Optical Diodes and Semiconductor Lasers

and

Jacquetta Marchand

Table of Contents

Chapter 1

Laser Diode

A packaged laser diode with penny for scale.
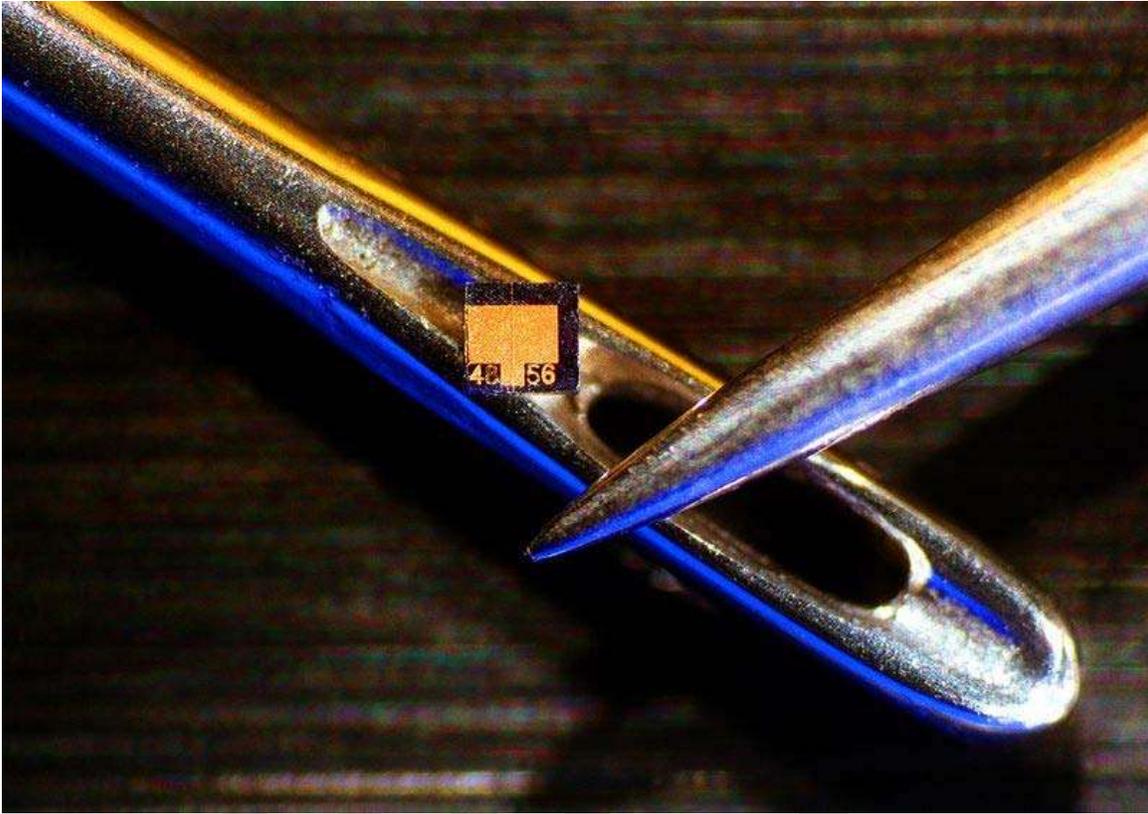
Image of the actual laser diode chip (shown on the eye of a needle for scale) contained within the package shown in the above image.

This is a visible light micrograph of a laser diode taken from a CD-ROM drive. Visible are the P and N layers distinguished by different colours. Also visible are scattered glass fragments from a broken collimating lens.

The **laser diode** is a laser where the active medium is a semiconductor similar to that found in a light-emitting diode. The most common type of laser diode is formed from a p-n junction and powered by injected electric current. The former devices are sometimes referred to as *injection laser diodes* to distinguish them from *optically pumped laser diodes*.
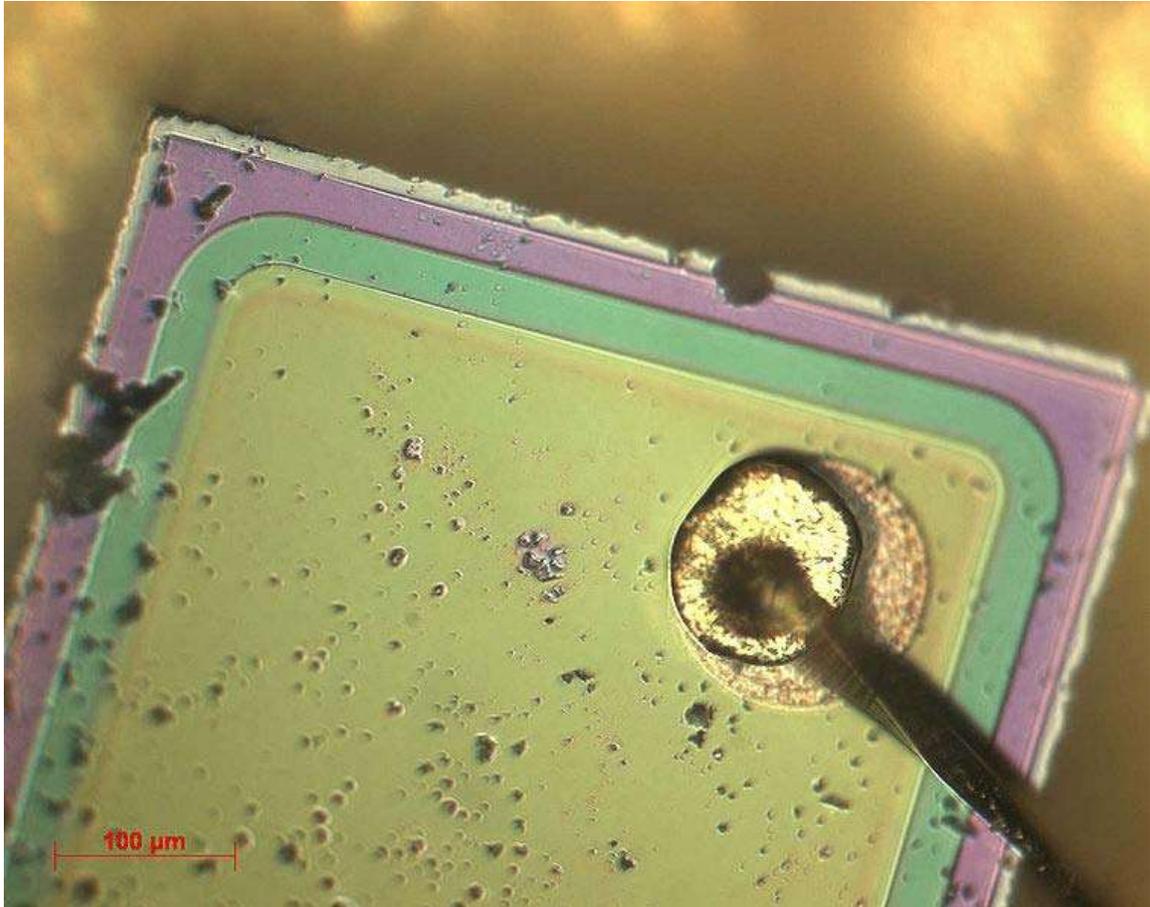
Theory of operation

A laser diode is formed by doping a very thin layer on the surface of a crystal wafer. The crystal is doped to produce an n-type region and a p-type region, one above the other, resulting in a *p-n* junction, or diode.

Laser diodes form a subset of the larger classification of semiconductor *p-n* junction diodes. Forward electrical bias across the laser diode causes the two species of charge carrier – holes and electrons – to be "injected" from opposite sides of the *p-n* junction into the depletion region. Holes are injected from the *p*-doped, and electrons from the *n-*

doped, semiconductor. (A depletion region, devoid of any charge carriers, forms as a result of the difference in electrical potential between n- and p-type semiconductors wherever they are in physical contact.) Due to the use of charge injection in powering most diode lasers, this class of lasers is sometimes termed "injection lasers," or "injection laser diode" (ILD). As diode lasers are semiconductor devices, they may also be classified as semiconductor lasers. Either designation distinguishes diode lasers from solid-state lasers.

Another method of powering some diode lasers is the use of optical pumping. Optically Pumped Semiconductor Lasers (OPSL) use a III-V semiconductor chip as the gain media, and another laser (often another diode laser) as the pump source. OPSL offer several advantages over ILDs, particularly in wavelength selection and lack of interference from internal electrode structures.

When an electron and a hole are present in the same region, they may recombine or "annihilate" with the result being spontaneous emission — i.e., the electron may re-occupy the energy state of the hole, emitting a photon with energy equal to the difference between the electron and hole states involved. (In a conventional semiconductor junction diode, the energy released from the recombination of electrons and holes is carried away as phonons, i.e., lattice vibrations, rather than as photons.) Spontaneous emission gives the laser diode below lasing threshold similar properties to an LED. Spontaneous emission is necessary to initiate laser oscillation, but it is one among several sources of inefficiency once the laser is oscillating.

The difference between the photon-emitting semiconductor laser and conventional phonon-emitting (non-light-emitting) semiconductor junction diodes lies in the use of a different type of semiconductor, one whose physical and atomic structure confers the possibility for photon emission. These photon-emitting semiconductors are the so-called "direct bandgap" semiconductors. The properties of silicon and germanium, which are single-element semiconductors, have bandgaps that do not align in the way needed to allow photon emission and are not considered "direct." Other materials, the so-called compound semiconductors, have virtually identical crystalline structures as silicon or germanium but use alternating arrangements of two different atomic species in a checkerboard-like pattern to break the symmetry. The transition between the materials in the alternating pattern creates the critical "direct bandgap" property. Gallium arsenide, indium phosphide, gallium antimonide, and gallium nitride are all examples of compound semiconductor materials that can be used to create junction diodes that emit light.
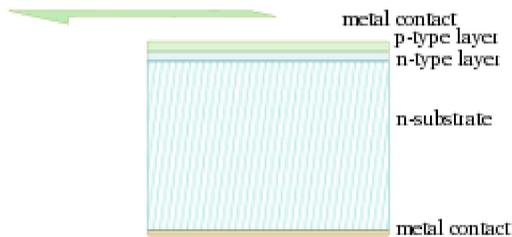
Diagram (not to scale) of a simple laser diode, such as shown above.

In the absence of stimulated emission (e.g., lasing) conditions, electrons and holes may coexist in proximity to one another, without recombining, for a certain time, termed the "upper-state lifetime" or "recombination time" (about a nanosecond for typical diode laser materials), before they recombine. Then a nearby photon with energy equal to the recombination energy can cause recombination by stimulated emission. This generates another photon of the same frequency, travelling in the same direction, with the same polarization and phase as the first photon. This means that stimulated emission causes gain in an optical wave (of the correct wavelength) in the injection region, and the gain increases as the number of electrons and holes injected across the junction increases. The spontaneous and stimulated emission processes are vastly more efficient in direct bandgap semiconductors than in indirect bandgap semiconductors; therefore silicon is not a common material for laser diodes.

As in other lasers, the gain region is surrounded with an optical cavity to form a laser. In the simplest form of laser diode, an optical waveguide is made on that crystal surface, such that the light is confined to a relatively narrow line. The two ends of the crystal are cleaved to form perfectly smooth, parallel edges, forming a Fabry–Pérot resonator. Photons emitted into a mode of the waveguide will travel along the waveguide and be reflected several times from each end face before they are emitted. As a light wave passes through the cavity, it is amplified by stimulated emission, but light is also lost due to absorption and by incomplete reflection from the end facets. Finally, if there is more amplification than loss, the diode begins to "lase".

Some important properties of laser diodes are determined by the geometry of the optical cavity. Generally, in the vertical direction, the light is contained in a very thin layer, and the structure supports only a single optical mode in the direction perpendicular to the layers. In the lateral direction, if the waveguide is wide compared to the wavelength of light, then the waveguide can support multiple lateral optical modes, and the laser is known as "multi-mode". These laterally multi-mode lasers are adequate in cases where one needs a very large amount of power, but not a small diffraction-limited beam; for example in printing, activating chemicals, or pumping other types of lasers.

In applications where a small focused beam is needed, the waveguide must be made narrow, on the order of the optical wavelength. This way, only a single lateral mode is

supported and one ends up with a diffraction-limited beam. Such single spatial mode devices are used for optical storage, laser pointers, and fiber optics. Note that these lasers may still support multiple longitudinal modes, and thus can lase at multiple wavelengths simultaneously.

The wavelength emitted is a function of the band-gap of the semiconductor and the modes of the optical cavity. In general, the maximum gain will occur for photons with energy slightly above the band-gap energy, and the modes nearest the gain peak will lase most strongly. If the diode is driven strongly enough, additional *side modes* may also lase. Some laser diodes, such as most visible lasers, operate at a single wavelength, but that wavelength is unstable and changes due to fluctuations in current or temperature.

Due to diffraction, the beam diverges (expands) rapidly after leaving the chip, typically at 30 degrees vertically by 10 degrees laterally. A lens must be used in order to form a collimated beam like that produced by a laser pointer. If a circular beam is required, cylindrical lenses and other optics are used. For single spatial mode lasers, using symmetrical lenses, the collimated beam ends up being elliptical in shape, due to the difference in the vertical and lateral divergences. This is easily observable with a red laser pointer.

The simple diode described above has been heavily modified in recent years to accommodate modern technology, resulting in a variety of types of laser diodes, as described below.

Types

The simple laser diode structure, described above, is extremely inefficient. Such devices require so much power that they can only achieve pulsed operation without damage. Although historically important and easy to explain, such devices are not practical.

Double heterostructure lasers
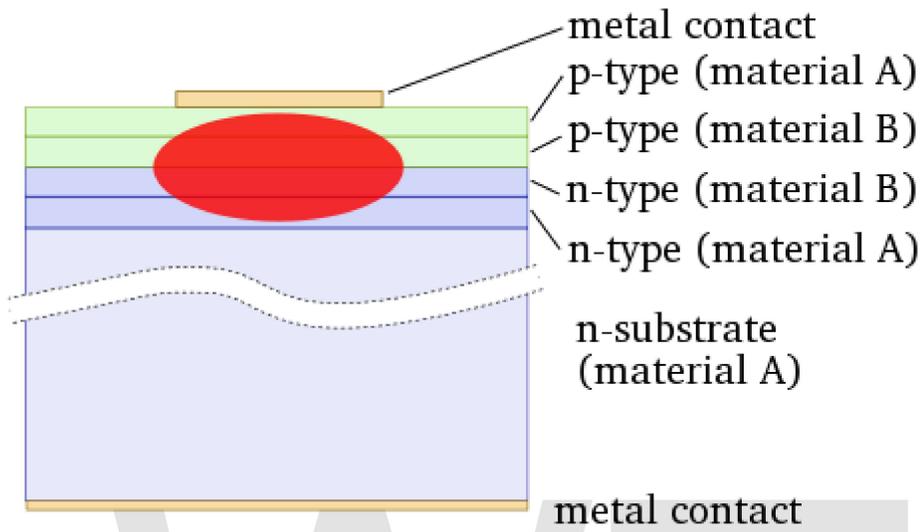
Diagram of front view of a double heterostructure laser diode (not to scale)

In these devices, a layer of low bandgap material is sandwiched between two high bandgap layers. One commonly-used pair of materials is gallium arsenide (GaAs) with aluminium gallium arsenide ($Al_xGa_{(1-x)}As$). Each of the junctions between different bandgap materials is called a *heterostructure*, hence the name "double heterostructure laser" or *DH* laser. The kind of laser diode described in the first part, may be referred to as a *homojunction* laser, for contrast with these more popular devices.

The advantage of a DH laser is that the region where free electrons and holes exist simultaneously—the active region—is confined to the thin middle layer. This means that many more of the electron-hole pairs can contribute to amplification—not so many are left out in the poorly amplifying periphery. In addition, light is reflected from the heterojunction; hence, the light is confined to the region where the amplification takes place.

Quantum well lasers

metal contact
p-type (material A)
quantum well (material B)
n-type (material A)
n-substrate (material A)
metal contact
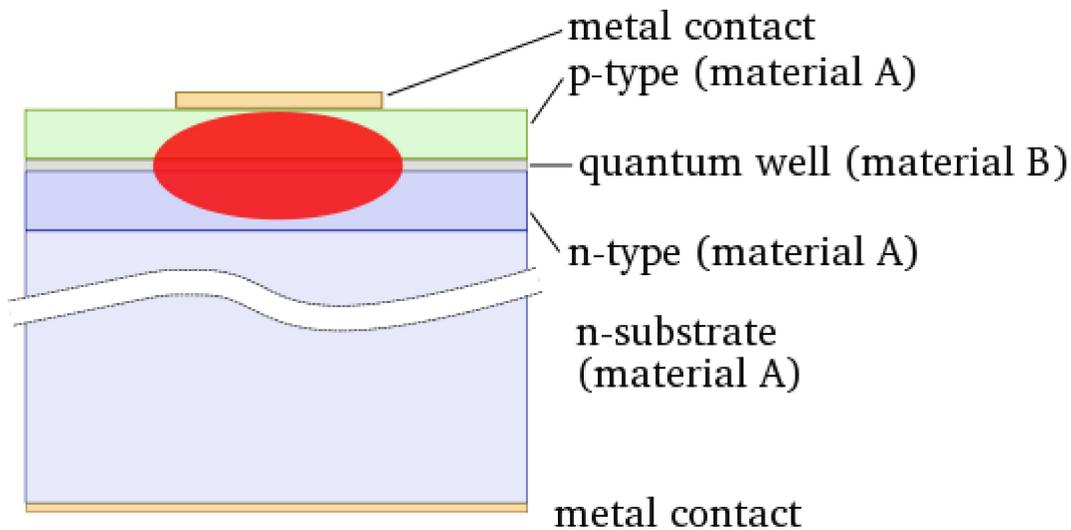
Diagram of front view of a simple quantum well laser diode (not to scale)

If the middle layer is made thin enough, it acts as a quantum well. This means that the vertical variation of the electron's wavefunction, and thus a component of its energy, is quantized. The efficiency of a quantum well laser is greater than that of a bulk laser because the density of states function of electrons in the quantum well system has an abrupt edge that concentrates electrons in energy states that contribute to laser action.

Lasers containing more than one quantum well layer are known as *multiple quantum well* lasers. Multiple quantum wells improve the overlap of the gain region with the optical waveguide mode.

Further improvements in the laser efficiency have also been demonstrated by reducing the quantum well layer to a quantum wire or to a "sea" of quantum dots.

Quantum cascade lasers

In a quantum cascade laser, the difference between quantum well energy levels is used for the laser transition instead of the bandgap. This enables laser action at relatively long wavelengths, which can be tuned simply by altering the thickness of the layer. They are heterojunction lasers.

Separate confinement heterostructure lasers

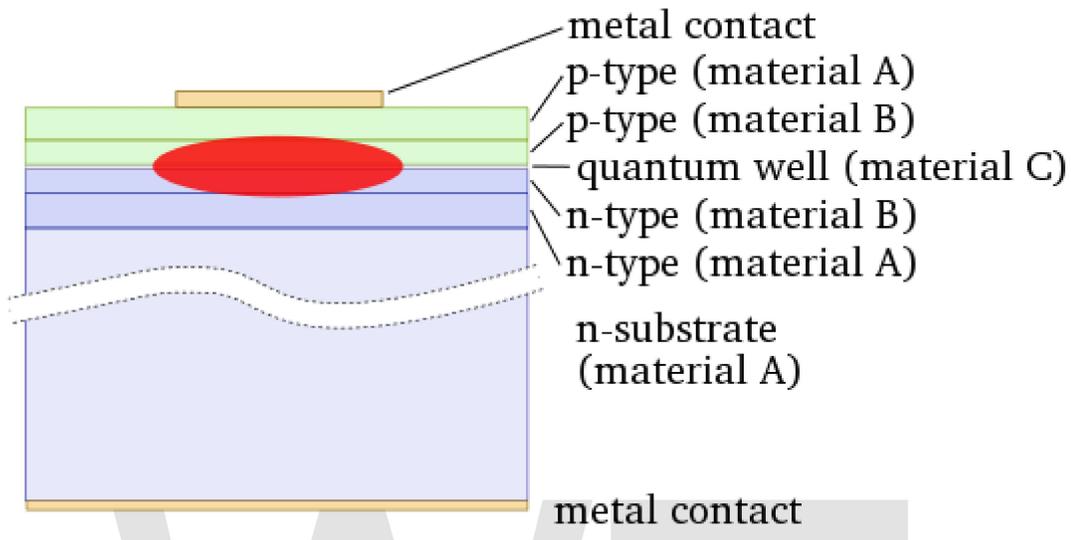Diagram of front view of a separate confinement heterostructure quantum well laser diode

The problem with the simple quantum well diode described above is that the thin layer is simply too small to effectively confine the light. To compensate, another two layers are added on, outside the first three. These layers have a lower refractive index than the centre layers, and hence confine the light effectively. Such a design is called a separate confinement heterostructure (SCH) laser diode.

Almost all commercial laser diodes since the 1990s have been SCH quantum well diodes.

Distributed feedback lasers

Distributed feedback lasers (DFB) are the most common transmitter type in DWDM-systems. To stabilize the lasing wavelength, a diffraction grating is etched close to the p-n junction of the diode. This grating acts like an optical filter, causing a single wavelength to be fed back to the gain region and lase. Since the grating provides the feedback that is required for lasing, reflection from the facets is not required. Thus, at least one facet of a DFB is anti-reflection coated. The DFB laser has a stable wavelength that is set during manufacturing by the pitch of the grating, and can only be tuned slightly with temperature. DFB lasers are widely used in optical communication applications where a precise and stable wavelength is critical.
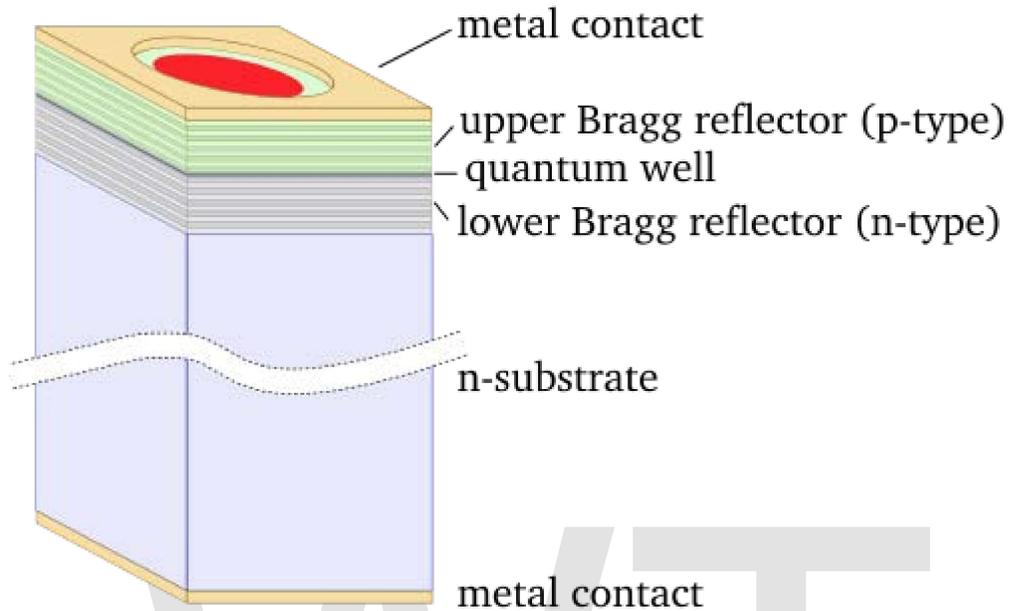
VCSELs

Diagram of a simple VCSEL structure

Vertical-cavity surface-emitting lasers (VCSELs) have the optical cavity axis along the direction of current flow rather than perpendicular to the current flow as in conventional laser diodes. The active region length is very short compared with the lateral dimensions so that the radiation emerges from the surface of the cavity rather than from its edge as shown in the figure. The reflectors at the ends of the cavity are dielectric mirrors made from alternating high and low refractive index quarter-wave thick multilayer.

Such dielectric mirrors provide a high degree of wavelength-selective reflectance at the required free surface wavelength λ if the thicknesses of alternating layers d_1 and d_2 with refractive indices n_1 and n_2 are such that $n_1 d_1 + n_2 d_2 = \lambda/2$ which then leads to the constructive interference of all partially reflected waves at the interfaces. But there is a disadvantage: because of the high mirror reflectivities, VCSELs have lower output powers when compared to edge-emitting lasers.

There are several advantages to producing VCSELs when compared with the production process of edge-emitting lasers. Edge-emitters cannot be tested until the end of the production process. If the edge-emitter does not work, whether due to bad contacts or poor material growth quality, the production time and the processing materials have been wasted. Additionally, because VCSELs emit the beam perpendicular to the active region of the laser as opposed to parallel as with an edge emitter, tens of thousands of VCSELs can be processed simultaneously on a three inch Gallium Arsenide wafer. Furthermore, even though the VCSEL production process is more labor and material intensive, the

yield can be controlled to a more predictable outcome. However, they normally show a lower power output level.

VECSELs

Vertical external-cavity surface-emitting lasers, or VECSELs, are similar to VCSELs. In VCSELs, the mirrors are typically grown epitaxially as part of the diode structure, or grown separately and bonded directly to the semiconductor containing the active region. VECSELs are distinguished by a construction in which one of the two mirrors is external to the diode structure. As a result, the cavity includes a free-space region. A typical distance from the diode to the external mirror would be 1 cm.

One of the most interesting features of any VECSEL is the small thickness of the semiconductor gain region in the direction of propagation, less than 100 nm. In contrast, a conventional in-plane semiconductor laser entails light propagation over distances of from 250 µm upward to 2 mm or longer. The significance of the short propagation distance is that it causes the effect of "antiguiding" nonlinearities in the diode laser gain region to be minimized. The result is a large-cross-section single-mode optical beam which is not attainable from in-plane ("edge-emitting") diode lasers.

Several workers demonstrated optically pumped VECSELs, and they continue to be developed for many applications including high power sources for use in industrial machining (cutting, punching, etc.) because of their unusually high power and efficiency when pumped by multi-mode diode laser bars.

Electrically pumped VECSELs have also been demonstrated. Applications for electrically pumped VECSELs include projection displays, served by frequency doubling of near-IR VECSEL emitters to produce blue and green light.

External-cavity diode lasers

External-cavity diode lasers are tunable lasers which use mainly double heterostructures diodes of the $Al_xGa_{(1-x)}As$ type. The first external-cavity diode lasers used intracavity etalons and simple tuning Littrow gratings. Other designs include gratings in grazing-incidence configuration and multiple-prism grating configurations.

Failure modes

Laser diodes have the same reliability and failure issues as light emitting diodes. In addition they are subject to *catastrophic optical damage* (COD) when operated at higher power.

Many of the advances in reliability of diode lasers in the last 20 years remain proprietary to their developers. The reliability of a laser diode can make or break a product line. Moreover, "reverse engineering" is not always able to reveal the differences between more-reliable and less-reliable diode laser products.

At the edge of a diode laser, where light is emitted, a mirror is traditionally formed by cleaving the semiconductor wafer to form a specularly reflecting plane. This approach is facilitated by the weakness of the crystallographic plane in III-V semiconductor crystals (such as GaAs, InP, GaSb, etc.) compared to other planes. A scratch made at the edge of the wafer and a slight bending force causes a nearly atomically perfect mirror-like cleavage plane to form and propagate in a straight line across the wafer.

But it so happens that the atomic states at the cleavage plane are altered (compared to their bulk properties within the crystal) by the termination of the perfectly periodic lattice at that plane. Surface states at the cleaved plane, have energy levels within the (otherwise forbidden) bandgap of the semiconductor.

Essentially, as a result when light propagates through the cleavage plane and transits to free space from within the semiconductor crystal, a fraction of the light energy is absorbed by the surface states whence it is converted to heat by phonon-electron interactions. This heats the cleaved mirror. In addition the mirror may heat simply because the edge of the diode laser—which is electrically pumped—is in less-than-perfect contact with the mount that provides a path for heat removal. The heating of the mirror causes the bandgap of the semiconductor to shrink in the warmer areas. The bandgap shrinkage brings more electronic band-to-band transitions into alignment with the photon energy causing yet more absorption. This is thermal runaway, a form of positive feedback, and the result can be melting of the facet, known as *catastrophic optical damage*, or COD.

In the 1970s this problem, which is particularly nettlesome for GaAs-based lasers emitting between 1 μm and 0.630 μm wavelengths (less so for InP based lasers used for long-haul telecommunications which emit between 1.3 μm and 2 μm), was identified. Michael Ettenberg, a researcher and later Vice President at RCA Laboratories' David Sarnoff Research Center in Princeton, New Jersey, devised a solution. A thin layer of aluminum oxide was deposited on the facet. If the aluminum oxide thickness is chosen correctly it functions as an anti-reflective coating, reducing reflection at the surface. This alleviated the heating and COD at the facet.

