mks

MKS INSTRUMENTS HANDBOOK

Principles and Applications in Photonics Technologies

by the Office of the CTO

1/2019

MKS Instruments Handbook

Principles and Applications in Photonics Technologies

by the Office of the CTO

This page intentionally left blank.

MKS Instruments Handbook

Principles and Applications in Photonics Technologies

by the Office of the CTO

Office of the CTO (OCTO)

Josh Ding, Ph.D., Chief Technology Officer, Flow Solutions

Mark Gitin, Ph.D., Senior Vice President & General Manager, Light & Motion Division

Jim Kafka, Ph.D., Chief Technology Officer, Lasers

Chris Le Tiec, Ph.D., Chief Technology Officer, Ozone Products,
Plasma & Reactive Gas Solutions

John TC Lee, Ph.D., President & Chief Operating Officer

Aaron Radomski, Ph.D., Fellow, Chief Technology Officer, Power Solutions

Erik Saarmaa, J.D., Chief IP Counsel, Intellectual Property

Eric Snyder, Vice President, Quality & Reliability

Philip Sullivan, Ph.D., Fellow, Pressure & Vacuum Measurement and
Integrated Process Solutions

Eric Taranto, Vice President & General Manager of Business Units,
Analytical & Control Solutions, Power & Reactive Gas, and Power Solutions

Ruben Zadoyan, Ph.D., Chief Technology Officer, Newport® Products

Former OCTO Members

Robin Swain, Ph.D., General Manager, Integrated Solutions Business

Contributing Editors

Joel Hales, Ph.D., Consultant

Harald Weiss, Technical Marketing Group, Inc.

Tom Whidden, Ph.D., TechLink Writers Group

Contributing Authors

Amit Arora, Ed Arriola, Nissim Asida, Ph.D., Juerg Aus der Au, Ph.D., Herman Chui, Ph.D.,
Christian Dini, Peter Dwyer, Ph.D., Beda Espinoza, Jarid Feldner, Andrei Florean Ph.D.,
Joel Hales, Ph.D., Alain Jeanne, Ph.D., Arnd Krueger, Ph.D., Kobi Lasri, Ph.D.,
Sean LeCave, John Park, Ph.D., Tyler Parker, Sebastian Peck, Vidi Saptari, Ph.D.,
Marijke VanSpyk, Ph.D., Alistair Wallace, Ph.D.

Special Thanks

Bill Casey, Anne Devine, Jane Lawlor

Foreword

Photonics is the science of generating, detecting and manipulating light. With the invention of the Laser in 1960, photonics became an enabling technology for applications such as fiber optic communications, bioimaging, laser surgery, LED lighting, security, flexible electronics, wafer inspection and metrology, lithography for semiconductor device manufacturing, and many other uses that enable the lives we live today. For many decades, MKS has been at the forefront of the photonics revolution with innovations in lasers, opto-mechanical components, vibration and motion control, and laser characterization.

Surround the Workpiece℠ is the MKS strategy that serves the needs of Advanced Markets that require laser-based solutions. Our goal is to provide customers with the key components, systems, and services to enable the successful implementation of these solutions. For our Semiconductor market customers, the workpiece is a semiconductor wafer. For customers in Research and Development or Life and Health Sciences, the workpiece can be a biological sample. The workpiece may represent a subsystem such as an optical network for the communications market or an imaging system for our remote sensing customers. And for Industrial Manufacturing customers, it can be a printed circuit board, a precision glass or metal sheet, or a lithium ion battery.

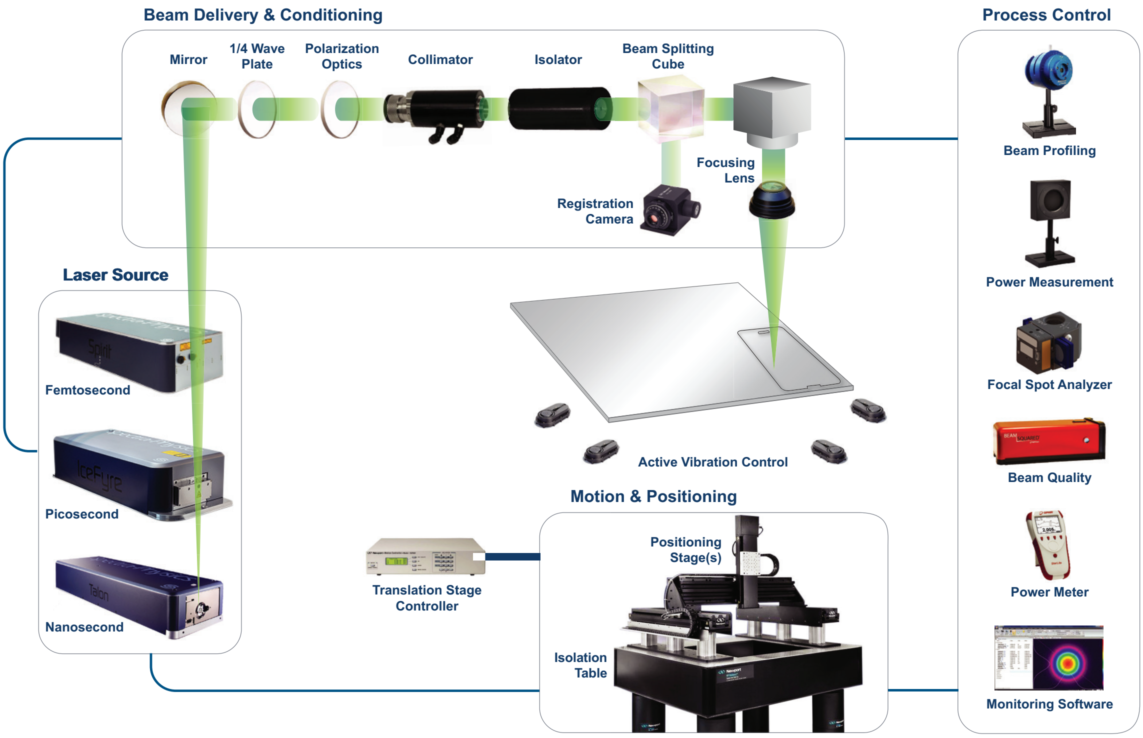


Figure 1. Surround the Workpiece solution strategy. An example of a laser micromachining system including beam delivery, motion and positioning, and monitoring of the laser beam delivery.

An example of this strategy is our unique offering for the laser machining market. In a typical laser machining setup (as shown in Figure 1), laser light with a predefined wavelength, pulse width and energy is generated, then carefully managed through a set of advanced optical and opto-mechanical components to precisely direct the light to the workpiece. The workpiece itself needs to be moved and positioned with high accuracy so that the light is delivered to the proper location. These motion stages are, in turn, placed on active vibration control tables that compensate in real-time for any external vibrations. Since the laser light is arguably the most important element of the system, stringent process control is required to ensure its integrity. For process control, MKS provides a complete suite of products ranging from laser beam profiling and power measurement to beam quality assessment.

The purpose of this book is to introduce the reader to the basics of photonics and to describe the key products and applications that Surround the Workpiece. As with our first book, MKS Instruments Handbook of Semiconductor Devices and Process Technology, we hope you find this Handbook informative and useful.

Gerald G. Colella
Chief Executive Officer

About MKS Instruments

MKS Instruments, Inc. is a global provider of instruments, subsystems and process control solutions that measure, monitor, deliver, analyze, power and control critical parameters of advanced manufacturing processes to improve process performance and productivity for our customers. Our products are derived from our core competencies in pressure measurement and control, flow measurement and control, gas and vapor delivery, gas composition analysis, residual gas analysis, leak detection, control technology, ozone generation and delivery, power, reactive gas generation, vacuum technology, lasers, photonics, sub-micron positioning, vibration control and optics. We also provide services relating to the maintenance and repair of our products, installation services and training. Our primary served markets include the semiconductor, industrial technologies, life and health sciences, research and defense. Additional information can be found at www.mksinst.com.

This page intentionally left blank.

Table of Contents

This page intentionally left blank.

Chapter 1

Introduction to Photonics

This page intentionally left blank.

I. Light Sources

The choice of a light source for an application depends on the desired characteristics of the radiation. These attributes can include the power or intensity of the light, its spatial distribution and how it will propagate, whether the light is constant as a function of time or is emitted as a burst, and the spectrum of the light which determines its specific color or range of colors. In this section, laser light sources are discussed initially, followed by light-emitting diodes (LEDs), and lamp sources. In each category, details are given regarding the mechanisms of light generation, the unique characteristics of the output radiation, the wide variety of light sources available, and the particular applications that are targeted.

A. Lasers

Lasers are devices that produce light with properties very different than those of other light sources, e.g., incandescent bulbs or LEDs. These unique characteristics enable a remarkably wide range of applications. Laser light can travel large distances as a narrow beam without diverging, allowing it to be used in laser pointers, laser light shows, and even for communication between satellites. The light can also be focused to a very tight spot, enabling sub-cellular microscopic imaging, reading/writing large amounts of data to/from DVDs and Blu-ray discs, and photolithography, which is critical in the production of modern microelectronics. Furthermore, if this tightly-focused light is confined to very short bursts or pulses, high-intensity lasers can be used for a variety of micromachining applications, including cutting/marking materials such a ceramics, glass, and metals as well as safe ablation of human tissue. Finally, laser light can have a very narrow spectrum or singular color that enables high-resolution spectroscopy and optical fiber communication.

A laser is a source of coherent light. It contains an optical oscillator that increases the amplitude of an optical field while maintaining its phase. This coherent amplification is achieved through Light Amplification by Stimulated Emission of Radiation (LASER). The process of amplification is the result of stimulating an atom such that it emits an identical photon to one already present. This “clone” photon has the same phase, frequency, direction, and polarization as the original photon, meaning they are coherent. It is this coherent amplification of light that gives lasers their unique output characteristics.

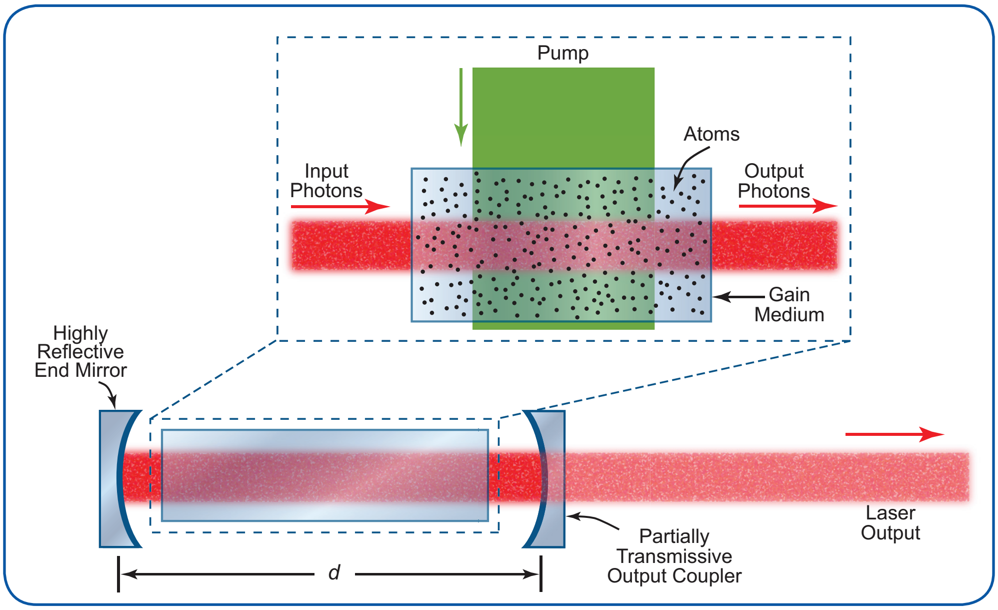


Figure 2. Simplified diagram of a laser, including the gain medium, pump source, and optical resonator mirrors with cavity length (d) [1].

The three key components of a laser are a gain medium, a pump source, and a resonator. Figure 2 shows a pictorial representation of an operational laser. The pump source is the mechanism by which a population inversion (necessary for generating stimulated emission) is produced in the gain medium. The gain medium is chosen to achieve efficient lasing operation as well the desired characteristics for the laser output. The resonator mirrors allow for selective reflection of incoherent photons produced by spontaneous emission back into the gain medium. Stimulated emission replicates these reflected photons and this sequential reflection/amplification process gives rise to a large number of coherent photons. Finally, a portion of these photons are transmitted through a partially reflective mirror, delivering the output laser beam. How laser light is generated, its unique characteristics, as well as the various types of lasers, are the focus of this section.

1. Basic Light-Matter Interactions in Lasers

Light is generated from a laser as a result of the following process: electrons in a material move from an excited energy level to a lower-lying energy level and produce photons that contribute to the laser beam. Therefore, fundamental interactions between light and matter are the basis for the analysis of laser operation and the properties of laser light [2, 3]. This section provides an abridged description of the interactions between atoms/molecules in a laser material and the photons that make up the resulting laser light.

Energy Levels

Atomic energy levels are determined by the interactions of the electrons with the atomic nucleus and other electrons. As electrons move or transition from one energy level to another, specific amounts of energy are either absorbed or emitted. This is the main mechanism by which photons of light can interact with atoms of matter. These energy levels and their associated photon energies are uniquely dependent on the electronic structure of the atom [4]. Figure 3 shows a generic energy-level diagram for an atom. Molecules are made up of two or more atoms and their energy levels are dictated by the interatomic forces. In addition to electronic transitions, molecules also possess transitions associated with vibrational and rotational interactions which lead to a more complex set of energy levels compared to simple atoms (see Figure 3). For a large group of identical atoms/molecules where the constituents are isolated from one another, such as in a dilute gas, each atom/molecule has the same set of discrete energy levels. However, when atoms/molecules are brought close together, such as in liquids or solids, a variety of intermolecular interactions become increasingly important and the initially discrete energy levels associated with the isolated atoms or molecules gradually broaden into collections of numerous, densely spaced energy levels that form energy bands [1]. These valence and conduction bands and their separating energy gaps are shown in Figure 3 for an insulator and a semiconductor. Details regarding the theory of band formation can be reviewed in the appendix of [5].

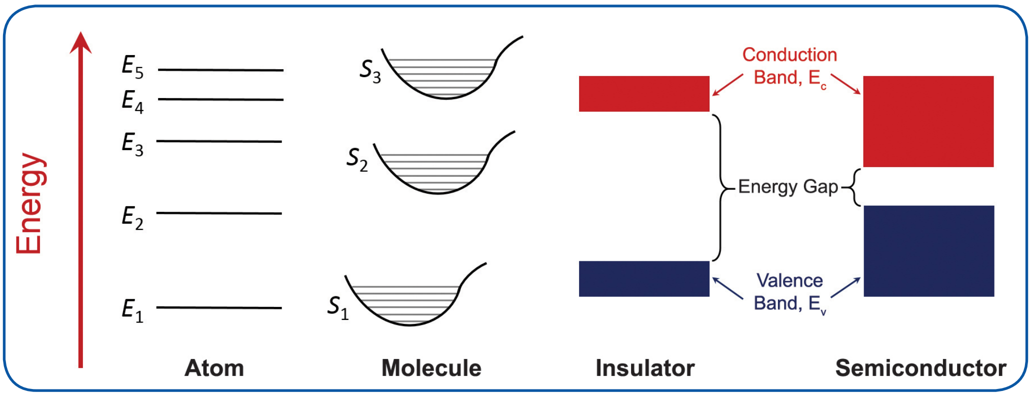


Figure 3. Evolution of energy-level diagrams for increasing interaction between atoms (left to right): an isolated atom, an isolated molecule, a solid insulator, and a solid semiconductor.

Radiation Processes

Regardless of the type of matter that makes up a material, there are three fundamental processes by which light interacts with atoms, creating upward and downward transitions between energy levels [1, 3]. Figure 4 shows the process known as spontaneous emission, which occurs when an atom is in an upper energy level (E_2) or excited state and decays spontaneously to a lower energy level (E_1), radiating a photon of light when it does so. Since each material has a unique set of energy levels, the emitted photons possess energies that are specific to that material. These photon energies are related to the frequency (ν) and wavelength (λ) of the light by:

$$\Delta E = h\nu = {}^{hc}/_{\lambda} \tag{1}$$

where h is Planck's constant and c is the speed of light. The second process shown in Figure 4 is absorption. Absorption occurs when the atom is initially in the lower energy level and a photon with energy $h\nu$ is absorbed by the atom raising it to the upper energy level. The final process in Figure 4 is stimulated emission. Emission is essentially the inverse process of absorption since it also requires a photon of energy $h\nu$ to be present, but the atom is initially in the upper energy level. The photon stimulates the atom to radiate a duplicate photon having the same characteristics as the original, including energy, direction, phase and polarization. These properties are discussed further in Section I.A.3. Stimulated emission produces photons at the expense of energy stored in atoms which gives rise to amplification of the light or optical gain. This phenomenon is central to the operation of a laser and will be discussed in detail in Section I.A.2.

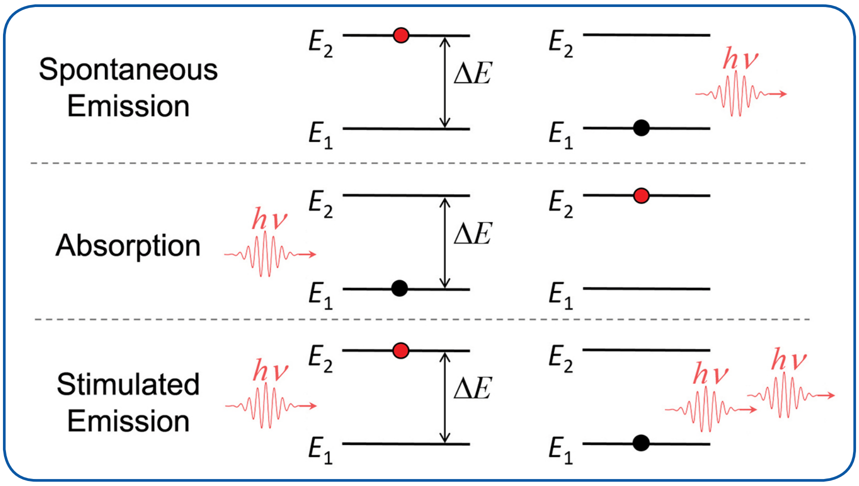


Figure 4. The three radiation processes that can occur when photons of light (hν) interact with atoms making up matter: system prior to interaction (left) and after interaction (right) [3].

Transition Cross Section

One term useful in quantitatively characterizing these atom-radiation interactions is the transition cross section $\sigma(\nu)$ [1]. This value describes the strength or probability that the interaction will occur and is applicable to all three of the radiation processes discussed above. As illustrated in Figure 5, the transition cross section is dependent on the frequency or wavelength of the photon of interest. It is centered on the resonance frequency (ν_0), where $\sigma(\nu)$ is largest, and drops as ν deviates from ν_0. A few key quantities of interest are the peak transition cross section, $\sigma_0 = \sigma(\nu_0)$, the linewidth ($\Delta\nu$) representing the full width at half maximum value (FWHM), and the area (S), which is called the transition strength or oscillator strength. A common way to write the transition cross section is in terms of its strength and profile:

$$\sigma(\nu) = Sg(\nu), \tag{2}$$

where the lineshape function g(ν) is a normalized function with unity area. This allows one to separate the strength of a transition from its frequency dependence. A number of processes contribute to the $\Delta\nu$ associated with g(ν) including homogenous and inhomogenous broadening; references [2] and [1] discuss these topics in detail. Since Equation (2) applies to stimulated emission, its importance in the laser amplification/gain process should be clear. A large transition strength can give rise to a large gain coefficient for lasing, while the lineshape helps determine the frequency response and gain bandwidth (see Section I.A.3 for more information).

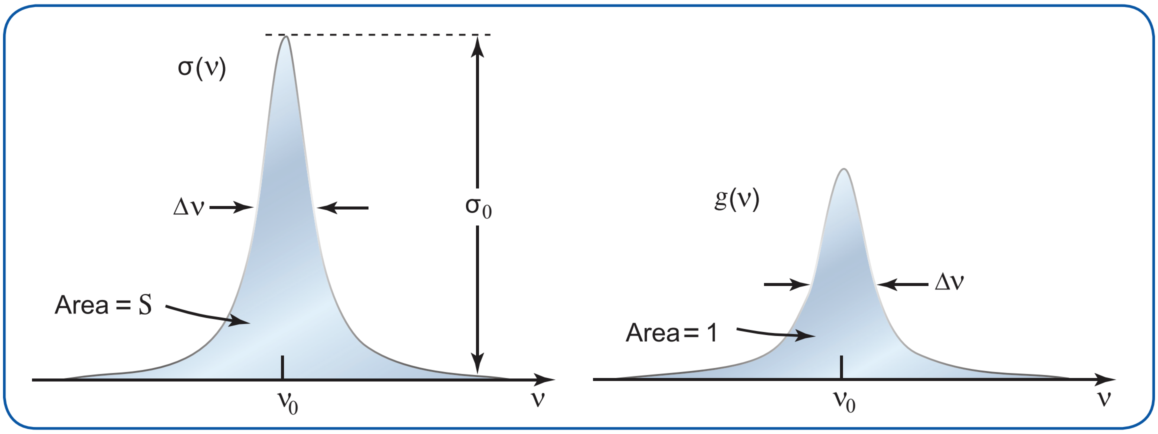


Figure 5. The transition cross section $\sigma(\nu)$ (left) and the lineshape function g(ν) (right) [1].

Population Inversion

The depictions of the radiation processes (Figure 4) illustrate that their impact is tied to the population of the atoms in the various energy levels. For a system in thermal equilibrium, the probability that an atom is in a particular energy level is given by the Boltzmann distribution. As shown in Figure 6, this probability exponentially decreases with increasing energy level. If one considers just two energy levels, E_2 and E_1 (see Figure 6), then the Boltzmann distribution gives the ratio of their respective populations, N_2 and N_1, as

$$\frac{N_2}{N_1} = exp\left(-\frac{\Delta E}{kT}\right), \tag{3}$$

where k is the Boltzmann constant and T is the temperature. Typical energy level differences (ΔE) for laser transitions are two orders-of-magnitude larger than kT at room temperature and so $N_2/N_1 << 1$. Since most of the population is in the lowest energy level, the process of absorption dominates, which is typically how light interacts with most matter. Furthermore, with little population in the upper energy level, stimulated emission is highly unlikely. However, if significant population could be transferred to the higher level such that N_2 exceeds N_1, a non-equilibrium condition known as a population inversion can be achieved. Population inversion is a prerequisite for laser action, as described in Section I.A.2, since stimulated emission is the means by which amplification is achieved.

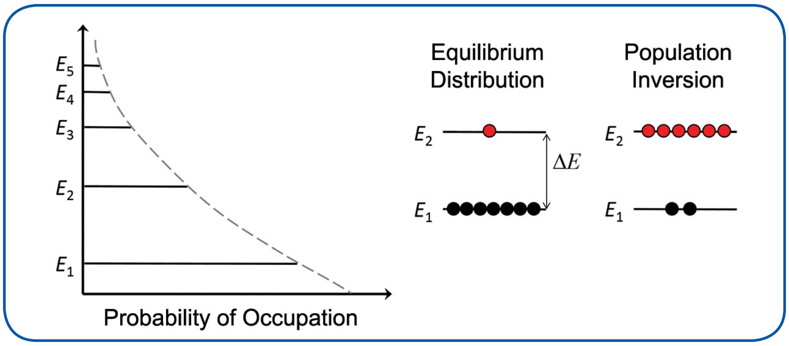


Figure 6. Probability that an atom occupies an energy level of an atom in thermal equilibrium (left, [1]). Population distributions for conditions of equilibrium and population inversion (right).

In addition to the references given in this section, online content discussing the fundamentals of light-matter interactions can be found in several references [6-9].

2. Required Components for Lasing

As discussed above, the critical components of a laser are a gain medium, a pump source, and a resonator. Table 1 lists the primary functions of these components as well as typical examples while Figure 2 illustrates these components in an operational laser. Details regarding the function and operation of these components are given below while in-depth discussions can be found in [1, 2, 10-14].

Component	Function	Examples
Gain Medium	Acts as medium for population inversion Determines laser emission properties	Atomic or molecular gases (Ne, Ar, CO_2) Ions in crystals or glasses (Nd, Er, Yb, Cr) Semiconductors (GaAs, InGaAsP)
Pump Source	Serves as energy source for inverting population	Electric discharge Flashlamp or arc lamp Other laser Electric current
Resonator	Provides feedback mechanism for amplification Selects spectral and spatial properties of light	Bulk mirrors in solid-state laser Cleaved or coated facets in laser diode Bragg reflectors in fiber laser

Table 1. Function and examples for the three components of lasers [2].

Gain Medium

Gain in a laser medium can be described by considering two energy levels with population densities, N_2 and N_1 (similar to Figure 6) and an associated transition cross section σ_{21}, as described by Equation (2). Since population exists in both levels, absorption and stimulated emission will occur within the medium. If I_0 is the intensity of a beam (definitions for laser beam intensity are given in Sections I.A.3 and II.A) before entering a material that has a length L, the light intensity I exiting the material is given by:

$$I = I_0 e^{\sigma_{21}(N_2 - N_1)L}. \tag{4}$$

Based on Section I.A.1, if N_1 is greater than N_2, absorption is the dominant process and the exponent in Equation (4) is negative, resulting in attenuation of the beam. In this case, the product of the population difference (ΔN) and the cross section is known as the absorption coefficient (α) and is a measure of the loss in intensity per unit length. Conversely, if N_2 exceeds N_1 a population inversion exists and stimulated emission will dominate over absorption. Equation (4) then possesses a positive exponent, giving rise to amplification. The product $\sigma\Delta N$ is then defined as the gain coefficient (G) representing the gain per unit length. In order to maximize the amplification, a laser gain medium should possess a large transition cross section at the wavelength of interest and be able to support a significant population inversion.

