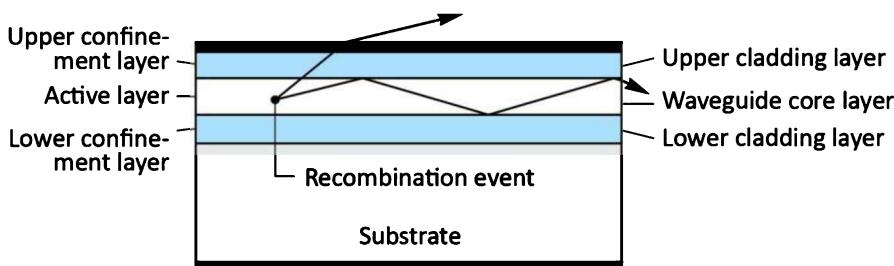


Fig. 36.8: Waveguide geometry guiding light rays with low angles of incidence in the core layer.

Since the light is guided by the waveguide, the light intensity emitted by the device linearly increases with the length of the waveguide. Thus, increasing the length of an edge-emitting LED allows one to obtain a higher light-output intensity. However, the electrical current required to drive the LED also increases with the stripe length.

Superluminescent diodes are edge-emitting LEDs that are pumped at such high current levels that ***stimulated emission*** occurs. In the stimulated emission process, one photon stimulates the recombination of an electron–hole pair and the emission of another photon. The photon created by the stimulated emission process has the same propagation direction, phase, and wavelength as the stimulating photon. Thus, SLDs have greater coherence compared with LEDs. In the stimulated emission regime, spontaneous emission towards the top surface of the LED is reduced and emission into waveguide modes is enhanced.

SLDs are quite similar to semiconductor laser structures with one important difference: SLDs lack the optical feedback provided by the reflectors of a semiconductor laser. Two typical SLD structures are shown in *Figure 36.9*. The SLD structure shown in *Figure 36.9* (a) has a reflective back-side reflector facet; the front-side facet, however, is coated with an anti-reflection (AR) coating. To prevent lasing, the front-side facet must have a reflectivity of $\leq 10^{-6}$ (Liu, 2000; Saul *et al.*, 1985). Exceeding the maximum allowable reflectivity can result in unwanted lasing of the device. Owing to the high-quality AR coating requirement, SLDs with an AR coating are expensive to manufacture.

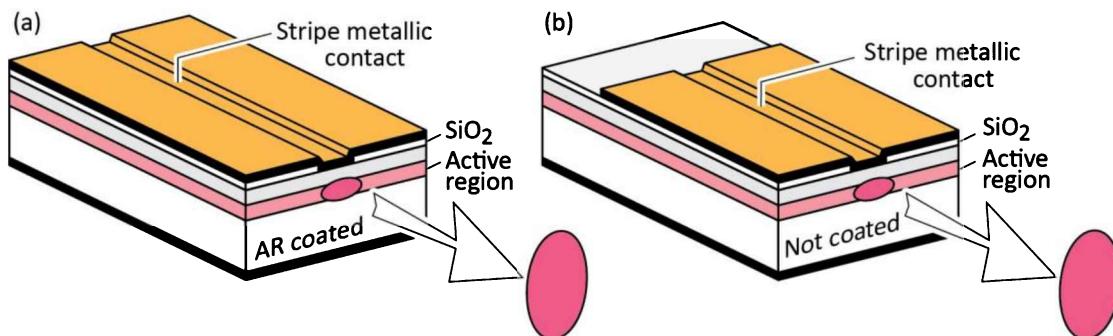


Fig. 36.9: Common structures of superluminescent diodes (SLDs). (a) SLD with cleaved facets coated with anti-reflection (AR) coatings. (b) SLD with cleaved, reflecting facets and stripe contact injecting current over the partial length of the device.

A lower-cost alternative SLD structure is shown in *Figure 36.9 (b)*. This structure uses a ***lossy region*** near the back-side facet of the diode. The lossy region is not covered by the top metal contact and thus is not pumped by the injection current. Practically no feedback occurs if the length of the lossy region is much longer than the absorption length of the core region, i.e.

$$\text{Length of lossy region} \gg \alpha^{-1}, \quad (36.1)$$

where α is the absorption coefficient of the core region. The absorption coefficient in III–V semiconductors near the band edge is $\alpha \approx 10^4 \text{ cm}^{-1}$. Thus, for lossy-region lengths exceeding several tens of micrometers, the optical feedback from the back-side facet is negligibly small.

In addition to absorption losses, diffraction losses occur in the region not pumped by the electrical current. Gain guiding occurs in the region injected by the electrical current but not in the lossy region. Thus, both absorption and diffraction losses prevent this type of SLD from lasing.

Emission spectra of an LED, SLDs and a laser are shown in *Figure 36.10*. The LED has a broad spontaneous emission spectrum. The spectrum of an SLD with a residual small facet reflectivity exhibits periodic oscillations in the emission spectrum due the Fabry–Perot cavity enhancement. An ideal SLD has a smooth spectrum and does not exhibit any oscillations. The spectral width of SLDs is narrower than that of LEDs due to increased coherence caused by stimulated emission. Also shown is the spectrum of a Fabry–Perot laser with several laser modes.

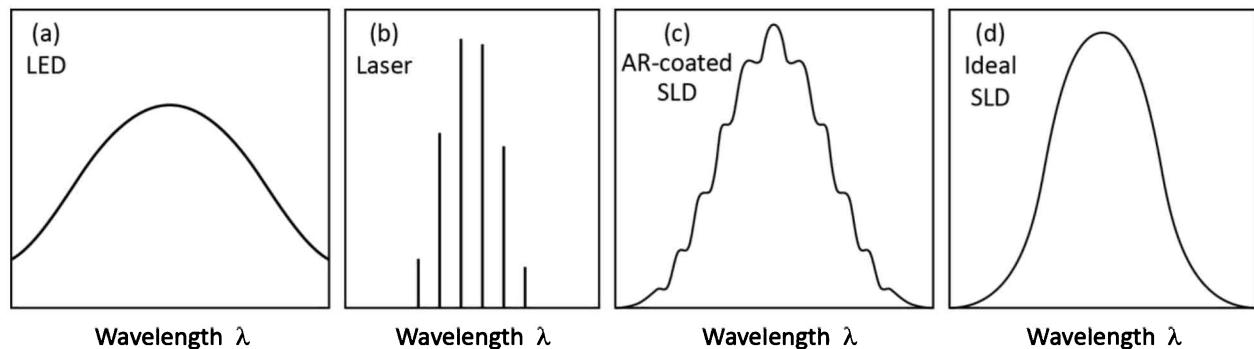


Fig. 36.10: Spectrum of (a) regular LED with a theoretical linewidth of $1.8kT$, (b) multimode semiconductor laser, (c) superluminescent diode (SLD) fabricated by AR coating of a multimode laser, (d) ideal SLD with linewidth less than kT (after Liu, 2000).

A comparison of the $L-I$ curves of an LED, an SLD, and a laser is shown in *Figure 36.11*. Surface-emitting LEDs with a small light-emitting region diameter, tend to have ***sublinear*** $L-I$ characteristics. At high injection current densities, the small active volume of surface-emitting

LEDs is swamped with carriers leading to saturation. Edge-emitting LEDs operating in the spontaneous emission regime have *linear L–I* characteristics, as expected for ideal LEDs. SLDs have a *superlinear L–I* characteristic due to stimulated emission. In the stimulated emission regime, an increasing number of photons are guided by the waveguide. The number of photons emitted into waveguide modes increases with injection current as stimulated emission becomes dominant. As for SLDs, semiconductor lasers have superlinear emission characteristics. However, the *L–I* curve of lasers exhibits a more distinct threshold than that of SLDs.

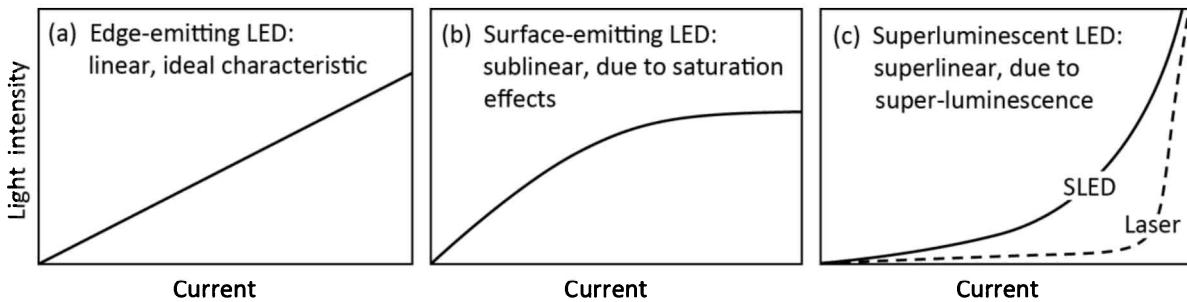


Fig. 36.11: Light-versus-current (*L–I*) characteristic of different LEDs. (a) Edge-emitting LED with little or no saturation effects. (b) Surface-emitting LED with small active area exhibiting saturation effects due to carrier overflow. (c) Superluminescent LED. Also shown is the *L–I* characteristic of a laser that exhibits a distinct threshold current.

36.7 References

- Burrus C. A. and Miller B. I. "Small-area double heterostructure AlGaAs electroluminescent diode sources for optical fiber transmission lines" *Optics Communications* **4**, 307 (1971)
- Liu Y. "Passive components tested by superluminescent diodes" February issue of *WDM Solutions* p. 41 (2000)
- Mitel Corporation. RCLEDs were first manufactured by the Mitel Corporation. RCLED emitting at 650 nm has Part # 1A466 (2000). Large-scale production of RCLEDs was started in 2001 by Osram Optosemiconductors Corporation, see Wirth R., Huber W., Karnutsch C., and Streubel K. "High-efficiency resonant-cavity LEDs emitting at 650 nm" *Compound Semiconductors* **8**, 49 (2002)
- Ostermayer Jr. F. W., Kohl P. A., and Burton R. H. "Photoelectrochemical etching of integral lenses on GaInPAs/InP light-emitting diodes" *Applied Physics Letters* **43**, 642 (1983)
- Saul R. H., Lee T. P., and Burrus C. A. "Light-emitting-diode device design" in *Lightwave Communications Technology* edited by W. T. Tsang, *Semiconductors and Semimetals* **22** Part C, (Academic Press, San Diego, 1985)
- Schubert E. F. and Hunt N. E. J. "Enhancement of spontaneous emission in microcavities" in *Vertical Cavity Surface Emitting Lasers* edited by C. Wilmsen, H. Temkin, and L. A. Coldren (Cambridge University Press, Cambridge, UK, 1999)
- Schubert E. F., Wang Y.-H., Cho A. Y., Tu L.-W., and Zydzik G. J. "Resonant-cavity light-emitting diode" *Applied Physics Letters* **60**, 921 (1992)
- Schubert E. F., Hunt N. E. J., Micovic M., Malik R. J., Sivco D. L., Cho A. Y., and Zydzik G. J. "Highly efficient light-emitting diodes with microcavities" *Science* **265**, 943 (1994)

- Streubel K. and Stevens R. "250 Mbit/s plastic fibre transmission using 660 nm resonant cavity light emitting diode" *Electronics Letters* **34**, 1862 (1998)
- Streubel K., Helin U., Oskarsson V., Backlin E., and Johansson A. "High brightness visible (660 nm) resonant-cavity light-emitting diode" *IEEE Photonics Technology Letters* **10**, 1685 (1998)
- Whitaker T. "Resonant cavity LEDs" *Compound Semiconductors* **5** (4), 32 (May 1999)

37

37 – Communication: LED modulation characteristics

LEDs are the most commonly used light source for local-area communication systems operating from very short (< 1 m) to medium distances (5 km). Typical bit rates are tens of Mbit/s up to about 1 Gbit/s.

LEDs are non-linear devices and as such the series resistance, shunt resistance, and capacitance depend strongly on the applied voltage. A thorough analysis must take into account these non-linearities. However, much can be learned by considering the LED as a linear device, even though some important aspects of the LED modulation behavior cannot be inferred from the linear model. In this chapter, we first analyze LEDs as linear devices and subsequently discuss modulation characteristics, including some non-linear modulation characteristics.