Since then, various other refinements have been employed. One approach is to create a so-called non-absorbing mirror (NAM) such that the final 10 μm or so before the light emits from the cleaved facet are rendered non-absorbing at the wavelength of interest.

In the very early 1990s, SDL, Inc. began supplying high power diode lasers with good reliability characteristics. CEO Donald Scifres and CTO David Welch presented new reliability performance data at, e.g., SPIE Photonics West conferences of the era. The methods used by SDL to defeat COD were considered to be highly proprietary and have still not been disclosed publicly as of June, 2006.

In the mid-1990s IBM Research (Ruschlikon, Switzerland) announced that it had devised its so-called "E2 process" which conferred extraordinary resistance to COD in GaAs-based lasers. This process, too, has never been disclosed as of June, 2006.

Reliability of high-power diode laser pump bars (employed to pump solid state lasers) remains a difficult problem in a variety of applications, in spite of these proprietary advances. Indeed, the physics of diode laser failure is still being worked out and research on this subject remains active, if proprietary.

Extension of the lifetime of laser diodes is critical to their continued adaptation to a wide variety of applications.

Applications of laser diodes
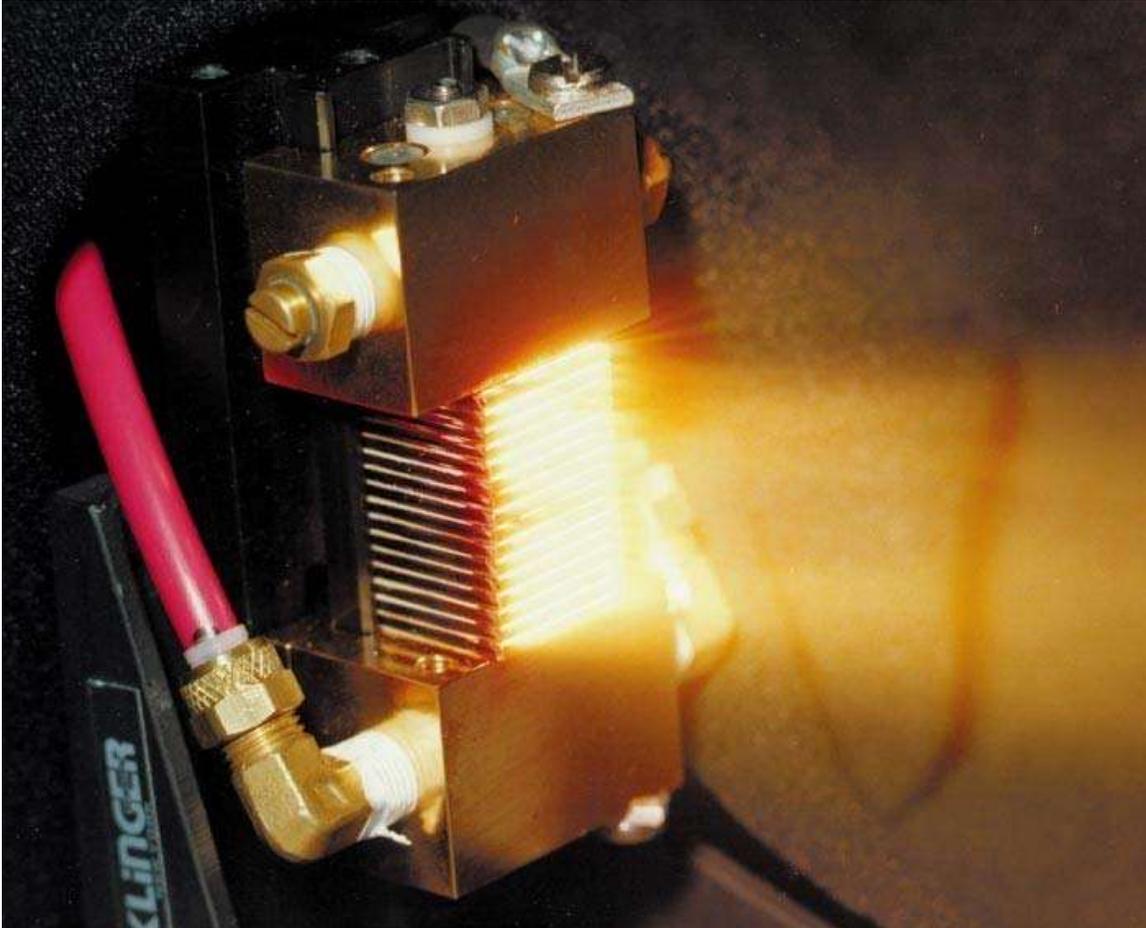
Laser diodes can be arrayed to produce very high power (continuous wave or pulsed) outputs. Such arrays may be used to efficiently pump solid state lasers for inertial confinement fusion or high average power drilling or burning applications.

Laser diodes are numerically the most common type of laser, with 2004 sales of approximately 733 million diode lasers, as compared to 131,000 of other types of lasers.

Laser diodes find wide use in telecommunication as easily modulated and easily coupled light sources for fiber optics communication. They are used in various measuring instruments, such as rangefinders. Another common use is in barcode readers. Visible lasers, typically red but later also green, are common as laser pointers. Both low and

high-power diodes are used extensively in the printing industry both as light sources for scanning (input) of images and for very high-speed and high-resolution printing plate (output) manufacturing. Infrared and red laser diodes are common in CD players, CD-ROMs and DVD technology. Violet lasers are used in HD DVD and Blu-ray technology. Diode lasers have also found many applications in laser absorption spectrometry (LAS) for high-speed, low-cost assessment or monitoring of the concentration of various species in gas phase. High-power laser diodes are used in industrial applications such as heat treating, cladding, seam welding and for pumping other lasers, such as diode pumped solid state lasers.

Applications of laser diodes can be categorized in various ways. Most applications could be served by larger solid state lasers or optical parametric oscillators, but the low cost of mass-produced diode lasers makes them essential for mass-market applications. Diode lasers can be used in a great many fields; since light has many different properties (power, wavelength and spectral quality, beam quality, polarization, etc.) it is interesting to classify applications by these basic properties.

Many applications of diode lasers primarily make use of the "directed energy" property of an optical beam. In this category one might include the laser printers, bar-code readers, image scanning, illuminators, designators, optical data recording, combustion ignition, laser surgery, industrial sorting, industrial machining, and directed energy weaponry. Some of these applications are emerging while others are well-established.

Laser medicine: medicine and especially dentistry have found many new applications for diode lasers. The shrinking size of the units and their increasing user friendliness makes them very attractive to clinicians for minor soft tissue procedures. The 800 nm – 980 nm units have a high absorption rate for hemoglobin and thus make them ideal for soft tissue applications, where good hemostasis is necessary.

Applications which may make use of the coherence of diode-laser-generated light include interferometric distance measurement, holography, coherent communications, and coherent control of chemical reactions.

Applications which may make use of "narrow spectral" properties of diode lasers include range-finding, telecommunications, infra-red countermeasures, spectroscopic sensing, generation of radio-frequency or terahertz waves, atomic clock state preparation, quantum key cryptography, frequency doubling and conversion, water purification (in the UV), and photodynamic therapy (where a particular wavelength of light would cause a substance such as porphyrin to become chemically active as an anti-cancer agent only where the tissue is illuminated by light).

Applications where the desired quality of laser diodes is their ability to generate ultra-short pulses of light by the technique known as "mode-locking" include clock distribution for high-performance integrated circuits, high-peak-power sources for laser-induced breakdown spectroscopy sensing, arbitrary waveform generation for radio-frequency

waves, photonic sampling for analog-to-digital conversion, and optical code-division-multiple-access systems for secure communication.

Common wavelengths

- **375 nm** – excitation of Hoechst stain, Calcium Blue, and other fluorescent dyes in fluorescence microscopy
- **405 nm** – InGaN blue-violet laser, in Blu-ray Disc and HD DVD drives
- **445 nm** – InGaN Deep blue laser multimode diode recently introduced (2010) for use in mercury free high brightness data projectors
- **473 nm** – Bright blue laser pointers, still very expensive, output of DPSS systems
- **485 nm** – excitation of GFP and other fluorescent dyes
- **510 nm** - Green diodes recently (2010) developed by Nichia for laser projectors.
- **532 nm** – AlGaAs-pumped bright green laser pointers, frequency doubled 1064 nm Nd:YAG laser or (more commonly in laser pointers) Nd:YVO$_4$ IR lasers (SHG)
- **593 nm** – Yellow-Orange laser pointers, DPSS
- **635 nm** – AlGaInP better red laser pointers, same power subjectively 5 times as bright as 670 nm one
- **640 nm** – High brightness red DPSS laser pointers
- **657 nm** – AlGaInP DVD drives, laser pointers
- **670 nm** – AlGaInP cheap red laser pointers
- **760 nm** – AlGaInP gas sensing: O$_2$
- **785 nm** – GaAlAs Compact Disc drives
- **808 nm** – GaAlAs pumps in DPSS Nd:YAG lasers (e.g. in green laser pointers or as arrays in higher-powered lasers)
- **848 nm** – laser mice
- **980 nm** – InGaAs pump for optical amplifiers, for Yb:YAG DPSS lasers
- **1064 nm** – AlGaAs fiber-optic communication
- **1310 nm** – InGaAsP fiber-optic communication
- **1480 nm** – InGaAsP pump for optical amplifiers
- **1512 nm** – InGaAsP gas sensing: NH$_3$
- **1550 nm** – InGaAsP fiber-optic communication
- **1625 nm** – InGaAsP fiber-optic communication, service channel
- **1654 nm** – InGaAsP gas sensing: CH$_4$
- **1877 nm** – GaSbAs gas sensing: H$_2$O
- **2004 nm** – GaSbAs gas sensing: CO$_2$
- **2330 nm** – GaSbAs gas sensing: CO
- **2680 nm** – GaSbAs gas sensing: CO$_2$

History

Coherent light emission from a semiconductor (gallium arsenide) diode (the first *laser* diode) was demonstrated in 1962 by two US groups led by Robert N. Hall at the General Electric research center and by Marshall Nathan at the IBM T.J. Watson Research Center. The priority is given to General Electric group who have obtained and submitted their

results earlier; they also went further and made a resonant cavity for their diode. The first visible wavelength laser diode was demonstrated by Nick Holonyak, Jr. later in 1962.

Other teams at MIT Lincoln Laboratory, Texas Instruments, and RCA Laboratories were also involved in and received credit for their historic initial demonstrations of efficient light emission and lasing in semiconductor diodes in 1962 and thereafter. GaAs lasers were also produced in early 1963 in the Soviet Union by the team led by Nikolay Basov.

In the early 1960s liquid phase epitaxy (LPE) was invented by Herbert Nelson of RCA Laboratories. By layering the highest quality crystals of varying compositions, it enabled the demonstration of the highest quality heterojunction semiconductor laser materials for many years. LPE was adopted by all the leading laboratories, worldwide and used for many years. It was finally supplanted in the 1970s by molecular beam epitaxy and organometallic chemical vapor deposition.

Diode lasers of that era operated with threshold current densities of 1000 A/cm^2 at 77 K temperatures. Such performance enabled continuous-lasing to be demonstrated in the earliest days. However, when operated at room temperature, about 300 K, threshold current densities were two orders of magnitude greater, or 100,000 A/cm^2 in the best devices. The dominant challenge for the remainder of the 1960s was to obtain low threshold current density at 300 K and thereby to demonstrate continuous-wave lasing at room temperature from a diode laser.

The first diode lasers were homojunction diodes. That is, the material (and thus the bandgap) of the waveguide core layer and that of the surrounding clad layers, were identical. It was recognized that there was an opportunity, particularly afforded by the use of liquid phase epitaxy using aluminum gallium arsenide, to introduce heterojunctions. Heterostructures consist of layers of semiconductor crystal having varying bandgap and refractive index. Heterojunctions (formed from heterostructures) had been recognized by Herbert Kroemer, while working at RCA Laboratories in the mid-1950s, as having unique advantages for several types of electronic and optoelectronic devices including diode lasers. LPE afforded the technology of making heterojunction diode lasers.

The first heterojunction diode lasers were single-heterojunction lasers. These lasers utilized aluminum gallium arsenide *p*-type injectors situated over *n*-type gallium arsenide layers grown on the substrate by LPE. An admixture of aluminum replaced gallium in the semiconductor crystal and raised the bandgap of the *p*-type injector over that of the *n*-type layers beneath. It worked; the 300 K threshold currents went down by 10× to 10,000 amperes per square centimeter. Unfortunately, this was still not in the needed range and these single-heterostructure diode lasers did not function in continuous wave operation at room temperature.

The innovation that met the room temperature challenge was the double heterostructure laser. The trick was to quickly move the wafer in the LPE apparatus between different "melts" of aluminum gallium arsenide (*p*- and *n*-type) and a third melt of gallium arsenide. It had to be done rapidly since the gallium arsenide core region needed to be

significantly under 1 μm in thickness. This may have been the earliest true example of "nanotechnology." The first laser diode to achieve *continuous wave* operation was a double heterostructure demonstrated in 1970 essentially simultaneously by Zhores Alferov and collaborators (including Dmitri Z. Garbuzov) of the Soviet Union, and Morton Panish and Izuo Hayashi working in the United States. However, it is widely accepted that Zhores I. Alferov and team reached the milestone first.

Chapter 2

Photodiode
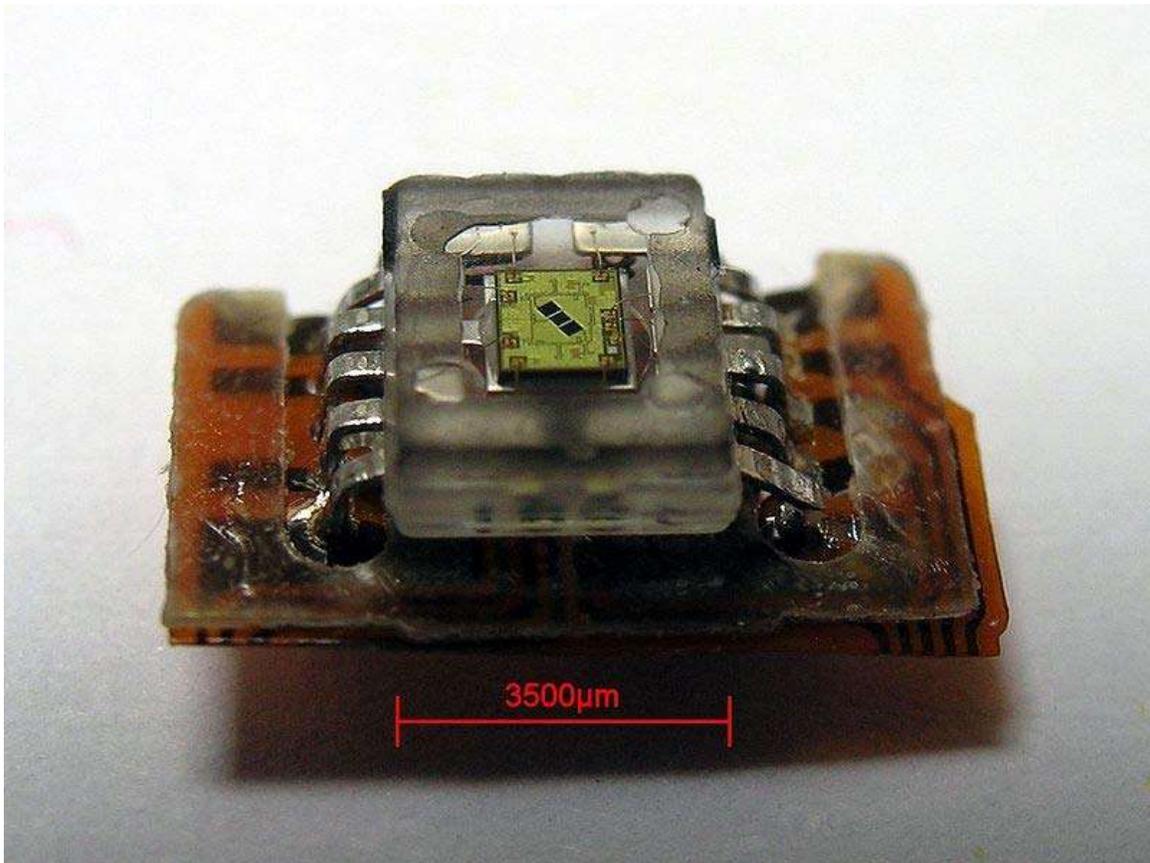
Photodetector from a CD-ROM Drive. 3 photodiodes are visible.

Symbol for photodiode.

A **photodiode** is a type of photodetector capable of converting light into either current or voltage, depending upon the mode of operation. The common, traditional solar cell used to generate electric solar power is a large area photodiode.

Photodiodes are similar to regular semiconductor diodes except that they may be either exposed (to detect vacuum UV or X-rays) or packaged with a window or optical fiber connection to allow light to reach the sensitive part of the device. Many diodes designed for use specifically as a photodiode will also use a PIN junction rather than the typical PN junction.

Principle of operation

A photodiode is a PN junction or PIN structure. When a photon of sufficient energy strikes the diode, it excites an electron, thereby creating a *free electron* and a (positively charged electron *hole*). This mechanism is also known as the photoelectric effect. If the absorption occurs in the junction's depletion region, or one diffusion length away from it, these carriers are swept from the junction by the built-in field of the depletion region. Thus holes move toward the anode, and electrons toward the cathode, and a photocurrent is produced. This photocurrent is the sum of both the dark current (without light) and the light current, so the dark current must be minimised to enhance the sensitivity of the device.

Photovoltaic mode

When used in zero bias or *photovoltaic mode*, the flow of photocurrent out of the device is restricted and a voltage builds up. This mode exploits the photovoltaic effect, which is the basis for solar cells – in fact, a traditional solar cell is just a large area photodiode.

Photoconductive mode

In this mode the diode is often reverse biased, dramatically reducing the response time at the expense of increased noise. This increases the width of the depletion layer, which decreases the junction's capacitance resulting in faster response times. The reverse bias induces only a small amount of current (known as saturation or back current) along its direction while the photocurrent remains virtually the same. For a given spectral distribution, the photocurrent is linearly proportional to the illuminance (and to the irradiance).

Although this mode is faster, the photoconductive mode tends to exhibit more electronic noise. The leakage current of a good PIN diode is so low (< 1nA) that the Johnson–Nyquist noise of the load resistance in a typical circuit often dominates.

Other modes of operation

Avalanche photodiodes have a similar structure to regular photodiodes, but they are operated with much higher reverse bias. This allows each *photo-generated* carrier to be

multiplied by avalanche breakdown, resulting in internal gain within the photodiode, which increases the effective *responsivity* of the device.

Phototransistors also consist of a photodiode with internal gain. A phototransistor is in essence nothing more than a bipolar transistor that is encased in a transparent case so that light can reach the *base-collector junction*. The electrons that are generated by photons in the base-collector junction are injected into the base, and this photodiode current is amplified by the transistor's current gain β (or h_{fe}). Note that while phototransistors have a higher responsivity for light they are not able to detect low levels of light any better than photodiodes. Phototransistors also have significantly longer response times.

Materials

The material used to make a photodiode is critical to defining its properties, because only photons with sufficient energy to excite electrons across the material's bandgap will produce significant photocurrents.

Materials commonly used to produce photodiodes include:

Material	Electromagnetic spectrum wavelength range (nm)
Silicon	190 – 1100
Germanium	400 – 1700
Indium gallium arsenide	800 – 2600
Lead(II) sulfide	<1000 – 3500

Because of their greater bandgap, silicon-based photodiodes generate less noise than germanium-based photodiodes, but germanium photodiodes must be used for wavelengths longer than approximately 1 μm.

Unwanted photodiodes

Since transistors and ICs are made of semiconductors, and contain P-N junctions, almost every active component is potentially a photodiode. Many components, especially those sensitive to small currents, will not work correctly if illuminated, due to the induced photocurrents. In most components this is not desired, so they are placed in an opaque housing. Since housings are not completely opaque to X-rays or other high energy radiation, these can still cause many ICs to malfunction due to induced photo-currents.

Features
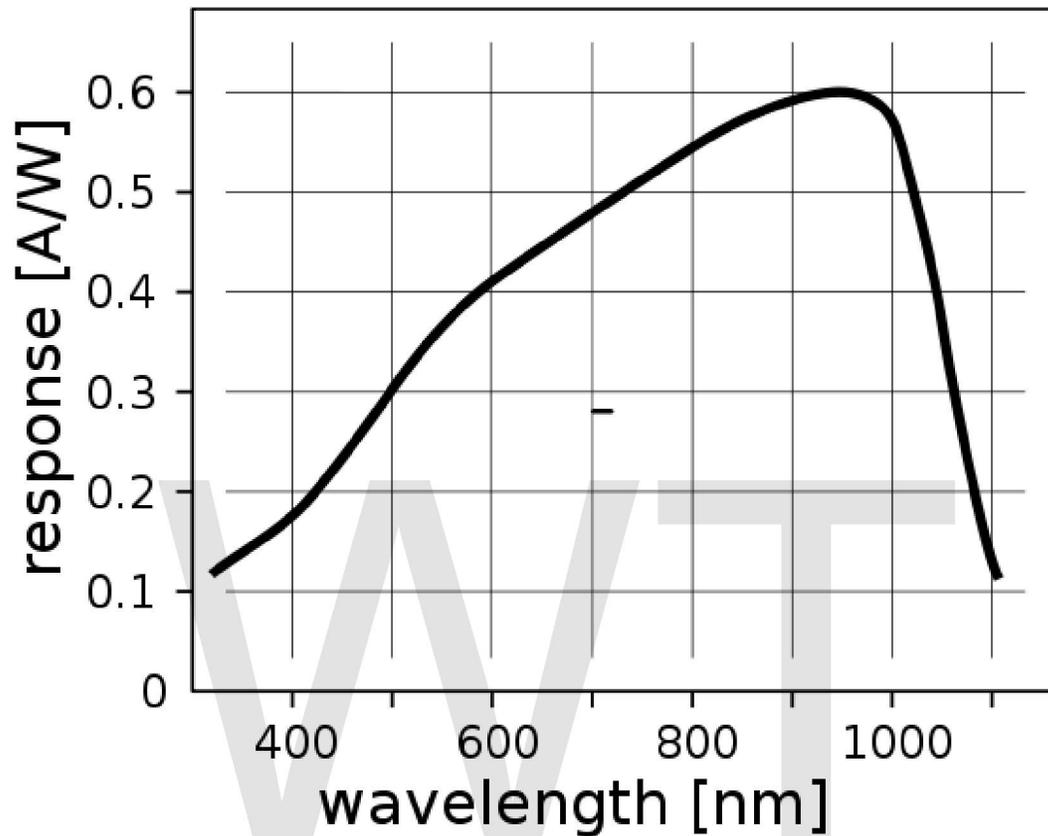
Response of a silicon photo diode vs wavelength of the incident light

Critical performance parameters of a photodiode include:

Responsivity
> The ratio of generated photocurrent to incident light power, typically expressed in A/W when used in photoconductive mode. The responsivity may also be expressed as a *Quantum efficiency*, or the ratio of the number of photogenerated carriers to incident photons and thus a unitless quantity.

Dark current
> The current through the photodiode in the absence of light, when it is operated in photoconductive mode. The dark current includes photocurrent generated by background radiation and the saturation current of the semiconductor junction. Dark current must be accounted for by calibration if a photodiode is used to make an accurate optical power measurement, and it is also a source of noise when a photodiode is used in an optical communication system.

Noise-equivalent power

(NEP) The minimum input optical power to generate photocurrent, equal to the rms noise current in a 1 hertz bandwidth. The related characteristic *detectivity* (D) is the inverse of NEP, 1/NEP; and the *specific detectivity* ($D^\star$) is the detectivity normalized to the area (A) of the photodetector, $D^\star = D\sqrt{A}$. The NEP is roughly the minimum detectable input power of a photodiode.

When a photodiode is used in an optical communication system, these parameters contribute to the *sensitivity* of the optical receiver, which is the minimum input power required for the receiver to achieve a specified *bit error ratio*.

Applications

P-N photodiodes are used in similar applications to other photodetectors, such as photoconductors, charge-coupled devices, and photomultiplier tubes.

Photodiodes are used in consumer electronics devices such as compact disc players, smoke detectors, and the receivers for remote controls in VCRs and televisions.

In other consumer items such as camera light meters, clock radios (the ones that dim the display when it's dark) and street lights, photoconductors are often used rather than photodiodes, although in principle either could be used.

Photodiodes are often used for accurate measurement of light intensity in science and industry. They generally have a better, more linear response than photoconductors.

They are also widely used in various medical applications, such as detectors for computed tomography (coupled with scintillators) or instruments to analyze samples (immunoassay). They are also used in pulse oximeters.

PIN diodes are much faster and more sensitive than ordinary p-n junction diodes, and hence are often used for optical communications and in lighting regulation.

P-N photodiodes are not used to measure extremely low light intensities. Instead, if high sensitivity is needed, avalanche photodiodes, intensified charge-coupled devices or photomultiplier tubes are used for applications such as astronomy, spectroscopy, night vision equipment and laser rangefinding.

Comparison with photomultipliers

Advantages compared to photomultipliers:

1. Excellent linearity of output current as a function of incident light
2. Spectral response from 190 nm to 1100 nm (silicon), longer wavelengths with other semiconductor materials
3. Low noise
4. Ruggedized to mechanical stress

5. Low cost
6. Compact and light weight
7. Long lifetime
8. High quantum efficiency, typically 80%
9. No high voltage required

Disadvantages compared to photomultipliers:

1. Small area
2. No internal gain (except avalanche photodiodes, but their gain is typically 10^2–10^3 compared to up to 10^8 for the photomultiplier)
3. Much lower overall sensitivity
4. Photon counting only possible with specially designed, usually cooled photodiodes, with special electronic circuits
5. Response time for many designs is slower

Photodiode array

A one-dimensional array of hundreds or thousands of photodiodes can be used as a position sensor, for example as part of an angle sensor. One advantage of photodiode arrays (PDAs) is that they allow for high speed parallel read out since the driving electronics may not be built in like a traditional CMOS or CCD sensor.