In practice, more than two energy levels are necessary to achieve a population inversion because a two-level scheme allows, at best, only equal populations of the upper and lower levels, thereby leading to no amplification. Alternatively, three- and four-level laser systems (as shown in Figure 7) essentially decouple the upper and lower levels. A three-level system pumps population from the ground state (level 1) into an excited state (level 3) which, rapidly decays to the level from which stimulated emission occurs (level 2). Level 2 is often referred to as the metastable level since it has a long lifetime. The lower level, where stimulated emission terminates (level 1), should be de-pumped rapidly to dispose of its population. These two requirements, i.e., long lifetime for upper level and short lifetime for lower level, help guarantee that a reasonably large population inversion can be maintained. Since level 1 is also the ground state, it possesses an inherently large population. Therefore, substantial pumping must be implemented to ensure that more than half of the atoms reach the metastable level 2. These disadvantages of a three-level system are avoided in a four-level system. Here, level 1 is situated above the ground state (level 0). Typically, rapid non-radiative decay depopulates this level, making a population inversion easier to achieve. State-level diagrams for a variety of laser media are given in Section I.A.4. Population densities also change dynamically depending on a variety of processes, including radiative and non-radiative decay pathways, pumping rates, as well as absorption and stimulated emission rates. Depictions of practical three- and four-level laser systems along with their decay pathways are shown in Figure 7. The analysis of population dynamics for such multi-level systems can be complicated and more details can be found in [1, 10, 11, 13].

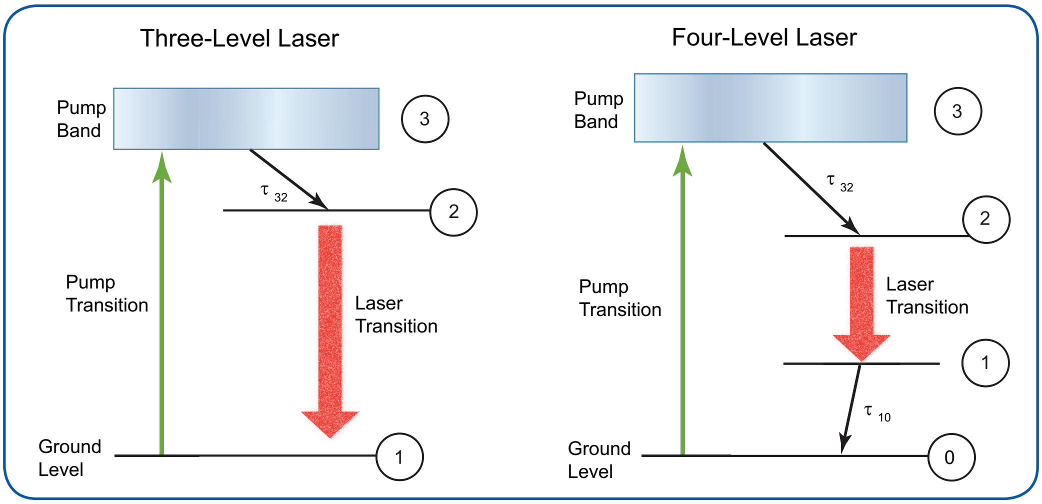


Figure 7. Energy level diagrams of a three- and four-level laser [10].

Pump Source

A variety of different pumping mechanisms are used to achieve population inversions and these are typically dictated by the gain medium. Specific lasers and their associated media are discussed in Section I.A.4. In all cases, the goal is to achieve pumping rates, i.e., atoms per unit time per unit volume being raised to a metastable level, sufficient for establishing sustainable stimulated emission. For gas lasers, the pump mechanism is typically via electric discharge where an electrical current is generated in a low-pressure gas. The electrons transfer their energy to atoms via collision and promote these atoms to higher energy levels. In semiconductor lasers, electrical pumping gives rise to charge carriers within a junction that provides the upper laser level population. Details of this process will be given in Section I.B.2.

Optical pumping is the most common mechanism for generating a population inversion in solid-state lasers and this is enabled using either a lamp or a separate laser. Flashlamps or arc lamps (see Section I.C for details) produce intense light emission and typically have a broad emission spectrum of wavelengths that can be matched to the absorption bands of the laser medium. To efficiently pump the laser medium, the flashlamp geometry can take on different configurations, including a helix wrapped around a laser rod or laser rods inserted into an elliptically-shaped or circularly-shaped elongated laser cavity (see Figure 8). Laser pumping is used when the pumping energy must be concentrated into a relatively small volume, delivered over a short period, or provided over a narrow bandwidth. Pumping via semiconductor lasers has several advantages, including increased system efficiency, better light coupling into the laser medium, and compact geometry. Figure 8 shows a typical end-pumping configuration for a semiconductor laser diode; however, other geometries such as side-pumping and close-coupling configurations also exist [1, 10, 14]. When a laser diode pumps a solid-state gain medium, it is referred to as a diode pumped solid-state (DPSS) laser system. DPSS lasers can be used to generate very large pumping rates for other solid-state lasers.

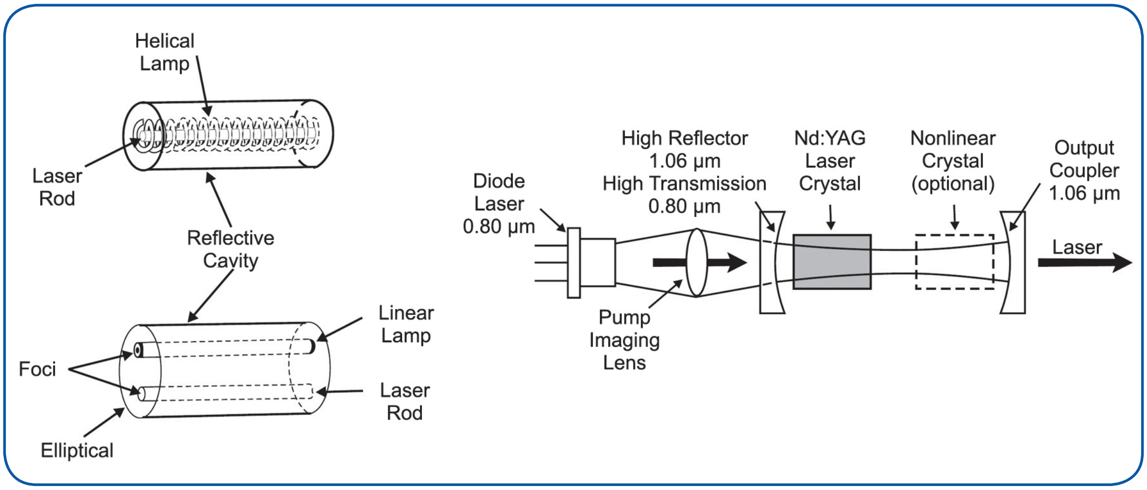


Figure 8. Examples for optically pumping a gain medium including flashlamp pumping (left) and semiconductor laser diode pumping (right). Figures are reprinted with permission from [3], SPIE Publications.

Resonator

In a laser, the gain medium is placed inside an optical resonator that provides feedback. This feedback mechanism allows for photons generated by stimulated emission to be reflected back into the laser medium for further amplification. A common example of a laser resonator or cavity uses two mirrors separated by a specific distance (*d*) (see Figure 2). Typically, one mirror (known as the end mirror) is highly reflective at the lasing wavelength, while the other mirror (known as the output coupler) is partially transmissive so that a portion of the light exits the cavity as the laser beam. One benefit of this feedback mechanism can be gleaned from Equation (4), which gives the optical gain in a laser as the product of the gain coefficient and the laser medium length (*GL*). For a given laser gain medium, extending *L* could be a means to generate sufficient gain for lasing but there are clearly practical limitations to simply elongating the gain medium.

An increase in the effective optical pathlength can provide a means for enhancing gain. This can be achieved by placing the laser medium inside an optical resonator. This is illustrated when comparing the upper two graphics in Figure 9, which show a gain medium with and without a resonator. While the additional passes allow for increased gain when compared to a single pass geometry, after just a few passes through the gain medium, the light begins to leak outside the cavity. For this optical feedback to work efficiently, the resonator must be designed to be stable such that light stays inside the cavity and do not leak out. The parallelism of the mirrors as well as their curvatures play a large role in the design of stable laser resonators. The lower part of Figure 9 shows a few typical resonator configurations while detailed discussions of resonators and their stability regimes can be found in [10, 12, 13]. In addition to the resonator-induced feedback leading to increased gain, it also decreases the beam divergence significantly (see upper portion of Figure 9). This decreased beam divergence is responsible for the high degree of collimation observed for most laser beams (see Section I.A.3 for more details).

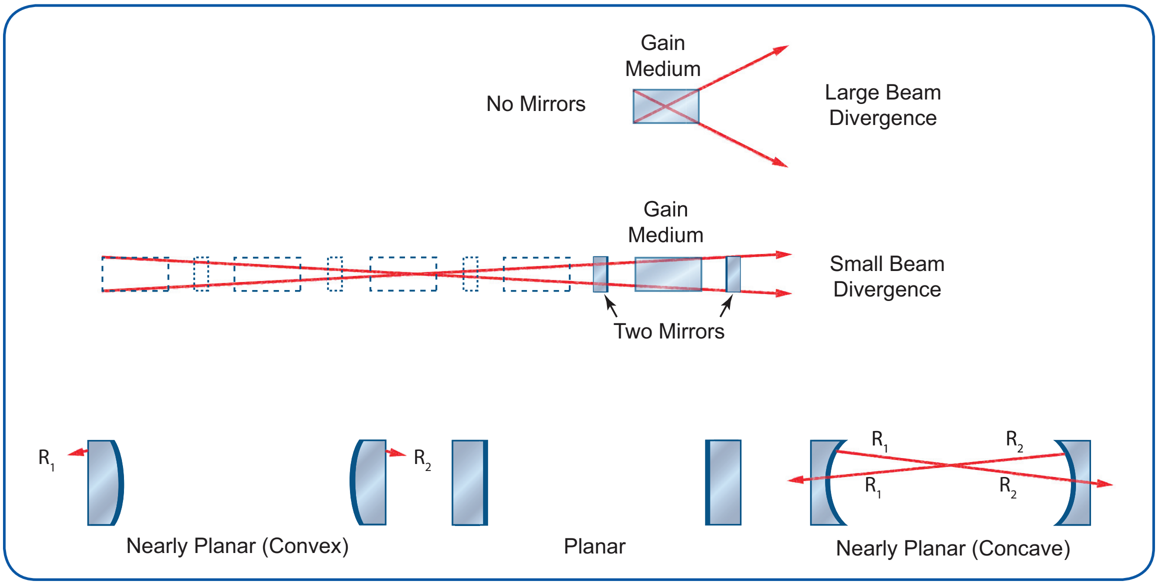


Figure 9. Depiction of single pass through an amplifier (top) and multiple passes through an amplifier when a resonator is present (middle) [15]. The dotted portions illustrate the effective increase in optical pathlength provided by the resonator. Common resonator configurations (bottom).

Since an amplifier with positive feedback is an oscillator, a laser is often called an optical oscillator. For laser oscillation or lasing to occur, two conditions must be satisfied that are governed by the gain medium and the resonator [1, 14]. The first condition dictates that the optical gain *GL* be greater than the total losses (α_r) for a single round trip through the cavity. Parasitic losses within a laser cavity typically come from absorption within the gain medium itself, optical scattering or absorption from any element within the cavity. Necessary losses, such as those associated with transmission through the output coupler, must also be considered. The second condition requires that light waves within the resonator must coherently add or constructively interfere with one another (see Section I.A.3 on coherence for details). Only light that possesses an integral number of oscillating waves that fit within the cavity length (d) will be supported. These standing waves are known as resonator modes and are depicted in Figure 10. Longitudinal modes differ from one another only in their oscillation frequency with a separation ($\Delta\nu$) given by:

$$\Delta\upsilon = \frac{c}{2d}. \tag{5}$$

This frequency separation is typically small (~ 0.5 GHz) compared to the laser gain bandwidth and any longitudinal mode that has sufficient gain can undergo lasing. There can be a single longitudinal mode or several hundred thousand modes depending on the gain bandwidth (see Section I.A.3). On the other hand, transverse modes follow different transverse paths through the amplifier and therefore emerge larger in size and in divergence. Figure 10 shows a variety of laser beam intensity distributions associated with different cylindrical transverse modes. These are defined by their transverse electromagnetic (TEM_{mn}) distributions with the lowest order mode being the TEM_{00}, which is discussed further in Section II.B.1.

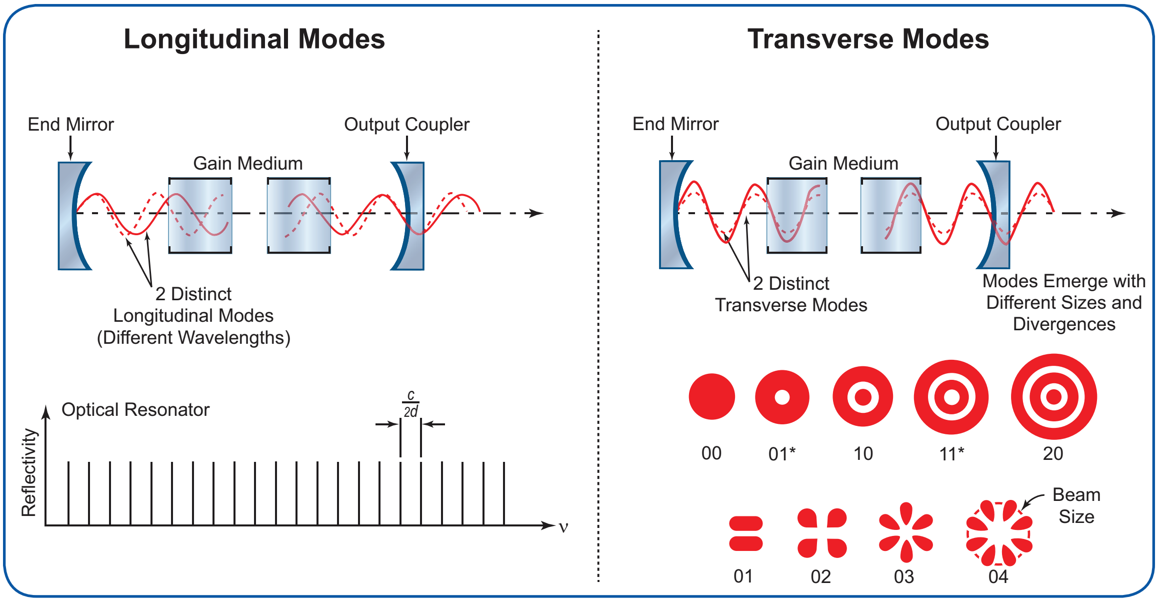


Figure 10. Depiction of longitudinal and cylindrical transverse modes produced by an optical cavity (top, [3]). The frequency spacing of longitudinal modes is shown in the lower left. The spatial patterns of the transverse modes are shown in the lower right [10].

Based on the discussion in this section, a laser can be considered a resonant optical amplifier whose output is fed back to the input with matching phase. Additional online content discussing the functional components of a laser and its principles of operation can be found in [7-9, 16].

3. Characteristics of Laser Output

The process of coherent amplification imbues laser light with a very unique set of characteristics. Typically, only one or a subset of these characteristics are most critical for a particular application, therefore each will be described separately; however, many of these characteristics are interrelated. While not exhaustive, the most common laser output characteristics include: wavelength, gain bandwidth, monochromaticity, spatial and temporal profiles, collimation, output power, coherence and polarization.

Wavelength

A large portion of the electromagnetic radiation spectrum is covered by a wide range of existing lasers (see [17]). The wavelength range extends from the ultraviolet (UV) to the mid-infrared (MIR) and does not account for other more exotic systems that provide access from the soft-X-ray spectral region (< 10 nm) to the far-infrared (FIR, > 100 μm). The lasing wavelength (or frequency ν_0) is determined by the laser gain medium, which provides the optical transition (see Figure 5 or Figure 7). The wide range of wavelengths possible is attributable to the large variety of available gain media. Furthermore, nearly all laser wavelengths can be converted or shifted to an alternative wavelength (see Section 1.A.6) and thus can reach from the UV to the MIR spectral region. This spectral agility is the reason laser systems can be employed for short-wavelength applications like lithography for semiconductor processing and long-wavelength applications like material processing and molecular spectroscopy.

Gain Bandwidth

The bandwidth of the laser gain medium (*B*) determines the range of wavelengths over which amplification can occur. This bandwidth is determined primarily by the bandwidth over which spontaneous emission occurs (see Figure 5). While various processes contribute to the broadening of the transition linewidth ($\Delta\nu$), several electronic transitions (which are also affected by rotational and vibrational motions) can overlap in frequency, leading to significantly wider bands, particularly for molecular or solid-state systems as shown in Figure 3. Typical bandwidths for select gain media are shown on the right side of Figure 11. Gas lasers, like the HeNe laser, typically have very narrow bandwidths on the order of 1 GHz owing to their atomic transitions. Conversely, solid-state lasers, such as the Ti:Al_2O_3 (sometimes referred to as Ti:Sapphire or Ti:Saph) laser, can have extremely wide bandwidths exceeding 100 THz. The gain bandwidth is also dictated by the total loss in the system (α_r) since a net gain is required for lasing (see Figure 11, left). Consequently, the actual gain bandwidth may be different than the spontaneous emission bandwidth. For instance, modulating the intracavity loss is a means for achieving laser wavelength tuning (see Section I.A.6). Furthermore, the gain bandwidth is not necessarily the same as the bandwidth of the exiting laser beam since that will also depend on the laser resonator as discussed below.

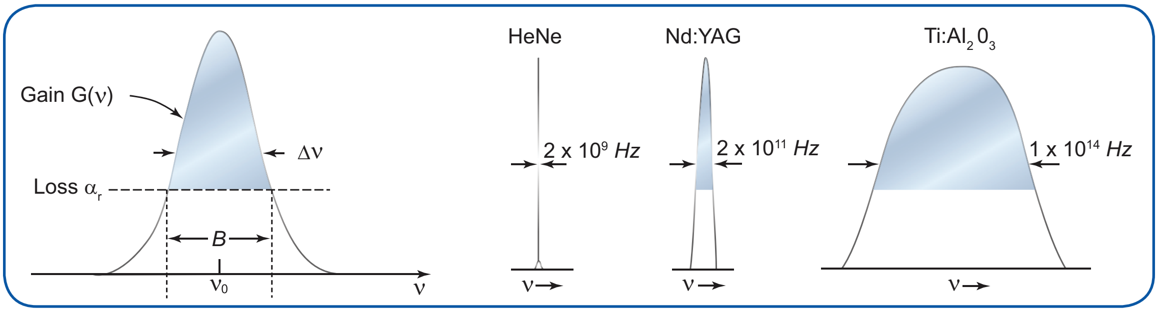


Figure 11. Laser oscillation can occur only at frequencies for which the gain coefficient is greater than the loss coefficient (filled-in region) (left, [1]). Laser gain bandwidths for the HeNe, Nd:YAG, and Ti:Al_2O_3 lasers (right, [3]).

Monochromaticity

Monochromaticity refers to color purity or, in the case of a laser, the spectral bandwidth of the laser (sometimes referred to as the laser linewidth). Figure 12 shows how a combination of the laser gain bandwidth and the laser cavity properties determines the bandwidth of the emerging laser beam. Any number of longitudinal modes can lase provided they lie within the window where the gain exceeds the loss. The number of these lasing modes (*N*) is given by the gain bandwidth divided by the resonator frequency spacing given by Equation (5):

$$N = \frac{B}{\Delta \upsilon}. \quad (6)$$

For example, if the entire spontaneous emission spectrum for a Ti:Al_2O_3 laser shown in Figure 11 is available for gain, *N* can exceed 200,000. This is the basis for implementing a process called mode-locking for generating short-pulse lasers in these types of systems (see Section I.A.5 for details). Alternatively, many gas lasers have a sufficiently narrow gain bandwidth for which only a few longitudinal modes are supported. Reducing the cavity length (or equivalently increasing the mode spacing) is a means of achieving lasing with a single longitudinal mode in any laser system. However, this would also place an upper limit on the length of the active medium and therefore limit the achievable gain. Alternatively, techniques exist for introducing frequency-selective elements inside the cavity that allow for single-mode selection without affecting the overall gain [1, 13]. An example is shown in Figure 12 where an etalon, i.e., a type of resonator, of length d_1 is placed inside the laser cavity and only one etalon mode fits within the gain bandwidth. The etalon mode is made to overlap with one of the laser cavity modes and only a single longitudinal mode lases. The spectral width of this mode is controlled by the reflectivity and stability of the cavity which is quantified in the cavity's quality factor *Q*. For more details about resonator mode linewidths and cavity *Q*-factors see [1, 10, 14]. Careful cavity designs can enable very narrow laser bandwidths (< 1 MHz) to be achieved. This highly monochromatic output is desirable for applications in remote sensing or as a frequency standard.

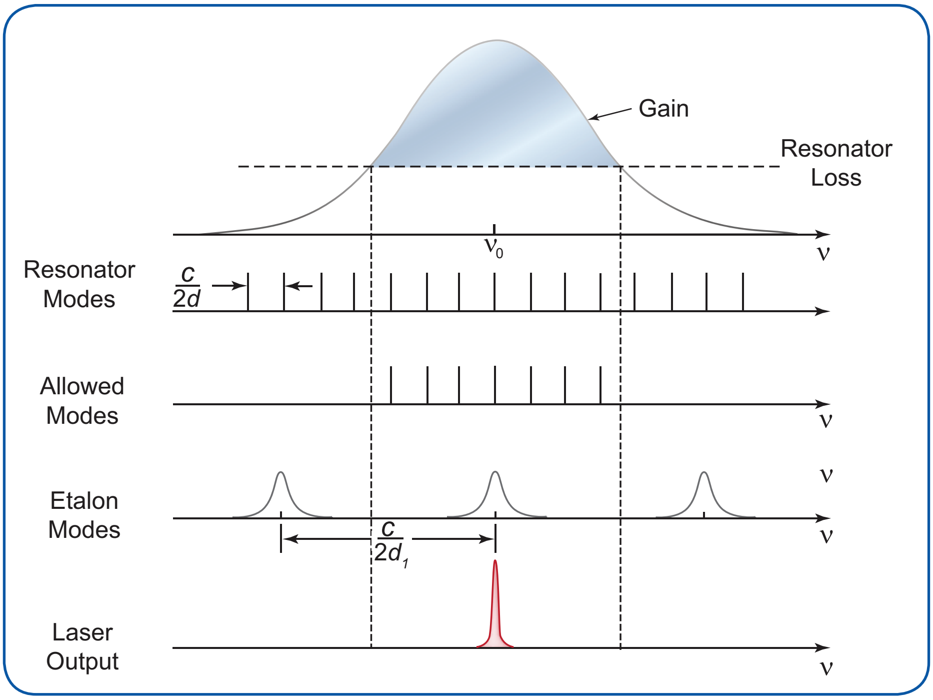


Figure 12. Oscillation can occur only for allowed resonator modes that lie under the gain bandwidth (upper portion of figure) and single longitudinal mode selection by the use of an intracavity etalon which is a type of resonator (lower portion of figure) [1].

Spatial and Temporal Profiles

Beams that emerge from a laser cavity have an intensity distribution that has both a transverse spatial profile as well as a temporal profile. The spatial profile is mainly determined by the transverse cavity mode (see Figure 10) and can be rotationally symmetric. Each transverse mode has a different spatial distribution; adjustments to the laser cavity mirrors along with insertion of an aperture inside the resonator can be used to selectively attenuate undesired modes. The lowest-order transverse mode (TEM_{00}), which travels down the central axis of the cavity, is often the desired mode because it propagates with the least beam divergence and can be focused to the tightest spot (see section below on collimation). A laser that operates in the fundamental or TEM_{00} mode will emit a transverse beam profile described by a Gaussian function. The specifics of this intensity profile as well as its evolution with distance will be described in Section II.B.1.

Certain lasers operate in continuous wave (or *CW*) mode where the temporal profile of their output power is constant over time (see section below on output power). Conversely, other lasers operate in pulsed mode such that their output power has a transient temporal profile. These lasers are often characterized by the shape and width of their temporal profiles (see Section II.E) and, since they typically emit a series of pulses, by their repetition rate, which is given in Hz. Pulsed lasers are useful for many different applications where *CW* laser properties are not sufficient. There are numerous pulse generation methods (detailed in Section I.A.5) which allow for pulse durations of microseconds (μs), nanoseconds (ns), picoseconds (ps), down to femtoseconds (fs) and below.

Collimation

Due to diffraction, light emitted from any source will diverge. That is, its transverse spatial profile will get larger as it travels or propagates a greater distance. Laser beams typically possess a much smaller divergence than any other source of light, which is another way of saying that a laser beam is highly collimated. Collimated light possesses photons which are highly directional and propagate parallel to one another. As described in Figure 9, a laser beam's high degree of collimation arises from the parallelism of the cavity mirrors, which forces the beam to be perpendicular to those mirrors. The TEM_{00} mode possesses the lowest divergence of all the transverse modes. The propagation of the Gaussian beam associated with this mode is shown in Figure 13 with a detailed description in Section II.B.1. Due to its negligible divergence, highly collimated laser light enables a wide variety of applications, including atmospheric sensing, adaptive optics for astronomical telescopes, and even lunar laser ranging where a beam is propagated to the Moon. Interestingly, a laser's high degree of directionality can also be used to focus it to a small spot size (see Figure 13). By sending a collimated TEM_{00} transverse mode through a focusing element such as a lens or curved mirror, the beam diameter can be reduced to a diameter that is approximately the size of the laser wavelength. Details regarding the physics of focused laser beams, particularly in the context of microscopy, can be found in [18]. This ability to tightly focus laser light to a small spot size with a large intensity is exploited for applications in high-resolution microscopy, nonlinear optics, photolithography, and even nuclear fusion.