37.1 Rise and fall times, 3-dB frequency, and bandwidth in linear circuit theory

A simple RC circuit is shown in *Figure 37.1 (a)*. When subjected to a step-function input pulse, the output voltage increases according to

$$V_{out}(t) = V_0 [1 - \exp(-t/\tau_1)] \quad (37.1)$$

where $\tau_1 = RC$ is the time constant of the RC circuit. When the input voltage returns to zero, the output voltage decreases according to

$$V_{out}(t) = V_0 \exp(-t/\tau_2) . \quad (37.2)$$

For an RC circuit, it is $\tau_1 = \tau_2$. The rise and fall times are defined as the time difference between the 10% and 90% points of the voltage, as shown in *Figure 37.1 (b)*. The rise and fall times of the signal are related to the time constants τ_1 and τ_2 by

$$\tau_r = (\ln 9) \tau_1 \approx 2.2 \tau_1 \quad \text{and} \quad \tau_f = (\ln 9) \tau_2 \approx 2.2 \tau_2 . \quad (37.3)$$

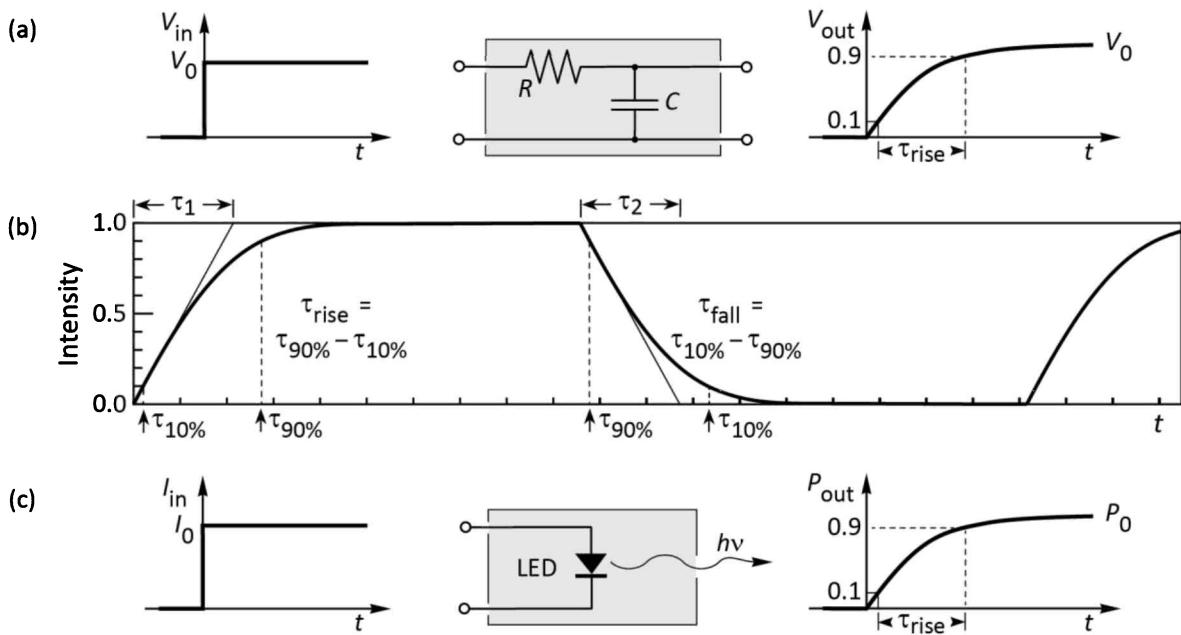


Fig. 37.1: (a) System response of a linear RC system with rise time τ_{rise} . (b) Rise and fall time of a signal with exponential time dependence having time constants τ_1 and τ_2 . (c) Light output power as a function of time for an LED with a rise time of τ_{rise} .

The *voltage* transfer function $H(\omega)$ is given by

$$H(\omega) = (1 + i\omega\tau)^{-1}. \quad (37.4)$$

The **bandwidth** of the system, Δf , corresponds to the frequency at which the power transmitted through the system is reduced to half of its low-frequency value. This condition can be written as $|H(2\pi f)|^2 = 0.5$. Thus, the bandwidth of the RC circuit is given by

$$\Delta f = f_{3\text{dB}} = \frac{1}{2\pi\tau} = \frac{\ln 9}{2\pi\tau_r} = \frac{\ln 9}{\pi(\tau_r + \tau_f)} \approx \frac{0.70}{(\tau_r + \tau_f)}. \quad (37.5)$$

The bandwidth is also called the **3-dB frequency**, since the power transmitted at this frequency is reduced by 3 dB compared with its low-frequency value.

Next, consider an LED with a rise time τ_r as illustrated in **Figure 37.1 (c)**. As a step-function input current is applied to the LED, the optical output power increases according to

$$P_{out}(t) = P_0 [1 - \exp(-t/\tau_r)]. \quad (37.6)$$

In analogy to Eqns. (37.1) and (37.4), the *power* transfer function is given by

$$H_{LED}^2(\omega) = (1 + i\omega\tau)^{-1}. \quad (37.7)$$

The absolute value of the power transfer function is reduced to half at the 3-dB frequency of the LED. Thus, the 3-dB frequency of an LED is given by

$$\Delta f = f_{3\text{dB}} = \frac{\sqrt{3}}{2\pi\tau} = \frac{\sqrt{3}\ln 9}{2\pi\tau_r} = \frac{\sqrt{3}\ln 9}{\pi(\tau_r + \tau_f)} \approx \frac{1.2}{(\tau_r + \tau_f)}. \quad (37.8)$$

Exercise: Derivation of equations. Derive Eqns. (37.3), (37.4), (37.5) and (37.8).

Solution:

Derivation of Eqn. (37.3): Consider an exponential decay with time constant τ , i.e. $\exp(-t/\tau)$. Assume that the time at which the amplitude has decreased to 90% and 10% of its maximum value are $t_{90\%}$ and $t_{10\%}$, respectively. Accordingly

$$e^{-t_{90\%}/\tau} = 0.9 \quad \text{and} \quad e^{-t_{10\%}/\tau} = 0.1.$$

Then

$$\frac{t_{90\%}}{\tau} - \frac{t_{10\%}}{\tau} = -\ln 0.9 + \ln 0.1 = \ln 9 \quad \text{or} \quad t_{90\%} - t_{10\%} = \tau \ln 9$$

what was to be shown. An analogous consideration applies to the rise time.

Derivation of Eqn. (37.4): The voltage transfer function of an RC voltage-divider circuit is given by

$$H(\omega) = \frac{V_{out}}{V_{in}} = \frac{I(i\omega C)^{-1}}{IR + I(i\omega C)^{-1}} = \frac{1}{1 + i\omega RC} = \frac{1}{1 + i\omega\tau}$$

where $\tau = RC$; this is what was to be shown.

Derivation of Eqn. (37.5): Using $|H(\omega_{3\text{dB}})|^2 = |(1 + i\omega_{3\text{dB}}\tau)^{-1}|^2 = 1/2$, and

$$|H(\omega)|^2 = \left| \frac{A}{A + iB} \right|^2 = \frac{|A|^2}{|A + iB|^2} = \frac{A^2}{A^2 + B^2} = \frac{1}{1 + \omega^2\tau^2}$$

we obtain $\omega_{3\text{dB}} = \tau^{-1} = (RC)^{-1}$ or $f_{3\text{dB}} = (2\pi\tau)^{-1}$, what is what was to be shown.

Derivation of Eqn. (37.8): Using $|H(\omega_{3\text{dB}})|^2 = |(1 + i\omega_{3\text{dB}}\tau)^{-1}|^2 = 1/2$, and

$$|H(\omega)|^2 = \left| \frac{A}{A + iB} \right|^2 = \frac{A}{\sqrt{A^2 + B^2}} = \frac{1}{\sqrt{1 + \omega^2\tau^2}}$$

we obtain

$$\frac{1}{2} = \frac{1}{\sqrt{1 + \omega_{3\text{dB}}^2\tau^2}}.$$

Solving for $\omega_{3\text{dB}}$ yields $\omega_{3\text{dB}} = 3^{1/2}/\tau$ or $f_{3\text{dB}} = 3^{1/2}/(2\pi\tau)$, what is what was to be shown.

Exercise: Rise and fall time and 3-dB frequency. Consider an LED with a rise time of 1.75 ns. Assume that the fall time of the LED is identical to the rise time. What is the 3-dB frequency of the device? Give the physical reasons as to why Eqn. (37.8) gives only an *approximate* value of the 3-dB frequency.

Solution: A 3-dB frequency of 343 MHz is expected on the basis of Eqn. (37.8). In practice, the 3-dB frequency can be lower or higher than the calculated value since the rise and fall are frequently not exponential. As a practical rule, the numerical factor 1.2 in the numerator of Eqn. (37.8) can vary between 1.0 and 1.5.

37.2 Rise and fall time in the limit of large diode capacitance

In diodes used for solid-state lamp applications, the current-injected p-n junction area is large, sometimes as large as the entire LED die. Such diodes have a *large* capacitance. Denoting the diode capacitance as C and the overall series resistance of the drive circuit and the diode as R , the rise and the fall time of the diode are equal and these times are given by the RC time constant.

In communication LEDs, the current-injected active region is much smaller, so that the spontaneous lifetime rather than the diode capacitance limits the maximum modulation frequency. As a result, communication LEDs can be modulated. Since LEDs do not exhibit strictly exponential changes in power, as predicted by Eqn. (37.6), Eqn. (37.8) is only an approximation.

Consider an LED in which the p-n junction region extends over the entire area of the die. The LED has a small contact area that determines the size of the light-emitting spot. Such LEDs are used for communication applications. At zero bias, the capacitance of the LED is given by the *depletion* capacitance (space-charge capacitance) of the diode. Since the area of the diode is large (e.g. $250 \times 250 \mu\text{m}^2$), the capacitance is large and can amount to 200 to 300 pF.

As the diode is turned on, the current *crowds* in the area below the contact. When the p-n junction is forward biased, the capacitance of the LED is given by the *diffusion* capacitance. The relevant area is, however, not the entire LED-die area but just the region injected with the diode current.

The reduction of the diode capacitance increases the LED modulation bandwidth. The *depletion* capacitance can be reduced by mesa etching. However, the mesa should be larger than the contact size in order to avoid surface recombination effects.

There is no viable way to reduce the *diffusion* capacitance. The diffusion capacitance can be reduced by purposely introducing defects that act as luminescence killers. Such defects reduce the minority carrier lifetime and thereby the diffusion capacitance. Such LEDs can be modulated at several GHz. However, the light output intensity decreases as well so that the overall benefit of such lifetime killers is questionable.

37.3 Rise and fall time in the limit of small diode capacitance

Next, we discuss the rise and fall time in the limit of small diode capacitance. This consideration applies, for example, to surface-emitting communication LEDs that have a small-area active region. Consider an LED driven by a constant current that is switched on at $t = 0$. Electrons are injected into the active region and the carrier concentration builds up. At the same time, the optical output intensity of the LED increases. In the case of the monomolecular recombination model, the light output intensity is directly proportional to the injected minority carrier concentration.

The monomolecular rate equation is given by

$$\frac{I}{e A d} = \frac{dn_{active}}{dt} = \frac{n_{active}}{\tau} \quad (37.9)$$

where n_{active} is the carrier concentration in the active region, A is the current-injected area of the active region, and d is the thickness of the active region. The steady-state current flow of magnitude I causes a steady-state minority carrier concentration $n_{active} = I \tau / (eAd)$. The mean lifetime of the carriers is the spontaneous recombination lifetime τ .

Next, consider that the diode is initially in the “off” state and, starting at $t = 0$, the diode is injected with a constant current I . When the diode is in the “off” state, the minority carrier concentration in the active region is very low and we approximate the concentration with $n_{active} \approx 0$.

Solving the differential Eqn. (37.9) for the initial condition $n_{active} \approx 0$ at $t = 0$ for a constant injection current yields that the carrier concentration in the active region increases according to

$$n_{active}(t) = n_{active} (1 - e^{-t/\tau}) = \frac{I \tau}{e A d} (1 - e^{-t/\tau}). \quad (37.10)$$

The equation reveals that it takes the spontaneous recombination time τ to fill the active region with the steady-state carrier concentration. The light output intensity follows the minority carrier concentration directly. Thus, the rise time is given by the spontaneous recombination time.

A similar consideration applies to the fall time of the diode. Once the diode has been switched off, the decay constant for emission is, of course, the spontaneous recombination lifetime. Thus, the fall time of an LED is given by the spontaneous recombination lifetime.

In the case of an undoped active region, the monomolecular recombination equation no

longer applies and the bimolecular recombination equation must be used to describe the carrier dynamics. Also, the carrier lifetime is no longer a constant. In this case, the *shortest* lifetime, i.e. the lifetime that applies when the carrier concentration is at the *highest* level, can be used.

It should be noted that there are methods to *reduce* both the rise and the fall times *below* the limit of the spontaneous recombination lifetime. The rise time can be reduced by *current shaping*. The fall time can be reduced by *carrier sweep-out*. Both methods will be discussed below.

37.4 Voltage dependence of the rise and fall times

The measurement of the rise time and fall time of an LED is shown in **Figure 37.2** (Schubert *et al.*, 1996). As indicated in the figure, the rise and fall times are measured from the 10~90% values of the optical signal. The photocurrent of a p-n junction photodetector is used in the measurement. It must be ensured in the measurement that the rise and fall time of both the pulse generator and the photodiode are much faster than the LED rise and fall time. The measured rise and fall times include the time constant of the pulse generator, the LED, the detector, detector amplifier circuit, and the oscilloscope. However, the time constant of the LED is the longest and hence the dominant time constant. The time constants shown in **Figure 37.2** are *upper limits* to the true time constants of the LED.