Chapter 3

Organic Light-Emitting Diode
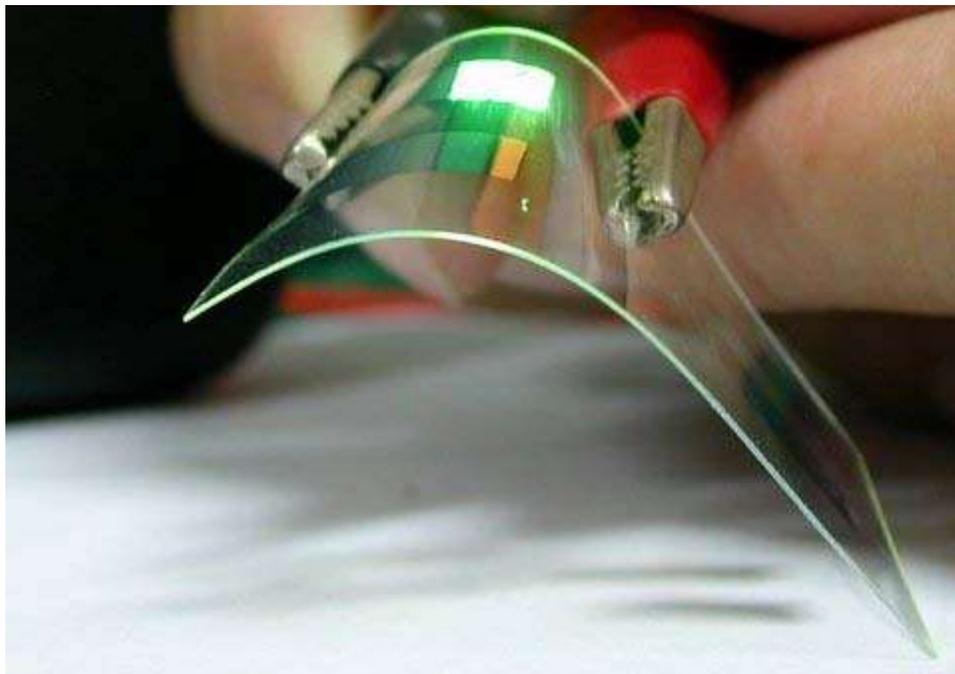
Demonstration of a flexible OLED device

A green emitting OLED device

An **organic light emitting diode** (**OLED**) is a light-emitting diode (LED) in which the emissive electroluminescent layer is a film of organic compounds which emit light in response to an electric current. This layer of organic semiconductor material is situated between two electrodes. Generally, at least one of these electrodes is transparent.
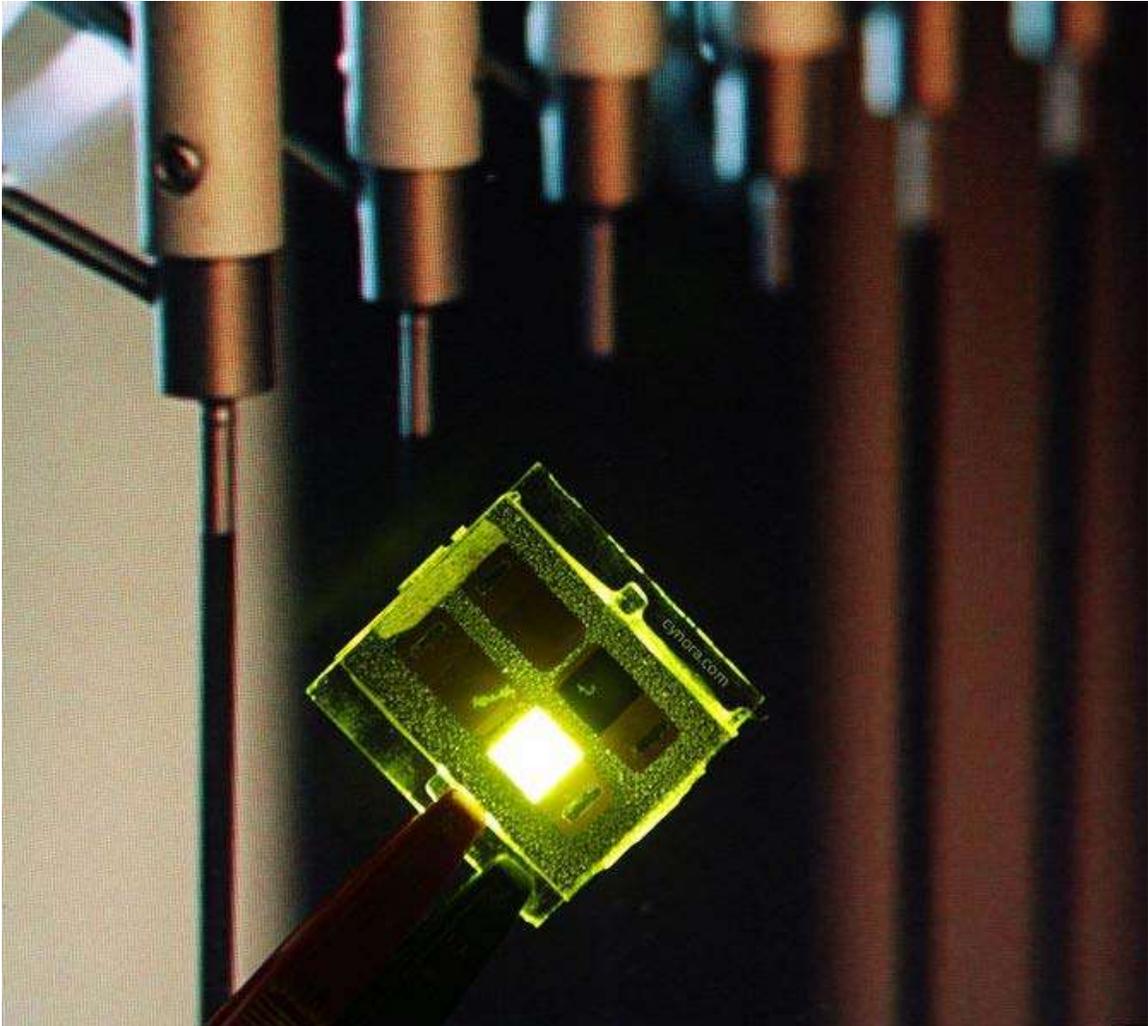
OLEDs are used in television screens, computer monitors, small, portable system screens such as mobile phones and PDAs, watches, advertising, information, and indication. OLEDs are also used in light sources for space illumination and in large-area light-emitting elements. Due to their early stage of development, they typically emit less light per unit area than inorganic solid-state based LED point-light sources.

An OLED display functions without a backlight. Thus, it can display deep black levels and can be thinner and lighter than liquid crystal displays. In low ambient light conditions such as dark rooms, an OLED screen can achieve a higher contrast ratio than an LCD

using either cold cathode fluorescent lamps or the more recently developed LED backlight.

There are two main families of OLEDs: those based upon small molecules and those employing polymers. Adding mobile ions to an OLED creates a Light-emitting Electrochemical Cell or LEC, which has a slightly different mode of operation.

OLED displays can use either passive-matrix (PMOLED) or active-matrix addressing schemes. Active-matrix OLEDs (AMOLED) require a thin-film transistor backplane to switch each individual pixel on or off, and can make higher resolution and larger size displays possible.

History

The first observations of electroluminescence in organic materials were in the early 1950s by A. Bernanose and co-workers at the Nancy-Université, France. They applied high-voltage alternating current (AC) fields in air to materials such as acridine orange, either deposited on or dissolved in cellulose or cellophane thin films. The proposed mechanism was either direct excitation of the dye molecules or excitation of electrons.

In 1960, Martin Pope and co-workers at New York University developed ohmic dark-injecting electrode contacts to organic crystals. They further described the necessary energetic requirements (work functions) for hole and electron injecting electrode contacts. These contacts are the basis of charge injection in all modern OLED devices. Pope's group also first observed direct current (DC) electroluminescence under vacuum on a pure single crystal of anthracene and on anthracene crystals doped with tetracene in 1963 using a small area silver electrode at 400V. The proposed mechanism was field-accelerated electron excitation of molecular fluorescence.

Pope's group reported in 1965 that in the absence of an external electric field, the electroluminescence in anthracene crystals is caused by the recombination of a thermalized electron and hole, and that the conducting level of anthracene is higher in energy than the exciton energy level. Also in 1965, W. Helfrich and W. G. Schneider of the National Research Council in Canada produced double injection recombination electroluminescence for the first time in an anthracene single crystal using hole and electron injecting electrodes, the forerunner of modern double injection devices. In the same year, Dow Chemical researchers patented a method of preparing electroluminescent cells using high voltage (500–1500 V) AC-driven (100–3000 Hz) electrically-insulated one millimetre thin layers of a melted phosphor consisting of ground anthracene powder, tetracene, and graphite powder. Their proposed mechanism involved electronic excitation at the contacts between the graphite particles and the anthracene molecules.

Device performance was limited by the poor electrical conductivity of contemporary organic materials. This was overcome by the discovery and development of highly conductive polymers.

Electroluminescence from polymer films was first observed by Roger Partridge at the National Physical Laboratory in the United Kingdom. The device consisted of a film of poly(n-vinylcarbazole) up to 2.2 micrometres thick located between two charge injecting electrodes. The results of the project were patented in 1975 and published in 1983.

The first diode device was reported at Eastman Kodak by Ching W. Tang and Steven Van Slyke in 1987. This device used a novel two-layer structure with separate hole transporting and electron transporting layers such that recombination and light emission occurred in the middle of the organic layer. This resulted in a reduction in operating voltage and improvements in efficiency and led to the current era of OLED research and device production.

Research into polymer electroluminescence culminated in 1990 with J. H. Burroughes *et al.* at the Cavendish Laboratory in Cambridge reporting a high efficiency green light-emitting polymer based device using 100 nm thick films of poly(p-phenylene vinylene).

Working principle
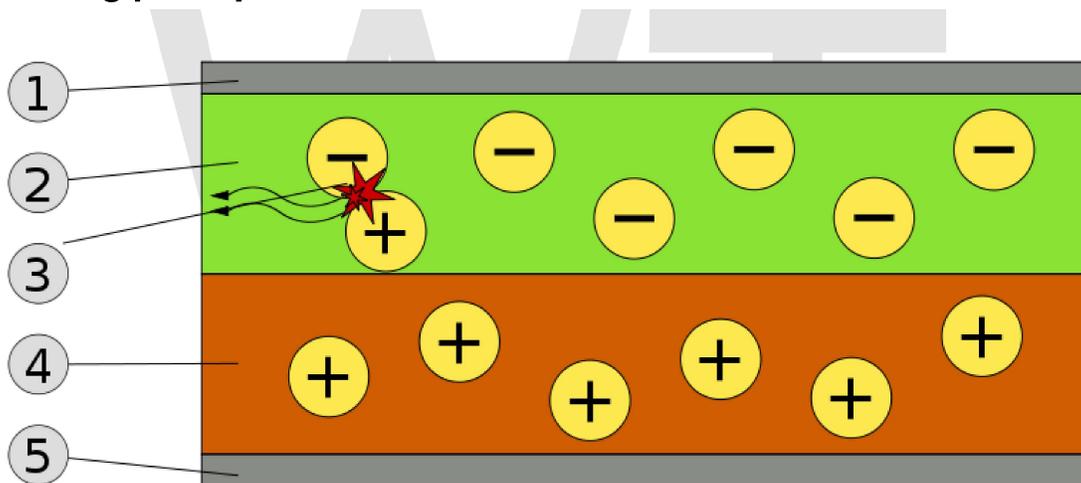
Schematic of a bilayer OLED: 1. Cathode (−), 2. Emissive Layer, 3. Emission of radiation, 4. Conductive Layer, 5. Anode (+)

A typical OLED is composed of a layer of organic materials situated between two electrodes, the anode and cathode, all deposited on a substrate. The organic molecules are electrically conductive as a result of delocalization of pi electrons caused by conjugation over all or part of the molecule. These materials have conductivity levels ranging from insulators to conductors, and therefore are considered organic semiconductors. The highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) of organic semiconductors are analogous to the valence and conduction bands of inorganic semiconductors.

Originally, the most basic polymer OLEDs consisted of a single organic layer. One example was the first light-emitting device synthesised by J. H. Burroughes *et al.*, which involved a single layer of poly(p-phenylene vinylene). However multilayer OLEDs can

be fabricated with two or more layers in order to improve device efficiency. As well as conductive properties, different materials may be chosen to aid charge injection at electrodes by providing a more gradual electronic profile, or block a charge from reaching the opposite electrode and being wasted. Many modern OLEDs incorporate a simple bilayer structure, consisting of a conductive layer and an emissive layer.

During operation, a voltage is applied across the OLED such that the anode is positive with respect to the cathode. A current of electrons flows through the device from cathode to anode, as electrons are injected into the LUMO of the organic layer at the cathode and withdrawn from the HOMO at the anode. This latter process may also be described as the injection of electron holes into the HOMO. Electrostatic forces bring the electrons and the holes towards each other and they recombine forming an exciton, a bound state of the electron and hole. This happens closer to the emissive layer, because in organic semiconductors holes are generally more mobile than electrons. The decay of this excited state results in a relaxation of the energy levels of the electron, accompanied by emission of radiation whose frequency is in the visible region. The frequency of this radiation depends on the band gap of the material, in this case the difference in energy between the HOMO and LUMO.

As electrons and holes are fermions with half integer spin, an exciton may either be in a singlet state or a triplet state depending on how the spins of the electron and hole have been combined. Statistically three triplet excitons will be formed for each singlet exciton. Decay from triplet states (phosphorescence) is spin forbidden, increasing the timescale of the transition and limiting the internal efficiency of fluorescent devices. Phosphorescent organic light-emitting diodes make use of spin–orbit interactions to facilitate intersystem crossing between singlet and triplet states, thus obtaining emission from both singlet and triplet states and improving the internal efficiency.

Indium tin oxide (ITO) is commonly used as the anode material. It is transparent to visible light and has a high work function which promotes injection of holes into the HOMO level of the organic layer. A typical conductive layer may consist of PEDOT:PSS as the HOMO level of this material generally lies between the workfunction of ITO and the HOMO of other commonly used polymers, reducing the energy barriers for hole injection. Metals such as barium and calcium are often used for the cathode as they have low work functions which promote injection of electrons into the LUMO of the organic layer. Such metals are reactive, so require a capping layer of aluminium to avoid degradation.

Single carrier devices are typically used to study the kinetics and charge transport mechanisms of an organic material and can be useful when trying to study energy transfer processes. As current through the device is composed of only one type of charge carrier, either electrons or holes, recombination does not occur and no light is emitted. For example, electron only devices can be obtained by replacing ITO with a lower work function metal which increases the energy barrier of hole injection. Similarly, hole only devices can be made by using a cathode comprised solely of aluminium, resulting in an energy barrier too large for efficient electron injection.

Material technologies

Small molecules
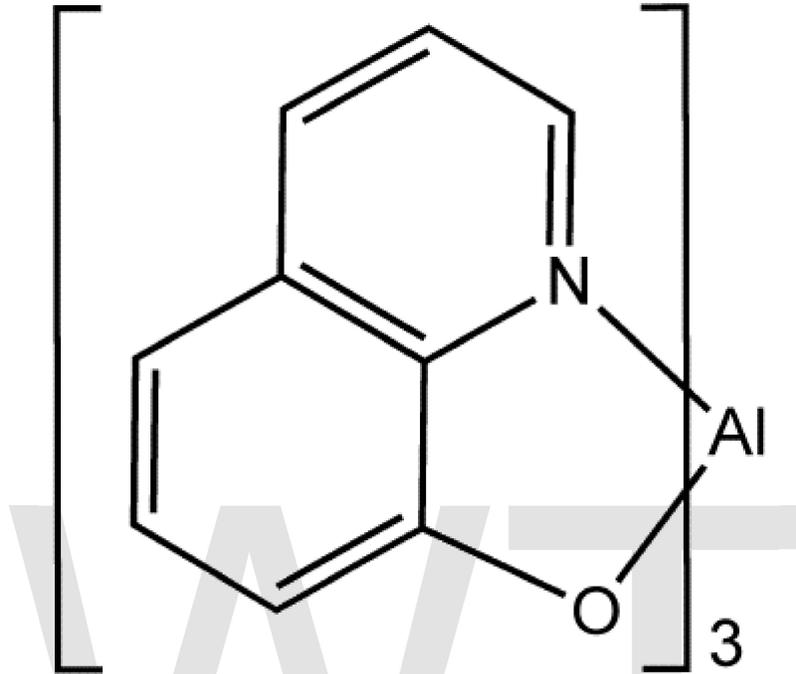
Alq$_3$, commonly used in small molecule OLEDs.

Efficient OLEDs using small molecules were first developed by Dr. Ching W. Tang *et al.* at Eastman Kodak. The term OLED traditionally refers specifically to this type of device, though the term SM-OLED is also in use.

Molecules commonly used in OLEDs include organometallic chelates (for example Alq$_3$, used in the organic light-emitting device reported by Tang *et al.*), fluorescent and phosphorescent dyes and conjugated dendrimers. A number of materials are used for their charge transport properties, for example triphenylamine and derivatives are commonly used as materials for hole transport layers. Fluorescent dyes can be chosen to obtain light emission at different wavelengths, and compounds such as perylene, rubrene and quinacridone derivatives are often used. Alq$_3$ has been used as a green emitter, electron transport material and as a host for yellow and red emitting dyes.

The production of small molecule devices and displays usually involves thermal evaporation in a vacuum. This makes the production process more expensive and of limited use for large-area devices than other processing techniques. However, contrary to polymer-based devices, the vacuum deposition process enables the formation of well controlled, homogeneous films, and the construction of very complex multi-layer structures. This high flexibility in layer design, enabling distinct charge transport and charge blocking layers to be formed, is the main reason for the high efficiencies of the small molecule OLEDs.

Coherent emission from a laser dye-doped tandem SM-OLED device, excited in the pulsed regime, has been demonstrated. The emission is nearly diffraction limited with a spectral width similar to that of broadband dye lasers.

Polymer light-emitting diodes
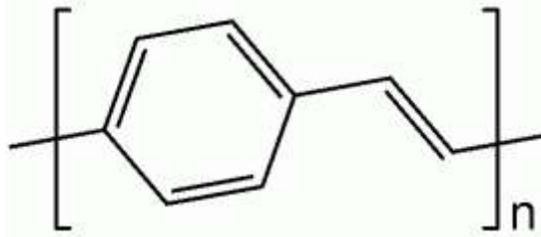
poly(*p*-phenylene vinylene), used in the first PLED.

Polymer light-emitting diodes (PLED), also light-emitting polymers (LEP), involve an electroluminescent conductive polymer that emits light when connected to an external voltage. They are used as a thin film for full-spectrum colour displays. Polymer OLEDs are quite efficient and require a relatively small amount of power for the amount of light produced.

Vacuum deposition is not a suitable method for forming thin films of polymers. However, polymers can be processed in solution, and spin coating is a common method of depositing thin polymer films. This method is more suited to forming large-area films than thermal evaporation. No vacuum is required, and the emissive materials can also be applied on the substrate by a technique derived from commercial inkjet printing. However, as the application of subsequent layers tends to dissolve those already present, formation of multilayer structures is difficult with these methods. The metal cathode may still need to be deposited by thermal evaporation in vacuum.

Typical polymers used in PLED displays include derivatives of poly(*p*-phenylene vinylene) and polyfluorene. Substitution of side chains onto the polymer backbone may determine the colour of emitted light or the stability and solubility of the polymer for performance and ease of processing.

While unsubstituted poly(p-phenylene vinylene) (PPV) is typically insoluble, a number of PPVs and related poly(naphthalene vinylene)s (PNVs) that are soluble in organic solvents or water have been prepared via ring opening metathesis polymerization.

Phosphorescent materials
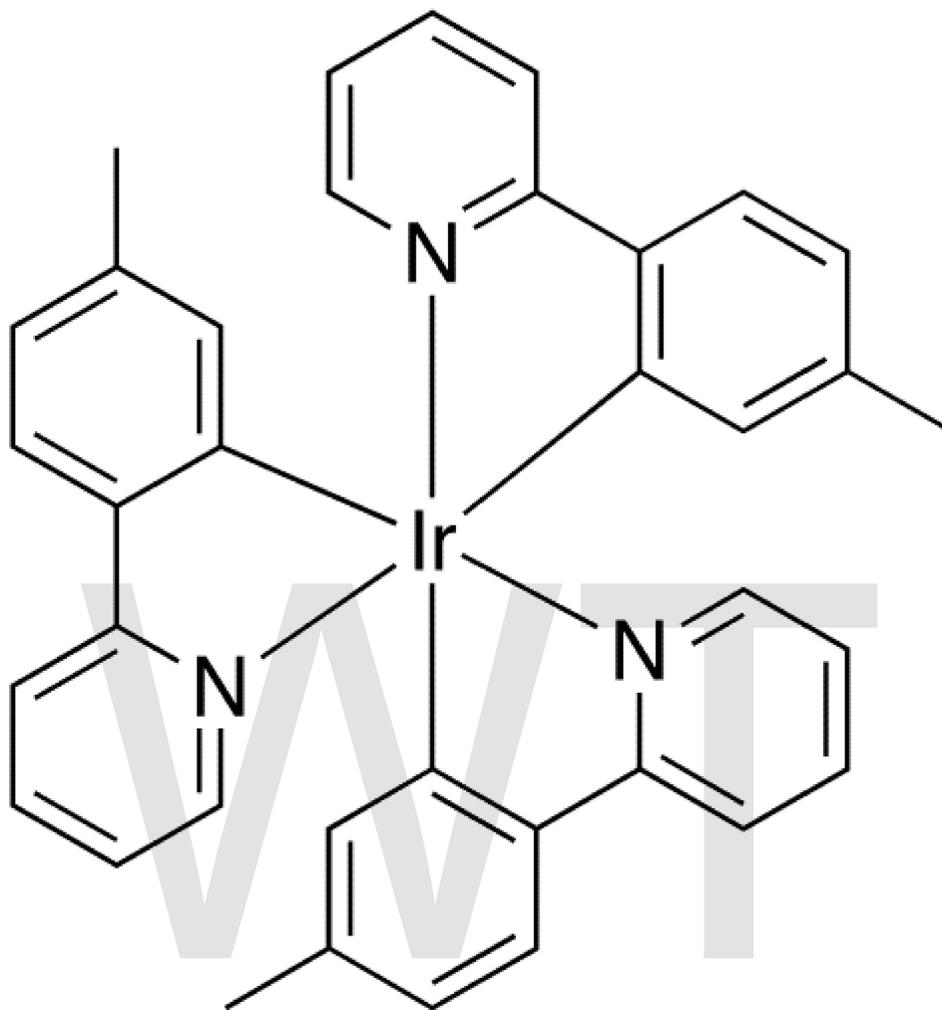
Ir(mppy)$_3$, a phosphorescent dopant which emits green light.

Phosphorescent organic light emitting diodes use the principle of electrophosphorescence to convert electrical energy in an OLED into light in a highly efficient manner, with the internal quantum efficiencies of such devices approaching 100%.

Typically, a polymer such as poly(n-vinylcarbazole) is used as a host material to which an organometallic complex is added as a dopant. Iridium complexes such as Ir(mppy)$_3$ are currently the focus of research, although complexes based on other heavy metals such as platinum have also been used.

The heavy metal atom at the centre of these complexes exhibits strong spin-orbit coupling, facilitating intersystem crossing between singlet and triplet states. By using these phosphorescent materials, both singlet and triplet excitons will be able to decay radiatively, hence improving the internal quantum efficiency of the device compared to a standard PLED where only the singlet states will contribute to emission of light.

Applications of OLEDs in solid state lighting require the achievement of high brightness with good CIE coordinates (for white emission). The use of macromolecular species like polyhedral oligomeric silsesquioxanes (POSS) in conjunction with the use of phosphorescent species such as Ir for printed OLEDs have exhibited brightnesses as high as 10,000 cd/m^2.

Device Architectures

Structure

- **Bottom or top emission:** Bottom emission devices use a transparent or semi-transparent bottom electrode to get the light through a transparent substrate. Top emission devices use a transparent or semi-transparent top electrode emitting light directly. Top-emitting OLEDs are better suited for active-matrix applications as they can be more easily integrated with a non-transparent transistor backplane.

- **Transparent OLEDs** use transparent or semi-transparent contacts on both sides of the device to create displays that can be made to be both top and bottom emitting (transparent). TOLEDs can greatly improve contrast, making it much easier to view displays in bright sunlight. This technology can be used in Head-up displays, smart windows or augmented reality applications. Novaled's OLED panel presented in Finetech Japan 2010, boasts a transparency of 60–70%.

- **Stacked OLEDs** use a pixel architecture that stacks the red, green, and blue subpixels on top of one another instead of next to one another, leading to substantial increase in gamut and color depth, and greatly reducing pixel gap. Currently, other display technologies have the RGB (and RGBW) pixels mapped next to each other decreasing potential resolution.

- **Inverted OLED:** In contrast to a conventional OLED, in which the anode is placed on the substrate, an Inverted OLED uses a bottom cathode that can be connected to the drain end of an n-channel TFT especially for the low cost amorphous silicon TFT backplane useful in the manufacturing of AMOLED displays.

Patterning technologies

Patternable organic light-emitting devices use a light or heat activated electroactive layer. A latent material (PEDOT-TMA) is included in this layer that, upon activation, becomes highly efficient as a hole injection layer. Using this process, light-emitting devices with arbitrary patterns can be prepared.

Colour patterning can be accomplished by means of laser, such as radiation-induced sublimation transfer (RIST).

Organic vapour jet printing (OVJP) uses an inert carrier gas, such as argon or nitrogen, to transport evaporated organic molecules (as in Organic Vapor Phase Deposition). The gas is expelled through a micron sized nozzle or nozzle array close to the substrate as it is being translated. This allows printing arbitrary multilayer patterns without the use of solvents.

Conventional OLED displays are formed by vapor thermal evaporation (VTE) and are patterned by shadow-mask. A mechanical mask has openings allowing the vapor to pass only on the desired location.

Backplane technologies

For a high resolution display like a TV, a TFT backplane is necessary to drive the pixels correctly. Currently, Low Temperature Polycrystalline silicon LTPS-TFT is used for commercial AMOLED displays. LTPS-TFT has variation of the performance in a display, so various compensation circuits have been reported. Due to the size limitation of the excimer laser used for LTPS, the AMOLED size was limited. To cope with the hurdle related to the panel size, amorphous-silicon/microcrystalline-silicon backplanes have been reported with large display prototype demonstrations.

Advantages

Demonstration of a 4.1" prototype flexible display from Sony

The different manufacturing process of OLEDs lends itself to several advantages over flat-panel displays made with LCD technology.

- **Lower cost in the future:** OLEDs can be printed onto any suitable substrate by an inkjet printer or even by screen printing, theoretically making them cheaper to produce than LCD or plasma displays. However, fabrication of the OLED substrate is more costly than that of a TFT LCD, until mass production methods lower cost through scalability. Roll-roll vapour-deposition methods for organic devices do allow mass production of thousands of devices per minute for minimal cost, although this technique also induces problems in that multi-layer devices can be challenging to make.

- **Light weight & flexible plastic substrates:** OLED displays can be fabricated on flexible plastic substrates leading to the possibility of flexible organic light-emitting diodes being fabricated or other new applications such as roll-up displays embedded in fabrics or clothing. As the substrate used can be flexible such as PET., the displays may be produced inexpensively.

- **Wider viewing angles & improved brightness:** OLEDs can enable a greater artificial contrast ratio (both dynamic range and static, measured in purely dark conditions) and viewing angle compared to LCDs because OLED pixels directly emit light. OLED pixel colours appear correct and unshifted, even as the viewing angle approaches 90° from normal.

- **Better power efficiency:** LCDs filter the light emitted from a backlight, allowing a small fraction of light through so they cannot show true black, while an inactive OLED element does not produce light or consume power.

- **Response time:** OLEDs can also have a faster response time than standard LCD screens. Whereas LCD displays are capable of between 2 and 8 ms response time offering a frame rate of +/-200 Hz, an OLED can theoretically have less than 0.01 ms response time enabling 100,000 Hz refresh rates.
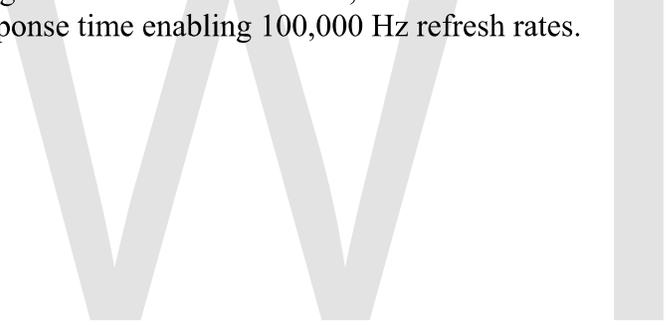
Disadvantages

LEP display showing partial failure

An old OLED display showing wear

- **Current costs:** OLED manufacture currently requires process steps that make it extremely expensive. Specifically, it requires the use of Low-Temperature Polysilicon backplanes; LTPS backplanes in turn require laser annealing from an amorphous silicon start, so this part of the manufacturing process for AMOLEDs starts with the process costs of standard LCD, and then adds an expensive, time-consuming process that cannot currently be used on large-area glass substrates.

- **Lifespan:** The biggest technical problem for OLEDs was the limited lifetime of the organic materials. In particular, blue OLEDs historically have had a lifetime of around 14,000 hours to half original brightness (five years at 8 hours a day) when used for flat-panel displays. This is lower than the typical lifetime of LCD, LED or PDP technology—each currently rated for about 25,000 – 40,000 hours to half brightness, depending on manufacturer and model. However, some manufacturers' displays aim to increase the lifespan of OLED displays, pushing their expected life past that of LCD displays by improving light outcoupling, thus achieving the same brightness at a lower drive current. In 2007, experimental OLEDs were created which can sustain 400 cd/m^2 of luminance for over 198,000 hours for green OLEDs and 62,000 hours for blue OLEDs.

- **Color balance issues:** Additionally, as the OLED material used to produce blue light degrades significantly more rapidly than the materials that produce other colors, blue light output will decrease relative to the other colors of light. This differential color output change will change the color balance of the display and is

much more noticeable than a decrease in overall luminance. This can be partially avoided by adjusting colour balance but this may require advanced control circuits and interaction with the user, which is unacceptable for some users. In order to delay the problem, manufacturers bias the colour balance towards blue so that the display initially has an artificially blue tint, leading to complaints of artificial-looking, over-saturated colors. More commonly, though, manufacturers optimize the size of the R, G and B subpixels to reduce the current density through the subpixel in order to equalize lifetime at full luminance. For example, a blue subpixel may be 100% larger than the green subpixel. The red subpixel may be 10% smaller than the green.