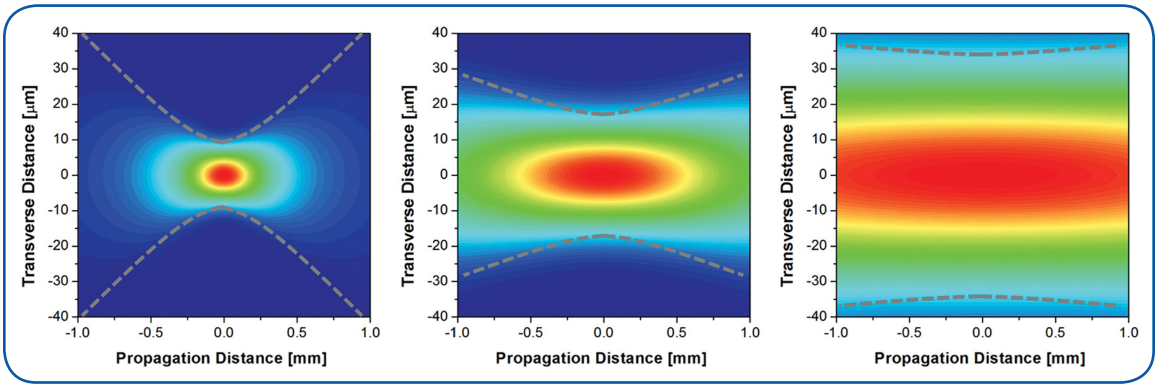


Figure 13. Propagation of a laser beam with a Gaussian distribution with large (left), moderate (middle), and small divergence (right). Dotted lines indicate transverse beam sizes at different propagation distances.

Output Power

The average output power of a laser is often quoted in units of milliwatts (mW), watts (W) or kilowatts (kW). For a *CW* laser whose output power is constant over time, the meaning of this value is straightforward. For a pulsed laser, an average output power can also be defined as the product of the pulse energy and the repetition rate (see Section II.A.1). Since the energy in each pulse is squeezed into a small amount of time, another quantity known as *peak output power* or peak power is proportional to the pulse energy divided by the width of the pulse (see Section II.E). To see how "powerful" a pulsed laser can be, consider a 100 W light bulb whose average and peak powers are necessarily the same. For a 10 ns laser with a pulse energy of 1 Joule (J) operating at 10 Hz, an average power of 10 W is generated, but the laser's peak power is 1,000,000 kW. *CW* lasers often have average powers much less than 100 W; however, it is the ability of confining that power into a collimated beam with a relatively small spatial distribution that distinguishes it from other lights sources such as lamps. Thus, quantities that account for the power per unit area (known as the intensity), or that also take into account divergence (a parameter known as radiance), are also important laser characteristics. These various radiometric quantities as well as specific examples will be discussed in more detail in Section II.A.

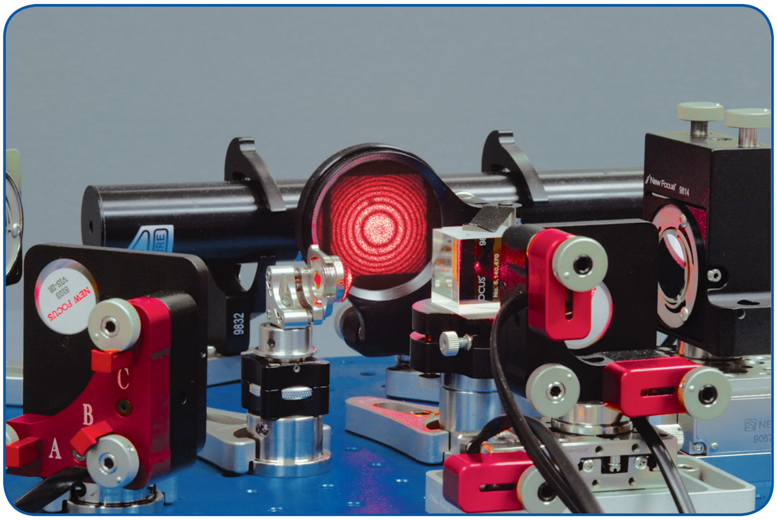


Figure 14. Interference fringes are created by splitting a beam of monochromatic light so that one beam strikes a fixed mirror and the other a movable mirror. When the reflected beams are brought back together, an interference pattern results. The central portion of the interference pattern shows granularity which is the result of laser speckle.

Coherence

Recall that during the stimulated emission process, laser photons are cloned such that they possess a fixed phase relation to one another. Coherence refers to the degree to which various portions of a single laser beam are in phase. This coherence is described in terms of both temporal coherence and spatial coherence. Temporal or longitudinal coherence determines how readily different beams can interfere with each other. Interference is a phenomenon that occurs when two waves occupy the same space and are coherent. This can result in either constructive or destructive interference depending on whether the amplitude of the resultant wave is either more or less intense, respectively, than the original waves. When a beam is divided such that each one travels a different distance before recombining, such as in an interferometer (see Figure 14), they will interfere with each other provided they are still coherent. The longitudinal coherence length describes the distance over which beams stay in phase and is inversely dependent on the laser bandwidth described in the monochromaticity section above. This temporal coherence is critical for applications such as holography and general interferometry. Spatial or transverse coherence refers to the phase relationship between different spatial portions of a laser beam after it has propagated a certain distance. The impact of spatial coherence is typically observed as laser speckle (see Figure 14) when laser light is reflected off a rough surface such that the waves of light from each portion of the surface interfere with one another, resulting in a granular pattern called speckle. For more information on temporal and spatial coherence and how these wave properties impact various interference process, see [1, 14].

Polarization

Polarization refers to the direction of oscillation for the transverse waves that make up the electromagnetic fields of light (details regarding polarization are discussed in Section III.A.4). Linear polarization refers to the oscillation of the field being confined to a single plane perpendicular to the propagation direction. Several laser-based applications require a linearly polarized source, including nonlinear frequency conversion (see Section A.1.6), certain forms of optical communication, and interferometry. Consequently, it is desirable for a laser to be linearly polarized. While this could be accomplished outside the laser cavity using a polarization-selective component (see Section III.A.4), there are issues with noise arising within the laser cavity from multiple polarizations competing for amplification. A single linearly polarized beam can be generated within the laser cavity if the laser gain medium is polarization dependent, which is the case for some solid-state lasers. An alternative method involves placing a polarization-selective component within the resonator. Figure 15 shows an example where a Brewster window (see Section III.A.4) induces additional loss in the polarization perpendicular to the plane of incidence, while the polarization in the plane of incidence experiences no loss. For more information on this topic, see [19].

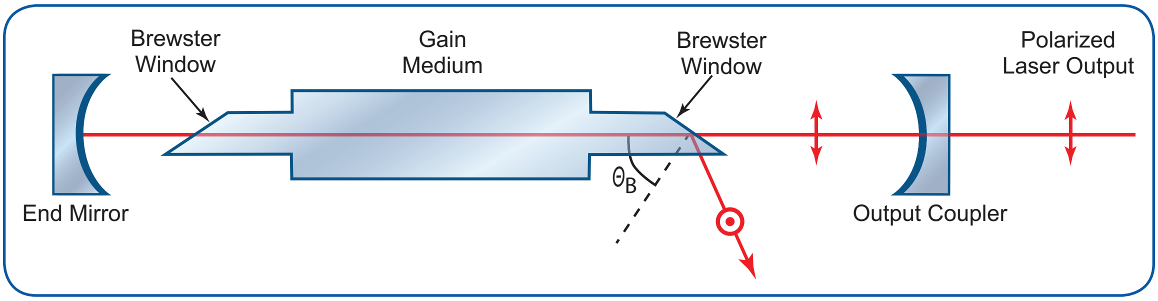


Figure 15. The use of Brewster windows in a gas laser provides a linearly polarized laser beam. Light polarized in the plane of incidence is transmitted without reflection loss through a window placed at the Brewster angle (Θ_B). The orthogonally polarized mode suffers reflection loss and therefore does not oscillate [1].

Additional online content discussing laser characteristics and the applications that they enable can be found in [16, 20-22].

4. Types of Lasers

Lasers are typically identified by their gain medium and are often classified by the radiating species that give rise to stimulated emission. These radiating species can include atoms and molecules in a dilute gas, organic molecules dissolved at relatively low concentration in liquid solutions, semiconductor materials, and dielectrics such as crystalline solids or glasses that are doped with a high concentration of ions. These laser categories are generally referred to as gas, liquid, semiconductor, and solid-state. As alluded to in Section I.A.1, the concentration of the radiating species plays a significant role in the formation of the medium's energy levels. In turn, these energy levels will dictate the optical pumping transition, configuration for population inversion, laser emission wavelength, and gain bandwidth (see Figure 7). While details regarding specific interactions leading to population inversion in these various laser media are discussed below, Figure 16 provides representative energy-level diagrams and population inversion processes for each laser type.

Lasing occurs in a wide variety of media with output characteristics that possess an even greater degree of variation. A list including all commercially-available systems as well as those specifically designed for targeted applications would easily fill up an entire book. For this section, the focus will be on the most common commercially-available lasers. Table 2 provides the names of these systems (based on their media) as well as their nominal characteristics and parameters. This table can be referred to during the subsequent discussions of the individual laser classes.

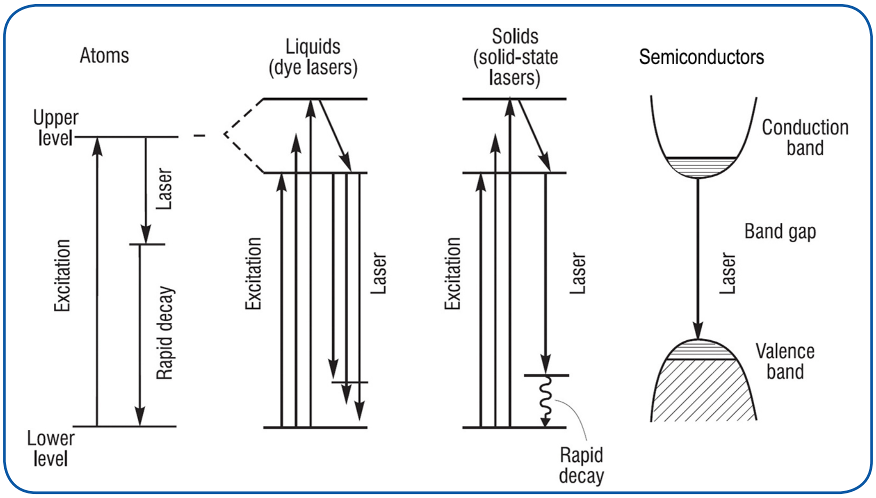


Figure 16. Typical inversion processes in gases, liquids, solids, and semiconductors [3] (reprinted with permission from SPIE Publications).

Laser	Type	Wavelength (µm)	CW or Pulsed	Output Power	Applications
ArF, Krf, XeCl, Xef	Gas (excimer)	0.193, 0.248, 0.308, 0.353	ns	10 W	UV lithography, laser surgery, LASIK, laser annealing
Nitrogen	Gas	0.337	ns	100 mW	Dye laser pumping, measuring air pollution
Dye	Liquid	0.4-1.0	CW-fs	1 W	Spectroscopy, laser medicine
GaN	Semiconductor	0.41	CW, ns	50 mW	Optical disc (Blu-ray) reading/recording
Argon-ion	Gas	0.488	CW	10 W	Microscopy, retinal phototherapy, lithography
HeNe	Gas	0.6328	CW	10 mW	Interferometry, holography, barcode scanning
AlGalnP, AlGaAs	Semiconductor	0.63-0.9	CW, ms	10 mW, 10 W	Optical disc (CD, DVD) reading/recording, laser pointers, solid-state laser pumping, machining
Ti:Saph	Solid-state	0.65-1.1	CW-fs	10 W	Spectroscopy, LIDAR, nonlinear frequency conversion, multiphoton microscopy
Yb:YAG	Solid-state	1.03	CW-ps	W-kW	Materials processing, optical refrigeration, LIDAR
Yb-glass	Fiber	1.03	CW-fs	W-kW	Materials processing, ultrashort pulse research, LIDAR
Nd:YAG	Solid-state	1.06	CW-ps	W-kW	Material processing, rangefinding, surgery, tattoo/hair removal, pumping other solid-state lasers
ND:glass	Fiber	1.06	CW-fs	W-kW	Material processing, pumping other solid-state lasers, extremely high power/energy systems for laser fusion
InGaAs, InGaAsP	Semiconductor	1.1-2.0	CW, ms	mW-W	Telecommunications, solid-state laser pumping, machining, medical
Er-glass	Fiber	1.53-1.56	CW	10 W	Optical amplifiers for telecommunications
Tm:YAG, Ho:YAG	Solid-state	2.0-2.1	µs, ns	W	Tissue ablation, kidney stone removal, dentistry, LIDAR
Cr:ZnSe	Solid-state	2.2-2.8	CW, fs	10 W	MWIR laser radar, missile countermeasures, ultrafast and high-resolution spectroscopy, frequency metrology
CO_2	Gas	10.6	CW, µs	kW	Material processing, surgery, dental laser, military lasers

Table 2. Characteristics and parameters for common lasers, in order of increasing wavelength [2, 17]. Output powers represent typical values only.

Gas Lasers

For gas lasers [1, 2, 14], population inversion is typically achieved by applying a voltage across a glass or ceramic tube that contains the gain medium which is either a low-pressure gas or gas mixture (see Figure 17). The voltage produces an electric field within the tube which induces an electrical current. These electrons collide with the gas atoms, thereby exciting them to higher energy levels that will serve as the upper laser level. The lower laser level typically decays to the ground state much faster than the upper level, thereby creating a population inversion between the two (see Figure 16). Since the radiating species are very dilute, the resulting laser transitions have very narrow spectral bandwidths and operate at well-defined wavelengths. Due to the wide variety of gaseous media, the range of operating wavelengths can vary from the UV for excimer lasers, through the visible (VIS) range for argon ion and HeNe lasers to the MIR range for CO_2 lasers. Gas lasers have historically been deployed in a wide variety of applications but have been largely supplanted by DPSS lasers and laser diodes except for specialized applications. The exceptions to this trend are for CO_2 and excimer lasers which still play significant roles in the laser processing and medical eye surgery markets.

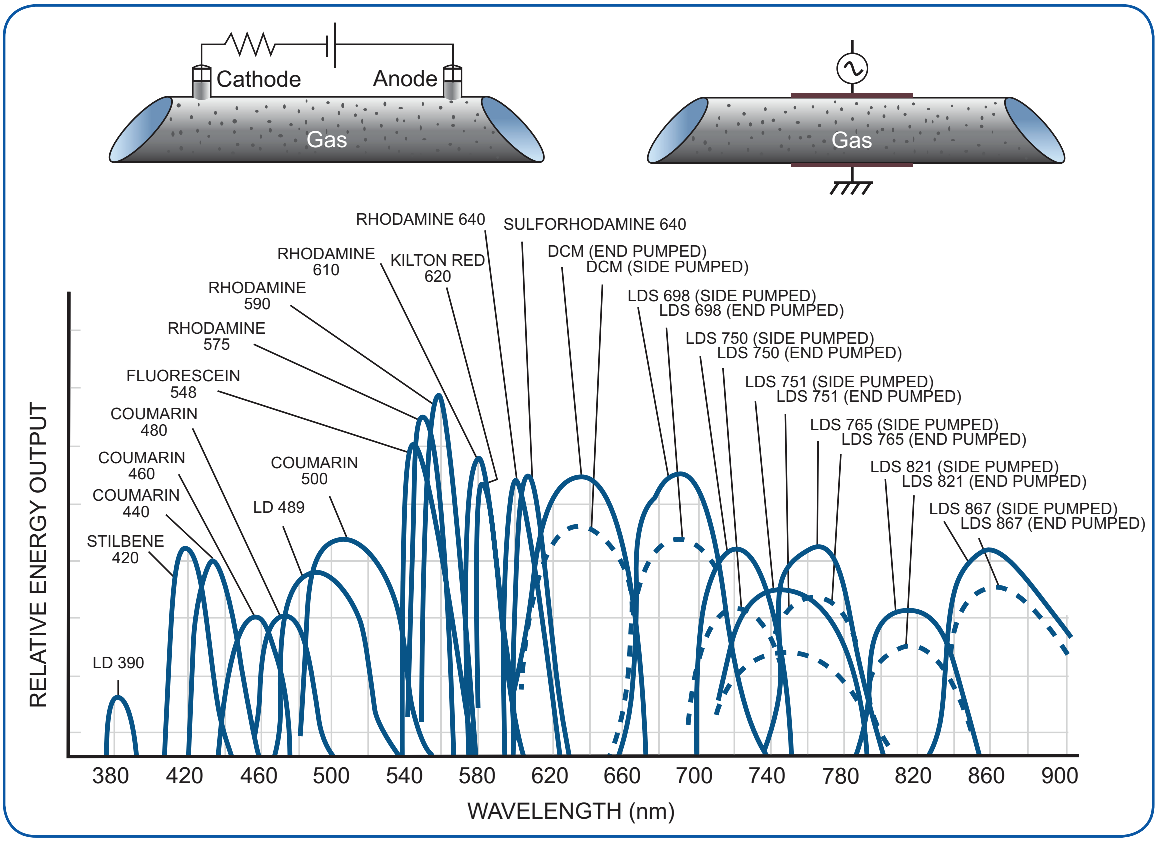


Figure 17. Direct current (upper left) and radio-frequency discharge currents (upper right) are used for pumping gas lasers [1]. Tuning curves of several dyes in a commercial dye laser with pulsed excitation (bottom).

Liquid Lasers

Certain organic dye molecules can act as radiating species for lasing since they have sufficiently long lifetimes in their upper energy levels and can therefore radiate energy from that level instead of losing energy due to collisions. To ensure the proper concentration of radiating species are present, the dye molecules (typically in powder form) are dissolved in a solvent at a concentration of about one part in ten thousand. Due to this solution form, the system is known as a liquid dye laser [1, 15]. Dye lasers are optically pumped by either flashlamps or other lasers. Each dye molecule, due to its overlapping electronic/rotational/vibrational transitions, has a wide homogeneously broadened gain spectrum on the order of 30-50 nm. By utilizing many different dye molecules, the laser can be tuned over a wide spectrum in the UV, VIS, and near-infrared (NIR) (see Figure 17). Combining this broad gain bandwidth with a frequency-selective element allows wide tunability coupled with a narrow spectral bandwidth. As a result, the dye laser has traditionally been used for various spectroscopic applications. Dye lasers require significant maintenance due to the decomposition of the dye when dissolved in its solvent. Therefore, DPSS lasers coupled with nonlinear frequency conversion (see Section I.A.6) have largely replaced dye lasers in many applications.

Semiconductor Lasers

A semiconductor laser is often referred to as a laser diode since it operates like a diode with current flowing in the forward direction of the junction. By injecting charge carriers into the region of space defined by the junction, recombination radiation can occur. Provided this current injection is strong enough, a population inversion can be achieved and stimulated emission will occur. Due to the large refractive index difference between the semiconductor material and air, the semiconductor crystal surface can possess sufficient reflectivity to act as its own resonator cavity. These two characteristics, electrical pumping and compact laser design, coupled with the maturity of the semiconductor manufacturing process, has enabled laser diodes to gain a number of advantages over other types of lasers, including high power and efficiency, small size, as well as compatibility with electronic components. Unsurprisingly, they are one of the most important classes of lasers in use today, not only because of their use in applications such as optical data storage and optical fiber communication, but also because they serve as pumping sources for solid-state lasers. Consequently, semiconductor laser diodes will be considered at some length in Section I.B.

Solid-State Lasers

The term solid-state laser refers to a laser whose gain medium consists of active ion species introduced as impurities in an optically transparent host material (typically crystals or glasses) [10, 11]. As detailed in Sections I.A.1 and I.A.2, materials for laser operation should possess strong and spectrally-narrow transition cross-sections, strong absorption bands for pumping, and a long-lived metastable state. Ions that have optical transitions between states of inner, incomplete electron shells generally exhibit these characteristics. However, these ions must be protected or shielded from other ions to prevent loss of these desired characteristics. This is accomplished by incorporating the ions in a solid host material whose lattice allows for ion doping levels sufficient for a gain medium while simultaneously shielding the ions from one another. According to Section I.A.2, solid-state lasers achieve their population inversion through optical pumping, which can be accomplished by using a flashlamp or direct pumping from another laser source such as a laser diode or a DPSS system.

Formula	Name
Crystals	
$Y_3Al_5O_{12}$	Yttrium Aluminum Garnet, YAG
$Gd_3Ga_5O_{12}$	Gadolinium Gallium Garnet, GGG
Al_2O_3	Sapphire
$LiSrAlF_6$	LiSAF
Mg_2SiO_4	Fosterite
$YLiF_4$	Yttrium Lithium Fluoride or YLF
YVO_4	Yttrium Vanadate, YVO
Glasses	
Silicate-based	e.g. $Si0_2$ or fused silica
Phosphate-based	

Figure 18. Table showing formulae and common names of various crystal and glass solid-state host media. Photograph shows typical solid-state media rods [23]. Image courtesy of Quantel Lasers.

The host material for a solid-state gain medium must possess both unique microscopic lattice properties and appropriate macroscopic mechanical, thermal, and optical properties [1, 2, 11]. These host materials can be organic matter, ceramics, crystals, and glasses but generally fall into the latter two categories. The most common crystal and glass solid-state host materials are given in Figure 18. Crystal hosts have several benefits, including narrower laser linewidths, lower laser thresholds which allow for lower doping levels, and higher thermal conductivities. Compared to crystals, glass hosts have many distinct merits: they have lower melting temperatures and can therefore be fabricated at a lower cost and in larger dimensions, they possess high optical quality, and can be doped homogeneously at higher concentrations. However, they have much lower thermal conductivity than crystals and therefore are utilized mainly for systems that operate at high peak powers and low repetition rate. Line-broadening behavior in solid-state lasers is typically dictated by the host medium with crystalline hosts generating homogeneous broadening and glass hosts giving inhomogeneous broadening.

The majority of dopant ions used in solid-state laser media are transition-metal and lanthanide-metal (or rare-earth) ions [1, 2, 14]. The most common transition-metals used are titanium (Ti) and chromium (Cr) while the lanthanides are neodymium (Nd), ytterbium (Yb), erbium (Er), thulium (Tm), and holmium (Ho). The dopant ions are generally dispersed throughout the host with a concentration (typically around one part in one hundred) depending on the dopant, host material, and application. These ions possess electrons that are surrounded by a "screen" (of electrons from the host material) that protect these electrons from interacting with neighboring dopant ions. Therefore, the ions can radiate their energy rather than decay by collisions similar to the way that organic dye molecules behave when in a liquid solvent. When these ions absorb light, the energy ends up in an excited energy level that serves as the upper laser level. By virtue of the screening provided by the host material, this upper level has a very long lifetime (the metastable level) before it radiatively decays. A population inversion can occur (see Figure 16) because the lower laser levels are being rapidly depopulated by collisions with the neighboring atoms as these levels are not protected the way the upper levels are.

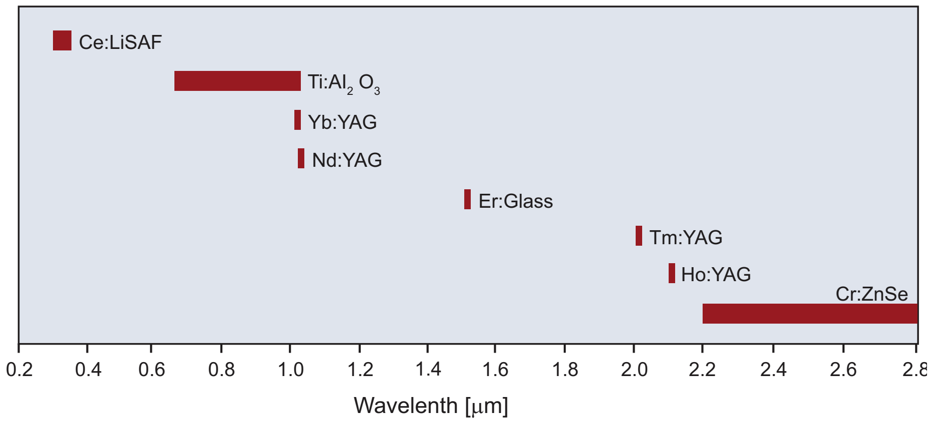


Figure 19. Laser gain bandwidths for common solid-state laser materials.

There are many varieties of solid-state lasers due to the range of host materials available and the numerous dopant ions. Some of the more common solid-state lasers are listed in Figure 19 along with their gain bandwidths. Clearly, solid-state media can possess large gain bandwidths with the possibility of broad wavelength tunability or narrow gain bandwidths with very monochromatic linewidths. In addition to the laser wavelength span shown in Figure 19, solid-state lasers can be frequency- or wavelength-tuned using methods discussed in Section A.1.6. Beyond this spectral agility, the variations in gain media enable systems to exhibit many different characteristics including very high output powers, lower powers with very good spatial beam quality, *CW* output with remarkable power stability, or ultrashort pulses with ps or fs durations. This flexibility enables solid-state lasers to be used in applications ranging from multiphoton microscopy to light detection and ranging (LIDAR) to material processing/marking/cutting and even laser fusion.