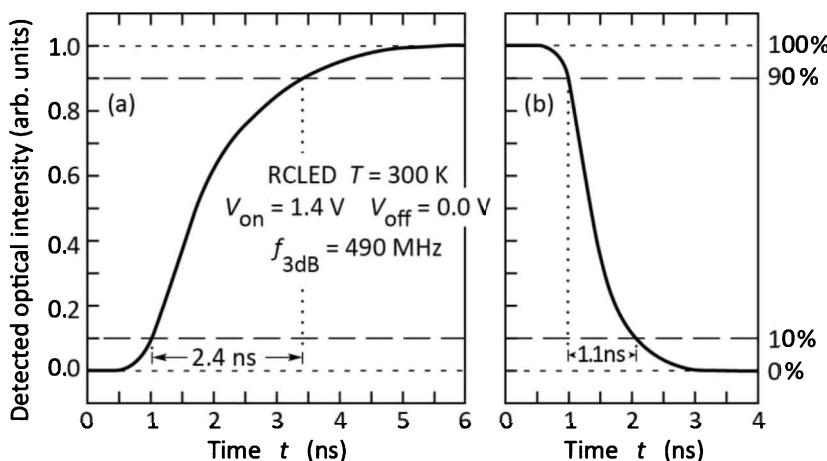


Fig. 37.2: Measured (a) rise time ($\tau_{\text{rise}} = 2.4\text{ ns}$) and (b) fall time ($\tau_{\text{fall}} = 1.1\text{ ns}$) of an RCLED. The photodetector used in the measurement is much faster than the LED, so that the measured times are essentially the rise and fall times of the LED.

Inspection of **Figure 37.2** reveals that the rise time is much longer than the fall time. The large difference between the rise and the fall times displayed in **Figure 37.2** is not expected based on the theoretical model discussed above.

To gain a better understanding of the difference between the rise time and the fall time, the times have been measured as a function of the diode bias conditions. In the “on” state, the diode is biased with a voltage of 1.4 V. However, a range of voltages can be chosen for the “off” state, since a p-n junction diode does not emit light for voltages even slightly below the turn-on voltage.

The experimental results are shown in **Figure 37.3**. Whereas the “on” voltage is kept constant at $V_{on} = 1.4$ V, the “off” voltage is varied from 0 to 1.0 V. Inspection of the figure reveals a strong voltage dependence of the fall time. The voltage dependence is caused by carrier sweep-out of the active region. In contrast, the rise time is practically independent of voltage.

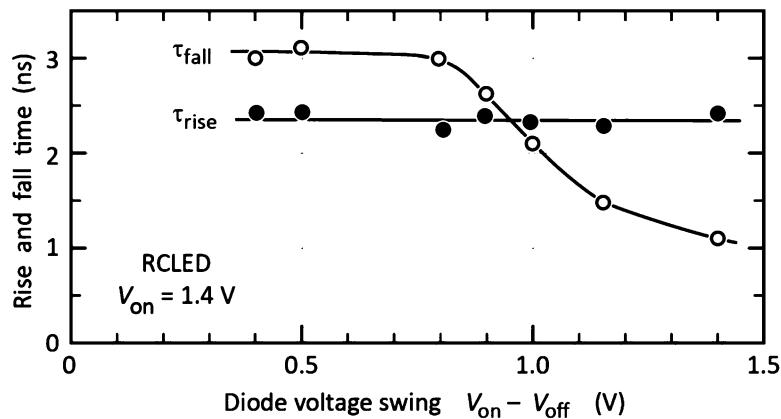


Fig. 37.3: Rise time and fall time as a function of the voltage swing. The fall time of the diode decreases with increasing voltage swing due to sweep-out of carriers out of the active region.

37.5 Carrier sweep-out of the active region

The voltage dependence of the fall time shown in **Figure 37.3** can be explained by voltage-dependent carrier sweep-out of the active region. **Figure 37.4** shows the active region band diagram in the “off” state for large (a) and small (b) voltage swings. For the case of a small voltage swing, carriers essentially remain in the active region until they recombine. As a result, it will take the spontaneous lifetime for the carriers to recombine and the light intensity to decay.

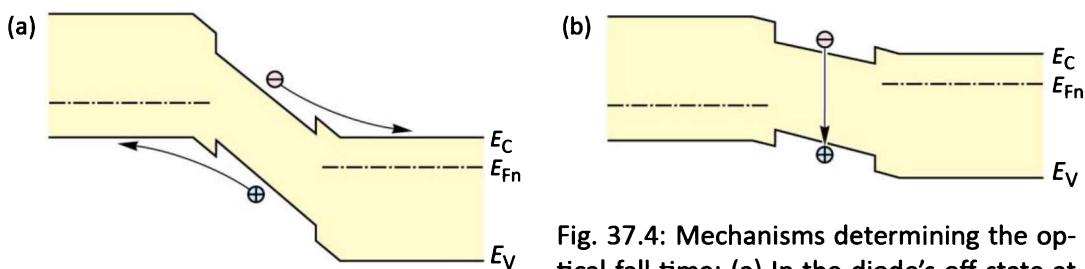


Fig. 37.4: Mechanisms determining the optical fall time: (a) In the diode’s off state at zero or reverse bias, carriers are swept out of the active region. The sweep-out time can be short, $\ll 1$ ns. (b) If the modulating-voltage amplitude is smaller, i.e., if the diode is, in its off state, forward-biased yet below threshold, carriers are *not* swept out, so that the intensity decay is determined by the spontaneous recombination lifetime (> 1 ns).

The situation is quite different for large voltage swings. At zero bias, the band diagram of the active region is highly sloped due to the built-in electric field of the p-n junction. As a result, free carriers are swept out of the active region into the neutral n- and p-type confinement regions of the semiconductor. The carrier sweep-out is most efficient for large voltage swings, i.e. when a high electric field is created in the space-charge region of the p-n junction. The sweep-out time can be much shorter than the spontaneous lifetime. Thus, the fall time is determined *not* by the spontaneous recombination lifetime but by the shorter sweep-out time. Considering the magnitude of the built-in electric field and the carrier mobility, the sweep-out time is estimated to be in the picosecond range.

Exercise: *Calculation of carrier sweep-out time.* Calculate the carrier sweep-out time for typical values of the electric field in the p-n junction depletion region, typical carrier velocity, and an active region thickness of $0.1\text{--}1.0 \mu\text{m}$.

Solution: The carrier sweep-out time can be very short. For typical diode parameters, the carrier sweep-out time is about 1 to 100 ps, i.e. much shorter than the spontaneous recombination time. As an example, let us assume that a carrier drifts with the drift-saturation velocity, which is about 10^7 cm/s , across a reverse-biased active region. The time needed to drift across a $1.0 \mu\text{m}$ thick active region is given by $1.0 \mu\text{m} / 10^7 \text{ cm/s} = 10 \text{ ps}$.

37.6 Current shaping

A common method to reduce the rise and the fall time is *current shaping* (Lee, 1975; Zucker, 1978; Saul *et al.*, 1985). The diagram of a current-shaping circuit is shown in **Figure 37.5 (a)**. The current-shaping circuit is essentially a capacitor and a resistor in series with the LED. The capacitor creates a current transient when the LED is switched on or off, as shown in **Figure 37.5 (b)**. During the switch-on transient, the excess current flowing through the capacitor helps in reaching the steady-state carrier concentration in the active region within a time shorter than the spontaneous lifetime. During the switch-off period ($V=0$), the capacitor biases the diode in the reverse direction, thus aiding the current sweep-out of the active region. Parameters entering the current transient include the power supply internal resistance, the resistor and capacitor of the current-shaping circuit, and the diode series resistance.

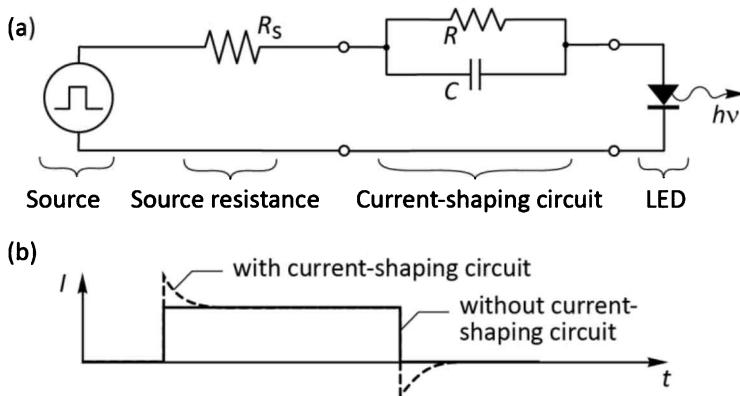


Fig. 37.5: (a) Illustration of an RC “current-shaping circuit” used to decrease the rise time of LEDs. (b) Diode current pulse versus time in the limit of small diode capacitance (solid line) and effect of current-shaping circuit on diode current (dashed line).

A reasonable design criterion for the current-shaping circuit is that the RC time constant of the circuit introduced by the RC current-shaping circuit should be equal to the spontaneous recombination lifetime and that the initial current at the beginning of the voltage pulse should be twice the magnitude of the steady-state current. In this case, the value of the resistor R of the current-shaping circuit is chosen to equal the differential resistance of the diode. The capacitance of the circuit can be chosen so that the RC time constant of the entire circuit coincides with the LED rise time (Schubert *et al.* 1996). As a result of the current-shaping circuit, the 3 dB bandwidth increases.

Note that the diode series resistance is strongly voltage-dependent, and therefore, linear circuit theory can only provide *estimates* rather than accurate values. Therefore, experimental or numerical methods are required to optimize the RC current-shaping circuit.

The current-shaping circuit requires an increased operating voltage and this reduces the overall efficiency of the drive circuit. However, the power-conversion efficiency is usually of little relevance due to the low overall power consumption of communication LEDs.

37.7 Three-dB frequency

The 3-dB frequency of an LED can be determined by measuring the frequency at which the detector signal decreases to one-half of the low-frequency value. The frequency response of the detector needs to be taken into account in the measurement. A 3-dB frequency of about 500 MHz was determined for the LED for which the rise and fall times are shown in *Figure 37.2*. Comparison of the 3-dB frequencies for different LED structures revealed that the highest 3-dB frequencies are attained with devices having low parasitic resistances and capacitances. This can be achieved by a small p-type contact area, a thick SiO_2 isolation layer on the p-type side of the

device, a small area of the bonding pad on the p-type side of the device, and a mesa etch limiting the p-n junction depletion capacitance at zero bias.

37.8 Eye diagram

The *eye diagram* allows one to estimate the overall performance of an optical communication system. The eye diagram is the receiver signal of a randomly generated digital signal. An eye diagram of an LED operating at 622 Mbit/s is shown in **Figure 37.6**. A data rate of 622 Mbit/s is used in the well-known “synchronous optical network” (SONET) standard. The figure indicates the level of the “1” state and of the “0” state and of the *decision level*, i.e. the boundary between what is interpreted by the receiver as a “0” and “1”. The figure also reveals the “eye”. An “open eye” such as the one shown in the figure, allows for a low bit-error rate. The “eye” shown in **Figure 37.6** is wide open, indicating that low bit-error-rate data transmission is possible at that frequency using LEDs. The “on” and “off” pulse-generator voltages of the diode were 1.4 and 1.1 V, respectively. A pulse-shaping *RC* circuit with $R = 20 \Omega$ and $C = 100 \text{ pF}$ was used for the measurement. Minimizing parasitic elements (e.g. bond pad capacitance) and employment of an *RC* pulse shaping circuit should make transmission rates of 1 Gbit/s possible.