- **Efficiency of blue OLEDs:** Improvements to the efficiency and lifetime of blue OLEDs is vital to the success of OLEDs as replacements for LCD technology. Considerable research has been invested in developing blue OLEDs with high external quantum efficiency as well as a deeper blue color. External quantum efficiency values of 20% and 19% have been reported for red (625 nm) and green (530 nm) diodes, respectively. However, blue diodes (430 nm) have only been able to achieve maximum external quantum efficiencies in the range between 4% to 6%.

- **Water damage:** Water can damage the organic materials of the displays. Therefore, improved sealing processes are important for practical manufacturing. Water damage may especially limit the longevity of more flexible displays.

- **Outdoor performance:** As an emissive display technology, OLEDs rely completely upon converting electricity to light, unlike most LCDs which are to some extent reflective; e-ink leads the way in efficiency with ~ 33% ambient light reflectivity, enabling the display to be used without any internal light source. The metallic cathode in an OLED acts as a mirror, with reflectance approaching 80%, leading to poor readability in bright ambient light such as outdoors. However, with the proper application of a circular polarizer and anti-reflective coatings, the diffuse reflectance can be reduced to less than 0.1%. With 10,000 fc incident illumination (typical test condition for simulating outdoor illumination), that yields an approximate photopic contrast of 5:1.

- **Power consumption:** While an OLED will consume around 40% of the power of an LCD displaying an image which is primarily black, for the majority of images it will consume 60–80% of the power of an LCD – however it can use over three times as much power to display an image with a white background such as a document or website. This can lead to reduced real-world battery life in mobile devices.

- **Screen burn-in:** Unlike displays with a common light source, the brightness of each OLED pixel fades depending on the content displayed. The varied lifespan of the organic dyes can cause a discrepancy between red, green, and blue intensity. This leads to image persistence, also known as burn-in.

- **UV sensitivity:** OLED displays can be damaged by prolonged exposure to UV light. The most pronounced example of this can be seen with a near UV laser (such as a Bluray pointer) and can damage the display almost instantly with more than 20 mW leading to dim or dead spots where the beam is focused. This is usually avoided by installing a UV blocking filter over the panel and this can easily be seen as a clear plastic layer on the glass. Removal of this filter can lead to severe damage and an unusable display after only a few months of room light exposure.

Manufacturers and commercial uses
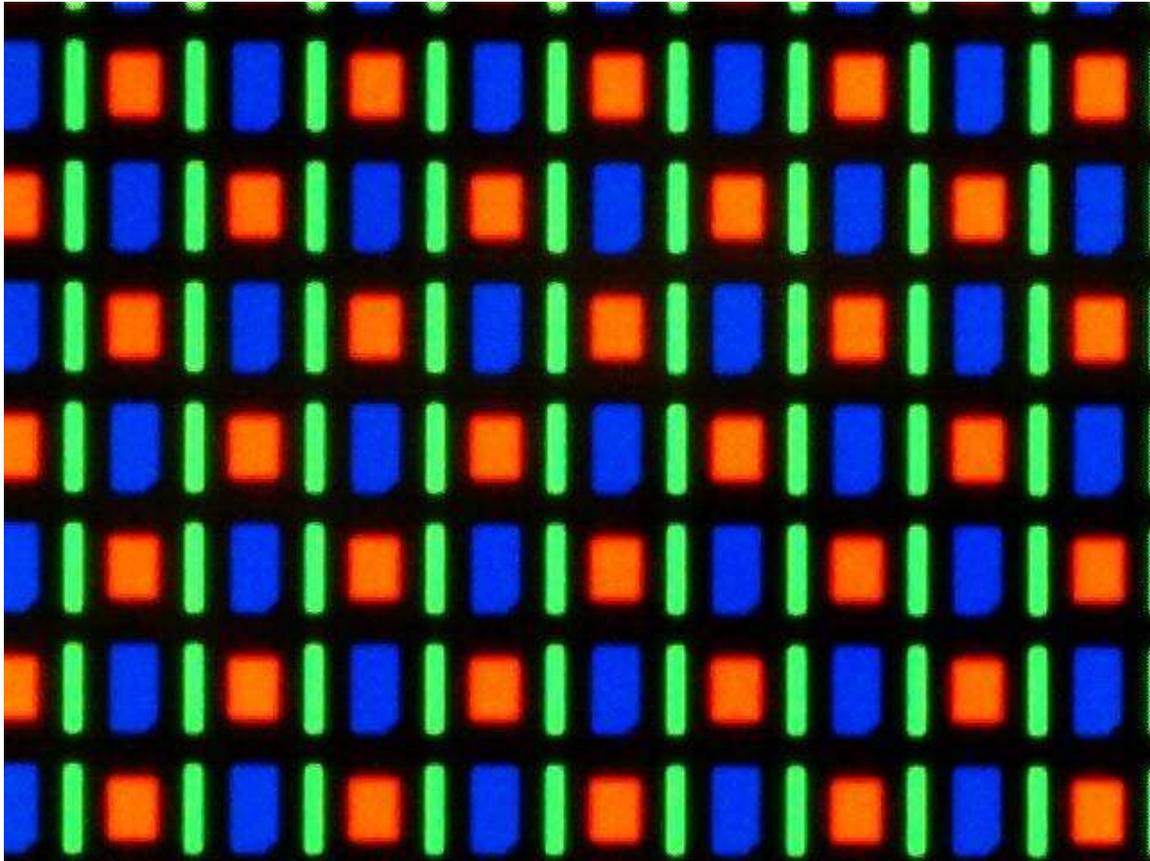
Magnified image of the AMOLED screen on the Google Nexus One smartphone using the RGBG system of the PenTile Matrix Family.

A 3.8 cm (1.5 in) OLED display from a Creative ZEN V media player

OLED technology is used in commercial applications such as displays for mobile phones and portable digital media players, car radios and digital cameras among others. Such portable applications favor the high light output of OLEDs for readability in sunlight and their low power drain. Portable displays are also used intermittently, so the lower lifespan of organic displays is less of an issue. Prototypes have been made of flexible and rollable displays which use OLEDs' unique characteristics. Applications in flexible signs and lighting are also being developed. Philips Lighting have made OLED lighting samples under the brand name 'Lumiblade' available online.

OLEDs have been used in most Motorola and Samsung colour cell phones, as well as some HTC, LG and Sony Ericsson models. Nokia has also recently introduced some OLED products including the N85 and the N86 8MP, both of which feature an AMOLED display. OLED technology can also be found in digital media players such as the Creative ZEN V, the iriver clix, the Zune HD and the Sony Walkman X Series.

The Google and HTC Nexus One smartphone includes an AMOLED screen, as does HTC's own Desire and Legend phones. However due to supply shortages of the Samsung-produced displays, certain HTC models will use Sony's SLCD displays in the future, while the Google and Samsung Nexus S smartphone will use "Super Clear LCD" instead in some countries.

Other manufacturers of OLED panels include Anwell Technologies Limited, Chi Mei Corporation, LG, and others.

DuPont stated in a press release in May 2010 that they can produce a 50-inch OLED TV in two minutes with a new printing technology. If this can be scaled up in terms of manufacturing, then the total cost of OLED TVs would be greatly reduced. Dupont also states that OLED TVs made with this less expensive technology can last up to 15 years if left on for a normal eight hour day.

The use of OLEDs may be subject to patents held by Eastman Kodak, DuPont, General Electric, Royal Philips Electronics, numerous universities and others. There are by now literally thousands of patents associated with OLEDs, both from larger corporations and smaller technology companies.

Samsung applications

By 2004 Samsung, South Korea's largest conglomerate, was the world's largest OLED manufacturer, producing 40% of the OLED displays made in the world, and as of 2010 has a 98% share of the global AMOLED market. The company is leading the world OLED industry, generating $100.2 million out of the total $475 million revenues in the global OLED market in 2006. As of 2006, it held more than 600 American patents and more than 2800 international patents, making it the largest owner of AMOLED technology patents.

Samsung SDI announced in 2005 the world's largest OLED TV at the time, at 21 inches (53 cm). This OLED featured the highest resolution at the time, of 6.22 million pixels. In addition, the company adopted active matrix based technology for its low power consumption and high-resolution qualities. This was exceeded in January 2008, when Samsung showcased the world's largest and thinnest OLED TV at the time, at 31 inches and 4.3 mm.

In May 2008, Samsung unveiled an ultra-thin 12.1 inch laptop OLED display concept, with a 1,280×768 resolution with infinite contrast ratio. According to Woo Jong Lee, Vice President of the Mobile Display Marketing Team at Samsung SDI, the company expected OLED displays to be used in notebook PCs as soon as 2010.

In October 2008, Samsung showcased the world's thinnest OLED display, also the first to be 'flappable' and bendable. It measures just 0.05 mm (thinner than paper), yet a Samsung staff member said that it is "technically possible to make the panel thinner". To achieve this thickness, Samsung etched an OLED panel that uses a normal glass substrate. The drive circuit was formed by low-temperature polysilicon TFTs. Also, low-molecular organic EL materials were employed. The pixel count of the display is 480 × 272. The contrast ratio is 100,000:1, and the luminance is 200 cd/m². The colour reproduction range is 100% of the NTSC standard.

In the same month, Samsung unveiled what was then the world's largest OLED Television at 40-inch with a Full HD resolution of 1920×1080 pixel. In the FPD International, Samsung stated that its 40-inch OLED Panel is the largest size currently possible. The panel has a contrast ratio of 1,000,000:1, a colour gamut of 107% NTSC, and a luminance of 200 cd/m² (peak luminance of 600 cd/m²).

At the Consumer Electronics Show (CES) in January 2010, Samsung demonstrated a laptop computer with a large, transparent OLED display featuring up to 40% transparency and an animated OLED display in a photo ID card.

Samsung's latest AMOLED smartphones use their Super AMOLED trademark, with the Samsung Wave S8500 and Samsung i9000 Galaxy S being launched in June 2010. In January 2011 Samsung announced their Super AMOLED Plus displays - which offer several advances over the older Super AMOLED displays - real stripe matrix (50% more sub pixels), thinner form factor, brighter image and a 18% reduction in energy consumption.

Sony applications

Sony XEL-1, the world's first OLED TV. (front)

Sony XEL-1 (side)

The Sony CLIÉ PEG-VZ90 was released in 2004, being the first PDA to feature an OLED screen. Other Sony products to feature OLED screens include the MZ-RH1 portable minidisc recorder, released in 2006 and the Walkman X Series.

At the Las Vegas CES 2007, Sony showcased 11-inch (28 cm, resolution 960×540) and 27-inch (68.5 cm, full HD resolution at 1920×1080) OLED TV models. Both claimed 1,000,000:1 contrast ratios and total thicknesses (including bezels) of 5 mm. In April 2007, Sony announced it would manufacture 1000 11-inch OLED TVs per month for market testing purposes. On October 1, 2007, Sony announced that the 11-inch model, now called the XEL-1, would be released commercially; the XEL-1 was first released in Japan in December 2007.

In May 2007, Sony publicly unveiled a video of a 2.5-inch flexible OLED screen which is only 0.3 millimeters thick. At the Display 2008 exhibition, Sony demonstrated a 0.2 mm thick 3.5 inch display with a resolution of 320×200 pixels and a 0.3 mm thick 11 inch display with 960×540 pixels resolution, one-tenth the thickness of the XEL-1.

In July 2008, a Japanese government body said it would fund a joint project of leading firms, which is to develop a key technology to produce large, energy-saving organic displays. The project involves one laboratory and 10 companies including Sony Corp.

NEDO said the project was aimed at developing a core technology to mass-produce 40 inch or larger OLED displays in the late 2010s.

In October 2008, Sony published results of research it carried out with the Max Planck Institute over the possibility of mass-market bending displays, which could replace rigid LCDs and plasma screens. Eventually, bendable, transparent OLED screens could be stacked to produce 3D images with much greater contrast ratios and viewing angles than existing products.

Sony exhibited a 24.5" prototype OLED 3D television during the Consumer Electronics Show in January 2010.

In January 2011, Sony announced the Next Generation Portable handheld game console (the successor to the PSP) will feature a 5-inch OLED screen.
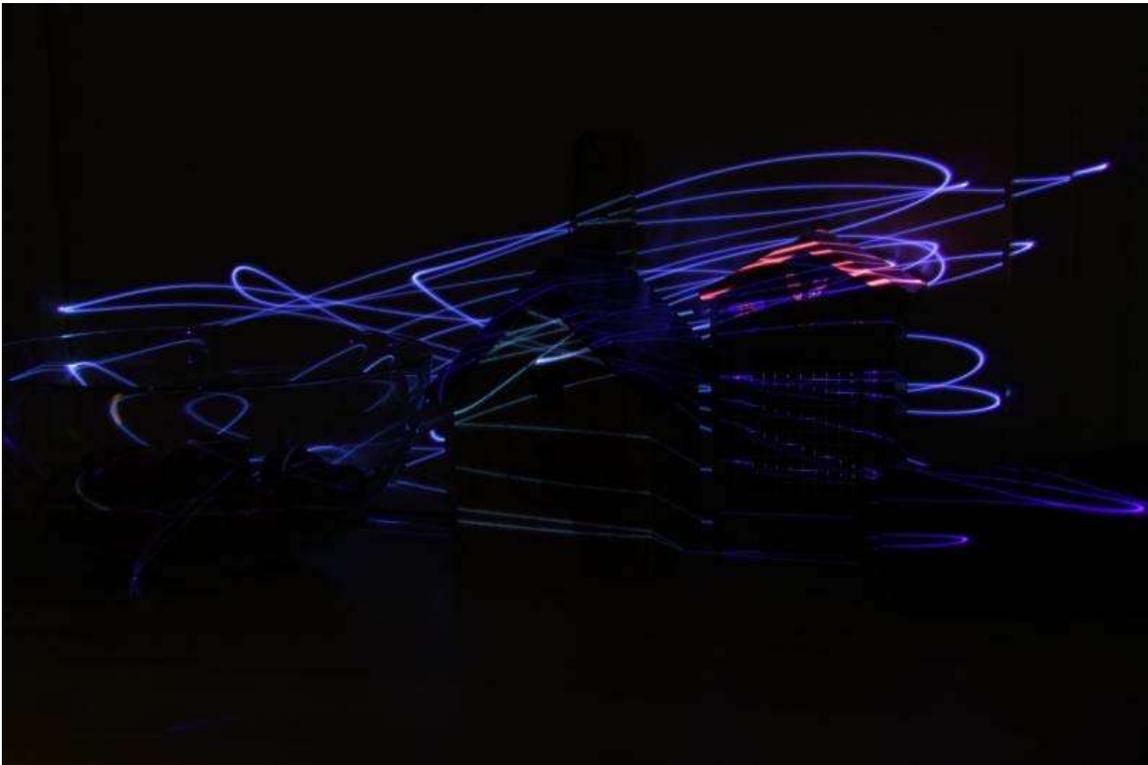
On February 17, 2011, Sony announced its 25" OLED Professional Reference Monitor aimed at the Cinema and high end Drama Post Production market.

LG applications

As of 2010, LG produces one model of OLED television, the 15 inch 15EL9500 and has announced a 31" OLED 3D television for March 2011.

Chapter 4

Blue Laser

Trails of a 20mW 405nm blacklight laser show clear fluorescence on some objects

A **blue laser** is a laser that emits electromagnetic radiation at a wavelength of between 360 and 480 nanometres, i.e. it emits "blue" light or light at the blue end of the spectrum. The light from "blue lasers" of shorter wavelengths is sometimes violet to the human eye, a distinctly different color. This is true, for example, of the most commercially-common "blue" lasers, the diode lasers used in Blu-ray applications, which emit 405 nm light that is distinctly violet.

Blue lasers are frequently semiconductor laser diodes based on gallium(III) nitride (violet color) or indium gallium nitride (often true-blue in color, but able to produce other colors,

as well). Both blue and violet lasers can also be constructed using frequency-doubling of infrared laser wavelengths from diode lasers or diode pumped lasers.

These new devices have applications in many areas ranging from optoelectronic data storage at high density to medical applications.

History

Inventing blue-laser technology

Red lasers can be built on gallium arsenide (GaAs) semiconductors, upon which a dozen layers of atoms are placed to form the part of the laser that generates light from quantum wells. Using methods similar to those developed for silicon, the substrate can be built free of the defects called dislocations, and the atoms laid down so the distance between the ones making up the ground and those of the quantum wells are the same.

But the best semiconductor for blue lasers is gallium nitride (GaN) crystals, which are much harder to manufacture, requiring higher pressures and temperatures, similar to the ones that produce synthetic diamonds, and the use of high-pressure nitrogen gas. The technical problems seemed insurmountable, so researchers since the 1960s have sought to deposit GaN on a base of readily available sapphire. But a mismatch between the structures of sapphire and gallium nitride created too many defects.

In 1992, Japanese inventor Shuji Nakamura invented the first efficient blue LED, and four years later, the first blue laser. Nakamura used the material deposited on the sapphire substrate, although the number of defects remained too high ($106–1010/cm^2$) to easily build a high-power laser.

In the early 1990s, the Institute of High Pressure Physics at the Polish Academy of Sciences in Warsaw (Poland), under the leadership of Dr. Sylwester Porowski developed technology to create gallium nitride crystals with high structural quality and fewer than 100 defects per square centimeter — at least 10,000 times better than the best sapphire-supported crystal.

In 1999, Nakamura tried Polish crystals, producing lasers with twice the yield and ten times the lifetime — 3,000 hours at 30 mW.

A further development of the technology has led to mass production of the device. Today, blue lasers use a sapphire surface covered with a layer of gallium nitride (this technology is used by Japanese company Nichia, which has an agreement with Sony), and blue semiconductor lasers use a gallium nitride mono-crystal surface (Polish company TopGaN).

After 10 years, Japanese manufacturers mastered the production of a blue laser with 60 mW of power, making them applicable for devices that read a dense high-speed stream of data from Blu-ray, BD-R, and BD-RE. Polish technology is cheaper than

Japanese but has a smaller share of the market. There is one more Polish high-tech company which creates gallium nitride crystal – Ammono, but this company does not produce blue lasers.

For his work, Nakamura received the Millennium Technology Prize awarded in 2006.

Until the late 1990s, when blue semiconductor lasers were developed, blue lasers were large and expensive gas laser instruments which relied on population inversion in rare gas mixtures and needed high currents and strong cooling.

Thanks to prior development of many groups, including, most notably, Professor Isamu Akasaki's group, Shuji Nakamura at Nichia Corporation and Sony Corporation in Anan (Tokushima-ken, Japan) made a series of inventions and developed commercially viable blue and violet semiconductor lasers. The active layer of the Nichia devices was formed from InGaN quantum wells or quantum dots spontaneously formed via self-assembly. The new invention enabled the development of small, convenient and low-priced blue, violet, and ultraviolet UV lasers, which had not been available before, and opened the way for applications such as high-density HD DVD data storage and Blu-ray discs. The shorter wavelength allows it to read discs containing much more information.

Variants

Blue and violet diode-pumped solid state (DPSS) laser modules

Blue laser pointers, which became available around 2006, have the same basic construction as DPSS green lasers. They most commonly lase at 473 nm (sometimes reported as 474 nm), which is produced by frequency doubling of 946 nm laser radiation from a diode-pumped Nd:YAG or Nd:YVO4 crystal. Neodymium-doped crystals usually produce a principal wavelength of 1064 nm, but with the proper reflective coating mirrors can be also made to lase at other non-principal neodymium wavelengths, such as the 946 nm transition used in blue-laser applications. For high output power BBO crystals are used as frequency doublers; for lower powers, KTP is used. Output powers available are up to 1000 mW, but this usually is the total output including the infrared. As with green DPSS lasers, use of a 1000 mW IR diode usually results in approximately 300 mW of visible blue light, even if the laser is reported at 1000 mW power.

Blue lasers can also be fabricated directly with InGaN semiconductors, which produce blue light without frequency-doubling. 445 nm blue laser diodes are currently available on the open market. The devices are brighter than the 405 nm laser diodes, since the longer wavelength is closer to the peak sensitivity of the human eye. Commercial devices like laser projectors have driven down the prices on these diodes, as of March 2010.

Violet lasers may be constructed directly with **GaN** (gallium nitride) semiconductors, as noted. However, a few higher-powered (120 mW) 404–405 nm "violet" laser pointers have become available which are not based on GaN, but also use DPSS frequency-doubler technology starting from 1 watt 808 nm gallium arsenide infrared diode lasers

being directly doubled, without a longer-wave neodymium laser interposed between diode laser and doubler-crystal. As with all high powered unfiltered infrared-driven DPSS lasers, such devices are able to pop balloons and light matches, but this is as a result of an unfiltered high-power infrared component in the beam.

Appearance

The violet 405 nm laser (whether constructed from GaN or frequency-doubled GaAs laser diodes) is not in fact blue, but appears to the eye as violet (reddish-purple), a color for which a human eye has a very limited sensitivity. When pointed at many white objects (such as commercial white typing or printing paper) the laser dot changes from violet, and appears blue. This color is actually due to blue fluorescence from brightening dyes added to the paper—the same effect as from a blacklight lamp.

For display applications, where the "true blue" color is required, a wavelength of 450–460 nm is required. With advances in production, and commercial sales of low cost laser projectors, 445 nm InGaN laser diodes have dropped in price.

(A last challenge in projection laser diodes is related to the construction of a "true green" InGaN laser (around 530 nm). Many companies have demonstrated devices working at only slightly shorter wavelengths: 480–500 nm.)

Applications

Areas of application of the blue laser include:

- High-definition Blu-ray players
- Telecommunications
- Information technology
- Environmental monitoring
- Electronic equipment
- Medical diagnostics
- Micro projectors and displays

Chapter 5

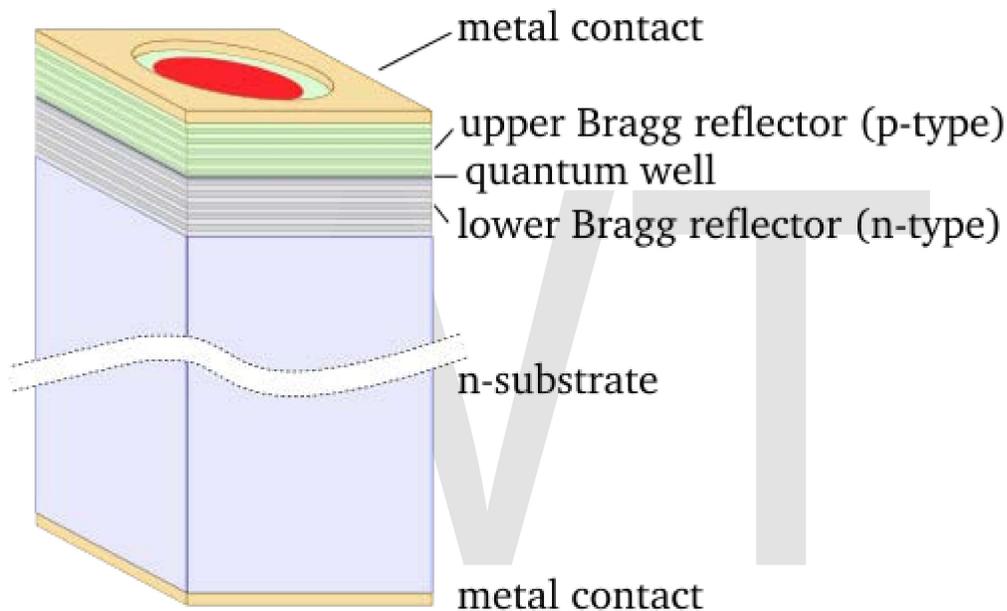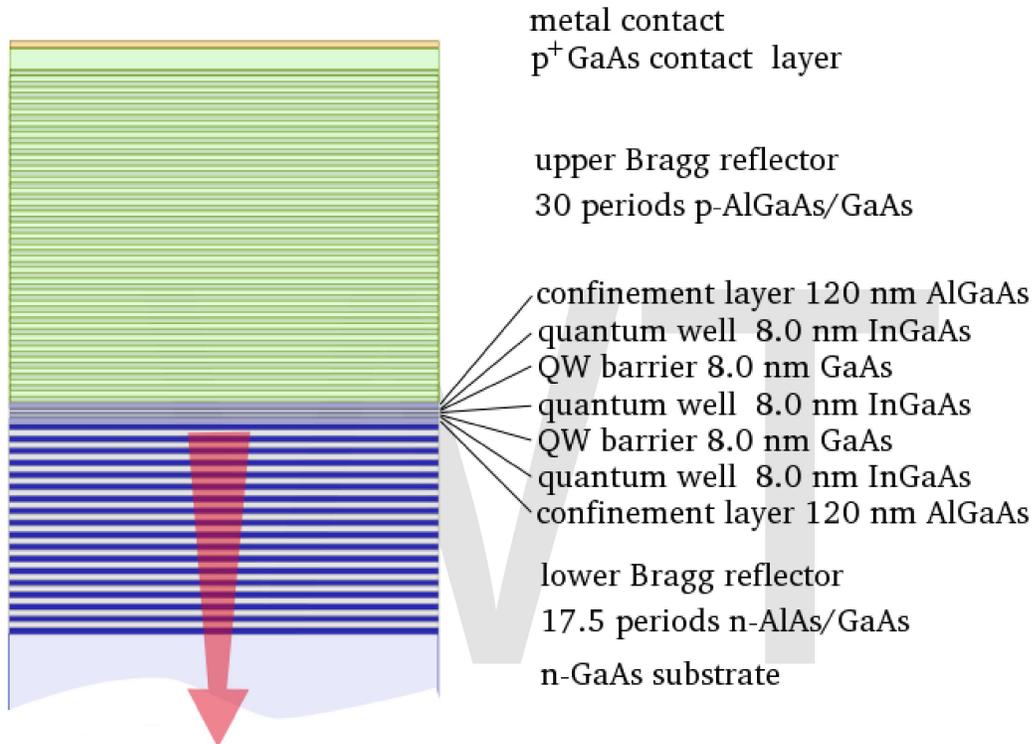# Vertical-Cavity Surface-Emitting Laser

Diagram of a simple VCSEL structure.

The **vertical-cavity surface-emitting laser**, or **VCSEL** is a type of semiconductor laser diode with laser beam emission perpendicular from the top surface, contrary to conventional edge-emitting semiconductor lasers (also *in-plane* lasers) which emit from surfaces formed by cleaving the individual chip out of a wafer.

There are several advantages to producing VCSELs when compared with the production process of edge-emitting lasers. Edge-emitters cannot be tested until the end of the production process. If the edge-emitter does not work, whether due to bad contacts or poor material growth quality, the production time and the processing materials have been wasted. VCSELs however, can be tested at several stages throughout the process to check for material quality and processing issues. For instance, if the vias have not been completely cleared of dielectric material during the etch, an interim testing process will flag that the top metal layer is not making contact to the initial metal layer. Additionally,

because VCSELs emit the beam perpendicular to the active region of the laser as opposed to parallel as with an edge emitter, tens of thousands of VCSELs can be processed simultaneously on a three inch Gallium Arsenide wafer. Furthermore, even though the VCSEL production process is more labor and material intensive, the yield can be controlled to a more predictable outcome.

Structure

metal contact
p^+ GaAs contact layer

upper Bragg reflector
30 periods p-AlGaAs/GaAs

confinement layer 120 nm AlGaAs
quantum well 8.0 nm InGaAs
QW barrier 8.0 nm GaAs
quantum well 8.0 nm InGaAs
QW barrier 8.0 nm GaAs
quantum well 8.0 nm InGaAs
confinement layer 120 nm AlGaAs

lower Bragg reflector
17.5 periods n-AlAs/GaAs

n-GaAs substrate

A realistic VCSEL device structure. This is a *bottom-emitting multiple-quantum-well VCSEL.*

The laser resonator consists of two distributed Bragg reflector (DBR) mirrors parallel to the wafer surface with an active region consisting of one or more quantum wells for the laser light generation in between. The planar DBR-mirrors consist of layers with alternating high and low refractive indices. Each layer has a thickness of a quarter of the laser wavelength in the material, yielding intensity reflectivities above 99%. High reflectivity mirrors are required in VCSELs to balance the short axial length of the gain region.

In common VCSELs the upper and lower mirrors are doped as p-type and n-type materials, forming a diode junction. In more complex structures, the p-type and n-type regions may be buried between the mirrors, requiring a more complex semiconductor

process to make electrical contact to the active region, but eliminating electrical power loss in the DBR structure.

In laboratory investigation of VCSELs using new material systems, the active region may be *pumped* by an external light source with a shorter wavelength, usually another laser. This allows a VCSEL to be demonstrated without the additional problem of achieving good electrical performance; however such devices are not practical for most applications.