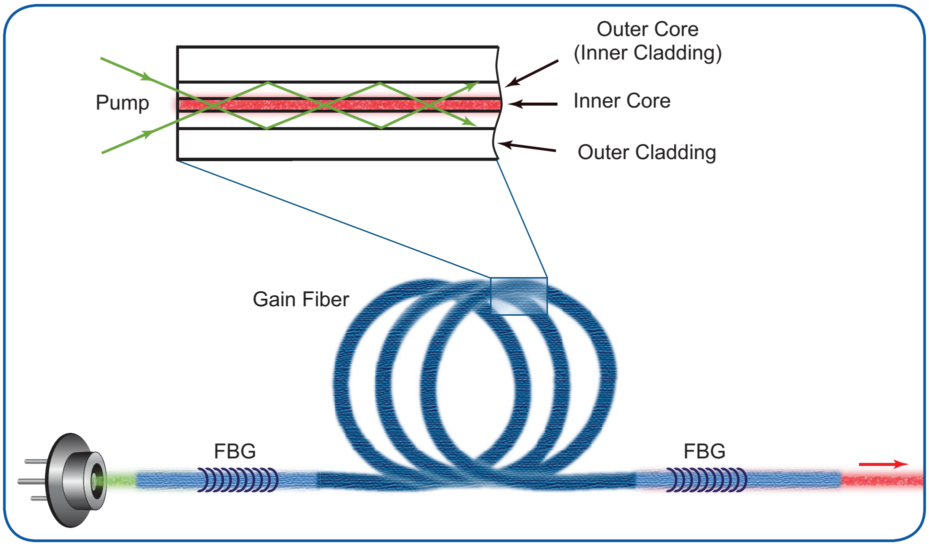


Figure 20. Simplified schematic of a laser-diode-pumped fiber laser with fiber Bragg gratings (FBGs) as reflectors [1]. Inset shows scheme for pumping via a double-clad fiber [11].

Fiber Lasers

When a solid-state gain medium is fabricated into an optical fiber and a resonator is integrated, a fiber laser is formed [1, 24]. This class of lasers is usually treated separately from the "bulk" solid-state lasers discussed above due to the unique light-guiding properties of fibers, e.g., strong spatial confinement over long distances; details given in Section III.B. A fiber laser is defined as a laser where the optical fiber is itself the gain medium which can be distinguished from having another type of laser or gain medium simply being coupled to an optical fiber. The laser-active ions are typically doped into the core of the optical fiber which is a glass host material. The most common dopant ions are the rare-earth elements of Yb, Er, and Tm whose laser gain bandwidths can be seen in Figure 19. Virtually all fiber lasers are optically pumped by laser diodes or other fiber lasers. In order to efficiently couple laser diodes with lower-quality spatial modes into the fibers, double-clad fibers are often implemented [25] as shown in Figure 20. There are different ways to integrate a resonator into a fiber laser geometry, including simply coupling the light out of the fiber and reflecting it off a normal cavity mirror, forming a fiber loop mirror, or through the use of fiber Bragg gratings (FBGs) (see Figure 20) which can be fabricated directly in a fiber (see Section III.D). Fiber lasers can achieve high output powers, operate in either pulsed or *CW* mode, generate high quality output beams, and can be operated over much of the NIR and MIR spectral region. This accounts for their use in a variety of applications including optical fiber communications, laser surgery, LIDAR and range-finding, and seeding of other more powerful lasers. Fiber lasers also dominate the high-power *CW* materials processing market.

Additional online content discussing the various types of lasers discussed above (except for semiconductor lasers which will be covered in Section I.B) can be found in [17, 24, 26-29].

5. Methods for Pulsed-Laser Operation

A laser operating in *CW* mode has constant output power as a function of time. This is the result of reaching a steady-state condition where balance is maintained between the laser cavity gain (dictated by the population inversion which is proportional to the pumping rate) and loss (which includes the cavity losses and the stimulated emission rate) [12, 13]. As discussed in Section I.A.3, it is often desirable to operate lasers in a pulsed mode since this can significantly increase the peak output power. Pulsed-laser operation can be achieved by sending a *CW* laser output through an external modulator that acts as a switch allowing light to transmit for only short time periods (see Figure 21, left). This simple method has many disadvantages. Since most of the light is blocked by the modulator, the method is very inefficient.

Furthermore, the peak power can never exceed the average power of the *CW* source. The duration of the pulse is also limited by the speed of the modulator. A more efficient method uses an *internal*, intracavity modulation process (see Figure 21, right) [1, 30]. By modulating the gain or loss in the cavity, the lasing process can be effectively switched on and off. Energy can be either (i) stored in the laser medium as a large population inversion that can be rapidly released to permit lasing or (ii) maintained in the resonator until it is allowed to escape. These methods permit pulsed-laser output with peak powers that significantly exceed those afforded by *CW* lasers. The most common methods for achieving laser pulses using internal modulation are discussed next.

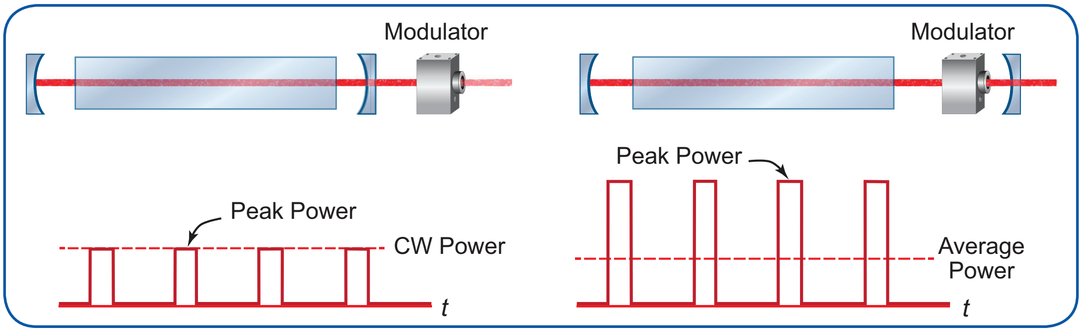


Figure 21. Comparison of pulsed laser outputs achievable with an external modulator (left) and an internal modulator (right) [1].

Gain Switching

For steady-state lasing, the gain remains at a threshold value because the stimulated emission depletes the population inversion as rapidly as it is produced by the pumping rate. However, if the gain medium is pumped much faster than the steady-state value, a transient effect called gain switching can occur [11, 14]. In this case, the population inversion (and hence the gain coefficient) builds up much faster than the rate of stimulated emission within the laser cavity. The photons in the cavity experience an enormous amount of gain leading to a rapid increase in laser intensity. This, in turn, causes significant stimulated emission that rapidly depletes the population inversion. The result is a short pulse of light (see Figure 22). Gain switching is a modulation approach where the gain is controlled by turning the pumping source on and off. This can be accomplished by flashlamp pumping, which yields pulses in the μs to ms range. However, gain switching is most commonly applied to semiconductor lasers since it is straightforward to modulate the electric current that is used for pumping. This can lead to pulses ranging from a few ns to tens of ps and with repetition rates up to several GHz. This method is often utilized to generate laser sources for optical telecommunications [30] where it is desirable to have high repetition rates to increase the amount of information transmitted per unit time.

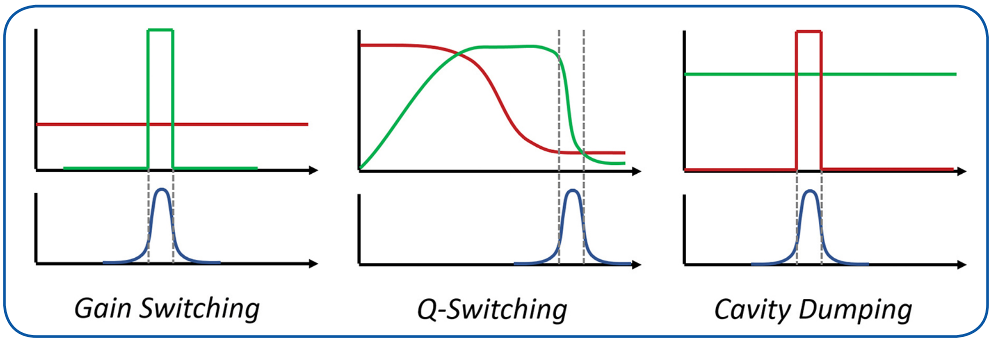


Figure 22. Schematic illustrating various modulation methods for producing pulsed lasers where loss (red), gain (green), and the laser output (blue) are shown as a function of time.

_Q_-Switching

Q-switching also involves storing energy in the laser gain medium but not by modulating the pump source. Here, the laser pumping process is allowed to build up a population inversion far in excess of the typical threshold value by ensuring that the cavity losses are large, which prevents lasing [11, 13]. Inhibiting the optical feedback is accomplished by adding a loss in the laser cavity. After a large inversion has been achieved, the cavity feedback is switched back on. The laser then experiences gain that greatly exceeds losses, and the stored energy is released as a short and intense light pulse (see Figure 22). The quality factor (*Q*) is the ratio of the energy stored in a cavity to its energy loss per cycle. Since this modulation approach involves switching the cavity *Q* from a low value to a high one, it is designated *Q*-switching. Devices used for *Q*-switching must be able to rapidly modulate the cavity *Q* to generate short pulses and are grouped into two categories, active and passive. Active devices require an external operation to induce modulation which include acousto-optical switches, electro-optical shutters, and rotating mirrors. Passive devices switch automatically based on the non-linear optical response of the element being used, e.g., saturable absorption in organic dyes or semiconductors. *Q*-switching results in ns laser pulses with very large pulse energies of mJ or greater. They often operate with repetition rates between a few Hz and many kHz [30].

Cavity Dumping

In contrast to the techniques above which store energy in the laser medium via a population inversion, cavity dumping stores energy in the photons within the resonator [1, 30]. The losses within the resonator are kept low for some time by keeping the cavity mirror transmittances negligible, effectively trapping the photons in the cavity and allowing an intense pulse to build up. This pulse is extracted by switching an intra-cavity element after one round trip and “dumping” the pulse out of the cavity (see Figure 22). The optical switch typically is an acousto-optic modulator or electro-optic shutter. One benefit to cavity dumping over *Q*-switching is that the latter requires that the pulse duration increase when the pulse repetition rate increases. However, cavity dumping allows for very high pulse repetition rates, e.g., several MHz, while maintaining pulse durations of a few ns. Cavity dumping can be combined with other pulse generation techniques in order to allow for extraction of higher pulse energies than would typically be available using other techniques.

Mode-Locking

The pulse-generation techniques described above produce pulses that are limited to a few ns. To achieve ultrafast pulses with durations down to a few fs, a technique known as mode-locking is utilized where the cavity losses are modulated periodically at the round-trip time of the laser pulse [1, 14, 15]. Unlike the other approaches which are based on transient effects within a laser cavity, mode-locking is a dynamic steady-state process. As shown in Figure 12, many longitudinal modes can participate in the lasing process and yet they are not necessarily in phase with one another upon arrival at a cavity mirror leading to random fluctuations in the output power. However, if these laser modes can be coupled together such that they are all brought into phase at the mirror, constructive and destructive interference can occur, resulting in the generation of an ultrashort pulse (see Figure 23). The coupling of these modes is achieved using a very fast intracavity shutter which operates at the intervals of the round trip of the laser pulse and effectively coordinates the time of arrival of these modes, thereby locking their phases. Like *Q*-switching, mode-locking devices can be active or passive [11]. Active devices that require external modulation include various amplitude and phase modulators. Passive devices rely on non-linear optical effects in suitable materials, including both slow and fast saturable absorption as well as intensity-dependent changes in the refractive index.

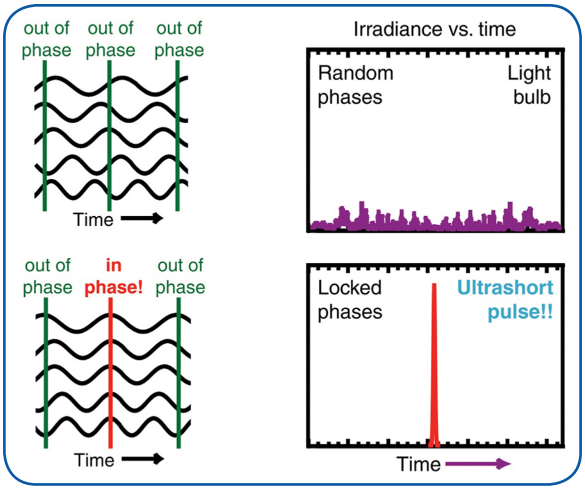


Figure 23. Locking the phases of the laser frequencies yields an ultrashort pulse.

Recall from Equation (6) that it is possible for all the longitudinal modes under the gain bandwidth to lase simultaneously, giving rise to a laser bandwidth of $N\Delta\nu$. If these modes are also locked together in phase, they interfere with each other to produce a series of intense pulses separated in time behaving like Fourier components of a periodic function [1, 15]. As shown in Figure 24, the separation of these pulses in time is simply the reciprocal of their frequency separation (i.e., $1/\Delta\nu$) which is approximately 1-10 ns based on typical cavity lengths. Furthermore, based on a time-bandwidth Fourier relationship, the pulse duration (τ) is proportional to the inverse of the laser bandwidth. Since the laser bandwidth is largely dictated by the gain bandwidth of the medium, the narrowest pulses are produced from lasers with the largest gain bandwidth. Mode-locked solid-state lasers typically possess pulse durations ranging from 30 fs to 30 ps, operate at large repetition rates e.g., MHz to GHz, and generate moderate pulse energies e.g., pJ to tens of nJ. Achieving larger pulse energies (e.g., up to a few mJ) is possible, at the expense of the repetition rate, e.g., typically kHz, by incorporating cavity dumping or regenerative amplifiers [30].

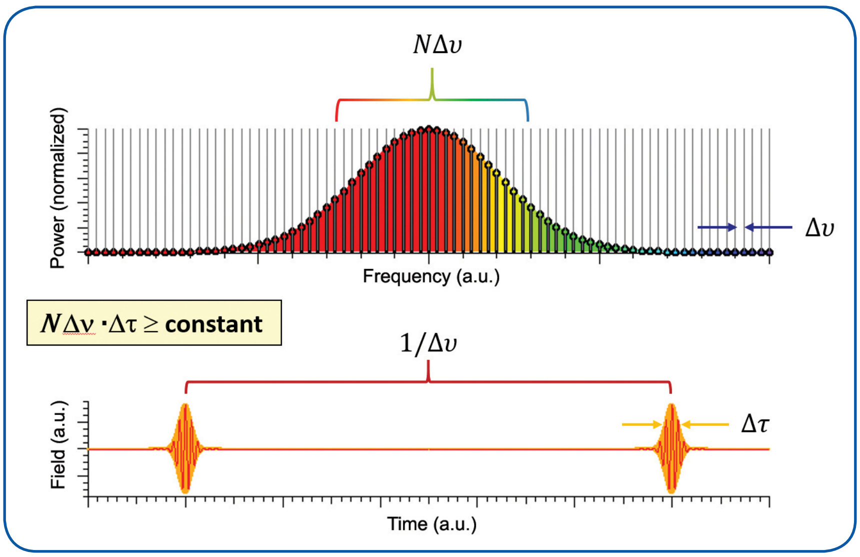


Figure 24. The frequency (top) and time (bottom) domain representation of a mode-locked laser with inset showing the time-bandwidth relationship.

Additional online content discussing the methods for generating pulsed lasers can be found in [30-35] (and references therein).

6. Spectral Tunability

As depicted in Figure 19, available laser gain media enable emission wavelengths to span a large portion of the electromagnetic spectrum. Despite this spectral coverage, applications that require access to a range of laser wavelengths typically do not utilize multiple laser systems due to cost and integration issues. Such applications rely on the output wavelength of a single laser system to be tuned over a spectral range. This is accomplished in two different ways. The first way to perform laser wavelength tuning uses a wavelength-selective element to choose a specific portion under a gain bandwidth to lase. This allows for tuning and comparable output characteristics, e.g., laser bandwidth, collimation, polarization, throughout the tuning range. A drawback is that emission wavelengths are constrained to the spectral bandwidth of the laser gain medium. The second method of wavelength tuning involves nonlinear frequency conversion. This approach introduces a nonlinear optical (NLO) crystal either outside or inside the laser cavity to generate new wavelengths or frequencies. This method allows for the generation of coherent light in spectral regions that range outside the laser gain bandwidth and can even extend to regions inaccessible to any laser system. The drawbacks include the need for high intensities to produce efficient output and a more complicated tuning method.

Laser Wavelength Tuning

Gain bandwidths for solid-state lasers can be very broad (see Figure 19). Introducing loss over a narrow and spectrally-adjustable region enables lasing to be restricted to a specific wavelength while also permitting tunability over the gain bandwidth [10, 11]. A number of methods can be utilized but all rely on reducing the optical losses at the desired wavelength. This can be accomplished by favoring the reflectivity of a specific wavelength off the end mirror in a resonator. The wavelength-dependent optical element can be a prism which refracts the light to the end mirror or a diffraction grating (see Section II.D.1), which also functions as the end mirror (see Figure 25). Rotation of either the prism or grating selects the wavelength. Alternatively, optical transmission can be maximized to reduce losses at a desired wavelength. This wavelength-adjustable transmission can utilize an intracavity etalon (see Figure 12, for example) or, more commonly a birefringent filter as shown in Figure 25 (see nonlinear frequency conversion section below for details). The filter consists of a single thin birefringent crystal (or several birefringent plates of increasing thickness) inclined at Brewster's angle (see Section III.A.4). The filter generates no reflection losses and therefore maximizes transmission for a specific wavelength based on the properties of the crystal and its angular orientation. Other wavelengths suffer losses at the filter and therefore are prevented from lasing. Wavelength tuning is achieved through rotation of the filter.

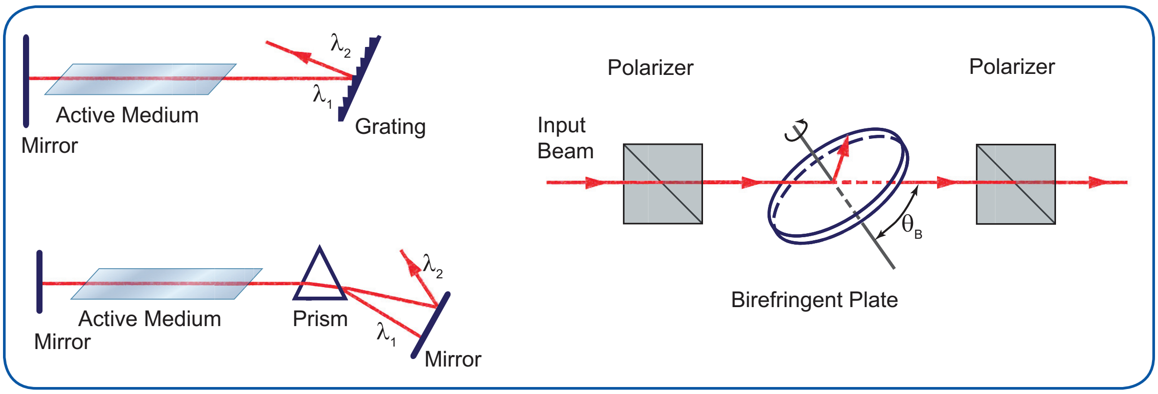


Figure 25. Laser tuning using the wavelength dispersive behavior of a diffraction grating or a prism (left). Use of a birefringent filter as a wavelength-selective element (right) [11].

Nonlinear Frequency Conversion

Nonlinear frequency conversion provides a means of extending the wavelength range of available laser sources. There are drawbacks to this approach, one of which is the requirement that the laser pump source must possess high intensity for efficient conversion. However, with the commercialization of high-power pulsed lasers, this requirement can now be readily met and nonlinear frequency conversion has become a viable approach for spectral tunability. Frequency conversion is an NLO phenomenon with the underlying light-matter interactions being quite complicated. Details of these interactions can be found in [36, 37]. Here, only a very basic description will be given of how an optical medium under intense optical fields can generate new frequencies.

When a weak optical field interacts with a material, it creates a polarization in the material made up of many electric dipoles. The strength of this polarization is *linearly* proportional to the magnitude of the optical field and the dipoles re-radiate light with the same frequency of the optical field. This is the regime of linear optics. When an intense optical field encounters a material, the polarization no longer responds linearly to the applied field. Instead, it becomes quadratically-dependent (or even higher order) on the field. This *nonlinear* dependence on the optical field is why intense light sources are required for nonlinear optics and why experimental observation of NLO phenomena did not occur until the laser was invented. Furthermore, dipoles can no longer reproduce the same frequency as the incident optical field. It is like music being played through an amplifier that cannot reliably reproduce the sound but instead distorts it. Thus, the distorted, re-radiated field contains additional frequencies compared to the original field. The nonlinear dependence of the material on the incident optical field results in a single frequency being converted into new frequencies.

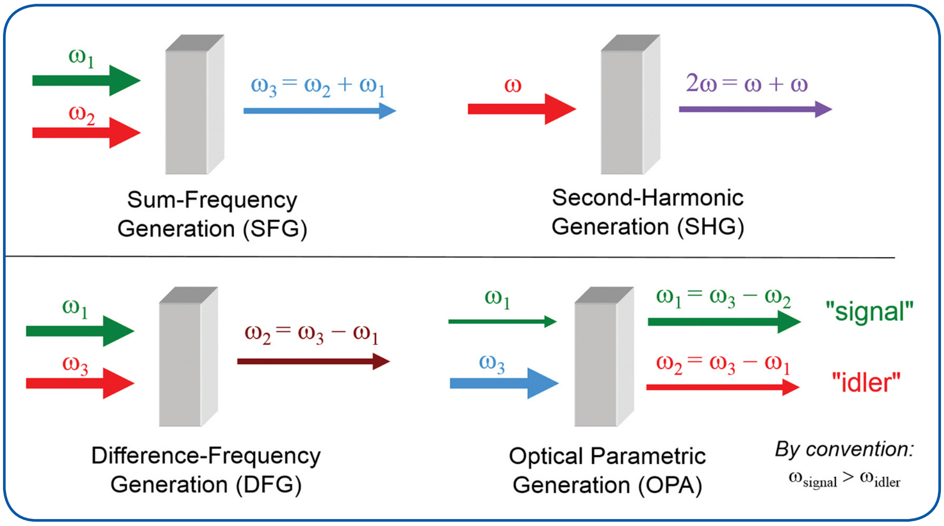


Figure 26. Three-wave mixing processes that allow for nonlinear frequency conversion. The colors of the frequencies generally describe their relative spectral positions, i.e., purple > blue > green > red > dark red, while the width of the arrows denotes the relative intensity of the beams.

For the case where the polarization is quadratically dependent on the field, these phenomena are known as second-order NLO effects. It is important to note that the two incident optical fields can be from the same laser source and thus possess the same frequency. However, in general, the two fields (sometimes referred to as waves) have different frequencies and, when they interact within the nonlinear medium, they produce a field with a new frequency. This process, known as three-wave mixing (depicted in Figure 26), can take on a variety of forms depending on the input frequencies and the desired output frequency. The upper row of Figure 26 shows the process of sum-frequency generation (SFG) where the input frequencies are added to produce higher frequency (shorter wavelength) light. Second harmonic generation (SHG) is a specific type of SFG that is commonly used since it allows for frequency-doubling of a single laser beam. By contrast, the lower row of Figure 26 shows difference frequency generation (DFG) where smaller frequency (longer wavelength) light is produced. Optical parametric amplification (OPA), a type of DFG, uses a pump beam to produce two waves of lower frequency that can span a wide spectral range. If the nonlinear medium that produces OPA is placed within a cavity, an optical parametric oscillator (OPO) can be formed, which effectively becomes an optically-pumped tunable source.

Once the conservation of energy condition is met (as shown by the frequency equations in Figure 26), a "phase-matching" condition must also be met to ensure efficient three-wave mixing. Phase-matching requires that the indices of refraction at the frequencies of the various waves be equal. This allows the three waves to be temporally and spatially overlapped over a long interaction length [2, 10]. An example of phase-matching is shown in Figure 27 for SHG. Due to dispersion, i.e., change in the refractive index with frequency, and the differences in frequencies of the waves, phase-matching cannot be achieved in a material with a single refractive index (see Figure 27, left). However, this can be accomplished by using a birefringent material, that is, a material that has different refractive indices for different input polarizations (see Figure 27, left). The polarizations for the waves must be orthogonal to one another and are referred to as the “ordinary” or *o*-wave and the “extraordinary” or *e*-wave, depending on the properties of the birefringent material. Birefringence naturally occur in certain crystalline materials and so NLO materials for three-wave mixing are typically crystals. The optimal phase-matching condition in an NLO crystal can be determined for a specific set of frequencies and polarizations. Wavelength (or frequency) tuning is then accomplished by either changing the orientation of the crystal through rotation or sometimes by adjusting its temperature in order to vary the index of refraction.

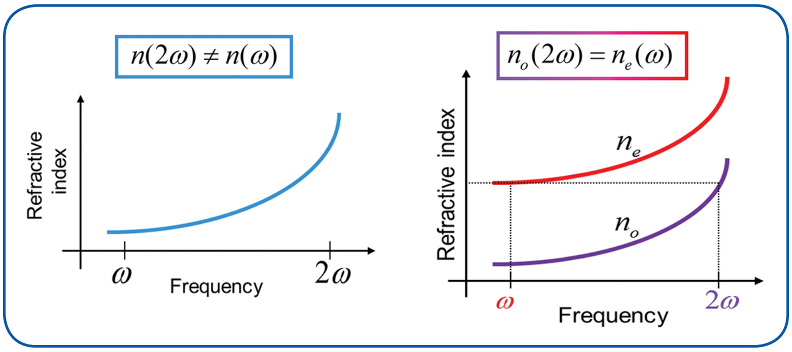


Figure 27. Dispersion and associated phase-matching conditions for a non-birefringent material (left) versus a birefringent crystal (right). The subscripts refer to the o-wave and e-wave.