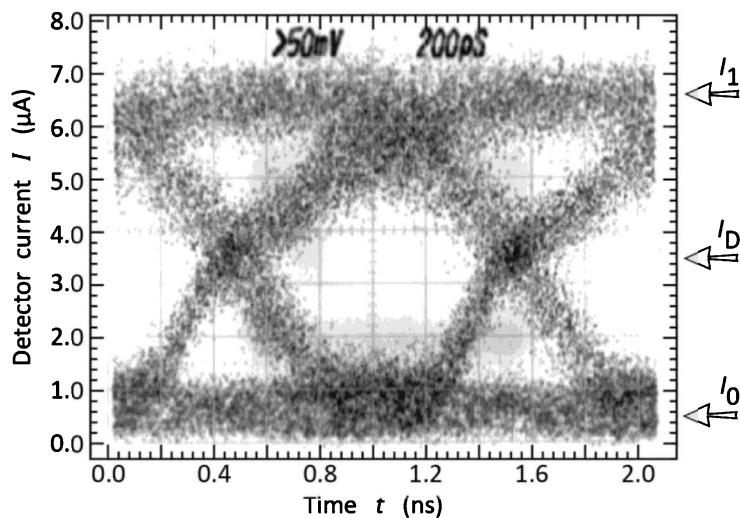


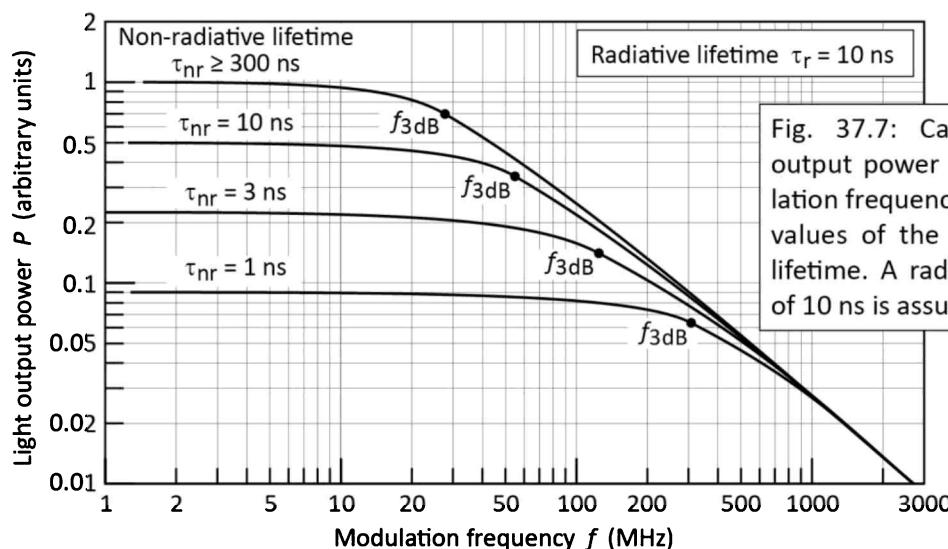
Fig. 37.6: “Eye diagram” of the received optical signal of an RCLED. The optical signal is measured as the photocurrent of a Si photodiode measured with a sampling oscilloscope. The data rate of the RCLED is 622 Mbit/s. Also indicated are the three current levels for the “0” state (I_0), “1” state (I_1), and decision current (I_D).

As the data rate increases, the “eye will close”, i.e. the photocurrent of the “0” and “1” level cannot be clearly distinguished. The “closing of the eye” causes an increase in bit-error rate.

37.9 Carrier lifetime and 3-dB frequency

Shortening the minority carrier lifetime through either very high doping of the active region or deliberate introduction of deep traps will increase the maximum modulation frequency. Deep traps have a two-fold effect. Firstly, they reduce the minority carrier lifetime, thereby increasing the 3-dB frequency. Secondly, they reduce the emission intensity and increase the heat generated inside the LED. Very high concentrations of a shallow dopant, on the other hand, will shorten the carrier lifetime and often, but not necessarily, hurt the device efficiency (Ikeda *et al.*, 1977).

We therefore analyze the effect of lifetime reduction on emission intensity and 3-dB frequency. According to Eqn. (37.8), the 3-dB frequency of an LED depends on the radiative lifetime according to $f_{3dB} = 3^{1/2}/(2\pi\tau)$, where $\tau^{-1} = \tau_r^{-1} + \tau_{nr}^{-1}$. The internal device efficiency is given by $\eta_{int} = \tau_{nr}/(\tau_r + \tau_{nr})$. In the limit of small non-radiative lifetimes, it is $f_{3dB} \propto \tau_{nr}^{-1}$ and $\eta_{int} \propto \tau_{nr}$. Thus, although the modulation bandwidth *can* be increased by the deliberate introduction of deep traps, the power-bandwidth product *cannot*. The relation between 3-dB frequency, output power, and radiative and non-radiative lifetime is shown in **Figure 37.7**. The 3-dB frequencies and intensity levels are calculated according to the equations stated above. At frequencies much higher than the 3-dB frequency, a linear decrease of the optical intensity on the log–log scale can be assumed (Wood, 1994).



If, however, one were to succeed in decreasing the minority carrier lifetime without affecting the efficiency, the modulation speed of an LED could be increased. Although modulation rates as high

as 1.7 GHz have been demonstrated for devices having a highly doped active region, the increase in bandwidth has been accompanied by a decrease in device efficiency (Chen *et al.*, 1999).

Chen *et al.* (1999) proposed very high Be doping ($N_A = 2 \times 10^{19}$ to $7 \times 10^{19} \text{ cm}^{-3}$) of a GaAs active region to shorten the lifetime without significantly degrading the internal quantum efficiency. For a test device that was doped at $2 \times 10^{19} \text{ cm}^{-3}$, the authors found a cutoff frequency of 440 MHz and an internal quantum efficiency of 25~30%. The authors found a cutoff frequency of 1.7 GHz and an internal quantum efficiency of 10% for a device doped at $7 \times 10^{19} \text{ cm}^{-3}$.

37.10 References

- Chen C. H., Hargis M., Woodall J. M., Melloch M. R., Reynolds J. S., Yablonovitch E., and Wang W. "GHz bandwidth GaAs light-emitting diodes" *Applied Physics Letters* **74**, 3140 (1999)
- Ikeda K., Horiuchi S., Tanaka T., and Susaki W. "Design parameters of frequency response of GaAs-AlGaAs DH LEDs for optical communications" *IEEE Transactions on Electron Devices* **ED-24**, 1001 (1977)
- Lee T. P. "Effect of junction capacitance on the rise time of LEDs and on the turn-on delay of injection lasers" *Bell System Technical Journal* **54**, 53 (1975)
- Saul R. H., Lee T. P., and Burrus C. A. "Light-emitting-diode device design" in *Lightwave Communications Technology* edited by W. T. Tsang, *Semiconductor and Semimetals*, **22** Part C (Academic Press, San Diego, 1985)
- Schubert E. F., Hunt N. E. J., Malik R. J., Micovic M., and Miller D. L. "Temperature and modulation characteristics of resonant-cavity light-emitting diodes" *IEEE Journal of Lightwave Technology* **14**, 1721 (1996)
- Wood D. *Optoelectronic Semiconductor Devices* p. 98 (Prentice Hall, New York, 1994)
- Zucker J. "Closed-form calculation of the transient behavior of (AlGa)As double heterojunction LEDs" *Journal of Applied Physics* **49**, 2543 (1978)

Appendices

Appendix 1 – Frequently used symbols

Symbol	Explanation	SI unit	Common unit
a_0	lattice constant	m	nm or Å
a_B	Bohr radius	m	nm or Å
A	area	m^2	cm^2
A	Einstein A coefficient	s^{-1}	s^{-1}
α	absorption coefficient	m^{-1}	cm^{-1}
α_0	absorption coefficient at $h\nu = 2E_g$	m^{-1}	cm^{-1}
α_{fc}	free-carrier absorption coefficient	m^{-1}	cm^{-1}
α_g	absorption coefficient at bandgap energy	m^{-1}	cm^{-1}
B	bimolecular recombination coefficient	m^3/s	cm^3/s
B	Einstein B coefficient	$\text{m}^3/(\text{J s}^2)$	$\text{m}^3/(\text{J s}^2)$
c	speed of light in vacuum	m/s	m/s
C	capacitance	F	F
CRI	color-rendering index	–	–
$C-V$	capacitance-versus-voltage characteristic	–	–
d	layer thickness	m	μm
D	diffusion constant of dopants or carriers	m^2/s	cm^2/s
D	electric flux density	C/m^2	C/cm^2
DOS	density of states, see N_c , N_v	–	–
D_n , D_p	diffusion constant of electrons, holes	m^2/s	cm^2/s
Δn , Δp	change in electron or hole concentration	m^{-3}	cm^{-3}
ΔE_C , ΔE_V	band discontinuity of CB or VB	J	eV
ΔE^*	color difference	–	–
e	elementary charge	C	C
\mathcal{E}	electric field intensity	V/m	V/cm
E	energy	J	eV
E_0	energy of lowest state of a quantum well	J	eV
E_C	energy of conduction band edge	J	eV
E_F	Fermi energy	J	eV
E_{Fi}	Fermi energy in intrinsic semiconductor	J	eV
E_{Fn}	quasi-Fermi energy in n-type region	J	eV
E_{Fp}	quasi-Fermi energy in p-type region	J	eV
E_g	energy of semiconductor bandgap	J	eV
E_{Ryd}	Rydberg energy	J	eV
E_T	energy of trap or deep level	J	eV
E_V	energy of valence band edge	J	eV
$\epsilon = \epsilon_r \epsilon_0$	dielectric permittivity	$\text{As}/(\text{Vm})$	$\text{As}/(\text{Vm})$

ϵ_0	absolute permittivity	A s/(V m)	A s/(V m)
ϵ_r	relative permittivity	—	—
η	efficiency	%	%
$\eta_{EQE} = EQE$	external quantum efficiency (<i>EQE</i>)	%	%
$\eta_{\text{injection}} = IE$	injection efficiency (<i>EQE</i>)	%	%
$\eta_{LEE} = LEE$	light-extraction efficiency (<i>LEE</i>)	%	%
$\eta_{LER} = LER$	luminous efficacy of radiation	lm/W	lm/W
$\eta_{LES} = LES$	luminous efficacy of a light source	lm/W	lm/W
$\eta_{IQE} = IQE$	internal quantum efficiency (<i>IQE</i>)	%	%
$\eta_{\text{power}} = PCE$	power-conversion efficiency	%	%
$\eta_{\text{radiative}} = RE$	radiative efficiency	%	%
f	frequency	Hz	Hz
F	finesse of cavity	—	—
ϕ	angle	°	°
ϕ_c	critical angle for total internal reflection	°	°
Φ	electrostatic potential	V	V
Φ	angle	°	°
Φ_{lum}	luminous flux	lm	lm
G	generation rate	m^{-3}/s	cm^{-3}/s
G_0	generation rate in equilibrium	m^{-3}/s	cm^{-3}/s
G_{excess}	excess generation rate	m^{-3}/s	cm^{-3}/s
h	Planck constant	J s	eVs
\hbar	Planck constant divided by 2π	J s	eVs
hv	photon energy	J	eV
I	optical intensity	W/m^2	W/cm^2
I	current	A	A
I_s	reverse saturation current	A	A
$I-V$	current-versus-voltage characteristic	—	—
J	current density	A/m^2	A/cm^2
J_s	saturation current density	A/m^2	A/cm^2
k	wave number ($2\pi/\lambda$)	m^{-1}	cm^{-1}
k	Boltzmann constant	J/K	eV/K
L	length	m	m
L_{cav}	cavity length	m	μm
L_n, L_p	diffusion length of electrons, holes	m	μm
$L-I$	light-power-vs-current characteristic	—	—
L^*	CIE uniform color space coordinate (1986)	—	—
λ	wavelength	m	nm
λ_{Bragg}	Bragg reflection wavelength	m	nm
m	mass	kg	kg
m_e	free-electron mass	kg	kg
m^*	effective mass	kg	kg
m_e^*, m_h^*	effective mass of electron, hole	kg	kg

m_{lh}^*, m_{hh}^*	effective mass of light hole, heavy hole	kg	kg
$\mu = \mu_r \mu_0$	magnetic permeability	V s/(A m)	V s/(A m)
μ_0	absolute magnetic constant	V s/(A m)	V s/(A m)
μ_r	relative magnetic constant	—	—
μ_n, μ_p	mobility of electrons, holes	$\text{m}^2/(\text{Vs})$	$\text{cm}^2/(\text{Vs})$
n-type	n-type semiconductor material	—	—
n ⁻ -type	lightly doped n-type material	—	—
n ⁺ -type	heavily doped n-type material	—	—
n	electron concentration	m^{-3}	cm^{-3}
n_0	equilibrium electron concentration	m^{-3}	cm^{-3}
n^{2D}	two-dimensional electron concentration	m^{-2}	cm^{-2}
n_{ideal}	diode ideality factor	—	—
n_p	electron concentration in p-type material	m^{-3}	cm^{-3}
\bar{n}	refractive index	—	—
\bar{n}_{air}	refractive index of air	—	—
\bar{n}_{gr}	group index	—	—
\bar{n}_s	refractive index of semiconductor	—	—
N_A, N_D	acceptor, donor concentration	m^{-3}	cm^{-3}
N_c	effective density of states at CB edge	m^{-3}	cm^{-3}
N_T	concentration of trap or deep level	m^{-3}	cm^{-3}
N_v	effective density of states at VB	m^{-3}	cm^{-3}
ν	frequency of optical radiation	Hz	Hz
p-type	p-type semiconductor material	—	—
p ⁻ -type	lightly doped p-type material	—	—
p ⁺ -type	heavily doped p-type material	—	—
p	momentum	kg m/s	kg m/s
p	hole concentration	m^{-3}	cm^{-3}
p_0	equilibrium hole concentration	m^{-3}	cm^{-3}
p^{2D}	two-dimensional hole concentration	m^{-2}	cm^{-2}
p_n	hole concentration in n-type material	m^{-3}	cm^{-3}
P	power	W	W
$P(\lambda)$	spectral power density	W/m	W/nm
Q	electrical charge	C	C
Q	cavity quality factor	—	—
r	optical field (amplitude) reflection coefficient	—	—
R	optical power reflection coefficient	—	—
R	resistance	Ω	Ω
R	recombination rate	m^{-3}/s	cm^{-3}/s
R_0	recombination rate in equilibrium	m^{-3}/s	cm^{-3}/s
R_{excess}	excess recombination rate	m^{-3}/s	cm^{-3}/s
R_s	series resistance	Ω	Ω
ρ	charge density	C/m^3	C/cm^3
ρ	resistivity	$\Omega \text{ m}$	$\Omega \text{ cm}$