VCSELs for wavelengths from 650 nm to 1300 nm are typically based on gallium arsenide (GaAs) wafers with DBRs formed from GaAs and aluminium gallium arsenide ($Al_xGa_{(1-x)}As$). The GaAs–AlGaAs system is favored for constructing VCSELs because the lattice constant of the material does not vary strongly as the composition is changed, permitting multiple "lattice-matched" epitaxial layers to be grown on a GaAs substrate. However, the refractive index of AlGaAs does vary relatively strongly as the Al fraction is increased, minimizing the number of layers required to form an efficient Bragg mirror compared to other candidate material systems. Furthermore, at high aluminium concentrations, an oxide can be formed from AlGaAs, and this oxide can be used to restrict the current in a VCSEL, enabling very low threshold currents.

Recently the two main methods of restricting the current in a VCSEL were characterized by two types of VCSELs: ion-implanted VCSELs and Oxide VCSELs.

In the early 1990s, telecommunications companies tended to favor ion-implanted VCSELs. Ions, (often hydrogen ions, H+), were implanted into the VCSEL structure everywhere except the aperture of the VCSEL, destroying the lattice structure around the aperture, thus inhibiting the current. In the mid to late 1990s, companies moved towards the technology of oxide VCSELs. The current is confined in an oxide VCSEL by oxidizing the material around the aperture of the VCSEL. A high content aluminium layer that is grown within the VCSEL structure is the layer that is oxidized. Oxide VCSELs also often employ the ion implant production step. As a result in the oxide VCSEL, the current path is confined by the ion implant and the oxide aperture.

The initial acceptance of oxide VCSELs was plagued with concern about the apertures "popping off" due to the strain and defects of the oxidation layer. However, after much testing, the reliability of the structure has proven to be robust. As stated in one study by Hewlett Packard on oxide VCSELs, "The stress results show that the activation energy and the wearout lifetime of oxide VCSEL are similar to that of implant VCSEL emitting the same amount of output power."

A production concern also plagued the industry when moving the oxide VCSELs from research and development to production mode. The oxidation rate of the oxide layer was highly dependent on the aluminium content. Any slight variation in aluminium would change the oxidation rate sometimes resulting in apertures that were either too big or too small to meet the specification standards.

Longer wavelength devices, from 1300 nm to 2000 nm, have been demonstrated with at least the active region made of indium phosphide. VCSELs at even higher wavelengths are experimental and usually optically pumped. 1310 nm VCSELs are desirable as the dispersion of silica-based optical fiber is minimal in this wavelength range.

Special forms

- Multiple active region devices (aka bipolar cascade VCSELs). Allows for differential quantum efficiency values in excess of 100% through carrier recycling
- VCSELs with tunnel junctions. Using a tunnel junction (n^+p^+), an electrically advantageous $n\text{-}n^+p^+\text{-}p\text{-}i\text{-}n$ configuration can be built that also may beneficially influence other structural elements (e.g. in the form of a *Buried Tunnel Junction* (BTJ)).
- Widely tunable VCSEL with micromechanically (MEMS) movable mirror
- *Wafer-bonded* or *wafer-fused* VCSEL: Combination of semiconductor materials that can be fabricated using different types of substrate wafers
- Monolithically optically pumped VCSELs: Two VCSELs on top of each other. One of them optically pumps the other one.
- VCSEL with longitudinally integrated monitor diode: A photodiode is integrated under the back mirror of the VCSEL.
- VCSEL with transversally integrated monitor diode: With suitable etching of the VCSEL's wafer, a resonant photodiode can be manufactured that may measure the light intensity of a neighboring VCSEL.
- VCSELs with external cavities, known as VECSELs or semiconductor disk lasers. VECSELs are optically pumped with conventional laser diodes. This arrangement allows a larger area of the device to be pumped and therefore more power can be extracted - as much as 30W. The external cavity also allows intracavity techniques such as frequency doubling, single frequency operation and femtosecond pulse modelocking.
- Vertical-cavity semiconductor optical amplifiers, known as VCSOAs . These devices are optimized as amplifiers as opposed to oscillators. VCSOAs must be operated below threshold and thus require reduced mirror reflectivities for decreased feedback. In order to maximize the signal gain, these devices contain a large number of quantum wells (optically pumped devices have been demonstrated with 21–28 wells) and as a result exhibit single-pass gain values which are significantly larger than that of a typical VCSEL (roughly 5%). These structures operate as narrow linewidth (tens of GHz) amplifiers and may be implemented as amplifying filters.

Characteristics

Because VCSELs emit from the top surface of the chip, they can be tested *on-wafer*, before they are cleaved into individual devices. This reduces the fabrication cost of the devices. It also allows VCSELs to be built not only in one-dimensional, but also in two-dimensional *arrays*.

The larger output aperture of VCSELs, compared to most edge-emitting lasers, produces a lower divergence angle of the output beam, and makes possible high coupling efficiency with optical fibers.

The high reflectivity mirrors, compared to most edge-emitting lasers, reduce the threshold current of VCSELs, resulting in low power consumption. However, as yet, VCSELs have lower emission power compared to edge-emitting lasers. The low threshold current also permits high intrinsic modulation bandwidths in VCSELs.

The wavelength of VCSELs may be tuned, within the gain band of the active region, by adjusting the thickness of the reflector layers.

While early VCSELs emitted in multiple longitudinal modes or in filament modes, single-mode VCSELs are now common.

High-power VCSELs

High-power vertical-cavity surface-emitting lasers can also be fabricated, either by increasing the emitting aperture size of a single device or by combining several elements into large two-dimensional (2D) arrays. There have been relatively few reported studies on high-power VCSELs. Large-aperture single devices operating around 100 mW were first reported in 1993. Improvements in the epitaxial growth, processing, device design, and packaging led to individual large-aperture VCSELs emitting several hundreds of milliwatts by 1998. More than 2 W continuous-wave (CW) operation at -10 degrees Celsius heat-sink temperature was also reported in 1998 from a VCSEL array consisting of 1,000 elements, corresponding to a power density of 30 W/cm^2. In 2001, more than 1 W CW power and 10 W pulsed power at room temperature were reported from a 19-element array. The VCSEL array chip was mounted on a diamond heat spreader, taking advantage of diamond's very high thermal conductivity. A record 3 W CW output power was reported in 2005 from large diameter single devices emitting around 980 nm.

In 2007, more than 200 W of CW output power was reported from a large (5 × 5mm) 2D VCSEL array emitting around the 976 nm wavelength, representing a substantial breakthrough in the field of high-power VCSELs. The high power level achieved was mostly due to improvements in wall-plug efficiency and packaging. In 2009, >100 W power levels were reported for VCSEL arrays emitting around 808 nm.

At that point, the VCSEL technology became useful for a variety of medical, industrial, and military applications requiring high power or high energy. Examples of such applications are:

- Medical/cosmetics: laser hair removal, laser wrinkle removal
- Infrared illuminators for military/surveillance
- Pumping of solid-state lasers and fiber lasers
- High-power/high-energy second harmonic generation (blue/green light)
- Laser machining: laser cutting, laser drilling, laser ablation, laser engraving

Applications

- Optical fiber data transmission
- Analog broadband signal transmission
- Absorption spectroscopy (TDLAS)
- Laser printers
- computer mouse
- Biological tissue analysis
- Chip scale atomic clock

History

The first VCSEL was presented in 1979 by Soda, Iga, Kitahara and Suematsu, but devices for CW operation at room temperature were not reported until 1988. The term VCSEL was coined in a publication of the Optical Society of America in 1987. Today, VCSELs have replaced edge-emitting lasers in applications for short-range fiberoptic communication such as Gigabit Ethernet and Fibre Channel.

Chapter 6

Quantum Well Laser and Hybrid Silicon Laser

Quantum well laser

A **quantum well laser** is a laser diode in which the active region of the device is so narrow that quantum confinement occurs. The wavelength of the light emitted by a quantum well laser is determined by the width of the active region rather than just the bandgap of the material from which it is constructed. This means that much shorter wavelengths can be obtained from quantum well lasers than from conventional laser diodes using a particular semiconductor material. The efficiency of a quantum well laser is also greater than a conventional laser diode due to the stepwise form of its density of states function.

Origin of the concept of quantum wells

In 1972, Charles H. Henry, a physicist and newly-appointed Head of the Semiconductor Electronics Research Department at Bell Laboratories, had a keen interest in the subject of integrated optics, the fabrication of optical circuits in which the light travels in waveguides.

In late 1972, while pondering the problems associated with waveguides, he had a sudden insight, a realization that a double heterostructure is a waveguide for electron waves, not just lightwaves. On further reflection, he saw that there is a complete analogy between the confinement of light by a slab waveguide and the confinement of electrons by the potential well that is formed from the difference in bandgaps in a double heterostructure.

Henry realized that there should be discrete modes (levels) in the potential well, and a simple estimate showed that if the active layer of the heterostructure is as thin as several tens of nanometres, the electron levels would be split apart by tens of milli-electron volts, which should be observable. This structure is now called a quantum well.

Henry then calculated how this quantization would alter the optical absorption edge of the semiconductor. His conclusion was that instead of the optical absorption increasing smoothly, the absorption edge of a thin heterostructure would appear as a series of steps.

In addition to Henry's contributions, the quantum well (or double-heterostructure laser, as it was originally known) was actually first proposed in 1963 by Herber Kroemer in Proceedings of the IEEE and simultaneously (in 1963) in the U.S.S.R by Zh. I. Alferov and R.F. Kazarinov . Alferov and Kroemer shared a Nobel Prize in 2000 for their work in semiconductor heterostructures.

Experimental verification of quantum wells

In early 1973, Henry proposed to R. Dingle, a physicist in his department, that he look for these predicted steps. The very thin heterostructures were made by W. Wiegmann using molecular beam epitaxy. The dramatic effect of the steps was observed in the ensuing experiment, published in 1974 .

Invention of the quantum well laser

After this experiment showed the reality of the predicted quantum well energy levels, Henry tried to think of an application. He realized that the quantum well structure would alter the density of states of the semiconductor, and result in an improved semiconductor laser requiring fewer electrons and electron holes to reach laser threshold. Also, he realized that the laser wavelength could be changed merely by changing the thickness of the thin quantum well layers, whereas in the conventional laser a change in wavelength requires a change in layer composition. Such a laser, he reasoned, would have superior performance characteristics compared to the standard double heterostructure lasers being made at that time.

Dingle and Henry received a patent on this new type of semiconductor laser comprising a pair of wide bandgap layers having an active region sandwiched between them, in which "the active layers are thin enough (e.g., about 1 to 50 nanometres) to separate the quantum levels of electrons confined therein. These lasers exhibit wavelength tunability by changing the thickness of the active layers. Also described is the possibility of threshold reductions resulting from modification of the density of electron states." The patent was issued on September 21, 1976, entitled "Quantum Effects in Heterostructure Lasers," U.S. Patent No. 3,982,297.

Quantum well lasers require fewer electrons and holes to reach threshold than conventional double heterostructure lasers. A well-designed quantum well laser can have an exceedingly low threshold current.

Moreover, since quantum efficiency (photons-out per electrons-in) is largely limited by optical absorption by the electrons and holes, very high quantum efficiencies can be achieved with the quantum well laser.

To compensate for the reduction in active layer thickness, a small number of identical quantum wells are often used. This is called a multi-quantum well laser.

Early demonstrations

The term "quantum well laser" was coined in the late 1970s by Nick Holonyak and his students at the University of Illinois at Urbana Champaign. They demonstrated the first quantum well laser in 1977. W.T. Tsang at Bell Laboratories in the late 1970s and early 1980s showed the promise of quantum well lasers by demonstrating that when the quantum well parameters are optimized, they have exceedingly low threshold current and very high efficiency in converting current-in to light-out, making them ideal for widespread use.

The story of the origin of the quantum well concept, its experimental verification, and the invention of the quantum well laser is told by Henry in more detail in the Foreword to "Quantum Well Lasers," ed. by Peter S. Zory, Jr.

Hybrid silicon laser

A **hybrid silicon laser** is a semiconductor laser fabricated from both silicon and group III-V semiconductor materials. The hybrid silicon laser was developed to address the lack of a silicon laser to enable fabrication of low-cost, mass-producible silicon optical devices. The hybrid approach takes advantage of the light-emitting properties of III-V semiconductor materials combined with the process maturity of silicon to fabricate electrically driven lasers on a silicon wafer that can be integrated with other silicon photonic devices.

Physics
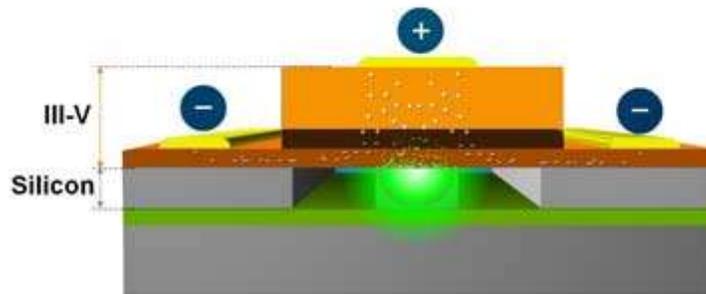
Schematic cross-section of a hybrid silicon laser

A hybrid silicon laser is an optical source that is fabricated from both silicon and group III-V semiconductor materials (e.g. Indium(III) phosphide, Gallium(III) arsenide). It comprises a silicon waveguide fused to an active, light-emitting, III-V epitaxial semiconductor wafer. The III-V epitaxial wafer is designed with different layers such that the active layer can emit light when it is excited either by shining light, e.g. a laser onto it; or by passing electricity through it. The emitted light from the active layer couples into the silicon waveguide due to their close proximity (<130 nm separation) where it can be guided to reflect off mirrors at the end of the silicon waveguide to form the laser cavity.

Fabrication

The hybrid silicon laser is fabricated by a technique called plasma assisted wafer bonding. Silicon waveguides are first fabricated on a silicon on insulator (SOI) wafer. This SOI wafer and the un-patterned III-V wafer are then exposed to an oxygen plasma before being pressed together at a low (for semiconductor manufacturing) temperature of 300C for 12hours. This process fuses the two wafers together. The III-V wafer is then etched into mesas to expose electrical layers in the epitaxial structure. Metal contacts are fabricated on these contact layers allowing electrical current to flow to the active region.

Uses

Intel suggests this light source could be used for optical communications when integrated with silicon photonics. Silicon manufacturing and fabrication is widely used in the electronic industry to mass-produce low-cost electronic devices. Silicon photonics uses these same electronic manufacturing technologies to make low cost integrated optical devices. One issue with using silicon for an optical device is that silicon is a poor light emitter and cannot be used to make an electrically pumped laser. This means that lasers have first to be fabricated on a separate III-V semiconductor wafer before being individually aligned to each silicon device, in a process that is both costly and time-consuming, limiting the total number of lasers that can be used on a silicon photonic circuit.

By using this wafer bonding technique many hybrid silicon lasers can be fabricated simultaneously on a silicon wafer, all aligned to the silicon photonic devices. Potential uses cited in the references below include fabricating many, possibly hundreds of hybrid silicon lasers on a die and using silicon photonics to combine them together to form high bandwidth optical links for personal computers, servers or back planes.

History

- Pulsed optically pumped lasing first demonstrated by UCSB
- Continuous wave optically pumped lasing demonstrated by Intel and UCSB
- Continuous wave electrically driven lasing demonstrated by UCSB and Intel

Chapter 7

Quantum Cascade Laser

Quantum cascade lasers (QCLs) are semiconductor lasers that emit in the mid- to far-infrared portion of the electromagnetic spectrum and were first demonstrated by Jerome Faist, Federico Capasso, Deborah Sivco, Carlo Sirtori, Albert Hutchinson, and Alfred Cho at Bell Laboratories in 1994.

Unlike typical interband semiconductor lasers that emit electromagnetic radiation through the recombination of electron–hole pairs across the material band gap, QCLs are unipolar and laser emission is achieved through the use of intersubband transitions in a repeated stack of semiconductor multiple quantum well heterostructures, an idea first proposed in the paper "Possibility of amplification of electromagnetic waves in a semiconductor with a superlattice" by R.F. Kazarinov and R.A. Suris in 1971.
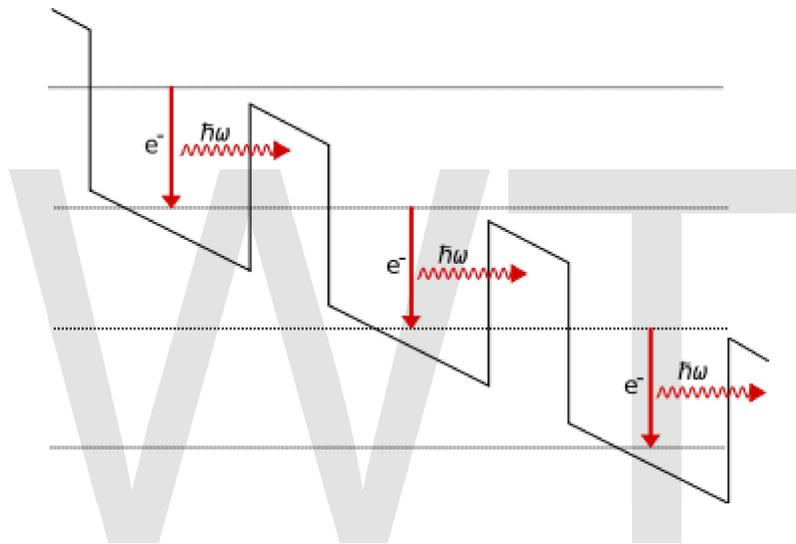
Intersubband vs. interband transitions
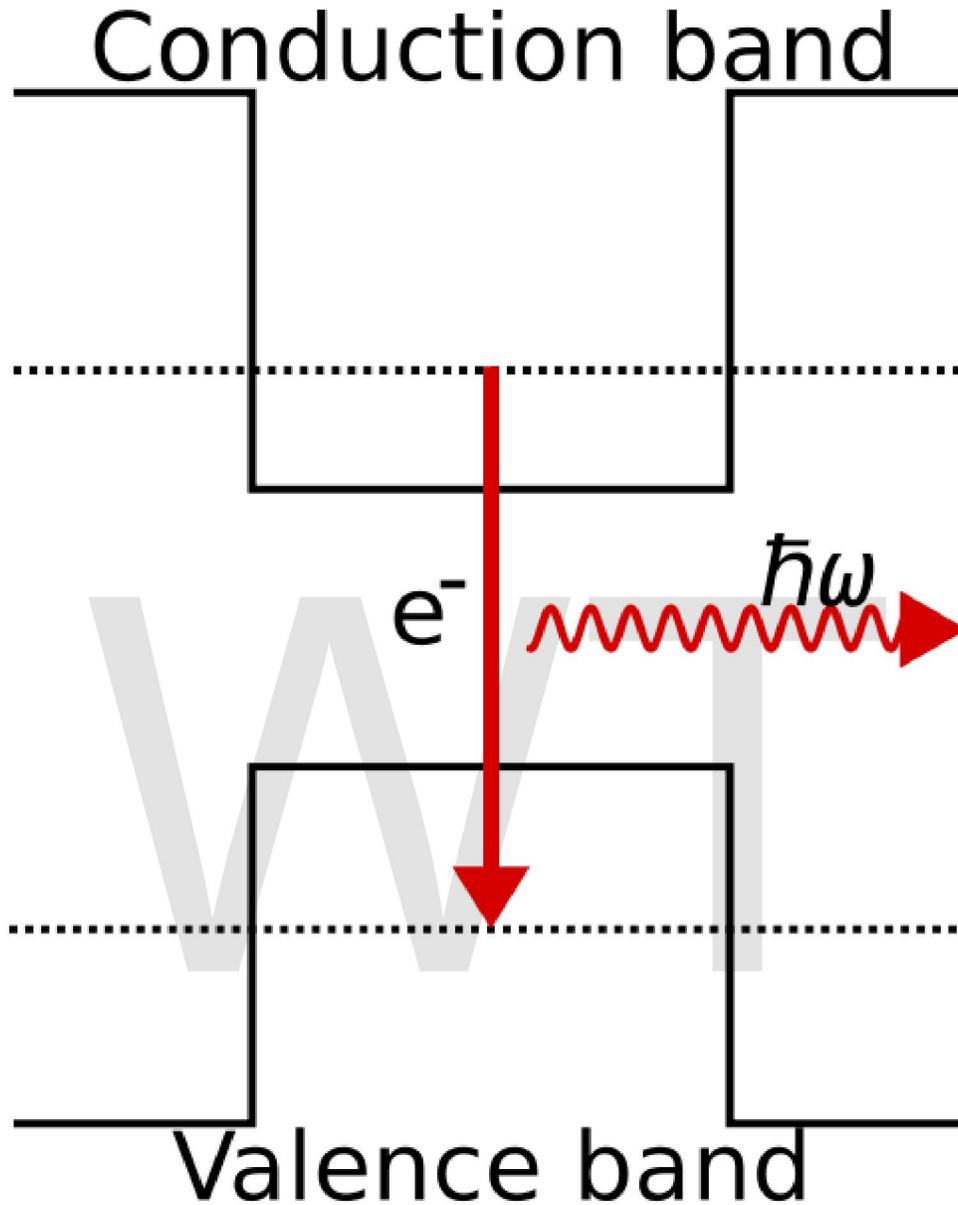
Interband transitions in conventional semiconductor lasers emit a single photon.

Within a bulk semiconductor crystal, electrons may occupy states in one of two continuous energy bands - the valence band, which is heavily populated with low energy electrons and the conduction band, which is sparsely populated with high energy electrons. The two energy bands are separated by an energy band gap in which there are no permitted states available for electrons to occupy. Conventional semiconductor laser diodes generate light by a single photon being emitted when a high energy electron in the conduction band recombines with a hole in the valence band. The energy of the photon and hence the emission wavelength of laser diodes is therefore determined by the band gap of the material system used.

A QCL however does not use bulk semiconductor materials in its optically active region. Instead it comprises a periodic series of thin layers of varying material composition forming a superlattice. The superlattice introduces a varying electric potential across the length of the device, meaning that there is a varying probability of electrons occupying different positions over the length of the device. This is referred to as one-dimensional multiple quantum well confinement and leads to the splitting of the band of permitted energies into a number of discrete electronic subbands. By suitable design of the layer thicknesses it is possible to engineer a population inversion between two subbands in the system which is required in order to achieve laser emission. Since the position of the energy levels in the system is primarily determined by the layer thicknesses and not the material, it is possible to tune the emission wavelength of QCLs over a wide range in the same material system.


In quantum cascade structures, electrons undergo intersubband transitions and photons are emitted. The electrons tunnel to the next period of the structure and the process repeats.

Additionally, in semiconductor laser diodes, electrons and holes are annihilated after recombining across the band gap and can play no further part in photon generation. However in a unipolar QCL, once an electron has undergone an intersubband transition and emitted a photon in one period of the superlattice, it can tunnel into the next period of the structure where another photon can be emitted. This process of a single electron causing the emission of multiple photons as it traverses through the QCL structure gives rise to the name *cascade* and makes a quantum efficiency of greater than unity possible which leads to higher output powers than semiconductor laser diodes.

Operating principles

Rate equations
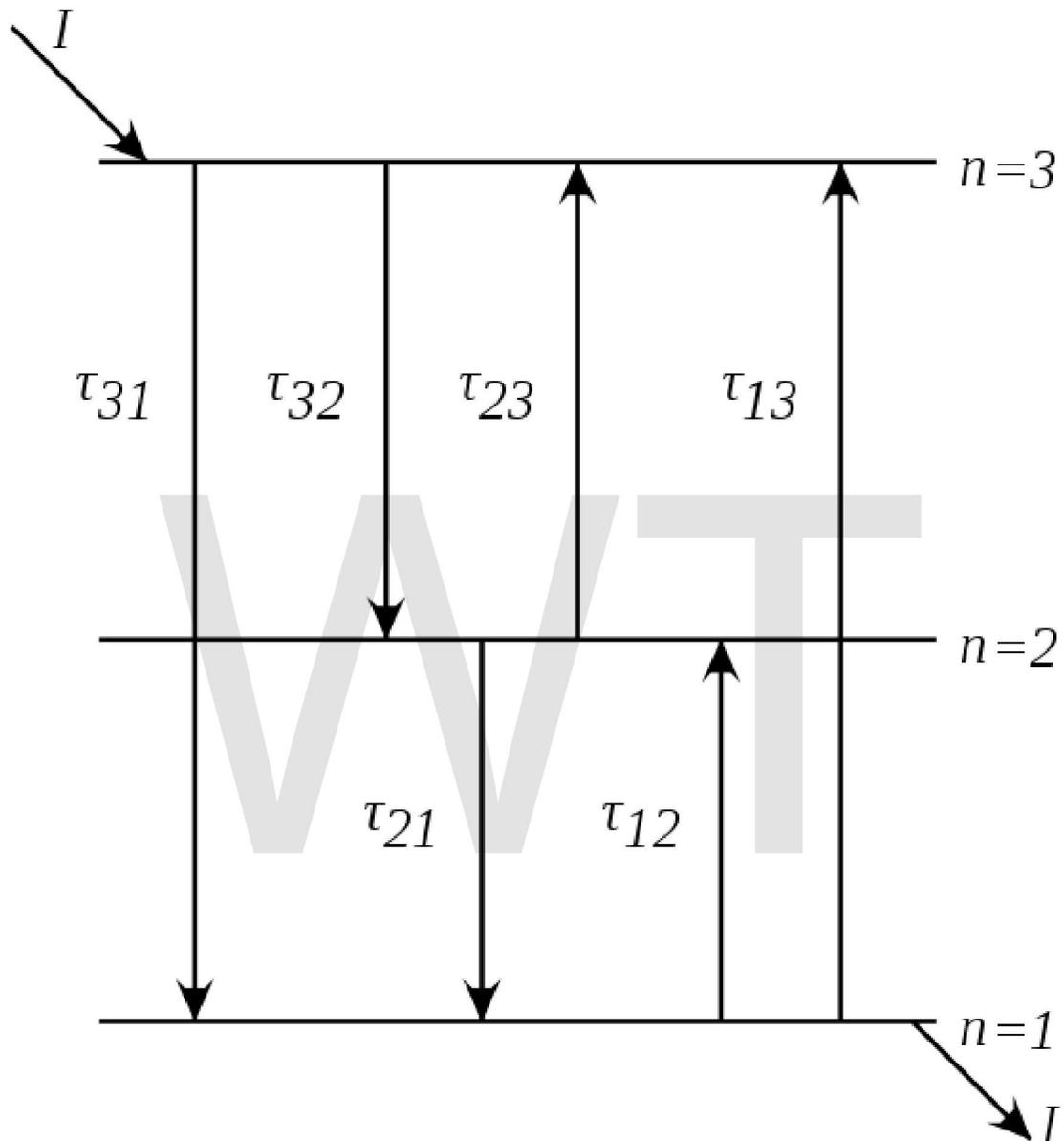
Subband populations are determined by the intersubband scattering rates and the injection/extraction current.