In addition to the phase-matching considerations for an NLO crystal, the efficiency of the nonlinear frequency conversion depends on the NLO coefficient of the material similar to the role the transition cross-section plays in the efficiency of a laser gain material. In addition to a large NLO coefficient, other characteristics such as laser damage threshold (see Section I.A.7), chemical stability, and optical quality and transparency play a critical role in the choice of the optimal NLO material [10, 38]. Some commonly-used crystalline materials for nonlinear frequency conversion are lithium niobate ($LiNbO_3$), potassium titanyl phosphate (KTP), potassium dihydrogen phosphate (KDP), beta-barium borate (BBO), and lithium triborate (LBO). Achieving the litany of requirements for nonlinear frequency conversion can be challenging. This challenge includes material constraints of the NLO crystal, optimal phase-matching conditions, and large pump laser intensities. However, the capability for extremely wide spectral tunability when implementing nonlinear frequency conversion has made this approach attractive. Figure 28 shows an example where the proper choice of NLO crystals can enable wavelength-tuning that spans from the UV to the MIR with a single pump wavelength of 800 nm from a commercially-available amplified fs laser.

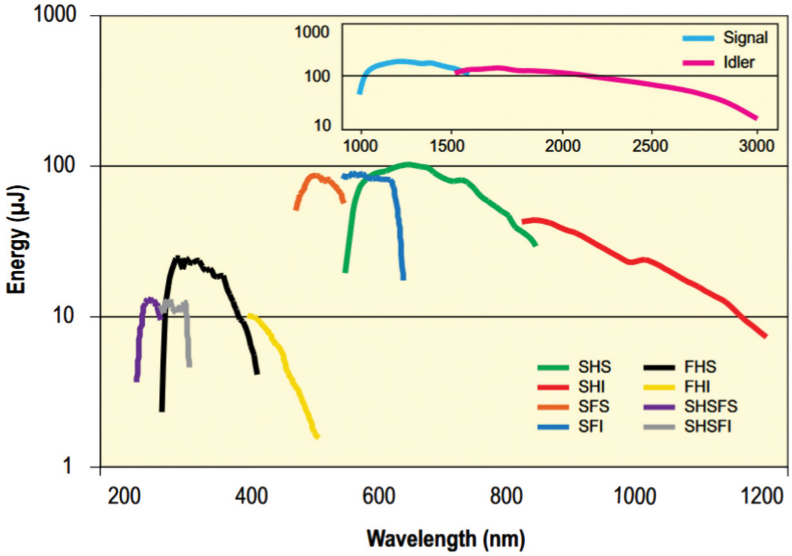


Figure 28. Wavelength tuning afforded by nonlinear frequency conversion when a single pump wavelength of 800 nm is used from an amplified fs laser source. SHS – second harmonic signal, SHI – second harmonic idler, SFS – sum frequency signal, SFI – sum frequency idler, FHS – fourth harmonic signal, FHI – fourth harmonic idler, SHSFS – second harmonic of sum frequency signal, SHSFI – second harmonic of sum frequency idler.

Additional online content discussing the methods for achieving spectral tunability with laser sources can be found in [39-41].

7. Classes of Lasers, Laser Safety, and Laser Damage

As detailed in Section I.A.3, lasers are distinguished from other light sources because lasers exhibit a high degree of collimation and have large optical intensities. Not only can these intensities remain large after long propagation distances, focusing elements can be used to drastically increase these intensities. While these unique attributes enable the myriad of laser-based applications, they also render laser radiation potentially hazardous to users. Such hazards come in the form of potential biological damage to the eye or skin, so laser safety is a critical aspect of laser usage. In this section, the classification of lasers based on their level of safety hazard is discussed along with general practices for maintaining laser safety. In addition, aspects of laser-induced damage to components within a laser system or an optical layout are also addressed.

Laser Classification

In order to give some guidance on proper handling and required safety precautions, lasers or laser systems are assigned to specific hazard classes [2, 42, 43]. Classification is based on the ability of the laser beam to cause biological damage to the eye or skin during use. Hazard level is determined based on exposure limits for the eye and must consider a variety of factors, including the distance of the user from the laser output, exposure duration, and possible use of optical instruments for viewing that could focus the radiation, e.g., binoculars, telescopes, microscopes. Furthermore, assignment to a laser safety class depends not only on the laser parameters, e.g., average or peak power, spatial beam size, lasing wavelength, but also on the accessibility of the hazardous area e.g., a high-power laser may be low-risk if the radiation is fully enclosed in a housing. The International Electrotechnical Commission (IEC) has published an internationally-accepted standard that outlines the safety of laser products (document 60825-1) and the American National Standards Institute (ANSI) has published a similar document (Z-136.1) that is primarily used in the United States. There are differences between the two laser standards/classifications (see [44], for details) but the general descriptions are similar. Table 3 describes these safety hazard classes along with some representative laser systems.

Class	Description	Examples
1	(safe) Laser radiation not dangerous under reasonable operating conditions including using optical instruments for viewing. Higher-power lasers can be included if radiation is fully-enclosed.	0.2-mW laser diode, enclosed 10-W Nd:YAG laser
1M	(low power) Same as Class 1 but with additional restriction that no optical instruments may be used which could focus radiation.	
2	(low power) Laser radiation limited to visible spectral range (400-700 nm) with maximum output power < 1 mW. Due to the blink response, not dangerous for eye in case of limited exposure (up to 0.25 s).	Alignment HeNe lasers, many laser pointers, supermarket scanners
2M	(low power) Same as Class 2 but with additional restriction that no optical instruments may be used. Power may be higher than 1 mW but beam diameter must be large enough to limit intensity to levels which are safe for short-time exposure.	
3R	(low power) Laser radiation may be dangerous for the eye with maximum output power < 5 mW (for visible radiation). Blink response of eye protects somewhat but use of optical instruments that could focus radiation is hazardous.	Some laser pointers
3B	(moderate power) Laser radiation dangerous for eye with direct viewing and for the skin under special conditions. Diffuse radiation, e.g., scattered from diffuse surface, is typically harmless. Maximum output power < 500 mW (for visible and non-visible raditaion).	100-mW CW frequency-doubled Nd:YAG laser
4	(high power) Laser radiation is very dangerous for the eye and for the skin. Even light from diffuse reflections may be hazardous for the eye. The radiation may cause fire or explosions.	40-W Q-switched Nd:YAG laser, 4-kW thin-disk laser in a non-encapsulated setup

Table 3. General description of international laser safety classes and examples of associated laser systems [42].

Laser Safety

Safety hazards related to laser usage are generally separated into radiation hazards (associated with interaction with the laser beam itself) and electrical hazards (those associated with the electronic equipment used to operate the laser). Since radiation hazards are unique to laser operation, those will be discussed here. While high-power lasers do have the capability of inflicting burns on the skin, the eye is the part of the body most vulnerable to radiation hazards [4, 10]. Eye injuries generally have more dire consequences, e.g., partial loss of vision, and can occur at lower laser powers. These low thresholds exist because the lens of the eye can focus collimated laser light to a tiny spot on the retina leading to much higher intensities than experienced at the pupil. To quantify how much exposure to laser light is hazardous, maximum permissible exposure (MPE) limits are tabulated in terms of the allowable exposure time for a given intensity. The MPE is dependent on the specific laser characteristics, e.g., wavelength, power, beam size, repetition rate, as well as the spectral transmission of the eye itself. The MPE values for various laser systems can be found in [4, 10].

Engineering Controls	Procedural Controls	Personal Protective Equipment
Protective housing and service panel	Laser safety officer	Eyewear
Door interlocks and remote control connector	Standard operating procedures	Clothing
Beam attenuators and beam shutters	Limitations on use by class	Gloves
Key switch or padlock	Entry limitations for visitors	
Warning lights, emission indicators	Education and training	
Beam enclosure	Maintenance and service manuals	
Controlled beam path	Marking of protective devices	
Laser controlled area	Warning signs and labels	
Beamstops		

Table 4. Control measures for Class 3b and Class 4 lasers [43].

The potential hazards associated with laser usage coupled with the ubiquity of laser applications require that laser safety control measures must be implemented in workplace and educational environments [2, 4, 10]. These control measures are summarized in Table 4. Engineering controls are specifically designed to restrict or reduce exposure to laser radiation. Obviously one should never look directly into a laser light source or at scattered laser light from any reflective surface, but misalignments leading to accidental reflections can and do occur during normal operation. Such controls include ensuring that the laser and associated optical components are situated below eye level, proper enclosures are provided for the beam, all unused beams are properly terminated with beam dumps or stops, and that laser interlocks are utilized. Procedural controls help guarantee that personnel involved with the use of hazardous laser systems are properly trained and familiar with safety protocols. Finally, the use of personal protection equipment (PPE) is critical since it functions as the final safeguard against injury. Proper clothing and gloves can ensure minimal skin exposure to laser radiation but protective eyewear or laser safety glasses or goggles are the most important type of PPE. Laser goggles should have the appropriate attenuation factor, e.g., transmission reduction, to ensure that the MPE is never exceeded given the maximum output power of the laser being used.

Laser Damage

The large intensities associated with laser output can lead to laser-induced damage of optical components [10, 45]. These can include components within a laser cavity itself, e.g., laser rod, resonator mirrors, modulators, the NLO crystals used for frequency conversion, any of the optical components (see Section III.A) that are used in optical systems, or power and energy sensors (see Section II.A). Optical damage can occur within the bulk of the material or at its surface but is more likely at the surface since there is a higher density of microscopic defects which can intensify the incident optical field. Optical components can also have anti-reflection coatings on them (see Section III.A.2) and so surface damage will often initiate here rather than at the surface of the bulk material. In any event, it is important to know the laser-induced damage threshold (LIDT) of a component since it can limit the overall operation of an optical system. The physical mechanisms that lead to laser-induced damage can be quite complicated and have dependencies on pulse duration and laser wavelength. Therefore, LIDT values are typically determined experimentally by evaluating the extent of damage induced at various intensity levels. The LIDT value for a material or component is defined at a specific wavelength or for a range of wavelengths. For *CW* lasers, this value is often given as a power density or intensity (in units of W/cm^2), whereas for pulsed lasers the value is usually given as an energy density or fluence (in units of J/cm^2). The quantities of fluence and intensity will be discussed in more detail in Section II.A.

Additional online content discussing laser safety and laser-induced damage can be found in [46-48].

B. Laser Diodes and LEDs

Semiconductor lasers, generally referred to as laser diodes, are one of the most important classes of lasers today. They are critical for optical communications and control applications and are finding widespread usage as pump sources for solid-state and fiber lasers. Laser diodes possess several unique attributes compared to other types of lasers. They are very small compared to other classes of lasers, can operate very efficiently with relatively low input powers, and are compatible with modern electronics. Many of these characteristics can be attributed to a gain medium that is electrically-pumped while also serving as the resonant cavity. Furthermore, laser diodes leverage mature semiconductor manufacturing processing for good quality control, large-scale production, and tailorability. Laser diodes are a sub-set of solid-state lasers. Therefore, many aspects of the light generation process and output characteristics discussed in the context of solid-state lasers remain valid. However, there are some fundamental differences for laser diodes in terms of the light-matter interactions and the optical emission properties. These differences are addressed in this section.

Light-emitting diodes (LEDs), like laser diodes, generate radiation via electrical current injection into a junction. LED light comes from spontaneous emission, whereas laser diode light arises from stimulated emission. Thus, LEDs generally have lower output powers and omnidirectional emission. LEDs are efficient and small like their laser counterparts but typically operate at lower drive currents and are less expensive. Consequently, they have found use in a myriad of everyday applications such as automotive and architectural lighting, remote controls, displays and detection. Despite their similarities, there are significant differences in the output characteristics of LEDs and laser diodes and these will be discussed in Section I.B.3.

1. Basic Semiconductor Physics

Recall from Section I.A.4 that the radiating species which make up most laser gain media are either molecules in a dilute gas or liquid, or ions doped in a crystalline or glass host. In a semiconductor, the constituent atoms form a crystalline or amorphous solid that becomes the gain medium. Laser diodes are typically made up of at least two different types of semiconductor materials, and the lasing action occurs at the interface between those two materials. This section gives a brief description of the physics of these interactions and how they enable the generation of radiation.

Semiconductor Materials

The atoms making up a semiconductor material have strong interatomic interactions which lead to the formation of valence and conduction bands as shown in Figure 3 and Figure 16. These bands are separated by a bandgap energy (E_g) which is analogous to the energy level separation shown in Figure 4. In semiconductors, an electron that is promoted to the conduction band from the valence band leaves behind an empty state known as a hole. When the electron returns from the conduction band to this empty state in the valence band, a process known as electron-hole recombination occurs. In analogy to Equation (1), this can result in a radiating photon with a wavelength corresponding to E_g. Clearly, the material and its associated bandgap play a critical role in determining the laser emission wavelength. However, the behavior of an electron in a periodic crystal lattice, i.e., a semiconductor material, requires that both energy conservation and momentum conservation be met during this recombination process [1, 12]. This is most often depicted in a diagram that shows the valence and conduction bands as a function of energy (E) and momentum (k). E-k diagrams for two semiconductors, silicon (Si) and gallium arsenide (GaAs), are shown in Figure 29. When the minima of the conduction and valence bands have the same momentum, as for GaAs, the semiconductor is a direct-gap material. Conversely, Si is an indirect-gap semiconductor where any transition between the bottom of the conduction band and the top of the valence band must include a phonon-assisted change in momentum. A consequence of this is that direct-gap semiconductors emit photons much more efficiently than indirect-gap semiconductors because non-radiative processes tend to dominate in the latter.

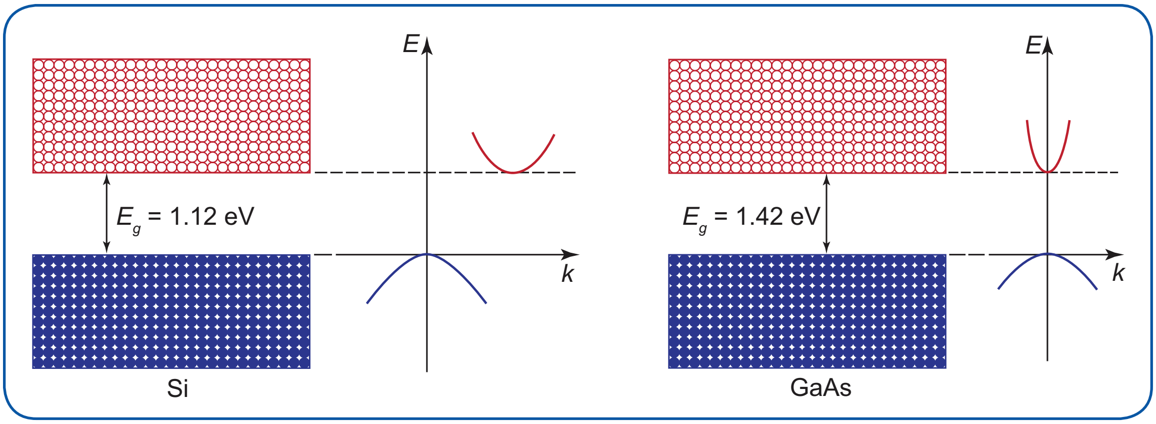


Figure 29. The E-k diagrams for Si, an indirect-bandgap semiconductor, and GaAs, a direct-bandgap semiconductor [1].

Laser diodes and LEDs almost exclusively use direct-gap semiconductors. Most direct-gap semiconductors are compound materials formed by combining elements from group III in the periodic table (Aluminum (Al), Gallium (Ga), Indium (In)) and group V (Nitrogen (N), Phosphorus (P), Arsenic (As), Antimony (Sb)) [1, 3, 11]. These III-V compound semiconductors can be either binary, ternary, or quaternary, depending on the number of elements involved. Changing the compositional mixing ratio of one of the elements creates more material flexibility with the concomitant increase in fabrication complexity. By virtue of this material flexibility, the bandgap energies and associated laser emission wavelengths can span a wide range of the electromagnetic spectrum as shown in Figure 30.

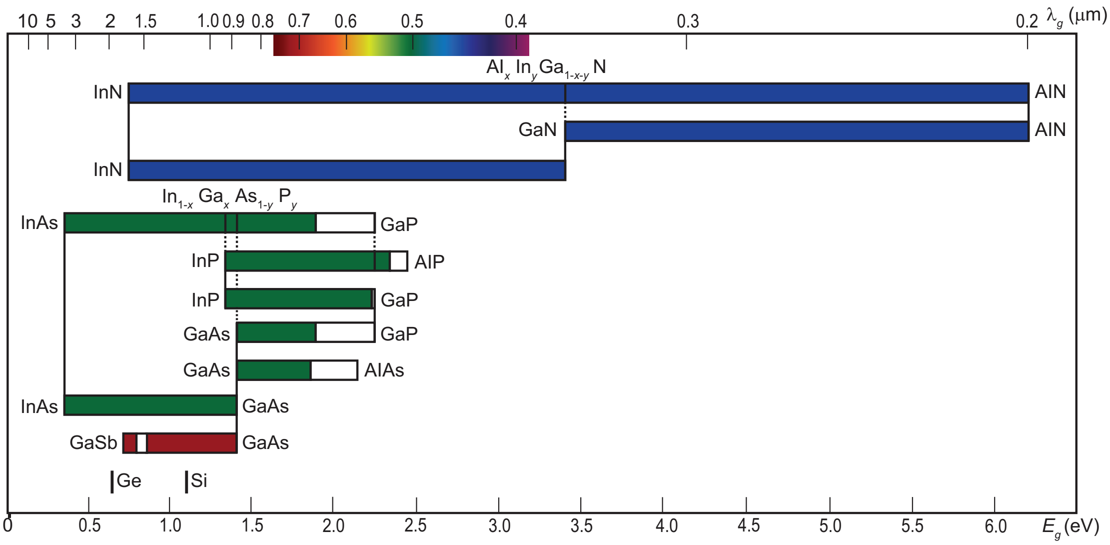


Figure 30. Bandgap wavelength and energy for selected elemental and III-V binary, ternary, and quaternary semiconductor materials. Successive rows, starting at the top, represent AlInGaN, AlGaN, InGaN, InGaAsP, AlInGaP, InGaP, GaAsP, AlGaAs, InGaAs, and GaAsSb. The shaded regions indicate compositions for which the materials are direct bandgap semiconductors [1].

The semiconductors discussed above are called intrinsic semiconductors. When impurity atoms known as dopants are added to intrinsic semiconductors, the doped materials are referred to as extrinsic semiconductors. The electrical and optical properties of extrinsic semiconductors can be significantly modified depending on the concentration of these dopants. Typically, the dopant atom is similar in size to the intrinsic semiconductor atom but has either one fewer or one greater electron in its valence shell. If the dopants have one greater electron, e.g., a dopant from group VI replacing a group V atom, the resulting material has a surplus of mobile electrons and is referred to as an *n*-type semiconductor. Conversely, doping with an atom having one fewer electron, e.g., a dopant from group II replacing a group III atom, results in an excess of holes and a *p*-type semiconductor is formed. *n*-type and *p*-type semiconductors are critical in the formation of the *p-n* junction (see below) which is the essential building block of an LED or laser diode.

As discussed in Section I.A.1, the probability of a radiation process occurring is dependent on the population in the participating energy levels. Therefore, the efficiency of electron-hole recombination depends on the concentration of carriers (electrons and holes) in the valence and conduction bands. Determining these concentrations requires knowledge of the density of states and their occupation probability, which can be a complicated calculation involving the *E-k* diagram of the particular semiconductor, the operating temperature, and the doping level [1, 12]. However, qualitatively, the distribution of carriers mimics that of the Boltzmann distribution (see Figure 6) in that at nominal operating temperatures, the valence band is nearly filled with electrons while the conduction band will be mostly empty. Consequently, in order achieve a significant population of electron-hole pairs for recombination, a pumping mechanism must be present.

***p-n* Junctions**

Increasing the number of electron-hole pairs in a semiconductor can be achieved through optical pumping as in other solid-state gain media. However, the most convenient pumping method is via electrical injection of charge carriers. This is accomplished by forming a *p-n* junction diode [1, 5, 11]. When *p*-type and *n*-type semiconductor materials are placed in physical contact, the area around the contact (known as the junction) behaves differently than either of the two source materials. The excess electrons and holes diffuse from their respective materials into the adjacent material and recombination occurs. A region on both sides of the junction becomes devoid of free carriers and is known as the depletion region (see Figure 31). An electric field is created across the depletion region by the fixed charges that are left behind following this carrier diffusion. This built-in field, which points from the *n*-side to the *p*-side of the junction, prevents further diffusion. This *p-n* junction is in a state of equilibrium with no current flowing across the diode. However, if an external potential is applied to the junction, the flow of carriers will be affected. If the junction is forward biased such that a positive potential is applied to the *p*-region, an electric field is produced that opposes the built-in field. Holes from the *p*-region are injected into the *n*-region while electrons are injected from the *n*-region to the *p*-region. These injected minority carriers recombine with the majority carriers in the destination region (see Figure 31). In effect, radiative recombination is achieved by electrically injecting charge carriers in the junction region. The junction can also be reverse biased, as shown in Figure 31, which is important in the operation of photodiodes (see Sections II.A.2 and II.C.1).

A homojunction is a *p-n* junction where both *p*-type and *n*-type regions are made of the same material. The first laser diode devices employed homojunctions but were inefficient. This was due to a relatively thick active region (i.e., ~ 1 μm) which is the region of the junction where radiative recombination occurs. The gain coefficient for a laser diode is proportional to the current density injected into the junction (see Section I.B.2). Therefore, reducing the thickness of the active region would allow a smaller volume to be pumped and allow comparable gain with a lower injected current density. Double heterostructure (DH) designs, where the active medium is sandwiched between *p* and *n* materials which are different from the active material, allow for reduction of the active region down to thicknesses of 0.1 μm. Threshold current densities can be reduced by nearly two orders of magnitude compared to homojunctions and have largely superseded homojunctions in designs for LEDs and laser diodes [1, 11]. DH devices achieve this thickness reduction by utilizing a narrower bandgap material for the active medium compared to the sandwiched materials. This provides energy barriers at the two junctions and forces injected electrons and holes to occupy a narrow active region. In addition to this carrier confinement, light confinement can also be achieved if the refractive index of the active layer is larger than the cladding layers. The layer acts as an optical waveguide (see Section III.D), ensuring that photons are confined to the region where the gain exists.

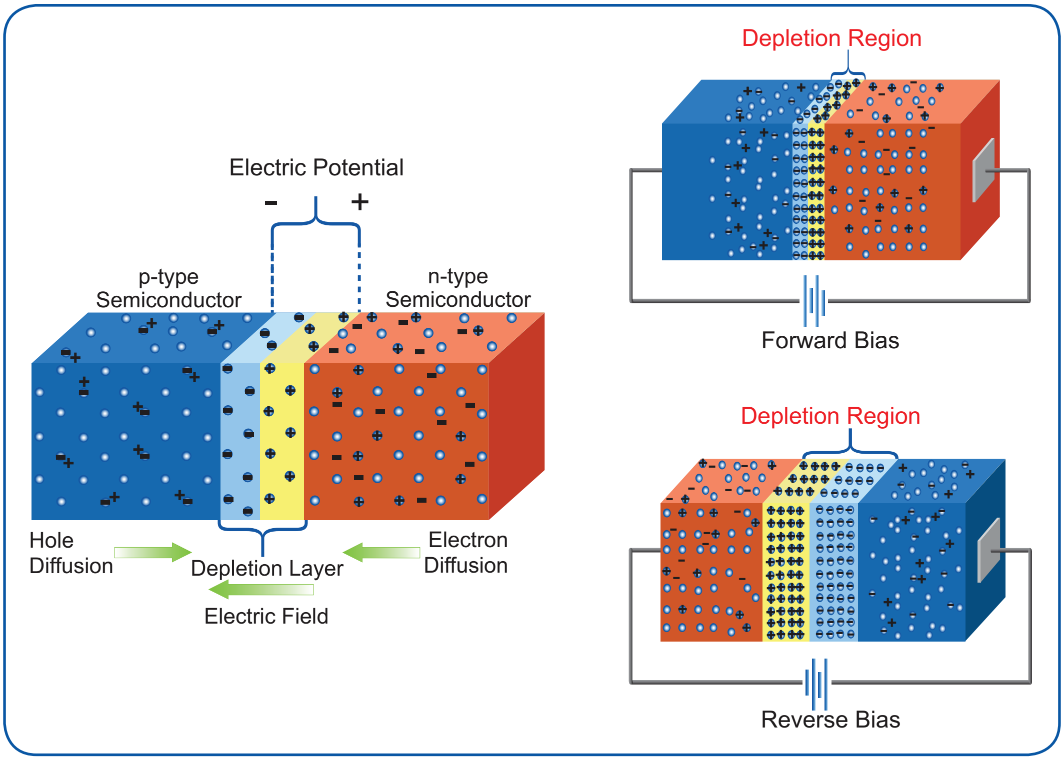


Figure 31. The p-n junction (left) and current flow characteristics of a p-n junction diode (right).

Quantum-Wells

DH devices with active layers that are reduced to thicknesses below 10 nm are referred to as quantum-well (QW) devices [1, 12, 49]. By further exploiting the relationship between threshold current density and active layer thickness, QW devices offer excellent performance and are commonly used in laser diodes and LEDs. In conventional DH devices, the active region is large enough that it acts as a bulk material where the conduction band and valence band are continuous. Since the active layer in a QW is smaller than the de Broglie wavelength for a thermalized electron, quantum effects become important [1]. This causes the valence and conduction bands to become step-function bands with discrete energies. In addition to lower threshold currents compared to their bulk DH counterparts, QW devices also benefit from narrower laser linewidths, increased modulation speeds, and a reduced dependence on temperature. There are a variety of device designs utilizing QWs, including multi-QWs, strained-layer QWs, as well as further spatial confinement reduction in the form of quantum wires and quantum dots.

Additional online content discussing the basic physics behind semiconductor materials used for laser diodes and LEDs can be found in [50-56].