ρ_c	specific contact resistance	$\Omega \text{ m}^2$	$\Omega \text{ cm}^2$
$\rho_{\text{DOS}}^{3\text{D}}$	three-dimensional density of states	m^{-3}/J	cm^{-3}/eV
$\rho_{\text{DOS}}^{2\text{D}}$	two-dimensional density of states	m^{-2}/J	cm^{-2}/eV
S	surface recombination velocity	m/s	cm/s
σ	conductivity	$(\Omega \text{ m})^{-1}$	$(\Omega \text{ cm})^{-1}$
t	time	s	s
t	layer thickness	m	μm
T	temperature	K	K or $^{\circ}\text{C}$
T	optical power transmission coefficient	—	—
T_c	carrier temperature	K	K or $^{\circ}\text{C}$
T_j	junction temperature	K	K or $^{\circ}\text{C}$
τ	recombination lifetime	s	s
τ_{cav}	recombination lifetime in a cavity	s	s
τ_n, τ_p	recombination lifetimes of electrons, holes	s	s
u, u'	CIE uniform chromaticity coordinate (1960, 1976)	—	—
u^*	CIE uniform color space coordinate (1986)	—	—
v	velocity	m/s	m/s
v_{gr}	group velocity	m/s	m/s
v_{ph}	phase velocity	m/s	m/s
v, v'	CIE uniform chromaticity coordinate (1960, 1976)	—	—
v^*	CIE uniform color space coordinate of (1986)	—	—
V	voltage	V	V
$V_D = V_{bi}$	diffusion voltage or built-in voltage	V	V
V_f	forward voltage	V	V
V_{f1}, V_{f2}, V_{f3}	forward voltage at a specified current	V	V
V_j	junction voltage	V	V
V_{th}	threshold voltage	V	V
$V(\lambda)$	eye sensitivity function	—	—
W	width	m	m
W_D	depletion layer width	m	μm
W_{DH}	double heterostructure width	m	μm
x	chromaticity coordinate	—	—
\bar{x}	CIE color-matching function (red)	—	—
X	CIE tristimulus value (red)	—	—
y	chromaticity coordinate	—	—
\bar{y}	CIE color-matching function (green)	—	—
Y	CIE tristimulus value (green)	—	—
\bar{z}	CIE color-matching function (blue)	—	—
Z	CIE tristimulus value (blue)	—	—

Appendix 2 – Physical constants

a_B	= 0.5292 Å	Bohr radius ($a_B = 0.5292 \times 10^{-10}$ m)
ϵ_0	= 8.8542×10^{-12} A s/(V m)	absolute dielectric constant
e	= 1.6022×10^{-19} C	elementary charge
c	= 2.9979×10^8 m/s	speed of light in vacuum
E_{Ryd}	= 13.606 eV	Rydberg energy
g	= 9.8067 m/s ²	acceleration on earth at sea level
G	= 6.6873×10^{-11} m ³ /(kg s ²)	gravitational constant ($F = G M m / r^2$)
h	= 6.6261×10^{-34} J s	Planck constant ($h = 4.1356 \times 10^{-15}$ eV s)
\hbar	= 1.0546×10^{-34} J s	$\hbar = h/(2\pi)$ ($\hbar = 6.5821 \times 10^{-16}$ eV s)
k	= 1.3807×10^{-23} J/K	Boltzmann constant ($k = 8.6175 \times 10^{-5}$ eV/K)
μ_0	= 12.566×10^{-7} V s/(A m)	absolute magnetic constant
m_e	= 9.1094×10^{-31} kg	free electron mass
N_{Avo}	= 6.0221×10^{23} mol ⁻¹	Avogadro number
$R = k N_{\text{Avo}}$	= 8.3145 J K ⁻¹ mol ⁻¹	ideal gas constant

Note: The *dielectric permittivity* of a material is given by $\epsilon = \epsilon_r \epsilon_0$ where ϵ_r and ϵ_0 are the *relative* and *absolute* dielectric constant, respectively. The *magnetic permeability* of a material is given by $\mu = \mu_r \mu_0$ where μ_r and μ_0 are the *relative* and *absolute* magnetic constant, respectively.

Useful conversions

1 eV	= 1.6022×10^{-19} C V = 1.6022×10^{-19} J
E	= $h\nu = hc/\lambda = 1239.5$ eV / (λ/nm)
kT	= 25.86 meV (at $T = 300$ K)
kT	= 25.25 meV (at $T = 20^\circ\text{C} = 293.15$ K)

Appendix 3 – Properties of III–V arsenides at 300 K

Quantity	Symbol	AlAs	GaAs	InAs	(Unit)
Crystal structure		Z	Z	Z	–
Gap: Direct (<i>D</i>) / Indirect (<i>I</i>)		<i>I</i>	<i>D</i>	<i>D</i>	–
Lattice constant	$a_0 =$	5.6611	5.6533	6.0584	Å
Bandgap energy	$E_g =$	2.168	1.42	0.354	eV
Intrinsic carrier concentration	$n_i =$	10	2×10^6	7.8×10^{14}	cm^{-3}
Effective DOS at CB edge	$N_c =$	1.5×10^{19}	4.4×10^{17}	8.3×10^{16}	cm^{-3}
Effective DOS at VB edge	$N_v =$	1.7×10^{19}	7.7×10^{18}	6.4×10^{18}	cm^{-3}
Electron mobility	$\mu_n =$	200	8500	33 000	cm^2/Vs
Hole mobility	$\mu_p =$	100	400	450	cm^2/Vs
Electron diffusion constant	$D_n =$	5.2	220	858	cm^2/s
Hole diffusion constant	$D_p =$	2.6	10	12	cm^2/s
Electron affinity	$\chi =$	3.50	4.07	4.9	V
Minority carrier lifetime	$\tau =$	10^{-7}	10^{-8}	10^{-8}	s
Electron effective mass	$m_e^* =$	$0.146 m_e$	$0.067 m_e$	$0.022 m_e$	–
Heavy hole effective mass	$m_{hh}^* =$	$0.76 m_e$	$0.45 m_e$	$0.40 m_e$	–
Relative dielectric constant	$\epsilon_r =$	10.1	13.1	15.1	–
Refractive index near E_g	$\bar{n} =$	3.2	3.4	3.5	–
Absorption coefficient at $\gtrsim E_g$	$\alpha =$	10^3	10^4	10^4	cm^{-1}

Note:

- D = Diamond; Z = Zincblende; W = Wurtzite; DOS = Density of states; CB = Conduction band; VB = Valence band.
- The Einstein relation relates the diffusion constant and mobility in a non-degenerately doped semiconductor: $D = \mu (kT/e)$
- Minority carrier diffusion lengths are given by $L_n = (D_n \tau_n)^{1/2}$ and $L_p = (D_p \tau_p)^{1/2}$
- The mobilities and diffusion constants apply to low doping concentrations ($\approx 10^{15} \text{ cm}^{-3}$). As the doping concentration increases, mobilities and diffusion constants decrease.

Appendix 4 – Properties of III–V nitrides at 300 K

Quantity	Symbol	AlN	GaN	InN	(Unit)
Crystal structure		W	W	W	–
Gap: Direct (<i>D</i>) / Indirect (<i>I</i>)		<i>D</i>	<i>D</i>	<i>D</i>	–
Lattice constant	$a_0 =$	3.112	3.191	3.545	Å
	$c_0 =$	4.982	5.185	5.703	Å
Bandgap energy	$E_g =$	6.28	3.425	0.77	eV
Intrinsic carrier concentration	$n_i =$	9.4×10^{-34}	1.9×10^{-10}	920	cm ⁻³
Effective DOS at CB edge	$N_c =$	6.2×10^{18}	2.3×10^{18}	9.0×10^{17}	cm ⁻³
Effective DOS at VB edge	$N_v =$	4.9×10^{20}	1.8×10^{19}	5.3×10^{19}	cm ⁻³
Electron mobility	$\mu_n =$	300	1500	3200	cm ² /Vs
Hole mobility	$\mu_p =$	14	30	–	cm ² /Vs
Electron diffusion constant	$D_n =$	7	39	80	cm ² / s
Hole diffusion constant	$D_p =$	0.3	0.75	–	cm ² / s
Electron affinity	$\chi =$	1.9	4.1	–	V
Minority carrier lifetime	$\tau =$	–	10^{-8}	–	s
Electron effective mass	$m_e^* =$	$0.40 m_e$	$0.20 m_e$	$0.11 m_e$	–
Heavy hole effective mass	$m_{hh}^* =$	$3.53 m_e$	$0.80 m_e$	$1.63 m_e$	–
Relative dielectric constant	$\epsilon_r =$	8.5	8.9	15.3	–
Refractive index near E_g	$\bar{n} =$	2.15	2.5	2.9	–
Absorption coefficient at $\gtrsim E_g$	$\alpha =$	3×10^5	10^5	6×10^4	cm ⁻¹

Note:

- D = Diamond; Z = Zincblende; W = Wurtzite; DOS = Density of states; CB = Conduction band; VB = Valence band.
- The Einstein relation relates the diffusion constant and mobility in a non-degenerately doped semiconductor: $D = \mu (kT/e)$
- Minority carrier diffusion lengths are given by $L_n = (D_n \tau_n)^{1/2}$ and $L_p = (D_p \tau_p)^{1/2}$
- The mobilities and diffusion constants apply to low doping concentrations ($\approx 10^{15}$ cm⁻³). As the doping concentration increases, mobilities and diffusion constants decrease.

Appendix 5 – Properties of III–V phosphides at 300 K

Quantity	Symbol	AlP	GaP	InP	(Unit)
Crystal structure		Z	Z	Z	–
Gap: Direct (<i>D</i>) / Indirect (<i>I</i>)		<i>I</i>	<i>I</i>	<i>D</i>	–
Lattice constant	$a_0 =$	5.4635	5.4512	5.8686	Å
Bandgap energy	$E_g =$	2.45	2.26	1.35	eV
Intrinsic carrier concentration	$n_i =$	0.044	1.6×10^0	1×10^7	cm^{-3}
Effective DOS at CB edge	$N_c =$	2.0×10^{19}	1.9×10^{19}	5.2×10^{17}	cm^{-3}
Effective DOS at VB edge	$N_v =$	1.5×10^{19}	1.2×10^{19}	1.1×10^{19}	cm^{-3}
Electron mobility	$\mu_n =$	60	110	4600	cm^2/Vs
Hole mobility	$\mu_p =$	450	75	150	cm^2/Vs
Electron diffusion constant	$D_n =$	1.6	2.8	120	cm^2/s
Hole diffusion constant	$D_p =$	11.6	1.9	3.9	cm^2/s
Electron affinity	$\chi =$	3.98	3.8	4.5	V
Minority carrier lifetime	$\tau =$	10^{-6}	10^{-6}	10^{-8}	s
Electron effective mass	$m_e^* =$	$0.83 m_e$	$0.82 m_e$	$0.08 m_e$	–
Heavy hole effective mass	$m_{hh}^* =$	$0.70 m_e$	$0.60 m_e$	$0.56 m_e$	–
Relative dielectric constant	$\epsilon_r =$	9.8	11.1	12.4	–
Refractive index near E_g	$\bar{n} =$	3.0	3.0	3.4	–
Absorption coefficient at $\gtrsim E_g$	$\alpha =$	10^3	10^3	10^4	cm^{-1}

Note:

- D = Diamond; Z = Zincblende; W = Wurtzite; DOS = Density of states; CB = Conduction band; VB = Valence band.
- The Einstein relation relates the diffusion constant and mobility in a non-degenerately doped semiconductor: $D = \mu (kT/e)$
- Minority carrier diffusion lengths are given by $L_n = (D_n \tau_n)^{1/2}$ and $L_p = (D_p \tau_p)^{1/2}$
- The mobilities and diffusion constants apply to low doping concentrations ($\approx 10^{15} \text{ cm}^{-3}$). As the doping concentration increases, mobilities and diffusion constants decrease.
- The minority carrier lifetime τ applies to doping concentrations of 10^{18} cm^{-3} . For other doping concentrations, the lifetime is given by $\tau = B^{-1}(n + p)^{-1}$, where $B_{GaP} \approx 10^{-13} \text{ cm}^3/\text{s}$.