QCLs are typically based upon a three-level system. Assuming the formation of the wavefunctions is a fast process compared to the scattering between states, the time independent solutions to the Schrödinger equation may be applied and the system can be modelled using rate equations. Each subband contains a number of electrons n_i (where i is the subband index) which scatter between levels with a lifetime τ_{if} (reciprocal of the average intersubband scattering rate W_{if}), where i and f are the initial and final subband

indices. Assuming that no other subbands are populated, the rate equations for the three level lasers are given by:

$$\frac{dn_3}{dt} = I_{\text{in}} + \frac{n_1}{\tau_{13}} + \frac{n_2}{\tau_{23}} - \frac{n_3}{\tau_{31}} - \frac{n_3}{\tau_{32}}$$

$$\frac{dn_2}{dt} = \frac{n_3}{\tau_{32}} + \frac{n_1}{\tau_{12}} - \frac{n_2}{\tau_{21}} - \frac{n_2}{\tau_{23}}$$

$$\frac{dn_1}{dt} = \frac{n_2}{\tau_{21}} + \frac{n_3}{\tau_{31}} - \frac{n_1}{\tau_{13}} - \frac{n_1}{\tau_{12}} - I_{\text{out}}$$

In the steady state, the time derivatives are equal to zero and $I_{\text{in}} = I_{\text{out}} = I$. The general rate equation for electrons in subband i of an N level system is therefore:

$$\frac{dn_i}{dt} = \sum_{j=1}^{N} \frac{n_j}{\tau_{ji}} - n_i \sum_{j=1}^{N} \frac{1}{\tau_{ij}} + I(\delta_{i0} - \delta_{iN})$$

,

Under the assumption that absorption processes can be ignored (which is valid at low temperatures), the middle rate equation gives

$$\frac{n_3}{\tau_{32}} = \frac{n_2}{\tau_{21}}$$

Therefore if $\tau_{32} > \tau_{21}$ (i.e. $W_{21} > W_{32}$) then $n_3 > n_2$ and a population inversion will exist. The population ratio is defined as

$$\frac{n_3}{n_2} = \frac{\tau_{32}}{\tau_{21}} = \frac{W_{21}}{W_{32}}$$

If all N steady-state rate equations are summed, the right hand side becomes zero, meaning that the system is underdetermined, and it is possible only to find the relative population of each subband. An additional equation is required to set the total number of carriers equal to the total number of dopant ions:

$$\sum_{i=1}^{N} n_i = N_{\text{2D}}$$

.
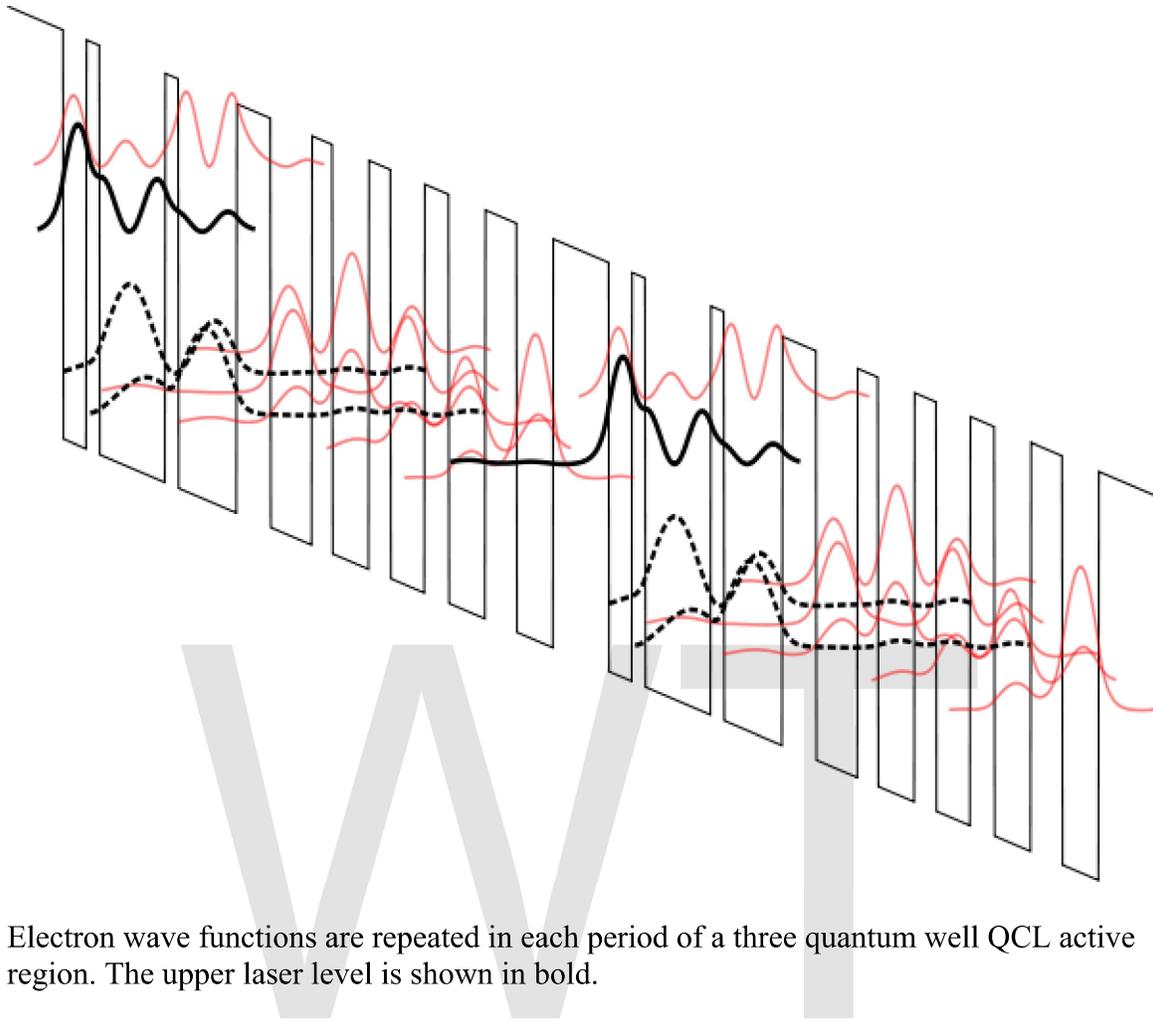
Electron wave functions are repeated in each period of a three quantum well QCL active region. The upper laser level is shown in bold.

Active region designs

The scattering rates are tailored by suitable design of the layer thicknesses in the superlattice which determine the electron wave functions of the subbands. The scattering rate between two subbands is heavily dependent upon the overlap of the wave functions and energy spacing between the subbands. The figure shows the wave functions in a three quantum well (3QW) QCL active region and injector.

In order to decrease W_{32}, the overlap of the upper and lower laser levels is reduced. This is often achieved through designing the layer thicknesses such that the upper laser level is mostly localised in the left-hand well of the 3QW active region, while the lower laser level wave function is made to mostly reside in the central and right-hand wells. This is known as a *diagonal* transition. A *vertical* transition is one in which the upper laser level is localised in mainly the central and right-hand wells. This increases the overlap and hence W_{32} which reduces the population inversion, but it increases the strength of the radiative transition and therefore the gain.

In order to increase W_{21}, the lower laser level and the ground level wave functions are designed such that they have a good overlap and to increase W_{21} further, the energy

spacing between the subbands is designed such that it is equal to the longitudinal optical (LO) phonon energy (~36 meV in GaAs) so that resonant LO phonon-electron scattering can quickly depopulate the lower laser level.

Material systems

The first QCL was fabricated in the InGaAs/InAlAs material system lattice-matched to an InP substrate. This particular material system has a conduction band offset (quantum well depth) of 520 meV. These InP-based devices have reached very high levels of performance across the mid-infrared spectral range, achieving high power, above room-temperature, continuous wave emission.

In 1998 GaAs/AlGaAs QCLs were demonstrated by Sirtori *et al.* proving that the QC concept is not restricted to one material system. This material system has a varying quantum well depth depending on the aluminium fraction in the barriers. Although GaAs-based QCLs have not matched the performance levels of InP-based QCLs in the mid-infrared, they have proven to be very successful in the terahertz region of the spectrum.

The short wavelength limit of QCLs is determined by the depth of the quantum well and recently QCLs have been developed in material systems with very deep quantum wells in order to achieve short wavelength emission. The InGaAs/AlAsSb material system has quantum wells 1.6 eV deep and has been used to fabricate QCLs emitting at 3 μm. InAs/AlSb QCLs have quantum wells 2.1 eV deep and electroluminescence at wavelengths as short as 2.5 μm has been observed.

QCLs may also allow laser operation in materials traditionally considered to have poor optical properties. Indirect bandgap materials such as silicon have minimum electron and hole energies at different momentum values. For interband optical transitions, carriers change momentum through a slow, intermediate scattering process, dramatically reducing the optical emission intensity. Intersubband optical transitions however, are independent of the relative momentum of conduction band and valence band minima and theoretical proposals for Si/SiGe quantum cascade emitters have been made.
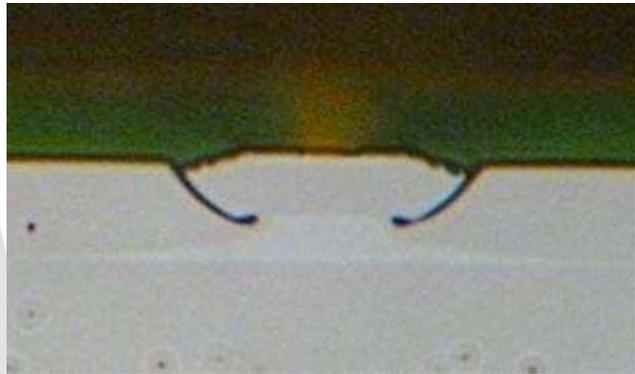
Emission wavelengths

QCLs currently cover the wavelength range from 2.75–250 μm (and extends to 355 μm with the application of a magnetic field).

Optical waveguides
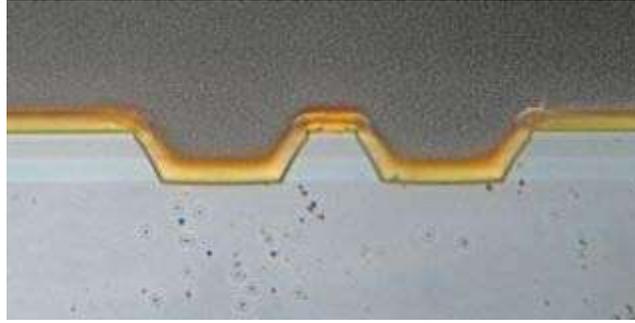
End view of QC facet with ridge waveguide. Darker gray: InP, lighter gray: QC layers, black: dielectric, gold: Au coating. Ridge ~ 10 um wide.

End view of QC facet with buried heterostructure waveguide. Darker gray: InP, lighter gray: QC layers, black: dielectric. Heterostructure ~ 10 um wide

The first step in processing quantum cascade gain material to make a useful light-emitting device is to confine the gain medium in an optical waveguide. This makes it possible to direct the emitted light into a collimated beam, and allows a laser resonator to be built such that light can be coupled back into the gain medium.

Two types of optical waveguides are in common use. A ridge waveguide is created by etching parallel trenches in the quantum cascade gain material to create an isolated stripe of QC material, typically ~10 um wide, and several mm long. A dielectric material is typically deposited in the trenches to guide injected current into the ridge, then the entire ridge is typically coated with gold to provide electrical contact and to help remove heat from the ridge when it is producing light. Light is emitted from the cleaved ends of the waveguide, with an active area that is typically only a few micrometers in dimension.

The second waveguide type is a buried heterostructure. Here, the QC material is also etched to produce an isolated ridge. Now, however, new semiconductor material is grown over the ridge. The change in index of refraction between the QC material and the overgrown material is sufficient to create a waveguide. Dielectric material is also deposited on the overgrown material around QC ridge to guide the injected current into

the QC gain medium. Buried heterostructure waveguides are efficient at removing heat from the QC active area when light is being produced.

Laser types

Although the quantum cascade gain medium can be used to produce incoherent light in a superluminescent configuration, it is most commonly used in combination with an optical cavity to form a laser.

Fabry–Pérot lasers

This is the simplest of the quantum cascade lasers. An optical waveguide is first fabricated out of the quantum cascade material to form the gain medium. The ends of the crystalline semiconductor device are then cleaved to form two parallel mirrors on either end of the waveguide, thus forming a Fabry–Pérot resonator. The residual reflectivity on the cleaved facets from the semiconductor-to-air interface is sufficient to create a resonator. Fabry–Pérot quantum cascade lasers are capable of producing high powers, but are typically multi-mode at higher operating currents. The wavelength can be changed chiefly by changing the temperature of the QC device.

Distributed feedback lasers

A distributed feedback (DFB) quantum cascade laser is similar to a Fabry–Pérot laser, except for a distributed Bragg reflector (DBR) built on top of the waveguide to prevent it from emitting at other than the desired wavelength. This forces single mode operation of the laser, even at higher operating currents. DFB lasers can be tuned chiefly by changing the temperature, although an interesting variant on tuning can be obtained by pulsing a DFB laser. In this mode, the wavelength of the laser is rapidly "chirped" during the course of the pulse, allowing rapid scanning of a spectral region.

External cavity lasers
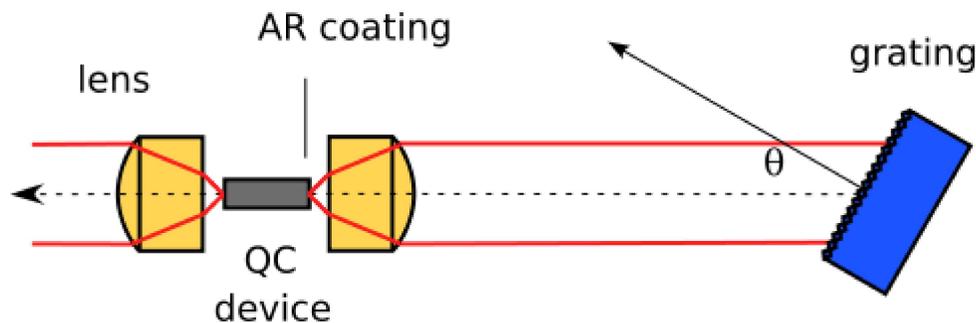
Schematic of QC device in external cavity with frequency selective optical feedback provided by diffraction grating in Littrow configuration.

In an external cavity (EC) quantum cascade laser, the quantum cascade device serves as the laser gain medium. One, or both, of the waveguide facets have an anti-reflection

coating that defeats the optical cavity action of the cleaved facets. Mirrors are then arranged in a configuration external to the QC device to create the optical cavity.

If a frequency-selective element is included in the external cavity, it is possible to reduce the laser emission to a single wavelength, and even tune the radiation. For example, diffraction gratings have been used to create a tunable laser that can tune over 15% of its center wavelength.

Growth

The alternating layers of the two different semiconductors which form the quantum heterostructure may be grown on to a substrate using a variety of methods such as molecular beam epitaxy (MBE) , metalorganic vapour phase epitaxy (MOVPE), or metalorganic chemical vapor deposition (MOCVD).

Applications

Distributed feedback (DFB) quantum cascade lasers were first commercialized in 2004, and broadly-tunable external cavity quantum cascade lasers first commercialized in 2006. The high optical power output, tuning range and room temperature operation make QCLs useful for spectroscopic applications such as remote sensing of environmental gases and pollutants in the atmosphere and homeland security. They may eventually be used for vehicular cruise control in conditions of poor visibility, collision avoidance radar, industrial process control, and medical diagnostics such as breath analyzers. QCLs are also used to study plasma chemistry.

Their large dynamic range, excellent sensitivity, and failsafe operation combined with the solid-state reliability should easily overcome many of the technological hurdles that impede existing technology in these markets. When used in multiple-laser systems, intrapulse QCL spectroscopy offers broadband spectral coverage that can potentially be used to identify and quantify complex heavy molecules such as those in toxic chemicals, explosives, and drugs.

Unguided QCL emission in the 3–5 μm atmospheric window could be used as a cheaper alternative to optical fibres for high-speed Internet access in built up areas.

Chapter 8

Diode-Pumped Solid-State Laser
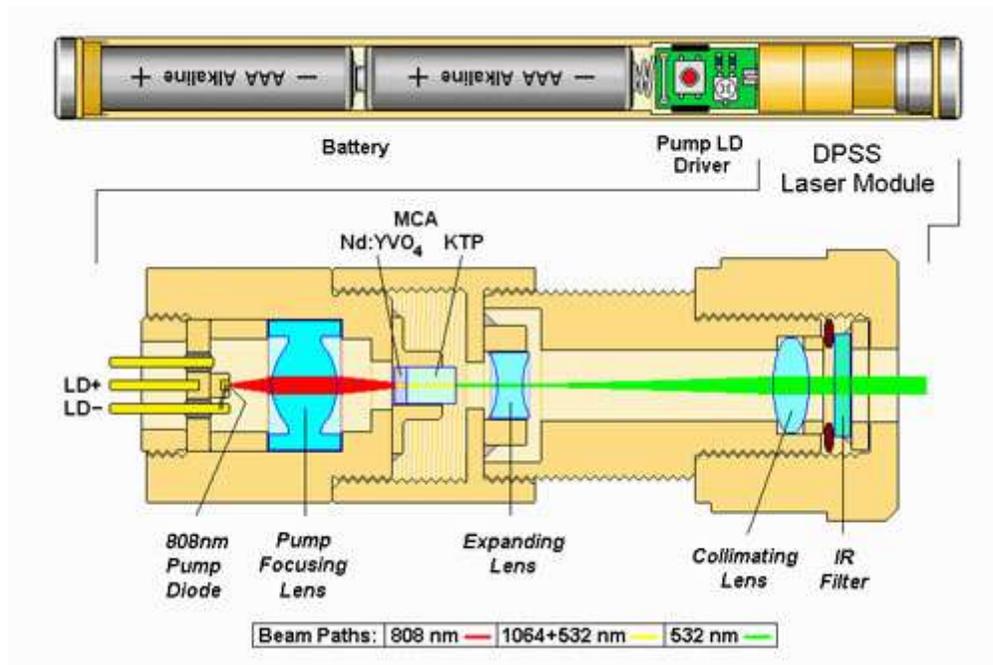
A frequency-doubled green laser pointer, showing internal construction. Cells and electronics lead to a laser head module. This contains a powerful 808 nm IR diode laser that pumps a Nd:YVO$_4$ laser crystal, that in turn outputs 1064 nm light. This immediately is doubled inside a non-linear KTP crystal, resulting in green light at the half-wavelength of 532 nm. This beam is expanded and infrared-filtered. In inexpensive lasers the IR filter is inadequate, or is omitted.

Diode-pumped solid-state (DPSS) lasers are solid-state lasers made by pumping a solid gain medium, for example, a ruby or a neodymium-doped YAG crystal, with a laser diode.

DPSS lasers have advantages in compactness and efficiency over other types, and high power DPSS lasers have replaced ion lasers and flashlamp-pumped lasers in many scientific applications, and are now appearing commonly in green and other color laser pointers.

Coupling

The wavelength of the laser diodes is tuned by means of temperature to produce an optimal compromise between the absorption coefficient in the crystal and energy efficiency (low as possible pump photon energy). As waste energy is limited by the thermal lens this means higher power densities compared to high-intensity discharge lamps.

High power lasers use a single crystal, but many laser diodes are arranged in strips (multiple diodes next to each other in one substrate) or stacks (stacks of substrates). This diode grid can be imaged onto the crystal by means of a lens. Higher brightness (leading to better beam profile and longer diode lifetimes) is achieved by optically removing the dark areas between the diodes, which are needed for cooling and delivering the current. This is done in two steps:

1. The "fast axis" is collimated with an aligned grating of cylindrical micro-lenses.
2. The partially collimated beams are then imaged at reduced size into the crystal. The crystal can be pumped longitudinally from both end faces or transversely from three or more sides.

The beams from multiple diodes can also be combined by coupling each diode into an optical fibre, which is placed precisely over the diode (but behind the micro-lens). At the other end of the fiber bundle, the fibers are fused together to form a uniform, gap-less, round profile on the crystal. This also permits the use of a remote power supply.

Some numbers

High power laser diodes are fabricated as bars with multiple single strip laser diodes next to each other.

Each single strip diode typically has an active volume of:

1 μm	2 mm	100 μm
Height	Depth	Width
fast axis	optical axis	slow axis

and depending on the cooling technique for the whole bar (100 to 200) μm distance to the next laser diode.

The end face of the diode along the fast axis can be imaged onto strip of 1 μm height. But the end face along the slow axis can be imaged onto a smaller area then 100 μm. This is due to the small divergence (hence the name: 'slow axis') which is given by the ratio of depth to width. Using the above numbers the fast axis could be imaged onto a 5 μm wide spot.

So to get a beam which is equal divergence in both axis, the end faces of a bar composed of 5 laser diodes, can be imaged by means of 4 (acylindrical) cylinder lenses onto an image plane with 5 spots each with a size of 5 mm x 1 mm. This large size is needed for low divergence beams. Low divergence allows paraxial optics, which is cheaper, and which is used to not only generate a spot, but a long beam waist inside the laser crystal (length = 50 mm), which is to be pumped through its end faces.

Also in the paraxial case it is much easier to use gold or copper mirrors or glass prisms to stack the spots on top of each other, and get a 5 x 5 mm beam profile. A second pair of (spherical) lenses image this square beam profile inside the laser crystal.

In conclusion a volume of 0.001 mm³ active volume in the laser diode is able to saturate 1250 mm³ in a Nd:YVO$_4$ crystal.

Common DPSS processes
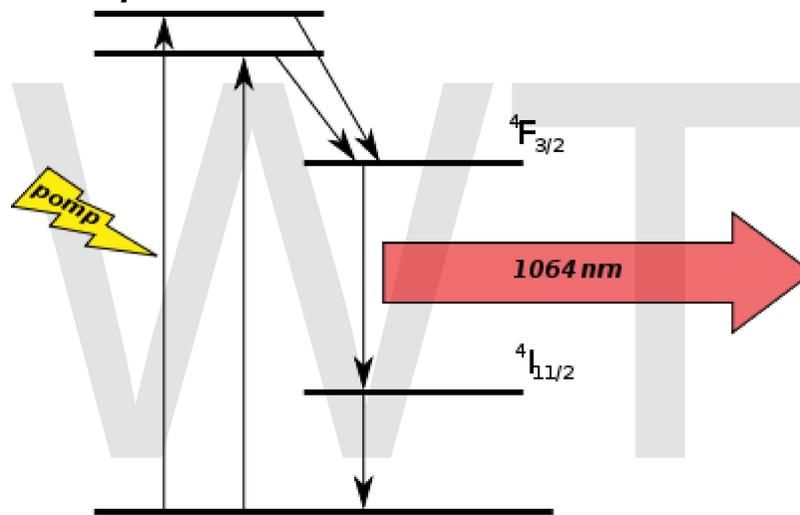
Neodymium ions in various types of ionic crystals, and also in glasses, act as a laser gain medium, typically emitting 1,064 nm light from a particular atomic transition in the neodymium ion, after being "pumped" into excitation from an external source. Selection of 946 nm transition light is possible, as well

The most common DPSS laser in use is the 532 nm wavelength **green** laser pointer. A powerful (>200 mW) 808 nm wavelength infrared GaAlAs laser diode pumps a neodymium-doped yttrium aluminium garnet (Nd:YAG) or a neodymium-doped yttrium orthovanadate (Nd:YVO$_4$) crystal which produces 1064 nm wavelength light from the main spectral transition of neodymium ion. This light is then *frequency doubled* using a nonlinear optical process in a KTP crystal, producing 532 nm light. Green DPSS lasers are usually around 20% efficient, although some lasers can reach up to 35% efficiency. In other words, a green DPSS laser using a 2.5 W pump diode would be expected to output around 500-900 mW of 532 nm light.

In optimal conditions, Nd:YVO$_4$ has a conversion efficiency of 60%, while KTP has a conversion efficiency of 80%. In other words, a green DPSS laser can theoretically have an overall efficiency of 48%.

In the realm of very high output powers, the KTP crystal becomes susceptible to optical damage. Thus, high-power DPSS lasers generally have a larger beam diameter, as the 1064 nm laser is expanded before it reaches the KTP crystal, reducing the irradiance from the infrared light. In order to maintain a lower beam diameter, a crystal with a higher damage threshold, such as LBO, is used instead.

Blue DPSS lasers use a nearly identical process, except that the 808 nm light is being converted by an Nd:YAG crystal to 946 nm light (selecting this non-principal spectral line of neodymium in the same Nd-doped crystals), which is then frequency-doubled to 473 nm by a beta barium borate (BBO) or lithium triborate (LBO) crystal. Because of the lower gain for the materials, blue lasers are relatively weak, and are only around 3-5% efficient. In the late 2000s, it was discovered that bismuth triborate (BiBO) crystals were more efficient than BBO and LBO and do not have the disadvantage of being hygroscopic, which degrades the crystal if it is exposed to moisture.

Violet DPSS lasers at 404 nm have been produced which directly double the output of a 1,000 mW 808 nm GaAlAs pump diode, for a violet light output of 120 mW (12% efficiency). These lasers out-perform 50 mW gallium nitride (GaN) direct 405 nm Blu-ray diode lasers, but the frequency-doubled violet lasers also have a considerable infrared component in the beam, resulting from the pump diode.

Yellow DPSS lasers use an even more complicated process: A 808 nm pump diode is used to generate 1,064 nm and 1,342 nm light, which is summed to become 593.5 nm. Due to their complexity, most yellow DPSS lasers are only around 1% efficient, and usually more expensive per unit of power.

Another method is to generate 1,064 and 1,319 nm light, which are summed to 589 nm. This process is more efficient, with about 3% of the pump diode's power being converted to yellow light.

Comparison to diode lasers

DPSS and diode lasers are two of the most common types of solid-state lasers. However, both types have their advantages and disadvantages.

DPSS lasers generally have a higher beam quality and can reach very high powers while maintaining a relatively good beam quality. Because the crystal pumped by the diode acts as its own laser, the quality of the output beam is independent of that of the input beam. In comparison, diode lasers can only reach a few hundred milliwatts unless they operate in multiple transverse mode. Such multi-mode lasers have larger a beam diameter and a greater divergence, which makes them less desirable. In fact, single-mode operation is essential in some applications, such as optical drives.

On the other hand, diode lasers are cheaper and more energy efficient. As DPSS crystals are not 100% efficient, some power is lost when the frequency is converted. DPSS lasers are also more sensitive to temperature and can only operate optimally within a small range. Otherwise, the laser would suffer from stability issues, such as hopping between modes and large fluctuations in the output power. DPSS lasers also require a more complex construction.

Chapter 9

Superluminescent Diode

A **superluminescent diode** (SLED or SLD) is an edge-emitting semiconductor light source based on superluminescence. It combines the high power and brightness of laser diodes with the low coherence of conventional light-emitting diodes. Its emission band is 5–100 nm wide.

History

In 1986 Dr. Gerard A. Alphonse at RCA Laboratories (now Sarnoff Corporation), invented the superluminescent diode. This light source was developed as a key component in the next generations of fibre optic gyroscopes, low coherence tomography for medical imaging, and external cavity tunable lasers with applications to fiber-optic communications. In 1989 the technology was transferred to GE-RCA in Canada, which became a division of EG&G. Superluminescent light emitting diodes are also called sometimes superluminescent diodes, superluminescence diodes or superluminescent LEDs.

Principles of operation

A superluminescent light emitting diode is, similar to a laser diode, based on an electrically driven pn-junction that, when biased in forward direction, becomes optically active and generates amplified spontaneous emission over a wide range of wavelengths. The peak wavelength and the intensity of the SLED depend on the active material composition and on the injection current level. SLEDs are designed to have high single pass amplification for the spontaneous emission generated along the waveguide but, unlike laser diodes, insufficient feedback to achieve lasing action. This is obtained very successfully through the joint action of a tilted waveguide and anti-reflection coated (ARC) facets.
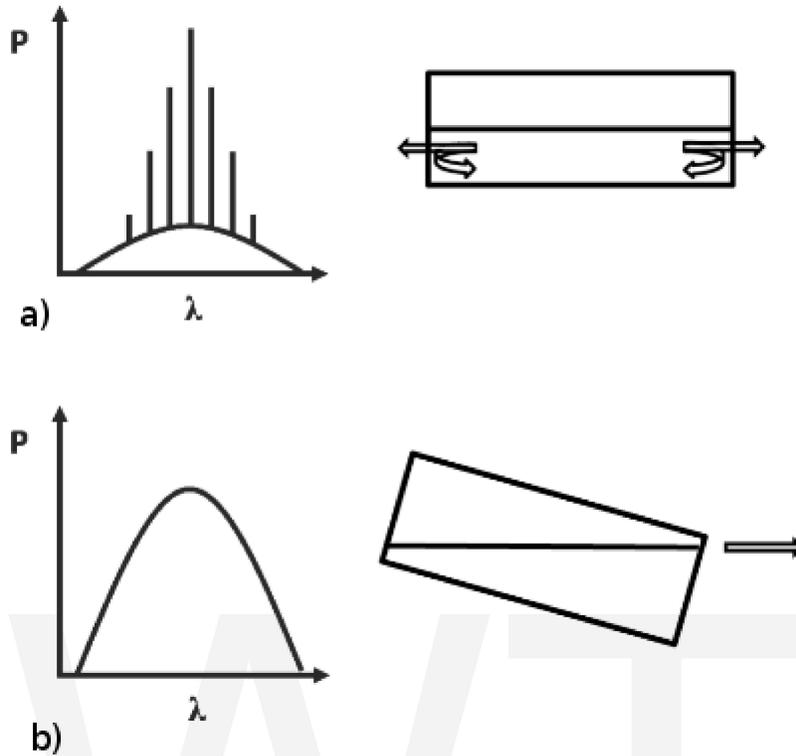
a) Facet feedback and wavelength resonances in the optical emission spectrum of a multimode Fabry-Perot laser; b) power spectral density of a superluminescent light emitting diode.

When an electrical forward voltage is applied an injection current across the active region of the SLED is generated. Like most semiconductor devices, a SLED consists of a positive (p-doped) section and a negative (n-doped) section. Electrical current will flow from the p-section to the n-section and across the active region that is sandwiched in between the p- and n-section. During this process, light is generated through spontaneous and random recombination of positive (holes) and negative (electrons) electrical carriers and then amplified when travelling along the waveguide of a SLED.