2. Optical Emission

Fundamental radiation processes that occur within semiconductor materials [1, 14] are shown in Figure 32. These processes are identical to the photon-atom interactions described in Figure 4. Figure 32 also shows the response of a biased *p-n* junction diode for each process. Absorption produces electron-hole pairs and is the fundamental process exploited in semiconductor photodiodes where the junction is typically reverse biased (see Section II.A.2). In contrast, electroluminescence (or electron-hole recombination radiation following electrical injection) occurs when a *p-n* junction is forward biased. This radiation is equivalent to spontaneous emission and leads to optical emission in an LED. This recombination radiation can also generate stimulated emission and, if the gain produced in the junction can overcome the optical losses (including absorption), lasing will occur in a laser diode. The evolution of this emission as a function of injection current, the mechanisms for efficient light extraction, and the typical device geometries employed for both LEDs and laser diodes are the subjects of this section.

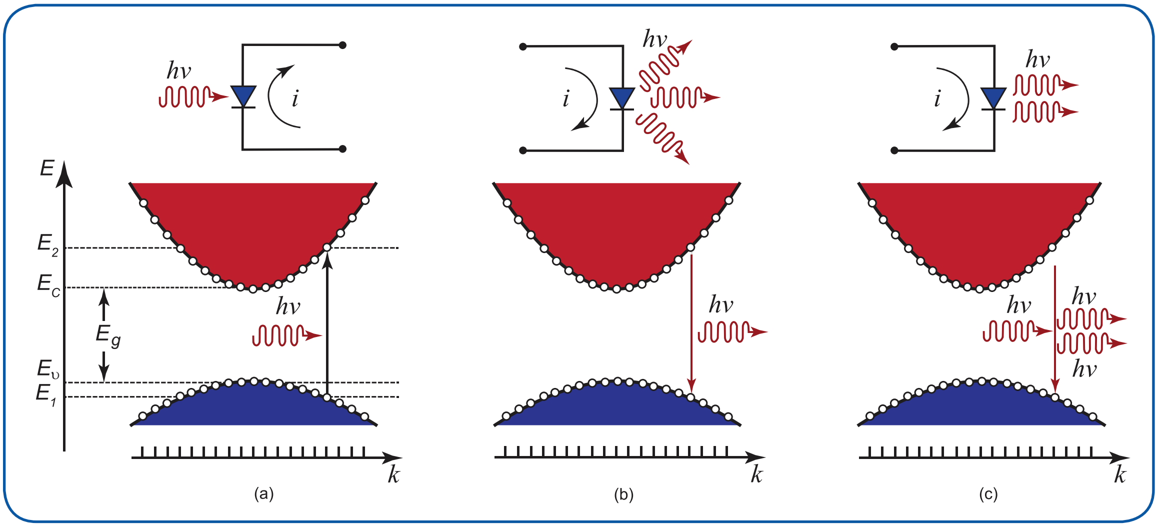


Figure 32. Fundamental radiation processes that occur within a semiconductor p-n junction. Generation of an electron-hole pair by absorption (a) governs the operation of a photodiode. Electron-hole recombination processes give rise to spontaneous emission (b) in an LED and stimulated emission (c) in a laser diode [1].

Light-Current Curve

Electrical injection in the *p-n* junction of a laser diode provides the population inversion necessary to produce lasing. Like other laser systems, once this population is large enough so that the gain exceeds the losses, stimulated emission dominates and the system reaches a steady-state where the output light increases linearly with injection rate. For a semiconductor laser diode with an operational wavelength of λ_0 (in µm), this can be expressed by a simple equation [1, 14, 49] which gives the steady-state output power (P_o in W) as a function of the current applied to the diode (i, in A):

$$P_o = \Re_d(i - i_t) \quad where \quad \Re_d = \eta_d \frac{1.24}{\lambda_0} \equiv \eta_e \eta_i \frac{1.24}{\lambda_0}. \tag{7}$$

This relationship is observed in the form of a light-current (*L-I*) curve in Figure 33. The current value at which the line intersects the x-axis and produces zero emission is the threshold current for lasing (i_t). The slope of this curve above threshold is known as the differential responsivity ($\Re_d$) and is typically expressed in units of W/A or mW/mA. This differential responsivity is determined by the external differential quantum efficiency (η_d), which essentially describes how efficient the laser system is in converting injected electrons into output photons. The efficiency is made up of an internal quantum efficiency (η_i), which is the ratio of generated photons (inside the junction) to injected electrons, and an extraction efficiency (η_e), which is associated with external losses (see section on extraction of light below). Equation (7) can be put in context with the generalized discussion of laser gain from Section I.A.2 where gain was given as the product of the population difference (ΔN) and the transition cross section (σ). ΔN is essentially i divided by the volume of the junction (V) while η_i plays the role of the σ. η_e accounts for the losses associated with the resonator cavity. Interestingly, i/V can also be expressed as a current density J divided by the thickness of the active region (l). In Section I.B.1, this reciprocal relationship between J and l was used to explain the benefit of moving from homojunction lasers to DH and QW devices.

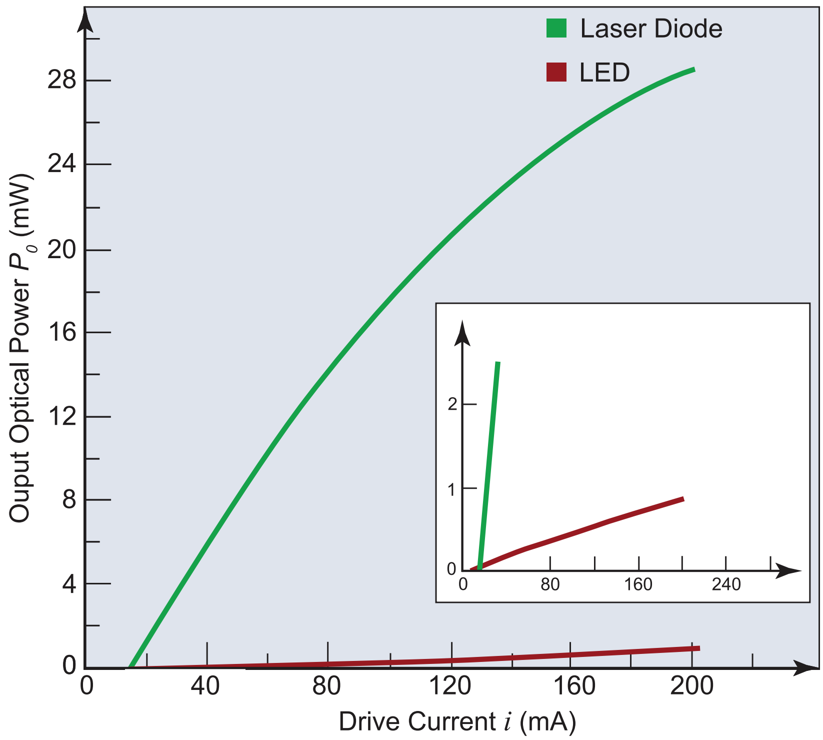


Figure 33. Representative L-I curves for an LED and laser diode operating at a wavelength of 1.6 µm consisting of InGaAsP / InP multiple QW structures. The inset provides an expanded view [1].

Equation (7) can also be used to describe the LED output power by setting $i_t = 0$ (since there is no threshold for spontaneous emission). $\Re_d$ and η_d are replaced by the responsivity ($\Re$) and the external efficiency (η_{ex}). Figure 33 shows a *L-I* curve for an LED and laser diode. Despite having the same architectures, the slope of the laser diode curve is significantly higher, indicating a larger value of $\Re_d$ compared to $\Re$. As detailed below, this superior performance is the result of a greater extraction efficiency for the laser diode. A laser diode exhibits the same response as an LED when it is operating below threshold; that is, it emits light spontaneously that propagates in all directions. However, when the laser reaches threshold, stimulated emission concentrates the light into a select number of spatial modes, giving rise to the sharp increase in the observed output power.

Extraction of Light

The overall efficiency of an LED (η_{ex}) is given by the product of η_i and η_e. While this device may be highly efficient at generating light within the semiconductor material, i.e., η_i approaching unity, the usefulness of the light source is predicated on efficiently extracting light from within the junction to outside the device [1, 57]. Spontaneous emission is isotropic, which implies that the photons generated in the junction of an LED are radiated uniformly in all directions. However, as shown in Figure 34, only a fraction of this light will emerge from an LED with a planar surface. In addition to losses associated with absorption in the material, the large refractive index of the semiconductor (typically between 2.9 and 3.6) gives rise to losses from Fresnel reflections (see Section III.A.1) and from total internal reflection (TIR, see Section III.D). The latter is particularly detrimental since any light outside of a cone defined by TIR (called the extraction cone) is not transmitted. Depending on the device design and junction geometry, virtually all the light can be either reflected or absorbed, yielding an extraction efficiency of less than 2%. Numerous mitigation approaches are used to enhance η_e. One approach employs an epoxy material formed into a parabolic shape (illustrated in Figure 34), which exploits index-matching and geometry to reduce the impact of reflection losses. Transparent contacts and substrates are used to reduce absorption losses while imparting a texture to a surface can minimize impacts due to TIR. Another way to increase η_e is by incorporating a microresonator that confines light within a very small spatial region. This also creates angular confinement of the light such that it lies within the extraction cone.

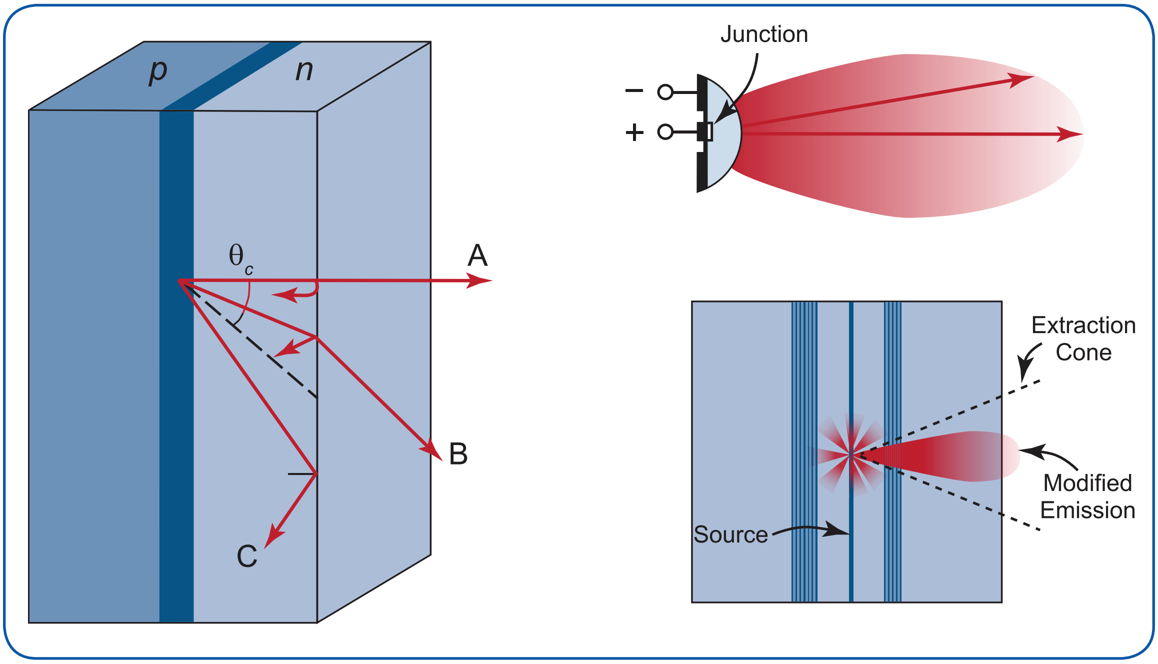


Figure 34. Optical emission from a surface-emitting LED (left) where light rays experience absorption and reflection (both A and B) while those outside of the critical angle undergo TIR (C) and are not extracted. Two methods for increased light extraction are shown on the right, including the use of a parabolic lens made of epoxy (upper right) and a microresonator that narrows the confinement into the light extraction cone (lower right) [1].

In Section I.A.2, the laser cavity end mirrors were a source of transmission loss. Since this loss represents useful laser light, the extraction efficiency in a laser diode is defined by the ratio of this loss to the total resonator loss [1, 57]. In this way, if the only losses are associated with laser light exiting the resonator mirror, the system would be considered highly efficient. Other loss sources include optical scattering from any element within the cavity, absorption due to impurities or charge carriers, as well as optical energy that lies outside the active region. Methods for addressing this optical confinement in different laser diode geometries are discussed below.

Edge-Emitting and Surface-Emitting

Semiconductor-based LEDs and laser diodes generally have two emission configurations: edge-emitting, where light exits from the edge of the active region of the device, or surface-emitting, where emission occurs through a face parallel to the plane of the active region [1, 11, 14]. This section will focus on the configurations for laser diodes (see Figure 35). The objective of laser diode design is to minimize the threshold injection current, which is accomplished by reducing the dimensions of the active region. In the direction of the layers, this is achieved using DH or QW architectures as discussed in Section I.B.1. The means for lateral confinement (in the plane of the junction) as well as the method for creating the resonator cavity are specific to the emission configuration. Details regarding edge-emitting lasers will be discussed first. These lasers produce a beam in the plane of the junction region and two methods are used to determine the lateral confinement of the gain. The first method, referred to as gain-guiding, uses fabrication techniques to limit the lateral extent of the gain, e.g., a narrow electrical contact that minimizes the region where current flows. The second method uses index-guiding, where a change in the index of refraction in the lateral direction creates optical confinement in the form of a waveguide (see Section III.D). The laser cavity in an edge-emitting laser can be produced in one of two ways. The first way is by cleaving the semiconductor crystal normal to the plane of the junction. The large refractive index of the semiconductor allows these surfaces to function as mirrors with sufficient reflectivity (~ 30%) to create an etalon. In this way, the length of the active region of the *p-n* junction becomes the resonator length. Figure 35 shows an edge-emitting laser diode with a DH architecture that utilizes gain-guiding for lateral confinement and an etalon structure for feedback. The second type of cavity uses fabrication techniques to produce distributed feedback (DFB) within the crystal or to create distributed Bragg reflector (DBR) structures at the ends of the crystal. Since these approaches typically produce single-frequency output, they will be discussed in Section I.B.3.

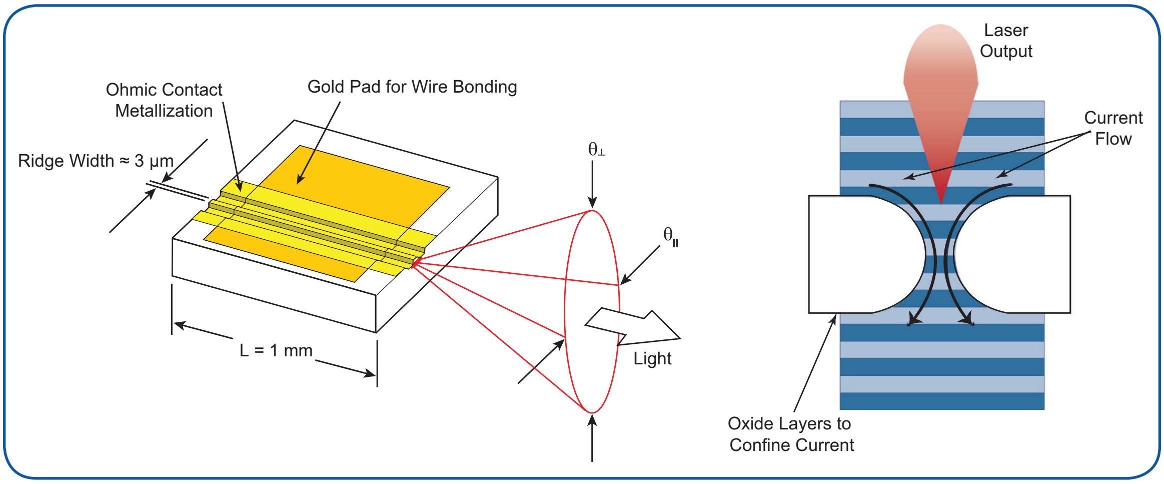


Figure 35. Typical structure and optical emission from an edge-emitting laser diode (left) and a surface-emitting laser diode (right, [14]).

Surface-emitting lasers produce a beam in a direction perpendicular to the junction region. A commonly-employed structure is known as a vertical cavity surface emitting laser (VCSEL). Here, mirrors in the form of DBRs clad the active layer which results in a vertical cavity giving rise to a beam that propagates normal to the junction plane. Since the length of the gain medium is defined by the thickness of the active region, much smaller gains are involved compared to edge-emitting lasers. As a result, mirror reflectivities must be very high. Lateral confinement of the gain region is accomplished using dielectric materials which narrow the cross-sectional area of current flow (see Figure 35). The lower optical gain for VCSELs generally results in lower output powers compared to edge-emitting lasers. However, their monolithic construction allows for high packing density of two-dimensional arrays and provides important practical advantages over their edge-emitting counterparts.

Additional online content discussing optical emission in both edge- and surface-emitting laser diodes and LEDs can be found in [53, 58-63].

3. Output Characteristics

Semiconductor lasers offer the benefits of efficiency, electrical pumping, and the ability to easily integrate with modern electronics which contribute to their ubiquitous usage. However, some of their output characteristics, particularly in terms of output power and beam quality, are considerably less than what can be achieved with other types of solid-state lasers [12]. For instance, the small and asymmetric junction region from which radiation is emitted leads to a laser beam that does not possess an ideal TEM_{00} mode and is highly divergent, spreading unequally in two directions. The lasing bandwidth is large enough that multimode operation is likely and the spectral purity is not as good as most solid-state lasers. Furthermore, it can be difficult to achieve high output powers with single-mode operation due to the possibility of optical damage and thermal management issues. Numerous approaches have been developed to mitigate these issues, leading to more desirable output characteristics for laser diodes. These topics, along with the differences between laser diodes and LEDs, are discussed in this section.

Output Power

Due to the large number of semiconductor materials and device architectures, the range of output powers for LEDs and laser diodes can be quite large. However, estimates of typical output powers can be gauged based on the product of the device responsivity and the nominal operating current. For low-power devices, laser diodes typically operate with a few 10's of mA to over 100 mA, whereas LEDs generally operate in the few mA to 20 mA range. For high-power devices, multimode laser diodes can operate with > 10 A, while high-power LEDs can easily exceed currents of 20 mA. The values of $\Re_d$ for a laser diode often fall in the range of 0.2 – 1 mW/mA (or W/A). Recall from Section I.B.2 that an LED has a lower extraction efficiency than a laser diode; therefore an LED's value of $\Re$ is necessarily smaller than that $\Re_d$. One final parameter that is often quoted for LEDs and laser diodes is the power-conversion efficiency (η_c) or wall-plug efficiency. η_c is the ratio of the emitted output power (given by Equation (7)) to the electrical input power. Laser diodes can readily exhibit values of η_c that exceed 50%, which is significantly greater than most other types of lasers [1].

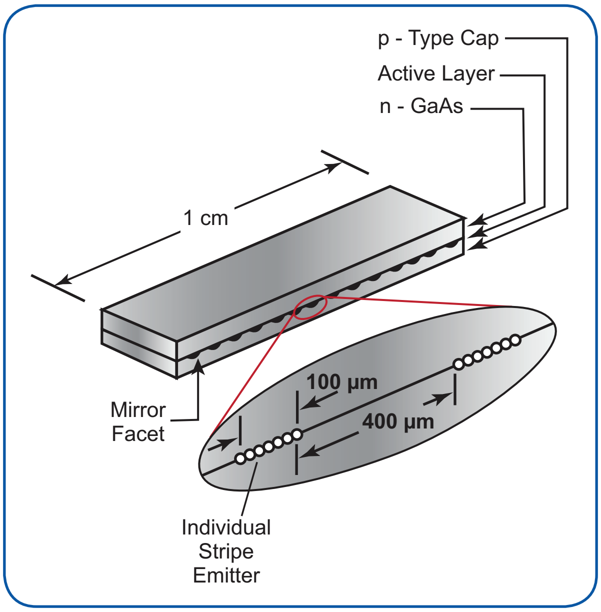


Figure 36. High-power laser diode bar consisting of serially repeated diode arrays.

There are many applications that require high laser output powers (> 10 W), including the optical pumping of solid-state lasers. The attributes of laser diodes make them intriguing candidates for such applications, but single devices cannot generate these types of powers [14, 49]. This limitation stems from the issues that arise when one tries to increase the gain per unit volume of a single device by increasing the width of the emission region and the injection current. Increasing the width leads to multiple transverse modes (see spatial profiles section below) which causes poor output beam quality, while large currents can lead to thermal loads in the active region that limit output power. Large output power densities at the cleaved facets can lead to optically-induced damage. One method to overcome these issues is to optically combine high-power single devices to achieve a high brightness output, which can be coupled into fibers for pumping DPSS or fiber lasers. A related method is to utilize an array of semiconductor laser diodes, which are fabricated adjacent to one another (see Figure 36). This method is feasible due to the large power-conversion efficiencies of the individual laser diodes. Depending on the spacing between the lasers, the combined output can either be coherent or incoherent with the latter generating larger output powers. These diode arrays can also be combined with one another to form two-dimensional arrays or a longer single array, typically referred to as a diode bar owing to its elongated rectangular structure (see Figure 36). These diode bars are ideal for DPSS lasers and can readily generate output powers in excess of 50 W.

Spatial Profiles

Spontaneous emission from an LED is isotropic but, when this radiation emanates from a planar surface (such as that shown in Figure 34), it is Lambertian [64]. This means that the intensity decreases with the cosine of the angle from the plane-normal, effectively diminishing as one goes off-axis. As discussed in Section I.B.2, the radial extent of the beam depends on the extraction cone. Describing the spatial profile of the beam emitted from laser diode is more complicated owing to the fact that the transverse modes are determined by the dielectric waveguide that makes up the active region of the diode [1, 12, 14]. As shown in Figure 35, this junction region for an edge-emitting laser resembles a stripe whose width is much larger than the depth. The depth (*l*) is the region over which carrier recombination occurs and, by virtue of the DH architecture, is typically much smaller than the laser wavelength (λ_0). This ensures that only a single mode will be emitted in the direction perpendicular to the junction plane (for more information on optical waveguides see Section III.D). The width (*w*) of the active region is typically larger than λ_0 and so multiple so-called lateral modes can exist in the direction parallel to the plane (see Figure 37). The width can be reduced to produce a single spatial mode using the gain-guiding and index-guiding methods discussed in Section I.B.2. Additional methods for ensuring single transverse mode output involve the use of an external cavity or appropriate anti-reflection coatings on the crystal facets.

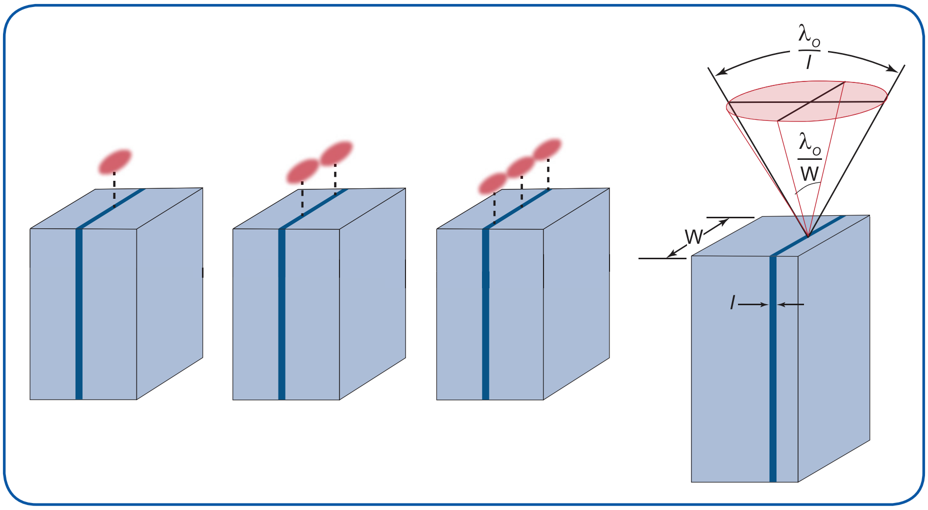


Figure 37. Illustration of spatial distributions consisting of multiple transverse modes for an edge-emitting laser diode (left) and the angular distribution for a single transverse mode (right) [1].

Even if a single transverse mode is achieved from the output of the laser diode, the divergence of this mode is significantly larger than nearly any other type of lasers [1, 11]. This divergence is the result of diffraction (since the cross-section of the active region is on the order of λ_0) and a short cavity length, which prevents the high degree of collimation typically enjoyed by other laser systems. The angular divergence of the beam is proportional to the ratio of λ_0 to the dimension of the active region (*l*). Given the small values of *l*, this can easily give rise to divergence angles exceeding 25° in the direction perpendicular to the junction (most solid-state lasers have angles much less than 1°). Since *w* is typically several times larger than *l*, the divergence in the direction parallel to the junction plane is significantly lower. This unequal divergence in orthogonal directions is called astigmatism and can make collimating such a beam quite difficult. However, optical systems have been developed to compensate for this astigmatic behavior (see Section III.A.3).

Spectral Characteristics

The spectral distribution from an LED is centered at the transition wavelength (λ_g) associated with E_g (see Section I.B.1). For a laser diode, the center wavelength (λ_0) typically occurs at λ_g as well since the gain bandwidth follows the spontaneous emission distribution. However, if a frequency-selective approach is used to isolate a single longitudinal mode (see below), continuous tuning of λ_0 under the gain bandwidth is possible. Finally, the carrier concentrations in the valence and conduction bands are dependent on the temperature of the semiconductor and the injection current (see Section I.B.1). Since the carrier concentration plays a role in determining the effective value of E_g and therefore λ_g, the emission wavelength can be shifted with either drive current or junction temperature [65]. These effects are described in more detail in Section I.B.4.