Appendix 6 – Properties of Si and Ge at 300 K

Quantity	Symbol	Si	Ge	(Unit)
Crystal structure		D	D	–
Gap: Direct (D) / Indirect (I)		I	I	–
Lattice constant	$a_0 =$	5.43095	5.64613	Å
Bandgap energy	$E_g =$	1.12	0.66	eV
Intrinsic carrier concentration	$n_i =$	1.0×10^{10}	2.0×10^{13}	cm^{-3}
Effective DOS at CB edge	$N_c =$	2.8×10^{19}	1.0×10^{19}	cm^{-3}
Effective DOS at VB edge	$N_v =$	1.0×10^{19}	6.0×10^{18}	cm^{-3}
Electron mobility	$\mu_n =$	1500	3900	cm^2/Vs
Hole mobility	$\mu_p =$	450	1900	cm^2/Vs
Electron diffusion constant	$D_n =$	39	101	cm^2/s
Hole diffusion constant	$D_p =$	12	49	cm^2/s
Electron affinity	$\chi =$	4.05	4.0	V
Minority carrier lifetime	$\tau =$	10^{-6}	10^{-6}	s
Electron effective mass	$m_e^* =$	$0.98 m_e$	$1.64 m_e$	–
Heavy hole effective mass	$m_{hh}^* =$	$0.49 m_e$	$0.28 m_e$	–
Relative dielectric constant	$\epsilon_r =$	11.9	16.0	–
Refractive index near E_g	$\bar{n} =$	3.3	4.0	–
Absorption coefficient at $\gtrsim E_g$	$\alpha =$	10^3	10^3	cm^{-1}

Note:

- D = Diamond; Z = Zincblende; W = Wurtzite; DOS = Density of states; CB = Conduction band; VB = Valence band.
- The Einstein relation relates the diffusion constant and mobility in a non-degenerately doped semiconductor: $D = \mu (kT/e)$
- Minority carrier diffusion lengths are given by $L_n = (D_n \tau_n)^{1/2}$ and $L_p = (D_p \tau_p)^{1/2}$
- The mobilities and diffusion constants apply to low doping concentrations ($\approx 10^{15} \text{ cm}^{-3}$). As the doping concentration increases, mobilities and diffusion constants decrease.
- The minority carrier lifetime τ applies to doping concentrations of 10^{18} cm^{-3} . For other doping concentrations, the lifetime is given by $\tau = B^{-1} (n + p)^{-1}$, where $B_{Si} \approx 5 \times 10^{-14} \text{ cm}^3/\text{s}$, and $B_{Ge} \approx 5 \times 10^{-13} \text{ cm}^3/\text{s}$.

Appendix 7 – Periodic Table of Elements

Periodic System of Elements
first conceived by Dimitri Mendeleev (1869)

The diagram shows the periodic table with various groups and categories highlighted:

- Groups:** IA, IIA, IIIA, IVA, VA, VIA, VIIA, VIII.
- Periods:** Rows 1 through 7.
- Elements:** Hydrogen (H), Helium (He), Lithium (Li), Beryllium (Be), Sodium (Na), Magnesium (Mg), Potassium (K), Calcium (Ca), Rubidium (Rb), Strontium (Sr), Cs, Ba, Fr, Ra, Ac, Ce, Pr, Nd, Pm, Sm, Eu, Tb, Dy, Ho, Er, Tm, Yb, Lu, Th, Pa, U, Np, Pu, Am, Cm, Cf, Es, Fm, Md, No, Lr.
- Properties:**
 - Explanation:** Shows the relationship between atomic number, atomic mass, and outer shell electron configuration for Na.
 - Transition metals:** Partially filled d-electron shell.
 - Coinage metals:** Cu, Ag, Cd, In, Sn, Sb, Te, Pb, Bi, Po.
 - Elemental semiconductors:** Ge, As, Sb, Te, Sn, Bi.
 - Metalloids:** Ge, As, Sb, Te, Sn, Bi.
 - Halogens:** F, Cl, Br, I, At.
 - Noble gases:** He, Ne, Kr, Xe, Rn.
 - Other non-metals:** All elements except metals.
- Alkaline-metals:** Li, Na, K, Rb, Cs.
- Alkaline-earth metals:** Be, Ca, Sr, Ba.
- Lanthanides:** Ce, Pr, Nd, Pm, Sm, Eu, Tb, Dy, Ho, Er, Tm, Yb.
- Actinides:** Th, Pa, U, Np, Pu, Am, Cm, Cf, Es, Fm, Md, No, Lr.
- Rare-earth elements:** Lu, Yb, Lu.

Note: s-electron shell can be occupied by at most 2 electrons; p-electron shell by at most 6 electrons; d-electron shell by at most 10 electrons; f-electron shell by at most 14 electrons; Noble Gases have 2 (He), 10 (Ne), 18 (Ar), 36 (Kr), 54 (Xe), and 86 (Rn) electrons; The Noble Gas valence electron configuration always consists of 8 valence electrons, except or He which has 2 valence electrons. The d and f electrons are closer to the nucleus than the outer s and p electrons so that the s and p electrons form the valence electrons. For this reason, the Noble Gas valence electron configuration includes 2 s electrons plus 6 p electrons = 8 electrons.

Index

Note:

- “f” means: “and following page”.
- “ff” means: “and following pages”.
- This index is extensive but not comprehensive. For comprehensive information, electronic search is recommended.

A

ABC model 6-1ff, 6-7ff, 22-3
 ABCCD model 6-1ff, 6-10ff
 ABC + $f(n)$ model 6-2f, 22-3
 Absorption
 below bandgap 15-1f
 free carrier 15-3f
 length 29-23
 Urbach 15-2
 AC LEDs 10-1ff
 bidirectional 10-2
 challenges 10-6f
 unidirectional 10-2
 Acheson process 1-1
 Activation of dopants, *see* Doping
 Activator, *see* Phosphor activator
 Active region 6-1f, 14-3ff
 Adaptive color vision, *see* Color vision
 Additive color mixing, *see* Color mixing
 Ag 12-4
 Al 12-4
 AlAs 12-15
 AlGaAs 2-9ff, 9-4f, 19-6ff
 AlGaN 20-1ff
 AlGaN P 2-18f, 9-2f, 12-15f, 15-15f, 17-11ff, 19-9ff, 19-16f
 AlInP 12-15
 Alternating current, *see* AC
 Aluminium, *see* Al
 Akasaki, Isamu 2-22ff
 Amano, Hiroshi 2-22ff
 ANSI color bins 25-10f
 Anti-reflection coating 15-17f
 a-plane, *see* Crystal structure
 Attenuation, *see* Optical fiber attenuation
 Au 12-4
 Auger coefficient 4-14f, 6-5f
 Auger recombination 4-13ff, 22-3, 22-12f

B

B-coefficient 4-2
 Backlighting 31-9ff

Bandgap

 direct 4-22ff
 direct-indirect transition 19-7, 19-10
 energy, *see* Energy
 indirect 4-22ff
 temperature dependence 34-6f
 Band-filling effect 34-4
 Bandwidth, *see* Modulation bandwidth
 Bandwidth, *see* Emission linewidth
 Bimolecular recombination 4-1ff, 4-6f, 5-6
 Black-body
 radiation 5-4f, 25-1ff
 radiation locus, *see* Planckian locus
 radiation source 27-2
 radiation spectrum 25-2, 25-4
 Bowing energy, 20-2
 Bowing parameter, *see* Bowing energy
 Bridge rectifier 10-3
 Brightness 23-12, 24-13
 Buffer layer 2-22f, 20-13ff
 Built-in voltage 7-1
 Buried contact 9-15

C

Candela, *see* Photometric units
 Carborundum 1-1ff, 1-8
 Carrier
 delocalization 22-11f
 leakage 6-6f, 8-7ff, 8-13ff, 22-4, 22-9f
 localization 20-17f, 22-11f
 loss, *see* Carrier leakage
 overflow 8-10ff
 sweep out of active region 37-7f
 temperature, *see* Temperature
 Cavity
 finesse 17-3ff
 quality factor 17-1, 17-3
 Ce = cerium, *see* Phosphor activator elements
 Characteristic temperature, *see* Temperature
 Chip shaping, *see* Shaping
 Chromaticity coordinates, *see* CIE
 Chromaticity diagram, *see* CIE

CIE

chromaticity diagram 24-2ff, 24-5ff, 24-12f, 25-5ff, 25-10f, 29-31, 30-8
 (x, y) chromaticity coordinates 24-4f
 (u', v') chromaticity coordinates 24-8f
color space 24-13ff
color space "Lab" 24-14ff
color space "Luv" 24-16f
color space "UVW" 24-16f
illuminant, *see* Standard illuminant
uniform chromaticity coordinates, *see* CIE
 (u', v') chromaticity coordinates
uniform color space 27-7

Circadian cycle, *see* Circadian rhythm

Circadian rhythm 23-13f

Circadian efficacy 23-14

Cladding layer 14-1

Color

- and wavelength 24-1f
- complementary 28-3
- gamut 26-4
- matching functions 24-2ff
- mixing 26-1ff
- primary 24-2ff
- purity 24-10f
- rendering 27-1ff
- rendering index (*CRI*) 27-1ff, 27-8, 27-9ff, 27-13f, 30-9
- samples 27-2ff
- saturation, *see* Color purity
- space, *see* CIE color space
- temperature, *see* Temperature
- vision 23-1ff, 24-1ff
 - adaptive 27-10ff
 - adaptive shift 27-10ff
 - mesopic 23-2
 - photopic 23-2
 - scotopic 23-2
- wheel 24-13, 24-15

Colorimetry 24-1ff

Communication

- free-space optical 35-13ff, 36-1
- LED, *see* LED
- optical 35-1ff

Complementary wavelengths, *see* Wavelengths

Composition grading 8-4ff

Compound semiconductors, *see* Semiconductors

Conductivity, *see* Thermal conductivity

Configuration coordinate model 29-4

Confinement layer 14-1

Cone cells, *see* Human eye

Conformal phosphor, *see* Phosphor

Contact

interdigitated contact geometry 9-14f
ohmic 7-17ff, 7-19ff, 7-21f
p-type reflective 12-1ff, 15-13
reflectivity 12-1ff, 12-4f, 12-19ff
resistance 21-6
shape 15-15
specific contact resistance 7-17ff
transparent 12-1ff, 12-6