The pn-junction of the semiconductor material of a SLED is designed in such a way that electrons and holes feature a multitude of possible states (energy bands) with different energies. Therefore, the recombination of electron and holes generates light with a broad range of optical frequencies, i.e. broadband light.

The output power performance of an ideal SLED can be described with a simple model, not taking spectral effects into account and considering both a uniform distribution of carrier densities and zero reflections from the facets.

$$P_{out} = \frac{h}{c} \cdot \nu \cdot \Pi \cdot R_{sp} \frac{exp[(g - \alpha)L] - 1}{g - \alpha}$$

Where h is the Planck constant, ν the optical frequency, Π the size of the optical mode, R_{sp} the spontaneous emission rate into the guided mode, g the modal gain, α the non-resonant optical losses, L the length of the active channel and c the velocity of light.

So the output power depends linearly on the spontaneous emission rate and exponentially on the optical gain. Obviously a high modal gain is required to obtain high optical output power.

Main characteristics
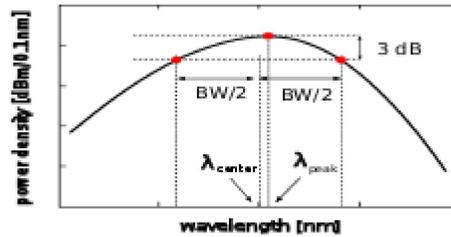
Power dependence of current
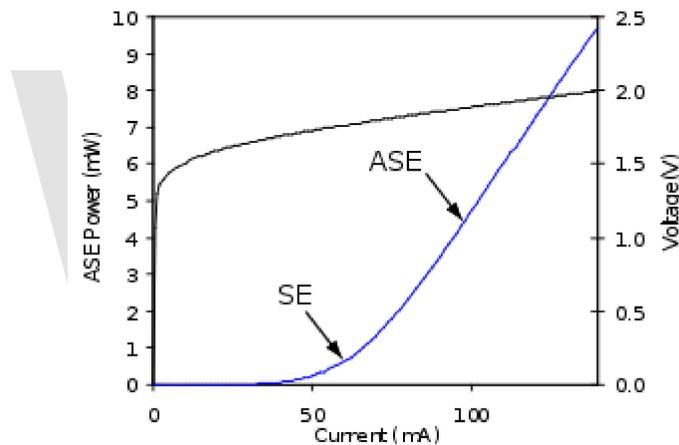
Typical dependence of the fibre-coupled optical power vs. injected current for a SLED module with a central wavelength of 1550 nm, a 3-dB bandwidth of 60 nm and a typical output power of 1.5 mW at 20°C.

The total optical power emitted by an SLED depends on the injected current (bias). Unlike laser diodes, the output intensity does not exhibit a sharp threshold but it gradually increases with current. A soft knee in the power vs. current curve defines a transition between a regime dominated by spontaneous emission (typical for surface emitting LEDs) and one that is dominated by amplified spontaneous emission (i.e. superluminescence). Even if the output power is based on spontaneous emission it has to be noted that the amplification mechanism affects the polarization state of the emitted radiation in a way which is related to the SLED structure and on the operating conditions.

The maximum value of the current that allows a safe operation of the device depends on the model and ranges between 70 mA (for low power SLED) and 500 mA for the most powerful devices.

Centre wavelength and optical bandwidth

Typical dependence of the optical power density versus wavelength for a Superluminescent diode module with a central wavelength of 1560 nm operated at 350 mA.

The optical power emitted by SLEDs is distributed over a wide spectral range. Two useful parameters that are related to the power density distribution at different wavelengths are the optical bandwidth (BW) and the peak wavelength, λ_{peak}. The first is defined as the full width at half maximum (FWHM) of the power density vs. wavelength curve at the nominal operating conditions while the latter corresponds to the wavelength having the highest intensity. The centre wavelength, λ_{centre} is defined as the central point between the two FWHM points of the spectral curve; it can be different from the peak wavelength since it is related to the spectrum asymmetry.

Typical values for SLED modules are for the BW between 5 nm and 100 nm with central wavelengths covering the range between 400 nm and 1700 nm. A trade off between maximum output power and bandwidth exists, however, the latter being larger for devices with lower output power.

Spectral ripple
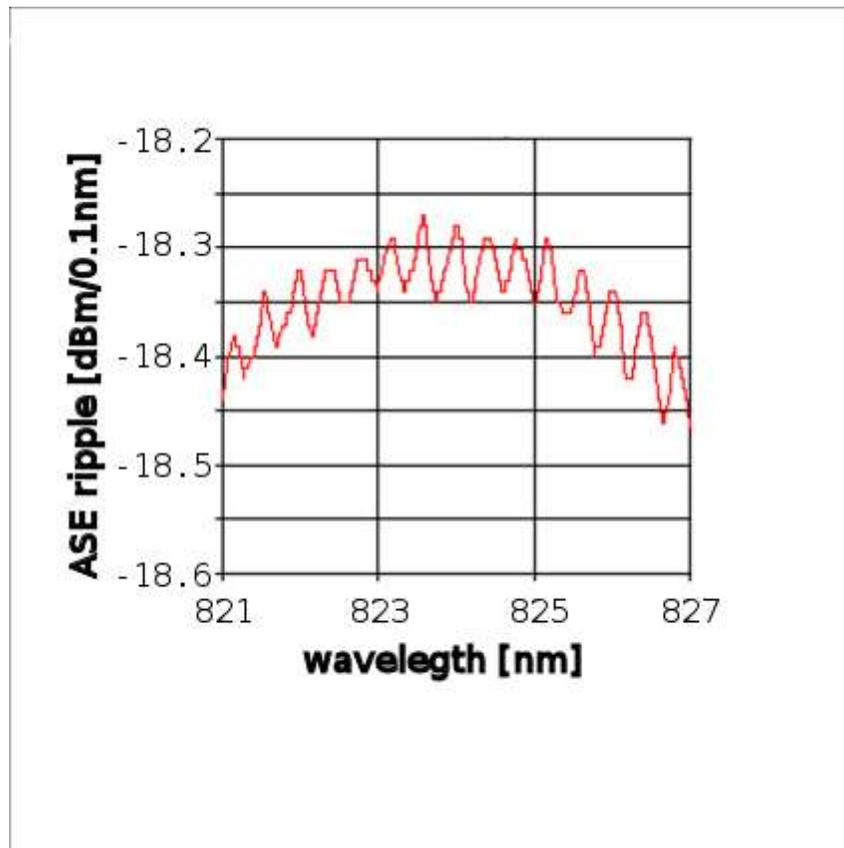
Typical spectral ripple of a 1300 nm SLED at its maximum output power.

The spectral ripple is the measure of the variation of the spectral power-density that can be observed for small change of the wavelength. It can be detected using high-resolution optical spectrum analyzers and can be ascribed to the residual reflectivity of the chip facets and of the coupling fibre. Spectral ripple is more evident in high-power devices and mainly around the peak wavelength where the device gain is higher. It is always present to some extent but undesirable since it has strong effects on the coherence properties of SLED.

Some SLEDs from certain manufacturers exhibit an extremely low value of the ripple even at the highest power levels. An excessive level of optical back-reflection can cause unexpected irregularities of the spectral distribution of SLEDs that have not to be confused with the ripple. During operation it is therefore important to carefully limit the feedback from any additional equipment.

Polarization

As described above, superluminescent light emitting diodes are based on the generation and on the amplification of spontaneous emission in a semiconductor waveguide. The structure and the material composition used for the SLED chip affect the gain that the

radiation experience during the propagation and lead to different amplification factors for different orientations of the electric field (polarization dependent gain). SLEDs operating in the wavelength range of 1300 and 1400 nm are mostly based on a bulk material and a chip structure both characterized by a low polarization dependence of the gain. On the contrary, devices operating in the 1550 and 1620 nm range make mostly use of a quantum well (QW) active region that has a strong polarization-dependent gain. The optical field emitted by the SLED chips, being a combination of unpolarized spontaneous emission and amplified radiation, has therefore a certain degree of polarization (DOP).

A useful quantity that describes the polarization characteristics of the SLED emission is the polarization extinction ratio (PER). This is the ratio between the maximum and the minimum intensities measured after a rotating linear polarizer.

The polarization extinction ratio of bulk chips is around 8–9 dB while it can be as high as 15–20 dB for QW chips. When SLED chips are coupled to pigtail fibres the pigtail bending and coiling will in general modify the polarization state at the fibre output. Modules provided with polarization maintaining (PM) fibre pigtails exhibit high values (>15 dB) of the polarization extinction ratio that are independent on the fibre bending. The polarization extinction ratio of the emission depends also on the bias (i.e. injected current level), having its highest value at the maximum driving current. On the contrary, the polarization state at the output of standard SM fibre pigtail is arbitrary but can be simply modified with a polarization controller and extinction ratios of about 10 dB can be easily achieved.

Relative intensity noise (RIN)

The optical power emitted by semiconductor active devices is always affected by fluctuations (intensity noise) that are induced by the spontaneous emission. When the emitted power is detected with a wide-bandwidth square-low detector the intensity noise will be converted into current fluctuations and the measured photocurrent will include a constant term, I_0, proportional to the mean optical intensity and a time dependent term, I_n, related to the intensity fluctuations.

The spectral distribution of the noise term in the photocurrent can be measured by means of an electrical spectrum analyzer over a radio frequency (RF) range that is limited by the electrical bandwidth of the detector used. The resultant noise spectrum is directly related to the optical intensity noise and in general depends on the RF frequency, ω.

From this measurement a useful parameter that provides quantitative information on the noise of the optical source can be evaluated: it is the relative intensity noise (RIN), that is the ratio between the power spectral density of the noise current, I_n, measured over a given bandwidth, and the square value of the average photocurrent, I_0

$$RIN(\omega) = <I_n^2(\omega)> / <I_0^2>$$

The RIN therefore represents the ratio between the noise power and the average power after detection; the measurement unit used is the dB/Hz. Typical values measured for SLEDs in a frequency range extending from DC up to 500 MHz are reported in the table.

Relativ intensity noise figures in (dB/Hz) of several SLED modules at different driving current levels

SLED center wavelength	100 mA	150 mA	200 mA	300 mA	400 mA	500 mA
1550 nm	−121.5		−123.5			
1550 nm		−124.5	−127.5	−128.0	−129.5	−130.0
1300 nm		−123.5	−125.0	−126.5	−127.0	−127.5
1300 nm		−124.0	−124.5			
1600 nm		−123.0		−123.0		

They depend on the injection current (more correctly on the output power) and on the RF frequency range. The highest measured values never exceed −119 dB/Hz for frequencies higher than 5 GHz, while the lowest value (around 127 dB/Hz) is attained by the most powerful SLEDs in the 1310 nm window and in the frequency range limited to values less than 500 MHz. The frequency dependence of RIN is thought to be related to spatial correlation effects induced by the gain saturation.

It has to be noted that, while the use of narrow band optical filters in front of a detector will usually result in the reduction of the detected noise, the relative intensity noise of SLEDs can exhibit an increase. This behaviour, present mainly in high power SLEDs, is similar to what is observed with multimode Fabry-Perot laser diodes where filtering makes evident the presence of mode partition noise (mostly at low RF frequencies) due to competition among several lasing modes.

Modulation characteristics

Intensity modulation of SLEDs can be easily achieved through direct modulation of the bias current. SLED modules do not include terminating resistors inside because, operating at relatively high currents, excessive cooling would be required to compensate for the heat dissipation of the resistor. In order to achieve the best performance some external network that reduces the impedance mismatch between the driver amplifier, that usually requires 50 Ohm loads, and the low impedance of the chip (a few Ohm) would be preferable. As shown in Fig. , response times of about 1 ns, extinction ratios of 27 dB and 3 dB bandwidths exceeding 200 MHz can be easily achieved.

Similar results can be obtained also for direct modulation of butterfly packaged SLEDs as shown in Fig. . Optically induced modulation allows to exploit the high speed modulation capabilities of the chip when they are not affected by package parasitics; as shown in Fig. , a 3 dB bandwidth exceeding 10 GHz also for packaged SLEDs can be achieved in this case.

Coherence length

SLEDs are optical sources with a rather wide optical bandwidth. In that they differ from both lasers, that have a very narrow spectrum, and white light sources, that exhibit a much larger spectral width. This characteristic mainly reflects itself in a low coherence of the source (which is the limited capability of the emitted light wave to maintain the phase over time). Some applications take advantage of the low coherence of SLEDs sources to achieve high spatial resolution in imaging techniques. The coherence length, L_c, is a quantity frequently used to characterize the coherence of the light source. It is related to the path difference between the two arms of an optical interferometer over which the light wave is still capable to generate an interference pattern. For sources having a Gaussian spectral distribution, the value of L_c is inversely proportional to the spectral width, BW, so that the full width at half maximum (FWHM) of the power spectral density can be related to Lc through the equation

$$L_c = \lambda^2 / BW,$$

where λ is the central wavelength of the emitted radiation. As an example, an SLED operating around 1300 nm and with an optical bandwidth of 100 nm is expected to have a coherence length of about 17 µm. From a practical point of view a definition independent on the spectral distribution (non-Gaussian spectrum) of the source is more suitable. If an optical interferometer is used for the coherence length evaluation a useful quantity is the FWHM value of the visibility, that is the relative amplitude $[(I_{peak} - I_{valley}) / (I_{peak} + I_{valley})]$ of the intensity variations evaluated as a function of the interferometer imbalance.

SLEDs exhibit a large spectral width even at the highest power levels so that corresponding FWHM values of the visibility less than 20 µm are easily achieved.

It is worth noting that the presence of an excessive spectral ripple in the power spectral density results in the presence of side lobes (see Fig.) in the visibility curve that can limit both the spatial resolution and the sensitivity of SLED based measurement systems. SLEDs of certain manufacturers have very low side lobes and allow measurements with high dynamic ranges.

Technical challenges

On the one hand SLEDs are semiconductor devices that are optimized to generate a large amount of amplified spontaneous emission (ASE). In order to do that, they incorporate high-power gain sections in which seeding spontaneous emission is amplified with high gain factors of 30 dB or more.

On the other hand SLEDs lack optical feedback, so that no laser action can occur. Optical feedback resulting from back-reflections of light from optical components such as e.g. connectors into the cavity is suppressed by means of tilting the facets relative to the waveguide, and can be suppressed further with anti-reflection coatings. The formation of

resonator modes and thus pronounced structures in the optical spectrum and/or to spectral narrowing are avoided.

It is therefore natural that even small amounts of back-reflections are amplified inside the SLED chip in a similar manner, producing optical power levels of several tens of milliwatts at the back facet, which may destroy the SLED device. SLEDs should be carefully protected against external optical feedback. Even small levels of feedback can reduce the overall emission bandwidth and the output power, or sometimes even lead to parasitic lasing, causing narrow spikes in the emission spectrum. Some devices may even be damaged by optical feedback. Note that the Fresnel reflection from a perpendicularly cleaved fiber end is already well above the level of feedback which can be tolerated. If back reflections cannot be avoided, an optical isolator must be installed directly behind the SLED module. The isolator provides a low insertion loss from the SLED to the fiber and a high insertion loss in the back direction. However, SLEDs from certain component manufacturers are on the market featuring intrinsically safe designs with high robustness against optical back reflections.

To a similar extent as laser diodes, superluminescent light emitting diodes are sensitive to electrostatic discharges and current spikes e.g. from ill-designed driver electronics. When selecting the current source to operate the SLED, special attention should be paid to low-noise specifications. Again certain suppliers are offering driver electronics especially designed to handle on the one hand the high-power, low-noise requirements and on the other hand protect the light sources against discharge and spikes. When treated carefully and operated well within the specifications, SLEDs can easily last for tens of thousands of hours of operation.

Availability of SLEDs

By means of the above mentioned optimized optical cavity design the SLEDs exhibit high output power, large bandwidth and low residual spectral ripple, making them an ideal light source for a number of applications. Based on the application's requirements and specifications, SLED devices are available in various packages or form factors covering a broad range of wavelengths and power levels. The packages include cooled 14-pin dual-in-line (DIL) and butterfly (BTF) modules or low-cost uncooled TOSA and TO-56 devices. The SLED modules includes indium phosphide (InP) based superluminescent light-emitting diodes operating in the high wavelength range (1100 nm to 1700 nm) as well as gallium arsenide (GaAs) based devices operating from 630 to 1100 nm. Usage of gallium nitride (GaN) based designs is breaking ground for SLEDs in the ultraviolet and blue spectral range.

SLEDs are commercially available from a number of suppliers, e.g. Denselight (Singapore), EXALOS (Switzerland), InPhenix (US), or Superlum (Russia). The product portfolio offered varies greatly from supplier to supplier by wavelength, power, and bandwidth.

Chapter 10

Single-Photon Avalanche Diode and Thermal Management of High-Power LEDs

Single-photon avalanche diode

In optoelectronics the term **Single-Photon Avalanche Diode** (**SPAD**) (also known as a **Geiger-mode APD** or **G-APD**) identifies a class of solid-state photodetectors based on a reverse biased p-n junction in which a photo-generated carrier can trigger an avalanche current due to the impact ionization mechanism. This device is able to detect low intensity signals (down to the single photon) and to signal the arrival times of the photons with a jitter of a few tens of picoseconds.

SPADs, like the Avalanche photodiode (APD), exploits the photon-triggered avalanche current of a reverse biased p-n junction to detect an incident radiation. The fundamental difference between SPAD and APD is that SPADs are specifically designed to operate with a reverse bias voltage well above the breakdown voltage (on the contrary APDs operate at a bias lesser than the breakdown voltage). This kind of operation is also called *Geiger mode* in literature, for the analogy with the Geiger counter.

SPAD operating principle

SPADs are semiconductor devices based on a p-n junction reverse-biased at a voltage V_a higher than V_B (Figure 1). "At this bias, the electric field is so high [higher than 3×10^5 V/cm] that a single charge carrier injected into the depletion layer can trigger a self-sustaining avalanche. The current rises swiftly [sub-nanosecond rise-time] to a macroscopic steady level in the milliampere range. If the primary carrier is photo-generated, the leading edge of the avalanche pulse marks [with picosecond time jitter] the arrival time of the detected photon." The current continues until the avalanche is quenched by lowering the bias voltage V_D down to or below V_B: the lower electric field is no longer able to accelerate carriers to impact-ionize with lattice atoms, therefore

current ceases. In order to be able to detect another photon, the bias voltage must be raised again above breakdown.
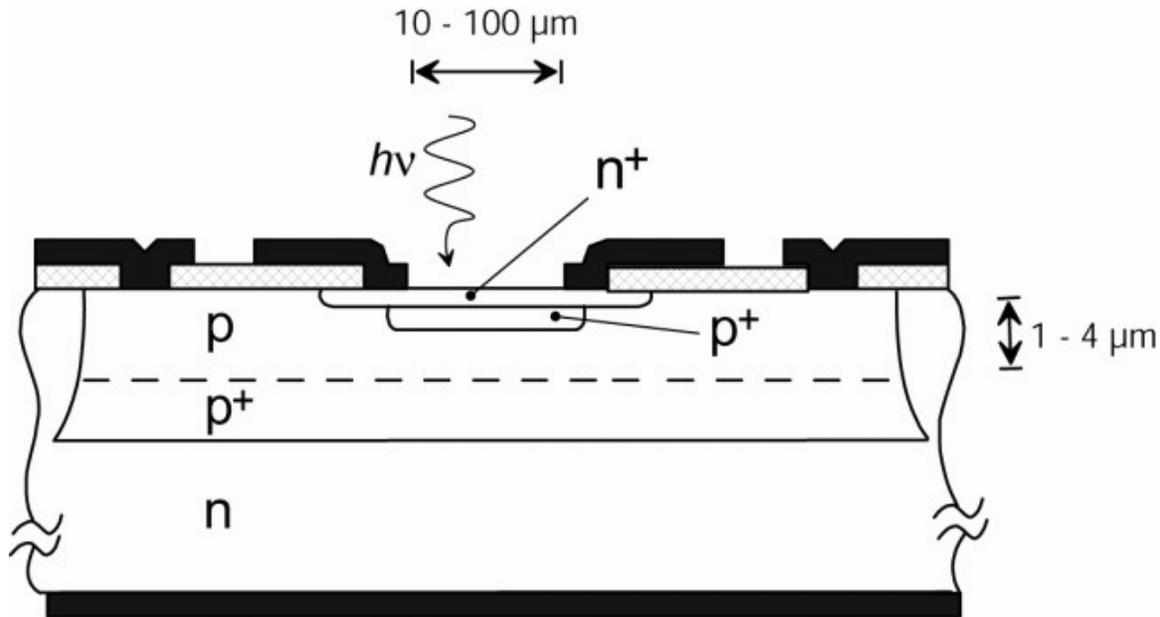
Figure 1 - Thin SPAD cross-section.

"This operation requires a suitable circuit, which has to:

1. sense the leading edge of the avalanche current;
2. generate a standard output pulse synchronous with the avalanche build-up;
3. quench the avalanche by lowering the bias down to the breakdown voltage;
4. restore the photodiode to the operative level. This circuit is usually referred to as a quenching circuit."

Passive quenching

The simplest quenching circuit is commonly called **Passive Quenching Circuit** and composed of a single resistor in series to the SPAD. This experimental setup has been employed since the early studies on the avalanche breakdown in junctions. The avalanche current self-quenches simply because it develops a voltage drop across a high-value ballast load R_L (about 100 kΩ or more). After the quenching of the avalanche current, the SPAD bias V_D slowly recovers to V_a, and therefore the detector is ready to be ignited again. A detailed description of the quenching process is reported in .

Active quenching

A more advanced quenching scheme is called **active quenching**. In this case a fast discriminator senses the steep onset of the avalanche current across a 50 Ω resistor and provides a digital (CMOS, TTL, ECL, NIM) output pulse, synchronous with the photon

arrival time. It then quickly reduces the bias voltage to below breakdown, then relatively quickly returns bias to above the breakdown voltage ready to sense the next photon.

Photon counting and timing

The intensity of the signal is obtained by counting (photon counting) the number of output pulses within a measurement time slot, while the time-dependent waveform of the signal is obtained by measuring the time distribution of the output pulses (photon timing). The latter is obtained by means of a Time Correlated Single Photon Counting (TCSPC) instrument.

Saturation

While the avalanche recovery circuit is quenching the avalanche and restoring bias, the SPAD cannot detect photons. Any photons that reach the detector during this brief period are not counted. As the number of photons increases such that the (statistical) time interval between photons gets within a factor of ten or so of the avalanche recovery time, missing counts become statistically significant and the count rate begins to depart from a linear relationship with detected light level. At this point the SPAD begins to saturate. If the light level were to increase further, ultimately to the point where the SPAD immediately avalanches the moment the avalanche recovery circuit restores bias, the count rate reaches a maximum defined purely by the avalanche recovery time (ten million counts per second or more). This can be harmful to the SPAD as it will be experiencing avalanche current nearly continuously.

Internal noise

Besides photon-generated carriers, thermally-generated carriers (through generation-recombination processes within the semiconductor) can also fire the avalanche process. Therefore, it is possible to observe output pulses when the SPAD is in complete darkness. The resulting average number of counts per second is called *dark count rate* and is the key parameter in defining the detector noise. It is worth noting that the reciprocal of the dark count rate defines the mean time that the SPAD remains biased above breakdown before being triggered by an undesired thermal generation. Therefore, in order to work as a single-photon detector, the SPAD must be able to remain biased above breakdown for a sufficiently long time (e.g., a few milliseconds, corresponding to a count rate well under a thousand counts per second, cps).

I-V characteristic
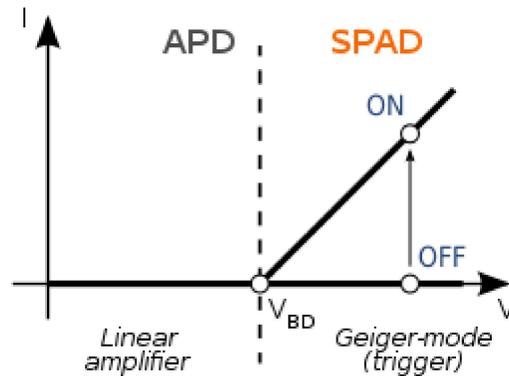
Figure 2: Current-voltage characteristic of a SPAD showing the off- and on-branch.

If a SPAD is observed by an analogue curve-tracer, it is possible to observe a bifurcation of the current-voltage characteristics beyond breakdown, during the voltage sweeps applied to the device. When the avalanche is triggered, the SPAD sustains the avalanche current (on-branch), instead when no carrier has been generated (by a photon or a thermal generation), no charge flows through the SPAD (off-branch). If the SPAD is triggered during a sweep above breakdown, a transition from the off-branch to the on-branch can be easily observed (like a "flickering").

APDs versus *SPADs*

Both APDs and SPADs are reverse biased semiconductor p-n junctions. However, APDs are biased close to, but below the breakdown voltage of the semiconductor. This high electric field provides an internal multiplication gain only on the order of few hundreds, since the avalanche process is not diverging as in SPADs. The resulting avalanche current intensity is linearly related to the optical signal intensity. A SPAD however operates with a bias voltage above the breakdown voltage. Because the device is operating in this unstable above-breakdown regime, a single photon (or a single dark current electron) can set off a significant avalanche of electrons. Practically, this means that in an APD a single photon produces only tens or few hundreds of electrons, but in a SPAD a single photon triggers a current in the mA region (billions of billions of electrons per second) that can be easily "counted".

Therefore, while the APD is a linear amplifier for the input optical signal with limited gain, the SPAD is a trigger device so the gain concept is meaningless.

Thermal management of high-power LEDs

High power light-emitting diodes (LEDs) are likely to replace other technologies such as incandescent and fluorescent bulbs in signaling, solid state lighting, and vehicle headlights because they save energy and extend the light's lifetime. LEDs that use from 500 milliwatts to as much as 10 watts in a single package have become standard, and researchers expect to use even more power in the future. Some of the electricity in an LED becomes heat rather than light. If that heat is not removed, the LEDs run at high temperatures, which not only lowers their efficiency, but also makes the LED more dangerous and less reliable. Thus, thermal management of high power LEDs is a crucial area of research and development.

Heat transfer procedure
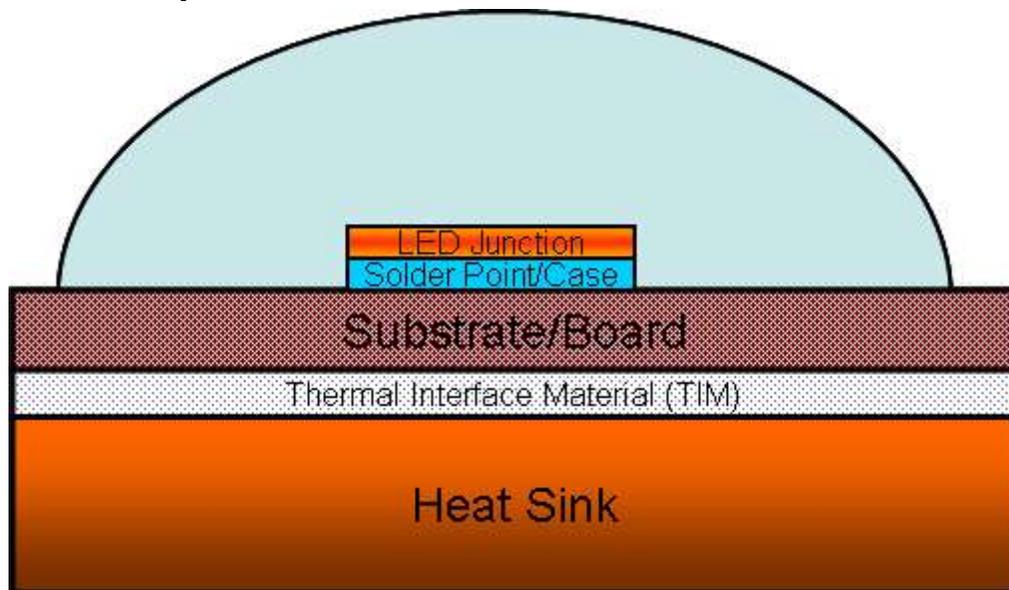
Typical LED package including thermal management design
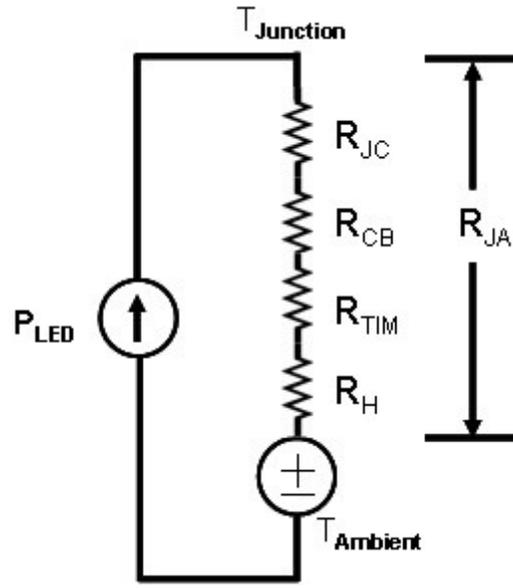
Typical thermal model of LED package. LED power dissipation is modeled as a current source; thermal resistance is modeled as a resistor; and the ambient temperature is modeled as a voltage source.