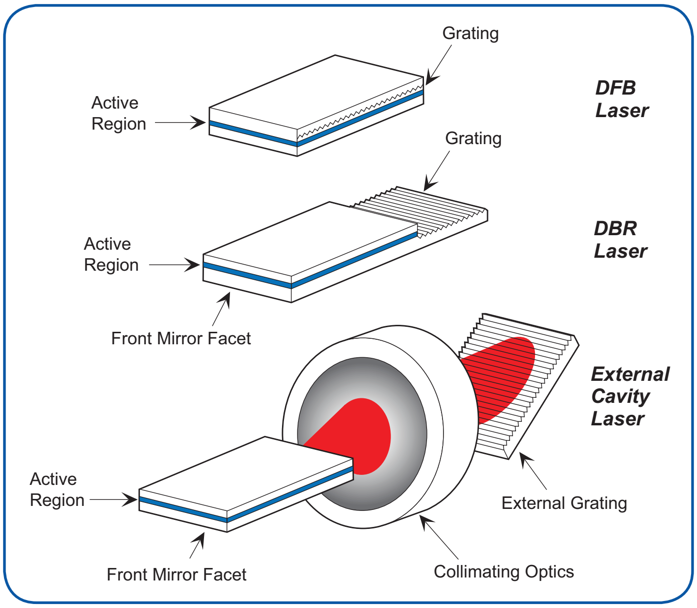


Figure 38. Frequency-selective feedback methods for generating single-longitudinal mode operation in a laser diode.

The width of an LED's spectral distribution is proportional to λ_g^2 [1], which leads to emission bandwidths in the VIS portion of the spectrum of ~10 nm while NIR emission bandwidths can approach 100 nm. As with other lasers, the spectral bandwidth of a laser diode depends on the overlap between the laser gain bandwidth and the laser cavity properties, which determine the longitudinal mode spacing (see Figure 12) [11, 14, 49]. As expected for a solid-state medium, the gain bandwidth of a semiconductor laser is relatively large due to the band-to-band transitions. The bandwidths are typically a few THz (corresponding to ~10 nm in the NIR), which are larger than most lasers but still smaller than many solid-state systems (see Figure 19). The cavities for most laser diodes are a few mm's in length and so the longitudinal mode separation is 50-100 GHz per Equation (5). This is much greater than most other types of lasers and implies that far fewer longitudinal modes will lase under the gain bandwidth. Since frequency-selective techniques have few modes among which to discriminate, narrow-frequency output (< 10 MHz) is more readily achievable by forcing operation in a single longitudinal mode. The four main frequency-selective approaches for accomplishing this are coupled cavity, frequency-selective feedback, injection locking, and geometry control. The geometry control approach simply increases the longitudinal mode spacing such that only a single mode resides under the gain bandwidth. The coupled cavity approach amounts to using an intracavity etalon which was illustrated previously in Figure 12. Injection locking uses a narrow-frequency laser whose output will be preferentially amplified in the semiconductor gain medium over the spontaneous emission spectrum. Frequency-selective feedback is the most common method (see Figure 38). It involves either the use of a grating external to the laser or a internally-fabricated grating that makes use of DFB or

DBR structures. DFB involves a periodic variation in either the gain or the index of refraction of the medium, which provides both the feedback and frequency-selection for the lasing process. DBR structures allow the periodic index structure to be present at the two cavity ends, and constructive interference occurs at a specific frequency.

Additional online content discussing the output characteristics of laser diodes and LEDs can be found in [51, 53, 55, 58, 63].

4. Drivers and Temperature Controllers

To assess the quality, performance, and characteristics of laser diodes, manufacturers often perform exhaustive testing which requires electro-optical, spectral and spatial characterization of the laser output [63, 66, 67]. As discussed in the previous sections, a laser diode's output is dependent on its injection current and temperature. Therefore, tightly controlling these parameters using current and temperature controllers is critical for extracting important operational parameters. An example of a laser diode test and characterization set-up is shown in Figure 39. Furthermore, laser diodes are expensive and have delicate electronic loads requiring controllers to be capable of protecting these devices while ensuring their output is stable. In general, the term "driver" typically refers to a current source while a "controller" often refers to an all-in-one current source and temperature-control mechanism. Current drivers and temperature controllers are discussed separately below. Chapter 5 – Photonics in Communications explores laser diode testing and characterization in more detail.

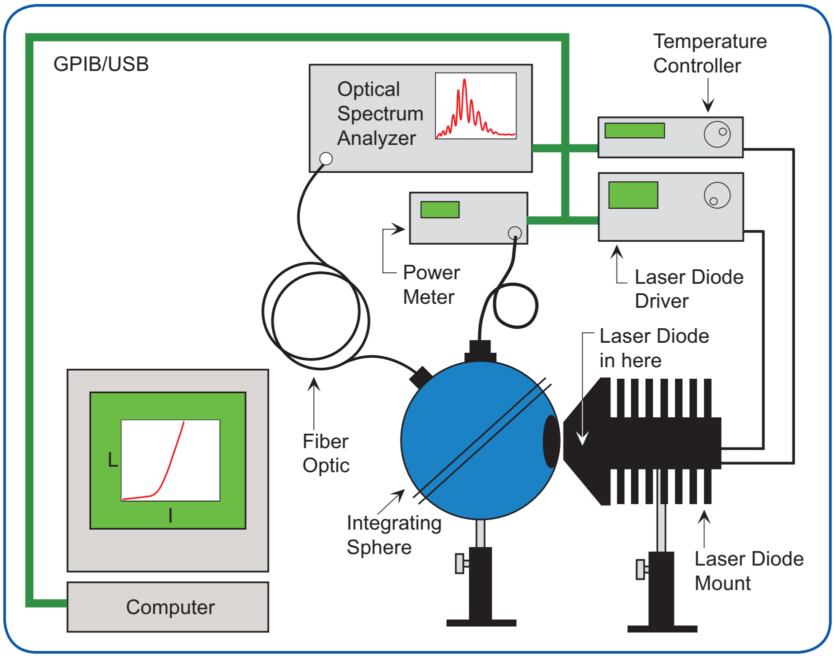


Figure 39. A typical computer-controlled laser diode test and characterization setup.

Current Drivers

The most important laser diode characteristic is how its light output power (*L*) responds to injected current (*I*). This is referred to as the *L-I* curve (see Figure 40). This curve can be used to determine a number of significant parameters that were discussed in Section I.B.2, including threshold current and threshold current density, differential responsivity, internal quantum efficiency, and external differential quantum efficiency [65, 68]. These values are usually listed in a laser diode's specification sheet so that a user can determine important operational parameters such as the current at which lasing begins, the drive current for a specific laser power, as well as the maximum current the device can take. Finally, as discussed in Section I.B.3, the drive current can influence the laser's center wavelength (see Figure 40), so precise control of the current is also important for spectral control of the diode output.

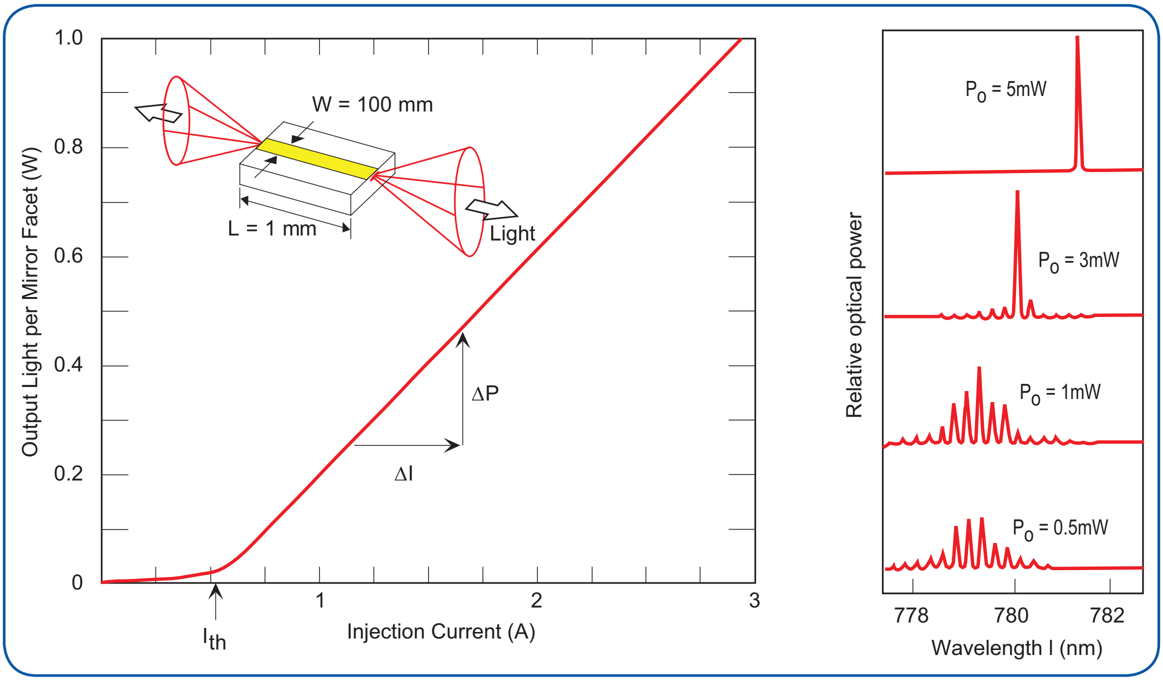


Figure 40. A typical L-I curve associated with a high-power laser diode (left) and effects of operating current level on the output spectrum of a laser diode with a single longitudinal mode (right).

Due to their sensitivity to injected current, laser diodes are typically driven by a stable current source [63, 67]. Other sources, e.g., voltage sources or generic power supplies, are too noisy for most applications and can generate voltage and current fluctuations and transients that may damage the laser diode. Specialized circuit designs have been developed to protect laser diodes from being damaged. The circuit designs typically include input AC power filtering and high-speed transient suppression circuits. Momentary internal transients may also occur when the output current is turned on or off. To guard against this, effective laser diode drivers short the output to the laser diode whenever the current is turned off. During the turn-on phase, the drivers ramp up the output current slowly to the desired set point value following a delay of several seconds. Another necessary feature of laser diode drivers is the independent current limit. which overrides any condition that may cause the output current to exceed the laser diode's maximum current rating. Features that maintain the stability of the operating current are crucial. Many controllers have a photodiode feedback mechanism built in for monitoring the output power. Since laser diodes generally emit light from both ends of their cavity, monitoring the rear facet output beam of the laser diode using a photodiode allows one to actively maintain the laser at a constant power level. Another advantage of laser diodes is that they can be directly modulated because of the linear dependence of the *L-I* curve above threshold (see Figure 40). An appropriate current driver can allow for modulation of the laser diode with frequencies up to several GHz or can be used to pulse the output [69, 70].

Temperature Controllers

Figure 41 shows that the threshold current and differential responsivity of a laser diode are strongly affected by the laser's temperature [65, 68]. The laser threshold will increase exponentially with temperature as $\exp(T/T_0)$, where T is the laser temperature and T_0 is the "characteristic temperature" of the laser (typically between 60 to 150°C). T_0 is a measure of the temperature sensitivity of the device with higher values implying that the device is more thermally stable. T_0 is an important laser diode characteristic and is commonly extracted from multiple *L-I* curves. As discussed in Section I.B.3, changes in temperature affect the bandgap of the semiconductor junction and therefore, the peak wavelength of the gain profile. This results in a linear relationship between temperature and the center wavelength of the laser diode (see Figure 41) with typical temperature tuning coefficients of 0.3 nm/°C. As a result, a temperature controller plays a key role in determining the laser wavelength.

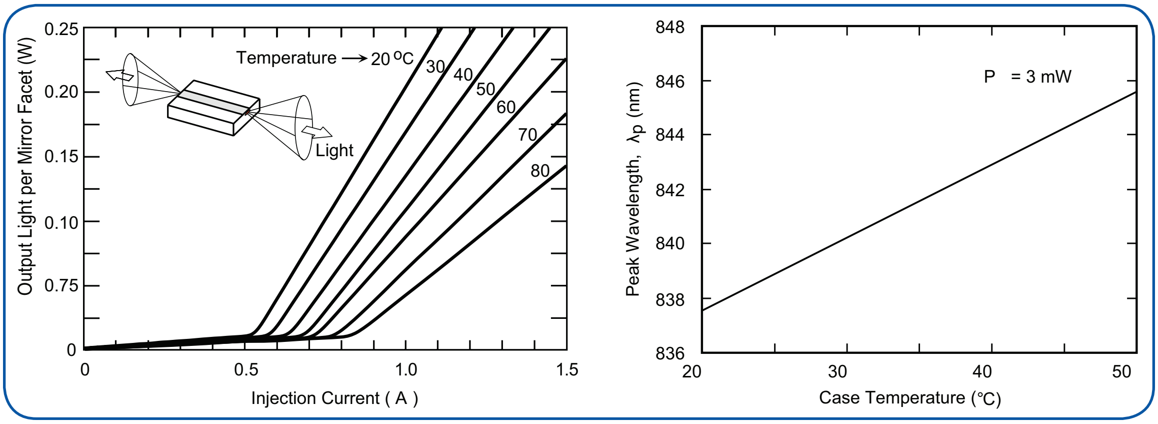


Figure 41. L-I curves for a laser diode operating at various temperatures (left) and effects of temperature on center wavelength (right).

Given the number of parameters that depend on laser diode temperature, it is important to set and maintain a stable temperature using a temperature controller [63, 71, 72]. Most laser diode applications use thermoelectric (TE) coolers to maintain a constant temperature. TE coolers rely on the Peltier Effect, whereby driving current through *p*- and *n*-type semiconductor materials will cause them to transfer heat. The most important point to consider when using TE coolers is that they are heat pumps. In other words, they pump heat from the laser, which generates heat, to the heat sink, which dissipates heat. To achieve this heat pump action, current must be driven to the TE cooler in the proper direction. These solid-state devices can heat or cool small thermal loads to more than 60°C from ambient and achieve temperature stabilities of better than 0.001°C. An accurate temperature sensor attached to the laser diode allows the TE cooler to properly regulate the device's temperature. One type of sensor is a thermistor which is a resistance device that exhibits a voltage drop proportional to temperature. It is the most commonly used sensor because it is inexpensive, accurate, highly sensitive and easy to work with. It is also the smallest type of sensor, which makes it ideal for integration into laser diode packages.

Additional online content discussing the output characteristics of laser diodes and LEDs can be found in [73-75].

C. Incoherent Light Sources

The incoherent light sources discussed in this section share several characteristics that distinguish them from lasers. Incoherent light gives rise to radiation, which is emitted from the source in all directions. Furthermore, unlike laser gain media where the radiating species are generated through optical or electrical excitation, the most common excitation mechanism for these sources is thermal excitation. This gives rise to spectrally broadband emission, which depends on the temperature of the source medium as described below. The broadband nature of these sources, coupled with its omnidirectional emission, makes them ideal for lighting in homes, workplaces, and vehicles. In research applications, the broadband output can be exploited for simulating solar radiation or can be spectrally filtered for applications such as spectroscopy or microscopy. These incoherent light sources are mainly categorized by the wavelength range and spectral shape of their output. These sources, which are detailed below, consist of deuterium light sources, arc lamp sources, quartz-tungsten halogen (QTH) sources, and IR emitters. LEDs are also incoherent light sources but have been described in detail earlier and will not be covered here. One key point is that LEDs have a narrower emission spectrum (see spectral characteristics in Section I.B.3) than the incoherent light sources discussed here and so broadband emission is typically achieved by using multiple LEDs with different center wavelengths.

Black Body Radiation

When charged particles in matter are heated, they gain kinetic energy and the resulting movement of these charged particles gives rise to electromagnetic radiation in the form of thermal energy. Thus, any material with a temperature above absolute zero emits thermal radiation. If the material system is in thermal equilibrium with its surroundings and is a perfect emitter, it is called a black body radiator [64, 76]. While most material systems are not true black bodies, they are often approximated as such since the laws governing the emission from a black body are simple and quantitative. Planck's Law describes the spectral distribution of radiant energy inside a black body. Spectra generated according to this law are typically given in units of spectral radiant exitance or spectral irradiance (see Section II.A for definitions). These spectra are smoothly varying curves with their distribution and output directly related to the temperature of the black body (see Figure 42). The inverse relationship between peak wavelength and temperature, known as Wien's Law, is also shown in Figure 42. Sources like the sun and the material systems making up the incoherent sources described below all have black body-like emission spectra. The temperature of the sun's surface is close to 6000 K and, as shown in Figure 42, this gives rise to a peak solar emission around 0.5 µm, which corresponds to green light. Even objects at room temperature emit thermal radiation but their peak emission wavelength is around 10 µm. Since this gives rise to no VIS radiation, this became the genesis of the term "black body".

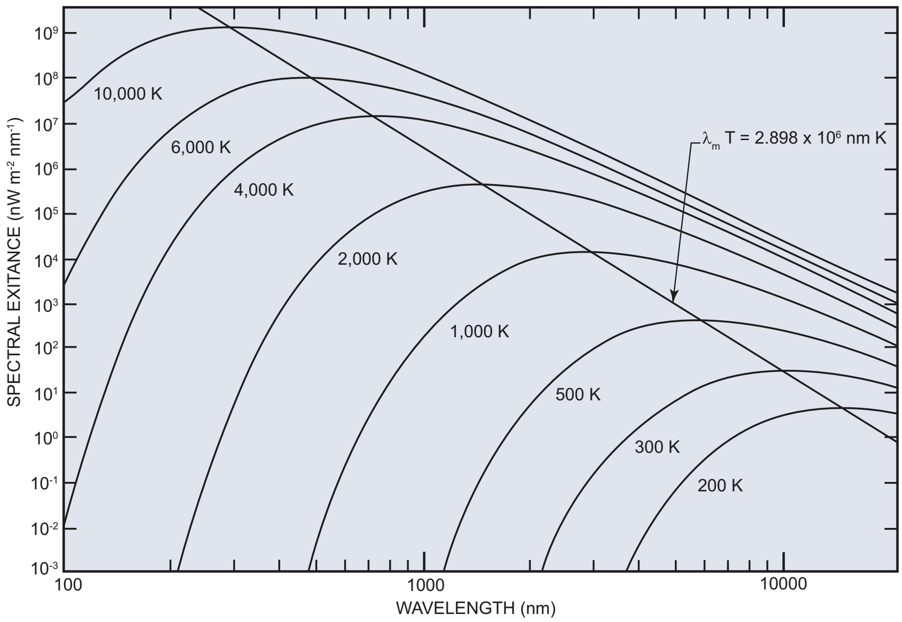
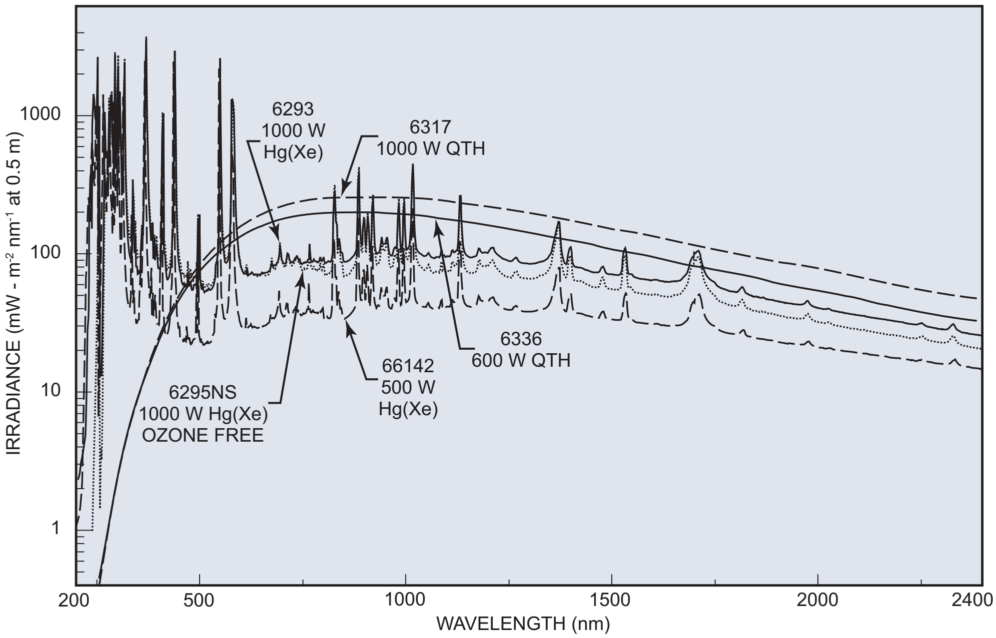


Figure 42. Spectral irradiance for various black bodies. Wien's Law is shown, which relates the peak wavelength (λ_m) to the black body temperature (T).

Types of Sources

Arc lamps operate by passing electricity through a discharge tube containing a high pressure gas. The electricity ionizes the gas and creates an arc that emits high-intensity light [43, 77]. These gases typically consist of either xenon or mercury-xenon mixtures (see Figure 43). A xenon arc lamp produces a black body-like emission spectrum corresponding to 6200 K which is a bright white light. The general characteristics of arc lamps are high irradiance output with a small source arc, intense UV output, and a spectrum that closely mimics natural sunlight [78]. These sources are therefore used as solar simulators and for lamps in movie projectors or searchlights. Arc lamps can also be made to emit extremely intense pulses of light instead of the typical *CW* emission for incoherent sources. These pulsed sources are often referred to as flashlamps and can be used for optically-pumping solid-state laser media. Finally, arc lamps generate strong, sharp emission peaks in addition to their black body-like emission (see Figure 43). These peaks are the result of spontaneous emission from atomic level transitions in the gases (see Figure 3). The resulting narrow and well-defined emission lines are therefore ideal for use as spectral calibration sources [79].



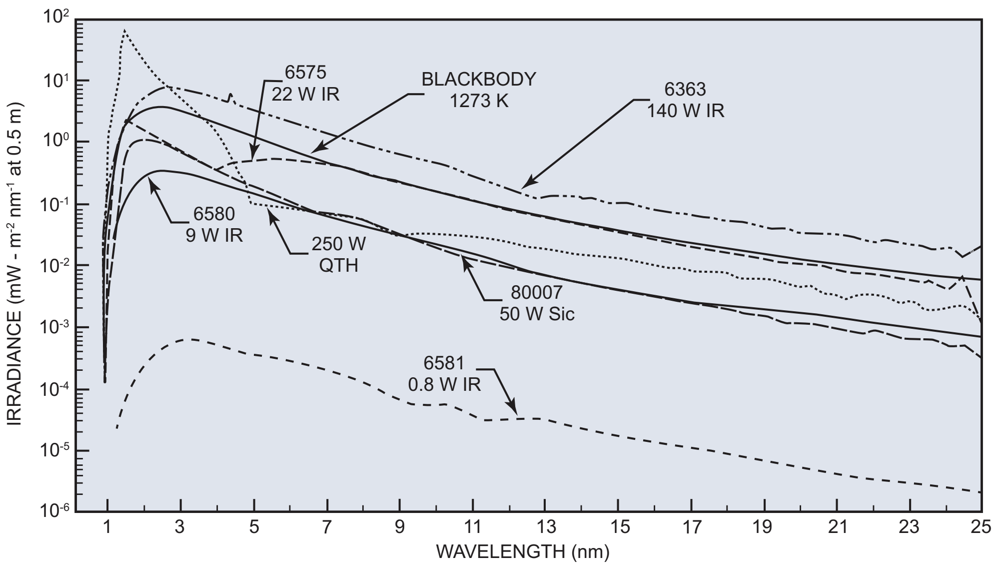


Figure 43. Spectral irradiance for different lamp types (top) and for typical infrared (IR) emitters (bottom).

A deuterium lamp is a type of arc lamp in which molecular deuterium is excited to a higher energy state prior to radiatively decaying to the ground state [80]. Consequently, the deuterium lamp is one of the few incoherent sources whose radiation is due to spontaneous emission as opposed to thermal radiation. Its emission spectrum is not black body-like but rather continuous and centered in the UV. Deuterium lamps possess the shortest output wavelength of all lamps, generate negligible output in the VIS and NIR spectral regions, and are the preferred sources for UV spectroscopy since they are both stable and have long lifetimes [78].

QTH lamps are a variation of the traditional incandescent lamps in which a filament of tungsten is heated to produce thermal radiation. The presence of halogen creates a regenerative cycle with the tungsten to enhance the overall lifetime and prevent the process of blackening [43, 81]. Since this process operates at a higher temperature than a conventional incandescent lamp, it must be housed in a bulb made of quartz since it has a high melting point. A QTH source produces a smooth and continuous spectrum from the near UV to well into the NIR (see Figure 43). These sources are extremely stable, possess high total visible output, and are easy and inexpensive to operate [78]. For these reasons, they are ideal as calibration sources when a known spectral irradiance is needed [79]. Furthermore, if a QTH source is coupled to a monochromator (see Section II.D.2), the system can also be used as a spectral calibration source [82].

Infrared (IR) emitters are useful light sources for IR spectroscopy [83]. IR emitters provide the necessary intensity at IR wavelengths that are not emitted by arc and QTH lamps, and IR emitters are more economical and have longer lifetimes. These sources function as almost perfect black bodies and generate broadband IR light from 1 µm – 25 µm (see Figure 43) with very efficient emission.

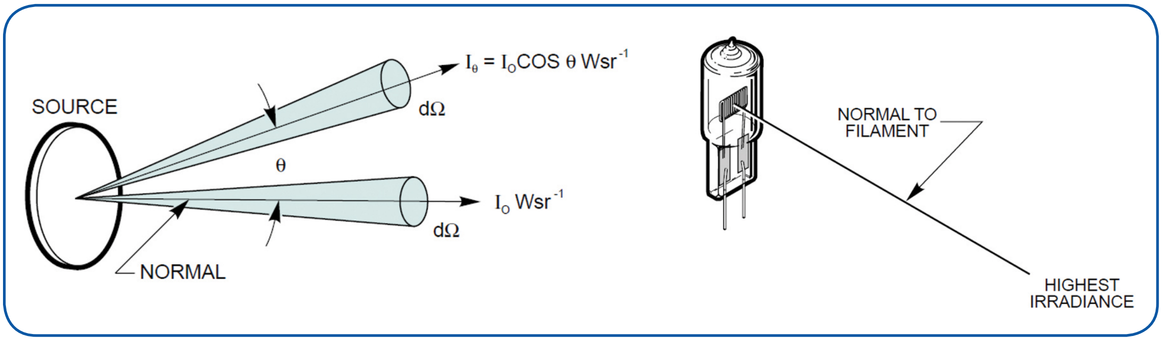


Figure 44. Lambert's cosine law indicating how intensity (I) depends on angle of observation (θ) from the normal (left). Orientation of a QTH lamp for maximum intensity, which is along the axis normal to the filament plane (right).