Correlated color temperature (CCT), *see* Color temperature

Coumarin, *see* Phosphor

Coupling, *see* Optical fiber coupling

c-plane, *see* Crystal structure

Cracking (of epitaxial films) 21-6f

Craford, M. George 2-6ff

Critical thickness, *see* Thickness

Crystal field theory 29-5

Chip on board, *see* Package COB

Chip scale package, *see* Package CSP

Chromaticity coordinates, *see* CIE

COB, *see* Chip on board

CRI, *see* Color rendering index

Critical thickness, *see* Thickness

Crystal structure

a-plane 20-6, 20-7

c-plane 20-6, 20-7

garnet 29-9, 29-17

inversion symmetry 20-2f

m-plane 20-6, 20-7

r-plane 20-6, 20-7

wurtzite 20-2ff

zincblende 20-2

CSP, *see* Chip scale package

Current

blocking layer 9-17f, 36-5

crowding 9-7ff, 9-10ff

shaping 37-8f

spreading 9-1ff, 9-7ff

spreading layer, *see* Current spreading

spreading length 9-7ff, 9-10ff, 9-13

-versus-voltage curve 1-5f, 7-1ff, 19-15

D

DBR, *see* Distributed Bragg reflector

Decay

exponential 4-7ff

Kohlrausch 4-8

luminescence 4-7ff

stretched exponential 4-7

- Deep levels 4-9ff, 4-13, 20-15ff
 Defect
 extended 4-12
 native 4-13
 point 4-12
 Depletion region 7-1ff
 Die shaping, *see* Shaping
 Diffuse reflector, *see* Reflector
 Diffuser 30-5f, 32-6
 Diffusion of light, *see* Light diffusion
 Diffusion 14-6f, 21-6
 Diffusion voltage, *see* Built-in voltage
 Diode
 ideality factor 7-9ff
 voltage, *see* Forward voltage
 Direct bandgap, *see* Bandgap
 Direct-indirect transition, *see* Bandgap
 Dislocation
 misfit 14-12ff
 in III-Nitrides 20-15ff
 Dispersion
 material 35-6ff
 multipath 35-14f
 modal 35-1f, 35-5f
 Distortion, *see* Multipath distortion
 Distributed Bragg reflector, *see* Reflector
 Dominant wavelength, *see* Wavelength
 Doping
 activation (de-passivation) of p-type doping 20-9ff
 active region 14-4ff
 confinement region 14-8ff
 high concentration 7-21, 14-7f
 Mg-doping 2-21f
 passivation of p-type doping 20-11ff
 p-type conductivity 21-6
 p-type doping 2-22f, 20-9ff
 Zn-doping 14-7f
 Double heterostructure 2-24, 14-1ff, 15-5, 19-7f
 Down-converting phosphor, *see* Phosphor
 Drift leakage 22-5ff
 Droop, *see* Efficiency droop
 Dye, *see* Phosphor
- E**
- EBL, *see* Electron blocking layer
 Efficacy 13-1
 luminous efficacy of radiation (*LER*) 13-8ff, 23-8, 23-10, 28-9ff
 luminous efficacy of source (*LES*) 13-10f, 23-
- Efficiency 13-1ff
 carrier injection (*IE*) 13-1, 13-4f
 droop 22-1ff, 22-14ff
 external quantum (*EQE*) 13-1, 13-6f, 21-5, 30-2
 forward voltage (*VfE*) 13-1, 13-7f
 internal quantum (*IQE*) 13-1, 13-5f, 14-1ff
 light extraction (*LEE*) 13-1, 13-6
 phosphor 13-11ff, 30-2ff
 power-conversion efficiency 13-1, 13-8
 radiative (*RE*) 4-22, 6-7ff, 13-1ff
 wall-plug, *see* Efficiency, power
 wavelength-conversion 13-11ff, 30-2f
 Einstein **A** and **B** coefficient 5-9ff
 Einstein model 5-10ff
 Electroluminescence, *see* Luminescence
 Electron affinity 7-19f
 Electron blocking layer 8-13f, 22-16f
 Electrophoresis, *see* Phosphor deposition
 Electrostatic discharge, *see* ESD
 Electrostatic discharge protection, *see* ESD protection
 Emission
 enhancement 16-11f, 16-13ff
 gaussian spectrum 28-8f
 linewidth 11-3f, 28-8f
 pattern 11-7ff
 isotropic 11-7ff
 lambertian 11-7ff
 spectrum 11-1ff, 19-13, 20-8, 21-2f
 spontaneous 16-1ff
 stimulated 5-9f, 36-7f
 temperature dependence 11-12f, 28-11ff, 34-3f
 wavelength 11-2ff
 Encapsulant 11-10f
 Energy
 bandgap 7-6ff, 8-16, 11-1f, *see also* Bandgap
 Epitaxial growth
 GaN 20-13ff
 Epoxy resin, *see* Resin, epoxy
 Equal energy point 24-7
 Er = erbium, *see* Phosphor activator elements
 Escape cone, *see* Light escape cone
 ESD 31-1, 33-1ff
 human body model (HBM) 33-1f
 IEC 61000-4-2, 33-1f
 integrated with LED chip 33-5f
 process 33-2f
 protection circuits 33-3ff
 transient voltage suppression diode 33-4

- Zener diode 33-4f
E
 Eu = europium, *see* Phosphor activator elements
 Exponential decay, *see* Decay
 External quantum efficiency (*EQE*), *see* Efficiency
 Eye diagram 37-10
 Eye, *see* Human eye
 Eye sensitivity function, *see* Human eye
- F**
 Fabry-Perot resonators 16-3ff
 transmittance 16-10
 Fall time, *see* Time
 Fermi's Golden Rule 5-1
 Fiber, *see* Optical fiber
 Finesse, *see* Cavity finesse
 Fluorescence 29-2
 Folberth, Otto Gert 2-5f
 Forward voltage, *see* Voltage
 Free carrier
 absorption, *see* Absorption
 concentration 6-1f
 Fresnel reflection, *see* Reflection
 Full conversion 3-7f, 30-13
- G**
 Ga-face 20-3
 GaAs 2-9ff, 12-15
 GaAsP 2-12ff, 19-1ff
 GaAsP:N 19-1ff
 GaN 2-19ff, 20-1ff
 GaInAs 17-6f
 GaInAsP 14-7
 GaInN 2-22ff, 19-11ff, 19-17, 20-1f
 Gamut, *see* Color gamut
 Gap 2-12ff, 9-3ff, 19-1f
 GaP:N 19-1f
 Garnet crystal structure, *see* Crystal structure
 Gaussian emission spectrum, *see* Emission
 Glass transition temperature, *see* Resin glass
 transition temperature
 Gold, *see* Au
 Grading, *see* Composition grading
 Green gap 19-16f
- H**
 Haitz observation 3-10f
 HBM, *see* Human body model
 Heat glow, *see* Incandescence
 Heat sink 31-11f
 Heisenberg, *see* Uncertainty principle
 Heterojunction 8-2
- High-injection conditions 7-2, 7-5f
 High-level excitation 4-6
 High-voltage LED, *see* HV LED
 History
 SiC LEDs 1-1ff
 III-V LEDs 2-1ff
 White LEDs 3-1ff
 Holonyak Jr., Nick 2-6f
 Homojunction 8-1f
 Human body model, *see* ESD HBM
 Human eye 23-1f
 cone cells 23-1
 ganglion cells 23-1, 23-13f
 rod cells 23-1
 retina 23-1, 23-13
 sensitivity function 23-6ff
 vision, *see* Color vision
 HV LEDs 10-1, 10-7ff
- I**
 Ideality factor, *see* Diode ideality factor
 IEC 61000-4-2 Standard 33-1f
 Illuminance 23-4
 Illuminants, *see* Standard illuminants
 In clusters 19-13
 Incandescence 1-2, 1-6, 25-1ff
 Indirect bandgap, *see* Bandgap
 Indium clusters, *see* In clusters
 Infrared 24-1f, 36-1
 Injection
 adiabatic 8-7
 non-adiabatic 8-16
 Interdigitated contact, *see* Contact
 Internal quantum efficiency (*IQE*), *see* Efficiency
 intra-f-electron-shell transition 29-5
 Inversion symmetry, *see* Crystal structure
 IR, *see* Infrared
 Isoelectronic impurity 2-13, 19-1ff
 Isotropic emitter 17-10
 IV curve, *see* Current-versus-voltage curve
- J**
 Junction temperature, *see* Temperature
- K**
 Kohlrausch decay, *see* Decay
- L**
 Lambert's cosine law 11-9, 12-24
 Lambertian emission pattern, *see* Emission pattern
 Lateral injection 9-16f

- Lattice
 constant 19-6, 19-10, 19-12, 20-1
 matching 14-11ff
- Law of mass action 6-3
- Leadframe, *see* Package leadframe
- Leakage, *see* Carrier leakage
- LED
 accent lighting 3-16
 Burrus type 36-2f
 calculator 2-17f
 color and wavelength 24-1f
 communication 36-1ff
 free space 36-1
 fiber optic 36-1f
 display 3-15, 3-17, 31-10f
 edge-emitting 35-2, 36-4, 36-7ff
 GaAs, *see* GaAs
 GaN, *see* GaN
 GaInN, *see* GaInN
 GaP, *see* GaP
 headlights 3-15
 IR 36-3f
 light bulb 3-17f
 lighting 3-9ff
 micro-cavity, *see* LED, resonant cavity
 packaging, *see* Packaging
 photonic crystal 18-1ff, 18-10f
 resonant cavity 16-2, 17-1ff, 36-6f
 side-view, *see* Package side-view
 single-mode 18-3ff
 SMD, *see* Package SMD
 SMT, *see* Package SMT
 super-luminescent 36-7ff
 thin-film 15-12ff
 thin-film flip-chip (TFFC) 15-12ff
 traffic light 2-24f
 tri-chromatic white 3-1f
 truncated inverted pyramid 15-8f
 visible spectrum 19-1ff
 white, *see* White LED
 watch 2-17,
 UV 21-1ff
- Lehovec, Kurt 1-8f
- Lens 35-11ff, 36-4
- Light
 emitting diode, *see* LED
 escape cone 11-4ff
 diffusion 26-5ff, 26-8f
 extraction efficiency (*LEE*), *see* Efficiency
 guide 31-10ff,
 scattering 29-24, 32-7f
- scattering, *see also* Light diffusion
 scattering (geometric optics type) 26-7
 scattering (Mie type) 26-7
 scattering (Rayleigh type) 26-7
 scattering length 26-8
 trapped 15-1ff, 15-6
- Lightness 24-13
- Linewidth, *see* Emission linewidth
- Loss, *see* Optical fiber attenuation
- Lossy region 36-9
- Lossev, Oleg 1-5ff
- Low-level excitation 4-2ff
- Lumen, *see* Photometric units
- Luminance 23-4
- Luminescence 29-1f
 cathodo- 29-1f
 chemi- 29-1
 electro- 1-2, 1-5, 1-8
 photo- 29-1
 super- 36-7ff
- Luminescence killers 4-9
- Luminous efficacy, *see* Efficacy
- Luminous flux 23-4, 23-9f
- Luminous intensity 23-3
- Lux, *see* Photometric units
- M**
- MacAdam ellipse 24-7f, 2-10
- Material dispersion, *see* Dispersion
- Matthews–Blakeslee law, *see* Critical thickness
- Median particle size, *see* Phosphor particle size
- Metal-insulator-semiconductor diode (MIS diode) 1-4, 2-21
- Metal reflector, *see* Reflector
- Metal-semiconductor diode (MS diode), *see* Schottky diode
- Micro-cavity LED, *see* LED
- Mie scattering, *see* Scattering
- Mineral diffuser, *see* Diffuser
- Minimum flicker method 23-9
- Minority carrier lifetime 4-4f
- Misfit dislocation, *see* Dislocation
- MIS diode, *see* Metal-insulator-semiconductor diode
- Metal insulator semiconductor diode 2-21
- Mg-doping, *see* Doping
- Mirror, *see* Reflector
- Mn = manganese, *see* Phosphor activator elements
- Modal dispersion, *see* Dispersion
- Mode density, *see* Optical mode density

- Modulation bandwidth 37-2ff
- Modulation of LEDs 37-1ff
- Mohs Hardness Scale 1-1
- Molding 31-3ff
 - injection 31-3ff, 32-3f
 - transfer 31-5ff, 32-4ff
- Momentum
 - carrier and photon 11-1f
 - conservation 11-2
- m-plane, *see* Crystal structure
- MQW, *see* Quantum well
- Multipath distortion 35-14f
- Multiple quantum well, *see* Quantum well
- Munsell color samples, *see* Color samples

- N**
- NA, *see* Numerical aperture
- N-face 20-3
- Nakamura, Shuji 2-23ff
- Nucleation layer, *see* Buffer layer
- Numerical aperture, *see* Optical fiber

- O**
- Ohmic contact, *see* Contact
- Optical
 - communication, *see* Communication
 - fiber 35-1ff
 - attenuation (loss) 35-3ff
 - coupling with lens 35-11ff
 - graded-index 35-1f
 - multimode 35-1f
 - numerical aperture 35-9ff
 - plastic 35-3f
 - PMMA 35-8
 - silica 35-3f
 - single mode 35-1f
 - mode density 16-7ff, 17-4
 - power-conversion efficiency, *see* Efficiency
 - scattering, *see* Light scattering
 - active impurity, *see* Phosphor activator

- P**
- Package 31-1ff,
 - 3 mm and 5 mm 31-2f
 - ceramic 31-7ff
 - COB = chip on board 31-12f
 - CSP = chip scale package 31-13
 - lead frame 31-3, 31-4f
 - low-power 31-2
 - resin, *see* Resin
 - side-view 31-9f
- SMD 31-3f
- SMT 31-3f
- surface mount device, *see* Package SMD
- Parallel resistor, *see* Resistor
- Parallel diodes 10-9
- Parallelepiped 15-7ff
- Parasitic diode 7-13, 7-15
- Parasitic resistance 7-11ff
- Partial conversion 3-7f, 30-6f, 30-13
- Particle size, *see* Phosphor particle size
- Passivation of dopants, *see* Doping
- PBG, *see* Photonic bandgap
- Phosphor 3-2ff, 28-1, 29-1ff, 30-4ff
 - activator 29-2
 - activator elements 29-3f
 - Ce 29-4, 29-9ff
 - Er 29-4
 - Eu 29-4, 29-11ff
 - Mn 29-4
 - Pr 29-4
 - Sb 29-4
 - Tb 29-4
 - Tm 29-4
 - agglomeration 29-24
 - aluminate 29-16
 - blends 29-28f, 30-5, 30-8f
 - body color 29-14f
 - CIE chromaticity diagram 29-31
 - conformal 30-18f
 - coumarin 29-6
 - deposition 29-27ff
 - dispensation 29-27
 - electrophoresis 29-28
 - preform 29-28
 - spraying 29-28
 - screen printing 29-28
 - spinning 29-28
 - distribution 30-18ff
 - conformal, *see* Phosphor conformal
 - proximate, *see* Phosphor proximate
 - remote, *see* Phosphor remote
 - down-converting 29-1
 - dye 29-6, 30-5
 - fluoride 29-16
 - host lattice 29-3
 - inorganic 29-7ff
 - nitride 29-12ff, 29-16, 29-17f, 29-21
 - organic 29-6f
 - oxide 29-16
 - particle size 29-23ff
 - distribution 29-25f

median 29-26
 pigment 29-6, 30-5
 proximate 29-29, 30-18ff
 quantum-splitting 29-30f
 quenching
 temperature quenching 29-20f
 concentration quenching 29-20, 29-21f
 remote 29-29, 30-18ff
 scattering 26-7f, 29-24
 sedimentation 29-25, 29-29f
 semiconductor 29-7f
 sensitizer 29-2
 silicate 29-14, 29-18
 structure 29-16ff
 sulfide 29-11f, 29-16
 synthesis 29-16ff
 synthesis using flux 29-19
 up-converting 29-1
 YAG:Ce 3-5ff, 29-9ff, 29-17, 30-6ff
 TAG:Ce 29-11
 Phosphorescence 29-2
 Photometric units 23-3ff
 candela 23-3
 lumen 23-4
 lux 23-4
 Photon recycling 18-1
 Photonic bandgap, *see* Photonic crystal
 Photonic crystal 16-1, 18-1ff, 18-5ff
 Photonic crystal LED, *see* LED
 Photopic vision, *see* Vision
 Pigment, *see* Phosphor
 Planck, Max 25-3
 Planckian locus 25-5ff, 25-10f, 28-13
 Planckian radiation, *see* Black-body radiation
 pn junction 7-1ff
 pn-junction displacement 14-6ff
 Polarization 20-2ff
 effect 20-2ff, 20-7ff
 effect on carrier leakage 22-9f
 matching 22-16
 piezoelectric 20-4f
 spontaneous 20-4f
 Polymerization reaction, *see* Resin
 polymerization reaction
 Power-conversion efficiency, *see* Efficiency
 Pr = praseodymium, *see* Phosphor activator elements
 Premature turn-on, *see* Sub-threshold turn-on
 Primary color, *see* Color
 p-type doping, *see* Doping
 p-type conductivity, *see* Doping