In order to maintain a low junction temperature to keep good performance of an LED, every method of releasing heat from LEDs should be considered. Conduction, convection, and radiation are the three means of heat transfer. Typically, LEDs are encapsulated in a transparent resin, which is a poor thermal conductor. Nearly all heat produced is conducted through the back side of the chip. Heat is generated from the PN junction by electrical energy that was not converted to useful light, and conducted to outside ambience through a long and extensive path, from junction to solder point, solder point to board, and board to the heat sink and then to the atmosphere. The heat path of tungsten light bulbs is almost all straight into the atmosphere, starting from filament to the glass and ending with the thermal resistance from glass to the atmosphere. A typical LED side view and its thermal model are shown in the figures.

The thermal resistance between two points is defined as the ratio of the difference in temperature to the power dissipated; the unit is $^\circ$C/W. From the LED junction to the thermal contact at the bottom of package, the thermal resistance is governed by the package design. It is referred to as the thermal resistance between junction and ambient (R_{JA}). Different components in the heat conduction path can be modeled as different thermal resistances. The total power dissipated by the LED (P_{LED}) is the product of the forward voltage and the forward current of the LED, which can be modeled as a current source. The ambient temperature is modeled as a voltage source. Therefore, the junction temperature (T_J) is the sum of the ambient temperature (T_A) and the product of the thermal resistance from junction to ambient and the power dissipated. By "thermic Ohm's Law", we have the equation as follows: $T_J = T_A + (R_{JA} \times P_{LED})$, and $R_{JA} = R_{JC} + R_{CB} + R_{TIM} + R_H$

Intuitively, you can see that the junction temperature will be lower if the thermal impedance is smaller and likewise, with a lower ambient temperature. To maximize the useful ambient temperature range for a given power dissipation, the total thermal resistance from junction to ambient must be minimized. The values for the thermal resistance vary widely depending on the material or component supplier. For example, R_{JC} will range from 2.6°C/W to 18°C/W, depending on the LED manufacturer. The thermal interface material's (TIM) thermal resistance will also vary depending on the type of material selected. Common TIMs are epoxy, thermal grease, pressure sensitive adhesive and solder. In the most cases, power LEDs will be mounted on metal-core printed circuit boards (MCPCB), which will be attached to a heat sink. Heat flows from the LED junction through the MCPCB to the heat sink by way of conduction, and the heat sink diffuses heat to the ambient surroundings by convection. So, we can also add convection to the thermal model at the end of the heat transmission path. In the package design, the surface flatness and quality of each component, applied mounting pressure, contact area, the type of interface material and its thickness are all important parameters to thermal resistance design.

Passive thermal designs

Some considerations for passive thermal designs to ensure good thermal management for high power LED operation include:

Adhesive

Adhesive is commonly used to bond LED and board, and board and heat sinks. Using a thermal conductive adhesive can further optimize the thermal performance.

Heat sink

Heat sinks provide a path for heat from the LED source to outside medium. Heat sinks can dissipate power in three ways: conduction (heat transfer from one solid to another), convection (heat transfer from a solid to a moving fluid, for most LED applications the fluid will be air), or radiation (heat transfer from two bodies of different surface temperatures through electromagnetic waves).

- **Material** – The thermal conductivity of the material that the heat sink is made from directly affects the dissipation efficiency through conduction. Normally this is aluminum, although copper may be used with an advantage for flat-sheet heat sinks.
- **Shape** - Thermal transfer takes place at the surface of the heat sink. Therefore, heat sinks should be designed to have a large surface area. This goal can be reached by using a large number of fine fins or by increasing the size of the heat sink itself.
- **Surface Finish** - Thermal radiation of heat sinks is a function of surface finish, especially at higher temperatures. A painted surface will have a greater emissivity than a bright, unpainted one. The effect is most remarkable with flat-plate heat

sinks, where about one-third of the heat is dissipated by radiation. Moreover, a perfectly flat contact area allows the use of a thinner layer of thermal compound, which will reduce the thermal resistance between the heat sink and LED source. On the other hand, anodizing or etching will also decrease the thermal resistance.

- **Mounting method**- Heat-sink mountings with screws or springs are often better than regular clips, thermal conductive glue or sticky tape.

## PCB (Printed Circuit Board)

- **MCPCB** - MCPCB (Metal Core PCB) are those boards which incorporate a base metal material as heat spreader as an integral part of the circuit board. The metal core usually consists of aluminum alloy. Furthermore MCPCB can take advantage of incorporating a dielectric polymer layer with high thermal conductivity for lower thermal resistance.
- **Separation** - Separating the LED drive circuitry from the LED board prevents the heat generated by the driver from raising the LED junction temperature.

Package type

- **Flip chip** - The concept is similar to flip-chip in package configuration widely used in the silicon integrated circuit industry. Briefly speaking, the LED die is assembled face down on the sub-mount, which is usually silicon or ceramic, acting as the heat spreader and supporting substrate. The flip-chip joint can be eutectic, high-lead, lead-free solder or gold stub. The primary source of light comes from the back side of the LED chip, and there is usually a built-in reflective layer between the light emitter and the solder joints to reflect the light emitted downwards up. Several companies have adopted flip-chip packages for their high-power LED, achieving bout 60% reduction in the thermal resistance of the LED while keeping its thermal reliability.

Chapter 11

PIN Diode
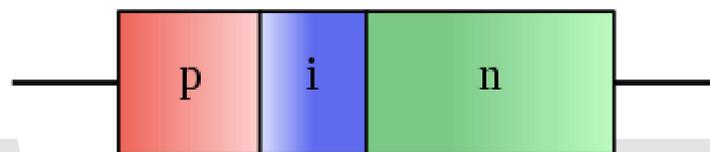
Layers of a PIN diode

A **PIN diode** is a diode with a wide, lightly doped 'near' intrinsic semiconductor region between a p-type semiconductor and an n-type semiconductor region. The p-type and n-type regions are typically heavily doped because they are used for ohmic contacts.

The wide intrinsic region is in contrast to an ordinary PN diode. The wide intrinsic region makes the PIN diode an inferior rectifier (one typical function of a diode), but it makes the PIN diode suitable for attenuators, fast switches, photodetectors, and high voltage power electronics applications.

Operation

A PIN diode operates under what is known as high-level injection. In other words, the intrinsic "i" region is flooded with charge carriers from the "p" and "n" regions. Its function can be likened to filling up a water bucket with a hole on the side. Once the water reaches the hole's level it will begin to pour out. Similarly, the diode will conduct current once the flooded electrons and holes reach an equilibrium point, where the number of electrons is equal to the number of holes in the intrinsic region. When the diode is forward biased, the injected carrier concentration is typically several orders of magnitude higher than the intrinsic level carrier concentration. Due to this high level injection, which in turn is due to the depletion process, the electric field extends deeply (almost the entire length) into the region. This electric field helps in speeding up of the transport of charge carriers from P to N region, which results in faster operation of the diode, making it a suitable device for high frequency operations.

Characteristics

A PIN diode obeys the standard diode equation for low frequency signals. At higher frequencies, the diode looks like an almost perfect (very linear, even for large signals) resistor. There is a lot of stored charge in the intrinsic region. At low frequencies, the charge can be removed and the diode turns off. At higher frequencies, there is not enough time to remove the charge, so the diode never turns off. The PIN diode has a poor reverse recovery time.

The high-frequency resistance is inversely proportional to the DC bias current through the diode. A PIN diode, suitably biased, therefore acts as a variable resistor. This high-frequency resistance may vary over a wide range (from 0.1 ohm to 10 kΩ in some cases; the useful range is smaller, though).

The wide intrinsic region also means the diode will have a low capacitance when reverse biased.

In a PIN diode, the depletion region exists almost completely within the intrinsic region. This depletion region is much larger than in a PN diode, and almost constant-size, independent of the reverse bias applied to the diode. This increases the volume where electron-hole pairs can be generated by an incident photon. Some photodetector devices, such as PIN photodiodes and phototransistors (in which the base-collector junction is a PIN diode), use a PIN junction in their construction.

The diode design has some design tradeoffs. Increasing the dimensions of the intrinsic region (and its stored charge) allows the diode to look like a resistor at lower frequencies. It adversely affects the time needed to turn off the diode and its shunt capacitance. PIN diodes will be tailored for a particular use.

Applications

PIN diodes are useful as RF switches, attenuators, and photodetectors.

RF and Microwave Switches

A PIN Diode RF Microwave Switch. Picture courtesy of Herley

Under zero or reverse bias, a PIN diode has a low capacitance. The low capacitance will not pass much of an RF signal. Under a forward bias of 1 mA, a typical PIN diode will have an RF resistance of about 1 ohm, making it a good RF conductor. Consequently, the PIN diode makes a good RF switch.

Although RF relays can be used as switches, they switch very slowly (on the order of 10 milliseconds). A PIN diode switch can switch much more quickly (e.g., 1 microsecond).

The capacitance of an off discrete PIN diode might be 1 pF. At 320 MHz, the reactance of 1 pF is about 500 ohms. In a 50 ohm system, the off state attenuation would be about 20 dB -- which may not be enough attenuation. In applications that need higher isolation, switches are cascaded to improve the isolation. Cascading three of the above switches would give 60 dB of attenuation.

PIN diode switches are used not only for signal selection, but they are also used for component selection. For example, some low phase noise oscillators use PIN diodes to range switch inductors.

RF and Microwave Variable Attenuators

A RF Microwave PIN diode Attenuator. Picture courtesy of Herley

By changing the bias current through a PIN diode, it's possible to quickly change the RF resistance.

At high frequencies, the PIN diode appears as a resistor whose resistance is an inverse function of its forward current. Consequently, PIN diode can be used in some variable attenuator designs as amplitude modulators or output leveling circuits.

PIN diodes might be used, for example, as the bridge and shunt resistors in a bridged-T attenuator.

Limiters

PIN diodes are sometimes used as input protection devices for high frequency test probes. If the input signal is within range, the PIN diode has little impact as a small capacitance. If the signal is large, then the PIN diode starts to conduct and becomes a resistor that shunts most of the signal to ground.

Photodetector and photovoltaic cell

The PIN photodiode was invented by Jun-ichi Nishizawa and his colleagues in 1950.

PIN photodiodes are used in fibre optic network cards and switches. As a photodetector, the PIN diode is reverse biased. Under reverse bias, the diode ordinarily does not conduct (save a small dark current or I_s leakage). A photon entering the intrinsic region frees a carrier. The reverse bias field sweeps the carrier out of the region and creates a current. Some detectors can use avalanche multiplication.

The PIN photovoltaic cell works in the same mechanism. In this case, the advantage of using a PIN structure over conventional semiconductor junction is the better long wavelength response of the former. In case of long wavelength irradiation, photons penetrate deep into the cell. But only those electron-hole pairs generated in and near the depletion region contribute to current generation. The depletion region of a PIN structure extends across the intrinsic region, deep into the device. This wider depletion width enables electron-hole pair generation deep within the device. This increases the quantum efficiency of the cell.

Typically, amorphous silicon thin-film cells use PIN structures. On the other hand, CdTe cells use NIP structure, a variation of the PIN structure. In a NIP structure, an intrinsic CdTe layer is sandwiched by n-doped CdS and p-doped ZnTe. The photons are incident on the n-doped layer unlike a PIN diode.

Example Diodes

SFH203 or BPW43 are cheap general purpose PIN diodes in 5 mm clear plastic case with bandwidth over 100 MHz. They are used in RONJA telecommunication systems and other circuitry applications.
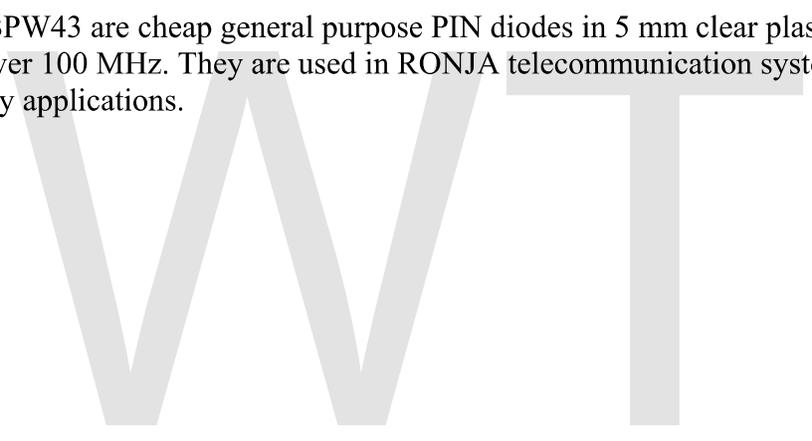
Chapter 12

Phosphorescent Organic Light-Emitting Diode, Avalanche Photodiode and Laser Diode Rate Equations

Phosphorescent organic light-emitting diode

Phosphorescent organic light-emitting diodes (PHOLED) are a type of organic light-emitting diode (OLED) that use the principle of phosphorescence to obtain higher internal efficiencies than fluorescent OLEDs. This technology is currently under development by many industrial and academic research groups.

Method of operation
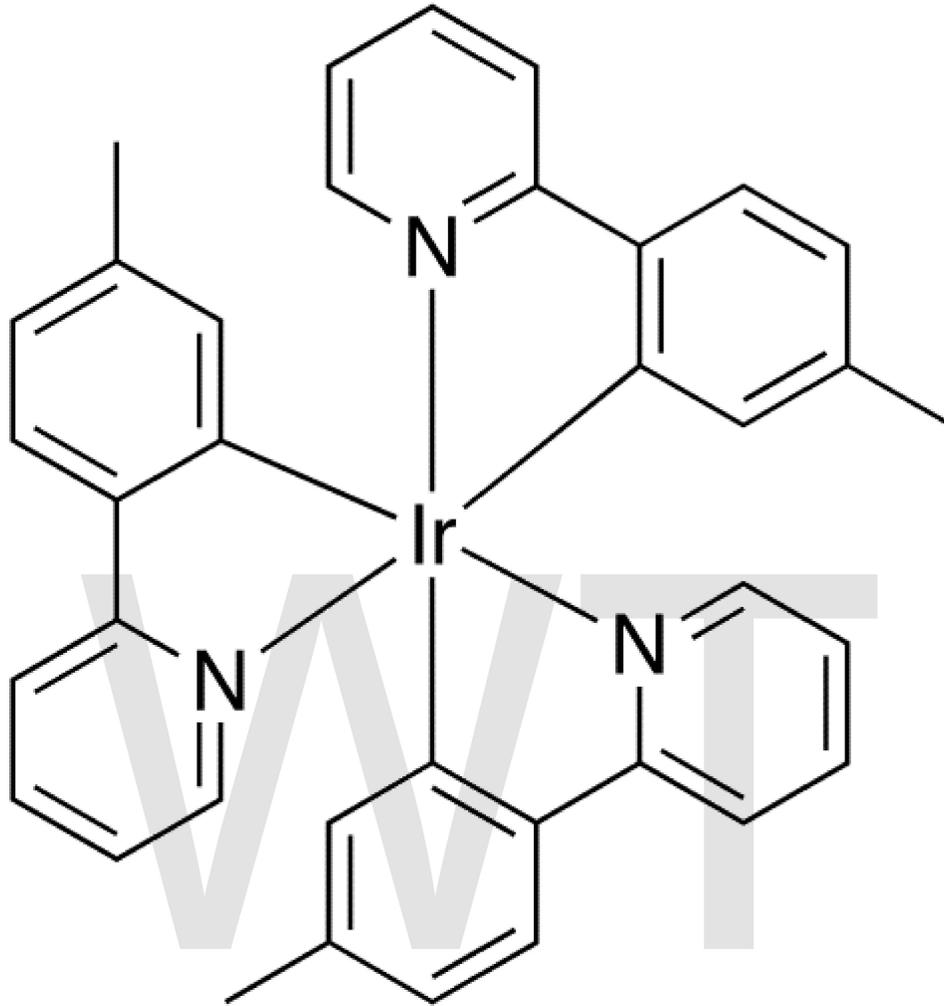
Ir(mppy)₃, an example of a phosphorescent dopant which emits green light.

Like all types of OLED, phosphorescent OLEDs emit light due to the electroluminescence of an organic semiconductor layer in an electric current. Electrons and holes are injected into the organic layer at the electrodes and form excitons, a bound state of the electron and hole.

Electrons and holes are both fermions with half integer spin. An exciton formed by the recombination of two such particles may either be in a singlet state or a triplet state, depending on how the spins have been combined. Statistically, there is a 25% probability of forming a singlet state and 75% probability of forming a triplet state. Decay of the excitons results in the production of light through spontaneous emission.

In OLEDs using fluorescent organic molecules only, the decay of triplet excitons is quantum mechanically forbidden by selection rules, meaning that the lifetime of triplet excitons is long and phosphorescence is not readily observed. Hence it would be expected

that in fluorescent OLEDs only the formation of singlet excitons results in the emission of useful radiation, placing a theoretical limit on the internal quantum efficiency (the percentage of excitons formed that result in emission of a photon) of 25%.

However, phosphorescent OLEDs generate light from both triplet and singlet excitons, allowing the internal quantum efficiency of such devices to reach nearly 100%.

This is commonly achieved by doping a host molecule with an organometallic complex. These contain a heavy metal atom at the centre of the molecule, for example platinum or iridium, of which the green emitting complex Ir(mppy)$_3$ is just one of many examples. The large spin-orbit interaction experienced by the molecule due to this heavy metal atom facilitates intersystem crossing, a process which mixes the singlet and triplet character of excited states. This reduces the lifetime of the triplet state, therefore phosphorescence is readily observed.

Applications

Due to their potentially high level of energy efficiency, even when compared to other OLEDs, PHOLEDs are being studied for potential use in large-screen displays such as computer monitors or television screens, as well as general lighting needs. One potential use of PHOLEDs as lighting devices is to cover walls with large area PHOLED light panels. This would allow entire rooms to glow uniformly, rather than require the use of light bulbs which distribute light unequally throughout a room. The United States Department of Energy has recognized the potential for massive energy savings via the use of this technology and therefore has awarded 200 000 US$ in contracts to develop PHOLED products for general lighting applications.

Challenges

One problem that currently hampers the widespread adoption of this highly energy efficient technology is that the average lifetimes of red and green PHOLEDs are often tens of thousands of hours longer than those of blue PHOLEDs. This may cause displays to become visually distorted much sooner than would be acceptable for a commercially viable device.

Avalanche photodiode

An **avalanche photodiode (APD)** is a highly sensitive semiconductor electronic device that exploits the photoelectric effect to convert light to electricity. APDs can be thought of as photodetectors that provide a built-in first stage of gain through avalanche multiplication. From a functional standpoint, they can be regarded as the semiconductor analog to photomultipliers. By applying a high reverse bias voltage (typically 100-200 V

in silicon), APDs show an internal current gain effect (around 100) due to impact ionization (avalanche effect). However, some silicon APDs employ alternative doping and beveling techniques compared to traditional APDs that allow greater voltage to be applied (> 1500 V) before breakdown is reached and hence a greater operating gain (> 1000). In general, the higher the reverse voltage the higher the gain. Among the various expressions for the APD multiplication factor *(M)*, an instructive expression is given by the formula

$$M = \frac{1}{1 - \int_0^L \alpha(x)\, dx}$$

where *L* is the space charge boundary for electrons and α is the multiplication coefficient for electrons (and holes). This coefficient has a strong dependence on the applied electric field strength, temperature, and doping profile. Since APD gain varies strongly with the applied reverse bias and temperature, it is necessary to control the reverse voltage to keep a stable gain. Avalanche photodiodes therefore are more sensitive compared to other semiconductor photodiodes.

If very high gain is needed (10^5 to 10^6), certain APDs (single-photon avalanche diodes) can be operated with a reverse voltage above the APD's breakdown voltage. In this case, the APD needs to have its signal current limited and quickly diminished. Active and passive current quenching techniques have been used for this purpose. APDs that operate in this high-gain regime are in Geiger mode. This mode is particularly useful for single photon detection provided that the dark count event rate is sufficiently low.

A typical application for APDs is laser rangefinders and long range fiber optic telecommunication. New applications include positron emission tomography and particle physics. APD arrays are becoming commercially available.

APD applicability and usefulness depends on many parameters. Two of the larger factors are: quantum efficiency, which indicates how well incident optical photons are absorbed and then used to generate primary charge carriers; and total leakage current, which is the sum of the dark current and photocurrent and noise. Electronic dark noise components are series and parallel noise. Series noise, which is the effect of shot noise, is basically proportional to the APD capacitance while the parallel noise is associated with the fluctuations of the APD bulk and surface dark currents. Another noise source is the excess noise factor, *F*. It describes the statistical noise that is inherent with the stochastic APD multiplication process.

Materials

In principle any semiconductor material can be used as a multiplication region:

- Silicon will detect in the visible and near infrared, with low multiplication noise (excess noise).

- Germanium (Ge) will detect infrared out to a wavelength of 1.7 μm, but has high multiplication noise.
- InGaAs will detect out to longer than 1.6 μm, and has less multiplication noise than Ge. It is normally used as the absorption region of a heterostructure diode, most typically involving InP as a substrate and as a multiplication layer. This materials system is compatible with an absorption window of roughly 0.9-1.7 μm. InGaAs exhibits a high absorption coefficient at the wavelengths appropriate to high-speed telecommunications using optical fibers, so only a few micrometres of InGaAs are required for nearly 100% light absorption. The excess noise factor is low enough to permit a gain-bandwidth product in excess of 100 GHz for a simple InP/InGaAs system, and up to 400 GHz for InGaAs on silicon. Therefore high speed operation is possible: commercial devices are available to speeds of at least 10 Gbit/s.
- Gallium nitride based diodes have been used for operation with ultraviolet light.
- HgCdTe based diodes operate in the infrared, typically out to a maximum wavelength of about 14 μm, but require cooling to reduce dark currents. Very low excess noise can be achieved in this material system.

Excess noise

As mentioned above, this is the noise due to the multiplication process at a gain, M and is denoted by $F(M)$ and can often be expressed as:

$$F = \kappa M + \left(2 - \frac{1}{M}\right)(1 - \kappa)$$

where κ is the ratio of the hole impact ionization rate to that of electrons. For an electron multiplication device it is given by the hole impact ionization rate divided by the electron impact ionization rate. It is desirable to have a large asymmetry between these rates to minimize $F(M)$, since $F(M)$ is one of the main factors that limit, among other things, the best possible energy resolution obtainable.

Performance limits

In addition to excess noise, there are limits to device performance associated with the capacitance, transit times and avalanche multiplication time. The capacitance increases with increasing device area and decreasing thickness. The transit times (both electrons and holes) increase with increasing thickness, implying a tradeoff between capacitance and transit time for performance. The avalanche multiplication time times the gain is given to first order by the gain-bandwidth product, which is a function of the device structure and most especially κ.

Laser diode rate equations

The laser diode rate equations model the electrical and optical performance of a laser diode. This system of ordinary differential equations relates the number or density of photons and charge carriers (electrons) in the device to the injection current and to device and material parameters such as carrier lifetime, photon lifetime, and the optical gain.

The rate equations may be solved by numerical integration to obtain a time-domain solution, or used to derive a set of steady state or small signal equations to help in further understanding the static and dynamic characteristics of semiconductor lasers.

The laser diode rate equations can be formulated with more or less complexity to model different aspects of laser diode behavior with varying accuracy.

Multimode rate equations

In the multimode formulation, the rate equations model a laser with multiple optical modes. This formulation requires one equation for the carrier density, and one equation for the photon density in each of the optical cavity modes:

$$\frac{dN}{dt} = \frac{I}{eV} - \frac{N}{\tau_n} - \sum_{\mu=1}^{\mu=M} G_\mu P_\mu$$
$$\frac{dP_\mu}{dt} = \Gamma_\mu (G_\mu - \frac{1}{\tau_p}) P_\mu + \beta_\mu \frac{N}{\tau_n}$$

where: N is the carrier density, P is the photon density, I is the applied current, e is the elementary charge, V is the volume of the active region, τ_n is the carrier lifetime, G is the gain coefficient (s^{-1}), Γ is the confinement factor, τ_p is the photon lifetime, β is the spontaneous emission factor, M is the number of modes modelled, μ is the mode number, and subscript μ has been added to G, Γ, and β to indicate these properties may vary for the different modes.

The first term on the right side of the carrier rate equation is the injected electrons rate (I/eV), the second term is the carrier depletion rate due to non-radiative recombination processes (described by the decay time τ_n) and the third term is the carrier depletion due to stimulated recombination, which is proportional to the photon density and medium gain.

In the photon density rate equation, the first term ΓGP is the rate at which photon density increase due to stimulated emission (the same term in carrier rate equation, with positive sign and multiplied for the confinement factor Γ), the second term is the rate at which photons leave the cavity, for internal absorption or exiting the mirrors, expressed via the decay time constant τ_p and the third term is the contribution of spontaneous emission from carrier non-radiative recombination.

The modal gain

G_μ, the gain of the μth mode, can be modelled by a parabolic dependence of gain on wavelength as follows:

$$G_\mu = \frac{\alpha N\left[1 - \left(2\frac{\lambda(t)-\lambda_\mu}{\delta\lambda_g}\right)^2\right] - \alpha N_0}{1 + \epsilon \sum_{\mu=1}^{\mu=M} P_\mu}$$

where: α is the gain coefficient and ϵ is the gain compression factor (see below). λ_μ is the wavelength of the μth mode, $\delta\lambda_g$ is the full width at half maximum (FWHM) of the gain curve, the centre of which is given by

$$\lambda(t) = \lambda_0 + \frac{k(N_{th} - N(t))}{N_{th}}$$

where λ_0 is the centre wavelength for $N = N_{th}$ and k is the spectral shift constant (see below). N_{th} is the carrier density at threshold and is given by

$$N_{th} = N_{tr} + \frac{1}{\alpha \tau_p \Gamma}$$

where N_{tr} is the carrier density at transparency.

β_μ is given by

$$\beta_\mu = \frac{\beta_0}{1 + (2(\lambda_s - \lambda_\mu)/\delta\lambda_s)^2}$$

where

β_0 is the spontaneous emission factor, λ_s is the centre wavelength for spontaneous emission and $\delta\lambda_s$ is the spontaneous emission FWHM. Finally, λ_μ is the wavelength of the μth mode and is given by

$$\lambda_\mu = \lambda_0 - \mu\delta\lambda + \frac{(n-1)\delta\lambda}{2}$$

where $\delta\lambda$ is the mode spacing.

Gain Compression

The gain term, G, cannot be independent of the high power densities found in semiconductor laser diodes. There are several phenomena which cause the gain to 'compress' which are dependent upon optical power. The two main phenomena are spatial hole burning and spectral hole burning.

Spatial hole burning occurs as a result of the standing wave nature of the optical modes. Increased lasing power results in decreased carrier diffusion efficiency which means that the stimulated recombination time becomes shorter relative to the carrier diffusion time. Carriers are therefore depleted faster at the crest of the wave causing a decrease in the modal gain.

Spectral hole burning is related to the gain profile broadening mechanisms such as short intraband scattering which is related to power density.

To account for gain compression due to the high power densities in semiconductor lasers, the gain equation is modified such that it becomes related to the inverse of the optical power. Hence, the following term in the denominator of the gain equation :

$$1 + \epsilon \sum_{\mu=1}^{\mu=M} P_\mu$$

Spectral Shift

Dynamic wavelength shift in semiconductor lasers occurs as a result of the change in refractive index in the active region during intensity modulation. It is possible to evaluate the shift in wavelength by determining the refractive index change of the active region as a result of carrier injection. A complete analysis of spectral shift during direct modulation found that the refractive index of the active region varies proportionally to carrier density and hence the wavelength varies proportionally to injected current.

Experimentally, a good fit for the shift in wavelength is given by:

$$\delta\lambda = k \left(\sqrt{\frac{I_0}{I_{th}}} - 1 \right)$$

where I_0 is the injected current and I_{th} is the lasing threshold current.