Q

Quality factor, *see* Cavity quality factor
 Quantum-confined Stark effect (QCSE) 20-7ff
 Quantum-splitting phosphor, *see* Phosphor
 Quantum well (QW) structure 2-23f, 19-8
 multiple quantum well (MQW) 14-3
 single quantum well (SQW) 14-3

R

Radiative recombination, *see* Recombination
 Radiant flux, *see* Optical power
 Radiometric units 23-3
 Rare earth element (REE) 29-4
 Rayleigh scattering, *see* Scattering
 RCLED, *see* LED
 Recombination
 doping dependence 5-7ff
 electron-hole 1-8
 non-radiative 4-1ff, 4-8ff, 4-13ff, 4-16ff, 4-20ff, 14-10f
 radiative 4-1ff, 4-20f, 5-1ff, 5-9ff, 6-4
 surface 4-16ff, 14-11
 temperature dependence 5-7ff

Reflection

Fresnel 11-10, 12-2f, 12-9, 12-11, 15-17f
 total internal 11-4f, 11-11, 12-1, 12-6ff, 15-6
 Reflectivity of test-color samples 27-3ff, 27-16ff
 Reflectivity of contact, *see* Contact reflectivity

Reflector 12-1ff

cup 31-2
 DBR 12-8ff
 DBR penetration depth 12-12f
 diffuse 12-22ff
 lambertian 12-22ff, 12-27
 metal 12-1ff
 mixed specular-diffuse 12-22ff,
 omnidirectional 12-19ff
 specular 12-22ff
 total internal 12-6ff

Refractive index

epoxy 32-10, 32-14
 graded-refractive index resin 32-15f
 group refractive index 35-6ff
 phase refractive index 35-6ff
 silicone 32-12
 filled resin 32-14f

Remote phosphor, *see* Phosphor

Resin 32-1ff
 advanced 32-14f
 acrylic resin 32-5
 degradation 32-16f

- epoxy 32-5, 32-8ff
- fillers 32-6ff
- glass transition temperature 32-3
- LCP (liquid crystal polymer) 32-4
- nylon 32-4
- polycarbonate 32-4
- polymer 32-2
- polymerization reaction 32-9
- polyurethane (PUR) 32-5
- PMMA (Poly (methyl methacrylate)) 32-3, 32-4
- PPA (polyphthalamide) 32-4
- PPS (polyphenylene sulfide) 32-4
- silicone 32-6, 32-11ff
- thermoplastic 31-4, 32-2, 32-3f
- thermosetting 31-6, 32-2, 32-4f
- types 32-2ff
- Resistor**
 - parallel 7-12, 7-15f
 - series 7-6, 7-12f, 7-15f, 8-3f, 10-7, 33-2
- Resonant cavity LED**, *see* LED
- Retina**, *see* Human eye
- Reverse saturation current**, *see* Saturation current
- Rise time**, *see* Time
- Rod cells**, *see* Human eye
- Round, Henry** 1-2ff
- Roughening**, *see* Surface roughening
- r-plane**, *see* Crystal structure

- S**
- Sapphire** 20-13ff
- Saturation current** 7-3, 7-16
- Sb = antimony**, *see* Phosphor activator elements
- Scattering**, *see* Light scattering
- Screen printing**, *see* Phosphor deposition
- Sedimentation of phosphor**, *see* Phosphor
- Semiconductors**
 - alloy 2-5ff
 - compound 2-1ff
 - wavelength converter, *see* Phosphor
 - III-V 2-1ff
 - II-VI 2-1
- Sensitizers**, *see* Phosphor sensitizer
- Schottky barrier** 7-20
- Schottky diodes** 1-4
- Schottky model** 7-20f
- Scotopic vision**, *see* Vision
- Series resistor**, *see* Resistor
- Shaping of chip** 15-6ff
- Shimizu, Yoshinori** 3-4ff, 3-8
- Shockley diode equation** 7-3
- Shockley-Read-Hall recombination** 4-8ff, 6-2f, 7-5, 22-3
- Shockley, Willian Bradford** 7-6
- Shunt**, *see* Parallel resistance
- SiC** 1-1ff
- Silicon carbide**, *see* SiC
- Silicone resin**, *see* Resin, silicone
- Siloxane**, *see* Silicone
- Silver**, *see* Ag
- Single-mode LED**, *see* LED
- SMD**, *see* Surface mount device
- SMT**, *see* Surface mount technology
- Snell's law** 11-4, 11-7, 11-10, 12-6
- Solar spectrum**, *see* Spectrum
- Solar blind range** 21-1f
- Solid-state lighting**, *see* White LED
- Solder bump bonding** 15-14
- Specific contact resistance**, *see* Contact
- Spectral linewidth**, *see* Emission linewidth
- Spectrum**
 - emission, *see* Emission spectrum
 - solar 25-3ff
 - width, *see* Emission linewidth
- Specular reflector**, *see* Reflector
- Spontaneous emission**, *see* Emission
- SQW**, *see* Quantum well
- SRH**, *see* Shockley-Read-Hall
- Standard**
 - ANSI, *see* ANSI color bins
 - IEC, *see* ESD IEC
 - illuminants (CIE) 25-5f, 25-10f
- Stokes shift energy** 3-3, 13-12, 29-5, 29-8, 30-2, 30-8, 30-10
- Stop band** 12-8, 12-10, 12-12ff, 12-17f
- Strain**
 - compressive 20-5
 - tensile 20-5
- Substrate**
 - transparent 15-15ff
- Sub-threshold turn-on** 7-12, 7-16
- Subtractive color mixing**, *see* Color mixing
- Super-luminescence**, *see* Luminescence
- Super-luminescent LED**, *see* LED
- Surface**
 - diffuse, *see* Reflection, diffuse
 - lambertian, *see* Reflection, lambertian
 - leakage current 9-19f
 - mount device, *see* Package SMD
 - mount technology, *see* Package SMT
 - recombination, *see* Recombination

- recombination velocity 4-18ff
- reconstruction 4-17
- roughening 15-10ff
- specular, *see* Reflection, specular
- states 4-17
- texturing, *see* Surface roughening

- T**
- TAG, *see* Phosphor
- TAG:Ce, *see* Phosphor
- Terbium aluminium garnet, *see* TAG
- Tb = terbium, *see* Phosphor activator elements
- Temperature
 - carrier 34-1ff
 - characteristic 11-12f
 - color temperature (CT) 25-6ff, 25-9f, 28-9ff, 29-15, 29-31, 30-9ff
 - color temperature bins 25-10f
 - correlated color temperature (CCT) 25-7ff, 30-9ff
 - glass transition, *see* Resin glass transition temperature
 - junction 34-1ff
 - junction, measurement 34-8ff
 - stability of epoxies 32-16
 - stability of silicones 32-16
- Terbium aluminum garnet, *see* TAG
- Test color samples, *see* Color samples
- Texturing, *see* Surface roughening
- Thermal conductivity 31-7
- Thermal consideration 34-11
- Thermal roll-over 22-10f
- Thermal time constant 34-11f
- Thickness
 - active region 14-3f
 - critical 14-13f
- Threading dislocation, *see* Dislocation
- Three dB frequency, *see* 3 dB frequency
- Threshold voltage 7-3
- Thresholdless laser, *see* Single-mode LED
- Time
 - fall 37-1ff, 37-4ff
 - rise 37-1ff, 37-4ff
- TIR, *see* Total internal reflection
- TLM, *see* Transfer length model
- Tm = thulium, *see* Phosphor activator elements
- Total internal reflection (TIR), *see* Reflection, total
- Transient voltage suppression diodes, *see* ESD
- Transfer length model 7-17ff
- Transmission line model, *see* Transfer length
- model
- Transparent contact, *see* Contact
- Transparent substrate, *see* Substrate
- Trapped light, *see* Light
- Traps, *see* Deep levels
- Trichromacy 24-3
- Trichromatic LED, *see* LED
- Tristimulus values 24-4f
- Truncated inverted pyramid LED, *see* LED
- Turn-on voltage, *see* Threshold voltage
- TVS diode, *see* Transient voltage suppression diode

- U**
- Ultraviolet, *see* UV
- Uncertainty principle 2-13
- Uniform chromaticity coordinates, *see* CIE
- Up-converting phosphor, *see* Phosphor
- Urbach absorption tail, *see* Absorption
- UV 21-1ff
 - extreme UV 21-1f
 - deep UV radiation 21-1f
 - near UV radiation 21-1f
 - UV-A radiation 21-1
 - UV-B radiation 21-1
 - UV-C radiation 21-1f
 - UV LED, *see* LED
 - vacuum UV 21-1f

- V**
- van Roosbroeck–Shockley model 5-3ff
- Varshni formula 34-6f
- Vertical cavity surface emitting laser (VCSEL) 17-5f
- Vertical transition 11-2
- Visible-spectrum LED, *see* LED
- Vision, *see* Color vision
- Voltage
 - forward voltage 7-6ff, 8-15f, 15-17, 19-15, 33-2, 34-4ff
 - forward voltage, one, V_{f1} , 7-16
 - forward voltage, two, V_{f2} , 7-16
 - forward voltage, three, V_{f3} , 7-16
 - forward voltage temperature dependence 19-16, 34-4ff
 - forward voltage temperature coefficient 34-7f, 34-13f

- W**
- Wake-sleep rhythm, *see* circadian rhythm
- Wall-plug efficiency, *see* Efficiency

- Wavelength**
complementary 28-3f
converter, *see* Phosphor
converting materials, *see* Phosphor
dominant 24-11
emission, *see* Emission wavelength
Waveguiding 12-4, 12-7f, 12-25, 15-6, 31-9ff, 36-7ff
Waver bonding 15-13, 15-15f
Welker, Heinrich 2-1f
White LED
di-chromatic 28-1f, 28-3ff, 30-1
full-conversion based, *see* Full conversion
high-CRI 30-11ff
history 3-1ff
lightbulb 3-2, 3-10, 3-17f, 30-10, 30-20
partial-conversion based, *see* Partial conversion
phosphor-based 30-1ff, 30-6f, 30-8f, 30-13ff
phosphor-blend-based 30-8ff
red-LED based 30-10f,
RGB LEDs 3-1f, 28-2f, 28-7ff
temperature dependence 28-11f, 28-13f
tetra-chromatic 28-1f, 28-14, 30-1
tri-chromatic 28-1f, 28-7ff, 30-1
violet-excitation-based 30-11ff
UV-excitation-based 30-13ff
Wien's displacement law 25-2
Window layer, *see* Current spreading layer
Wurtzite crystal structure, *see* Crystal structure
- X**
- Y**
YAG, *see* Phosphor
YAG:Ce, *see* Phosphor
Yellow luminescence in GaN 4-12f
Yttrium aluminium garnet, *see* YAG
- Z**
Zener diode, *see* ESD
Zincblende crystal structure, *see* Crystal structure
- Numerical values**
0.01 rule 28-13,
1-step (2-step, 3-step) MacAdam ellipse, *see* MacAdam ellipse
3dB frequency 37-2ff, 37-9f, 37-11f
1907, First LED 1-2
2007, LED centennial 1-9

End of book