The spatial emission properties of an incoherent source depend on the shape of its lamp. In a QTH lamp, the filament is effectively a planar surface, so its emission resembles that of an LED as described in Section I.B.3. Its emission is Lambertian [64], that is, its intensity decreases with the cosine of the angle from the normal leading to a reduction as one goes off axis (see Figure 44). This must be considered when orienting a lamp with respect to its target [84] since one typically wants to maximize the irradiance (see Figure 44). Conversely, the arc in an arc lamp is typically small enough that it resembles a point source. Consequently, its emission is isotropic, i.e., no dependence on the observation angle. This also has significant advantages in terms of collecting and collimating the light when using a lens system (see Section III.A.3 for more details).

Additional online content discussing incoherent light sources and their properties can be found in [85-87].

II. Laser and LED Characterization

Proper characterization of the output from a laser or LED is critical for performing any quantitative measurement, to ensure reliable and repeatable processing, or to enable one to choose the proper source for a specific application. Attributes of the source's output include average output power or pulse energy, the spatiotemporal distribution of the radiation and how it evolves as it propagates, as well as the light's spectral distribution. Each attribute requires its own characterization approach that utilizes a unique set of instruments that must deliver accurate, precise, and repeatable results. These approaches are the subjects of this section.

A. Power and Energy Sensors

Radiometry is the measurement of energy or power in electromagnetic radiation fields or light. The average output power is the most common radiometric measurement since many light sources, including *CW* lasers and LEDs, emit output power that is constant over time. For a pulsed source, the pulse energy is typically the radiometric unit of measure, although the average output power can be given as well. Since sources can have different spatial distributions and divergences, other parameters may be needed to fully characterize their outputs. For this reason, this section will discuss the most commonly-encountered radiometric quantities for measuring power and energy (see Table 5). There can be considerable confusion regarding the nomenclature of radiometric terms, which can lead to measurement errors if not properly understood. The discussion below aims to adhere to the International Commission on Illumination (CIE) system, which fits well with the SI system of units. For a more complete description of radiometry, its history, and its concepts, see [88].

Quantity	Usual Symbol	Typical Units
Power	φ	W
Energy	Q_e	J
Irradiance	E	W m^{-2}
Fluence	F	J m^{-2}
Radiant Intensity	I_r	W sr^{-1}
Radiance	L_r	W sr^{-1} m^{-2}

Table 5. Commonly used radiometric quantities.

The parameters in Table 5 are defined as follows:

- The average output power is defined as φ for a source with a continuous and stable output. For simplicity, φ is denoted as the power and is the radiometric quantity quoted most often.
- For a pulsed source, φ becomes a time-dependent quantity, i.e., $\varphi(t)$, with a peak amplitude and a temporal shape (see Section II.E). This amplitude is referred to as the *peak output power* or peak power. This pulsed quantity should not be confused with the average output power.
- The quantity Q_e is the energy within a pulse of light. This quantity can be measured provided the temporal response of the sensor is fast enough; otherwise, it can be determined based on φ and the repetition rate of the source (see temporal response in Section II.A.1).
- The irradiance (E) is essentially the power per unit area or power density. The terms exitance and intensity are often used synonymously with E since they have similar meanings and identical units of measure.

- The fluence (F) is associated with pulsed sources since it is the energy per unit area or energy density.
- The radiant intensity (I_r) and radiance (L_r) are derivatives of power and irradiance that also account for the divergence or spreading of the light source based on radiation into a solid angle. L_r is also important when determining light throughput in an optical system (see Section III.A.3).
- If the word "spectral" is used before any radiometric quantity, it implies consideration of the wavelength dependence of the quantity. The measurement wavelength should be given when a spectral radiometric value is quoted.

In order to ensure that a sensor or detector (the two terms will be used interchangeably in this section) can accurately measure a radiometric quantity such as power or energy, it needs to be calibrated using a detection calibration standard provided by one of the national standards laboratories such as National Institute of Standards and Technology (NIST, [85]) or Physikalisch-Technische Bundesanstalt (PTB, [89]). Typically, the output of a source such as a spectrally-filtered lamp or a laser is measured by a NIST/PTB traceable sensor and this calibration is transferred to a master sensor which is, in turn, used to calibrate the sensor under test. Errors associated with each of these steps, along with any additional errors associated with spectral or temperature corrections, are combined to determine the total error associated with the calibrated sensor. The sensors should be periodically calibrated to ensure these errors remain reliable. Such absolute errors are related to the sensor's ability to accurately measure the power or energy. This should not be confused with relative errors, which are related to the sensor's precision. These errors are based on the noise characteristics of the detector (see Section II.A.1).

1. Sensor Basics

All sensors that measure either power or energy from a laser or LED are typically described by a set of detector performance parameters. These parameters can be used as selection criteria when choosing the appropriate photodetector to characterize the source at hand. The three different types of sensors described in the following section each possess their own unique set of characteristics. These mainly result from the differences in their light-to-electrical signal conversion processes and are typically divided into thermal detectors (thermopile, pyroelectric) or photon detectors (photodiode). These detector performance parameters and how they differ for each type of sensor are the subjects of this section.

Spectral Responsivity

Responsivity is a measure of the transfer function between the input optical power and the output electrical signal from the detector [90]. Thermal detectors convert a temperature change into a voltage (see Section II.A.2); therefore, their responsivity is typically given in units of V/W. Photodiodes convert absorbed photons in a semiconductor into a current (see Section II.A.2); therefore, the responsivity is given in A/W. Spectral responsivity describes how the detector will respond as a function of wavelength or photon energy. This spectral response can be dependent on a material's transmission properties, e.g., a window in front of a sensor, or a material's absorptive properties, e.g., a coating on a sensor or the sensor material itself. Thermal detectors respond to heat and so one watt delivered by a UV photon produces the same response as one watt delivered by an IR photon. Provided the material that is being heated has uniform absorption, the spectral responsivity is flat. The response for an ideal thermal detector is shown in Figure 45 (responsivities for typical thermal detectors are shown in Section II.A.3). Photon detectors produce at most a single response element, i.e., an electron-hole pair per incoming photon. The energy carried by individual photons is inversely proportional to the wavelength, and so, for the same input power, there are fewer UV photons per watt compared to IR photons. Accordingly, photon detector responsivity is significantly lower in the UV than in the IR (see Figure 45 for ideal response). Furthermore, due to the presence of the bandgap in semiconductors, only photons with energies above the bandgap energy (E_g) will be absorbed (see Figure 32). Therefore, a photodiode will exhibit an abrupt increase in the responsivity for wavelengths just below the value associated with the bandgap (λ_g) while going to even shorter wavelengths will result in a decrease in responsivity like that described above. Figure 45 provides a representative spectral responsivity for a Si photodiode while actual responsivity curves for various photodiodes are given in Section II.A.3. Due to the strong wavelength-dependence of the responsivity for photodiodes, it is not uncommon for specification sheets to give a peak responsivity value at a wavelength and provide a relative spectral response curve.

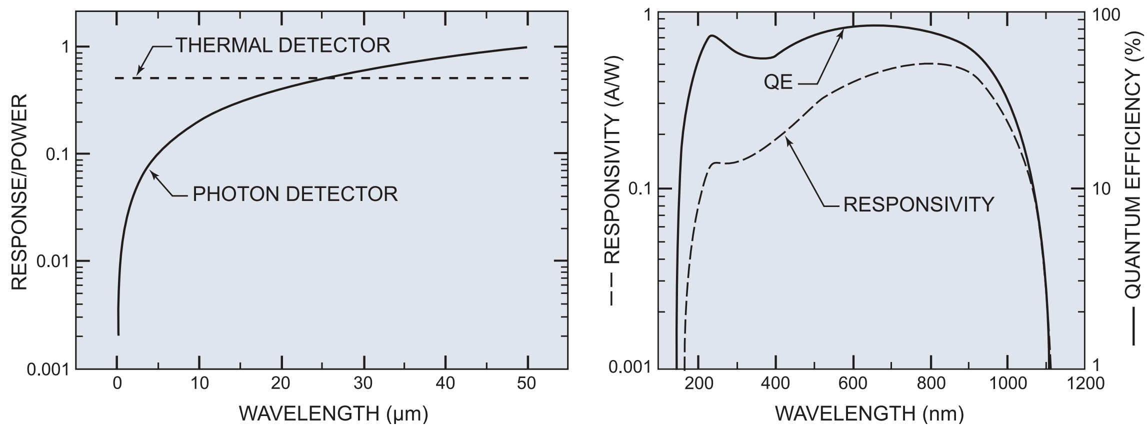


Figure 45. Relative spectral responsivities of perfect detectors (left). Si photodiode responsivity and quantum efficiency (QE) where the QE is the probability that a single photon will generate an electron-hole pair that contributes to the current (right).

Noise, Noise Equivalent Power, and Normalized Detectivity

A photodetector generates an output voltage or electric current that is proportional to the incident optical power. This output value generated by the device is a quantity whose value fluctuates above and below its average value. These fluctuations are generally regarded as noise and are typically quantified by the standard deviation of the value about its mean. There are numerous sources of noise in detectors, which are dependent on the type of sensor. A brief summary of the sources is given here while details can be found in [1, 90]. Sources of noise can generally be grouped into photon noise, detector noise, and circuit noise. Photon noise is due to the discrete nature of radiation, which is composed of photons arriving randomly in time. In photodiodes, absorbed photons produce charge carriers at random intervals giving rise to a variation in current that appears as noise. Detector noise can result from temperature fluctuations in thermal detectors or the random recombination of charge carriers in photodiodes. Fluctuations in a detector's internal resistance or in any resistance in series with the detector's terminals can give rise to circuit noise.

All sources of noise affect detector parameters such as the signal-to-noise ratio (SNR), the minimum detectable signal, and the detector sensitivity. The noise equivalent power (NEP) is the input power necessary to give an output signal that gives a SNR equal to one. Therefore, NEP is a measure of the minimum detectable signal for a sensor. NEP values are typically given in units of $W/Hz^{1/2}$ and must be stated at a specified wavelength, modulation frequency, detector area, temperature and detector bandwidth due to its dependencies on these parameters. Detectivity is the reciprocal of NEP and therefore gives a more intuitive figure-of-merit that is larger for more sensitive detectors. For most detectors, electrical noise power is proportional to the area of the detector (A_D) and its electrical bandwidth (Δf). Therefore, to compare different types of detectors independent of A_D and Δf, the normalized detectivity (D*, pronounced "D-star") is typically used. Figure 46 shows the formula for D* along with its spectral response for several common photodetectors. Two trends are immediately evident. Much like responsivity, D* is essentially constant across the entire wavelength range for thermal detectors while photodiodes show a strong wavelength dependence that is localized in a particular spectral region. However, photodiodes possess a much larger detectivity than their thermal counterparts.

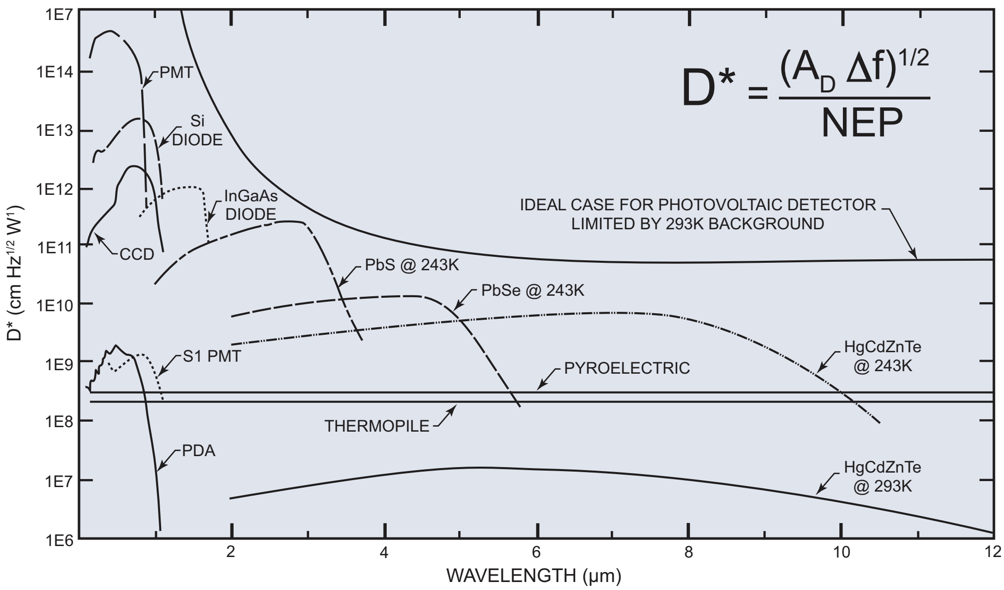


Figure 46. Approximate D values as a function of wavelength for various sensor types (PMT – photomultiplier tube, CCD – charge-coupled device, PDA – photodiode array).*

Linearity and Dynamic Range

In order to be practically useful, a detector's output should be linearly proportional to the input optical power. This linearity is typically defined for a range of input powers or energies based on when the response deviates from this linear response by a certain pre-determined amount (see Figure 47). This is called the dynamic range of a detector and essentially describes its usable range. The lower limit of the dynamic range is typically determined by the NEP or detectivity while the upper range may be device or external circuit limited. A summary of the typical dynamic ranges for detectors is given in Table 6 while specific examples are given in Section II.A.3. Photodiodes have very large dynamic ranges that can exceed 10^9 for properly designed circuits (see Figure 47 for example). Energy sensors, such as pyroelectrics and some photodiodes, often possess a reduced dynamic range compared to power sensors. To accurately measure the pulse energy, a detector must respond more rapidly (see temporal response below). Since detector response speed is often inversely proportional to detector sensitivity (see Section II.A.2), the dynamic range is limited at lower energies by reduced detectivity, while recombination can limit the range on the high end for photodiodes.

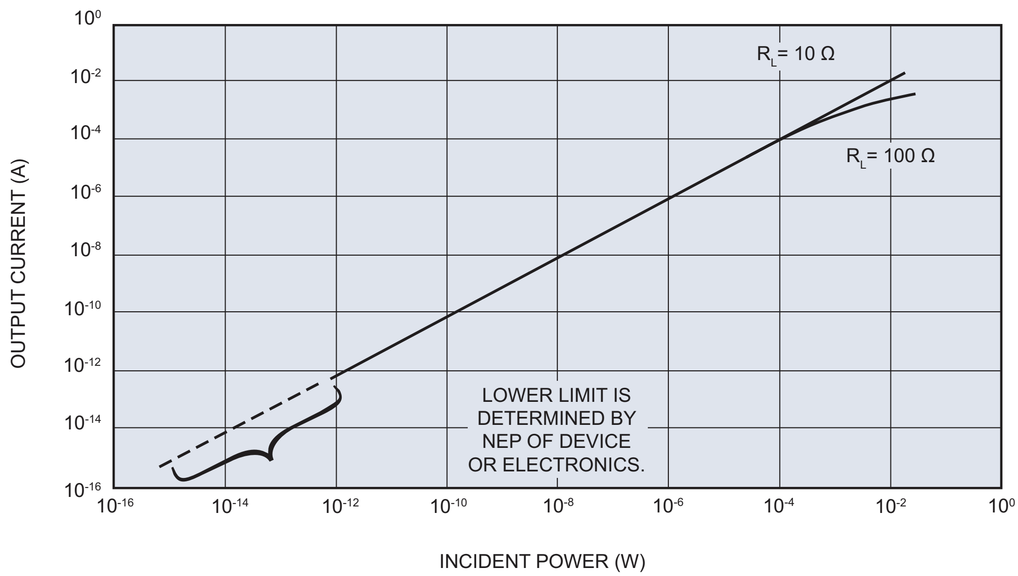


Figure 47. Linearity of a Si photodiode showing the dynamic range of the detector.

Temporal Response

A pulsed source of light exhibits an output power, which is a time-dependent quantity, i.e., $\varphi(t)$. Pulsed lasers possess a temporal shape and amplitude (see Section II.E) that is repeated in a pulsetrain with a certain repetition rate (see Figure 48). For an energy sensor to directly measure the pulse energy (Q_e), it needs to be able to discriminate between pulses at this repetition rate. This can be seen in the following equation where $\varphi(t)$ is integrated over a time interval (Δt) that encompasses an entire pulse:

$$Q_e = \int_{\Delta t} \varphi(t)\, dt. \tag{8}$$

Clearly the temporal response of the detector must be fast enough to initiate this time interval and decay rapidly enough to eliminate contributions from the following pulse (see Figure 48). Such rise and fall time constants are frequently different, since different physical parameters may cause them. These topics are covered in Section II.C.1 where the more stringent requirements of fast photoreceivers are introduced. For an energy sensor, if the response time is sufficient to provide the required Δt, the pulse energy can be measured directly. Power sensors typically do not possess sufficient temporal response to measure the pulse energy directly, i.e., Δt is much longer than the time between pulses. In this case, the pulse energy can be estimated by dividing the measured average power by the repetition rate in Hz as detailed in Figure 48. In this case, one must assume an average pulse energy within the measurement period. As discussed above, there is an inherent tradeoff between speed and sensitivity that results in differences in dynamic ranges between power and energy sensors.

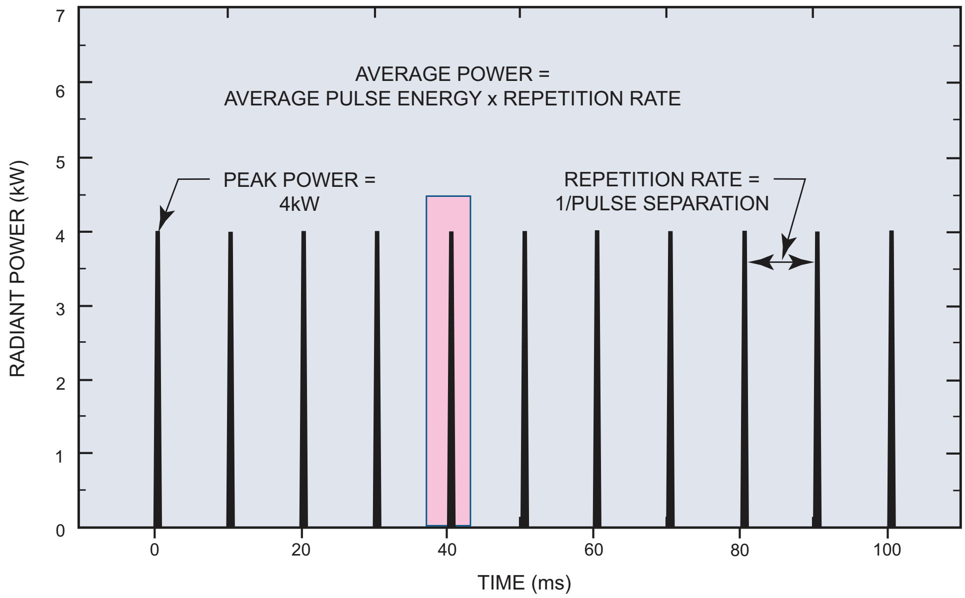


Figure 48. Typical laser pulsetrain, which indicates the time interval Δt (red rectangle) of an energy sensor capable of directly measuring pulse energy. A power sensor, which measures the average power, estimates the average pulse energy based on the repetition rate.

Aperture

The irradiance (*E*) is the power per unit area or power density. The irradiance emanating from any laser or LED has a certain spatial distribution (details are given in Sections II.B.1 and I.B.3, respectively, for each light source) which can be denoted as *E*(*s*). The active area of a detector (A_D) has a finite size that is typically dictated by its aperture. Therefore, the size of the *E*(*s*) with respect to A_D can affect the measured power or energy. The following equation quantifies this by showing the spatial integration of *E*(*s*) over the detector area:

$$\varphi = \int_{A_D} E(s)\, ds \quad or \quad Q_e = \int_{A_D} \int_{\Delta t} E(t,s)\, dtds. \qquad (9)$$

This results in a measured power or energy (the latter assumes a pulsed source such that the irradiance is also time-dependent). If the spatial extent of *E*(*s*) is larger than A_D, e.g., an expanded laser beam and a small detector, only a fraction of the power or energy will be measured, which can be estimated using a calculator found in [91]. Conversely, if *E*(*s*) is accommodated within A_D, the total power or energy will be measured; therefore, proper care must be taken to ensure that *E*(*s*) is properly delivered to the sensor. For collimated sources, such as lasers, this typically amounts to using the proper focusing geometry such that the beam is significantly smaller than A_D upon arrival. For highly divergent systems like laser diodes or for LEDs with omnidirectional output (see Section I.B.3), more sophisticated collection geometries must be used to ensure the total power or energy is measured. These topics are discussed in Section II.A.4.

Damage

As detailed in Section I.A.7, the large irradiances associated with laser output can lead to optical damage, particularly to the surfaces of power and energy sensors. When laser beams are focused into these detectors to ensure proper collection, the probability of damage increases. For power sensors, specification sheets often include a maximum recommended irradiance or power density that lies below the LIDT of the detector to ensure compliance. These values are typically in the 5-50 W/cm^2 range for pyroelectrics and photodiodes whereas they can be several orders-of-magnitude larger for thermopiles (see Section II.A.3 for examples). For energy sensors that measure pulsed lasers, the maximum values are typically given as an energy density or fluence (F). The fluence is essentially the temporally-integrated irradiance as shown below (the spatial dependence of E has been ignored):

$$F = \int_{\Delta t} E(t)\, dt. \tag{10}$$

These values are typically around 0.1 J/cm^2 for pyroelectric and photodiode energy sensors. The calculator found in [92] can be used to calculate the power density or fluence of a laser beam.

Additional online content discussing radiometry and the characteristics of power and energy sensors can be found in [93-97].

2. Types of Sensors

The sensors discussed in this section are differentiated by the way in which they convert the incident light into an electrical signal. Thermal detectors work by converting the incident radiation into an increase in temperature. The temperature change is measured either by a voltage generated at the junction of dissimilar metals or by the pyroelectric effect. In either case, the heat-sensitive element is coated with a black material to enhance the absorption of the radiation. The material is designed to possess a large and uniform absorption leading to good responsivity over a wide spectral range. This is the major advantage of thermal detectors. As a result of the time required to effect a temperature change, thermal detectors are generally slow. In a photodiode, photons are absorbed in a semiconductor *p-n* junction giving rise to mobile charge carriers. The electrical conductivity of the material increases in proportion to the incident optical power. Applying an electric field to the junction causes the carriers to be transported, resulting in a measurable electric current in the circuit. The detectivity of photodiodes is typically much larger than that of their thermal counterparts and the mechanism for conversion can be quite fast. The main disadvantage of photodiodes is that their responsivity is strongly dependent on the wavelength of the incident light. Table 6 lists the typical characteristics of the three different sensor types discussed in this section.

Thermopiles	Pyroelectrics	Photodiodes
Measure power from mW up to 30 kW	Measure energy from sub μJ up to 40 J	Measure power from fW up to 30 W
Response times of 1-3 seconds	Measure pulse widths up to 20 ms and repetition rates up to 25 kHz	Inexpensive and compact
Linearity of ±1%	Duty cycle - pulse width up to 25% of time between pulses	Linearity of ±0.5% below saturation
Relatively independent of beam size and position on detector	Accomodate beam sizes up to 96 mm diameter	Large dynamic range — over 9 orders of magnitude in one sensor
Broadband spectral response	Broadband spectral response	Large wavelength sensitivity — must specify exact wavelength
High laser damage threshold	Can measure low level *CW* powers with chopper	Can measure sub μJ energies with reduced dynamic range

Table 6. Typical characteristics of various sensor types.

Thermopiles

Thermopile sensors are based on thermocouples. A thermocouple consists of two dissimilar metals connected in series. To detect radiation, one metal junction is typically blackened to absorb the radiation. The temperature rise of this junction with respect to another non-irradiated junction generates a voltage. This effect is the basis of all thermocouple temperature sensors. The thermocouple materials used in thermopiles are usually bismuth and antimony, which have a relatively high thermoelectric coefficient (a measure of the magnitude of the induced voltage in response to the temperature difference). An individual thermocouple typically produces a low output voltage, which results in a low detectivity and limits its use as a sensing device. So, one way to increase the output voltage is to connect many thermocouple junctions (typically 20 to 120) in series. All the "hot" junctions are placed close together to collect the radiation.

This constitutes a thermopile. The typical operation of a thermopile is shown in Figure 49. The thermopile disc consists of a circular array of thermocouples. The optical radiation, e.g., laser beam, is absorbed by the disc generating the output voltage. The heat on the disc spreads radially across the thermopile on the reverse side of the disc where it is transferred to a heat sink that is convectively cooled by a fan or by water. These devices are quite sensitive in the IR due to their broadband absorption. Consequently, care must be taken to stabilize their field of view, since all near room-temperature objects, including people, emit significant IR. Due to the heating and cooling processes, the response time of a thermopile is quite slow, typically on the order of seconds. Due to this time constraint, thermopiles are used exclusively for measuring the power of *CW* or quasi-*CW* radiation sources. The response time can be shortened by proportionally reducing the thermal capacity of the material. However, the detectivity of the detector is inversely proportional to the heat capacity, resulting in a general tradeoff between temporal response and sensitivity.

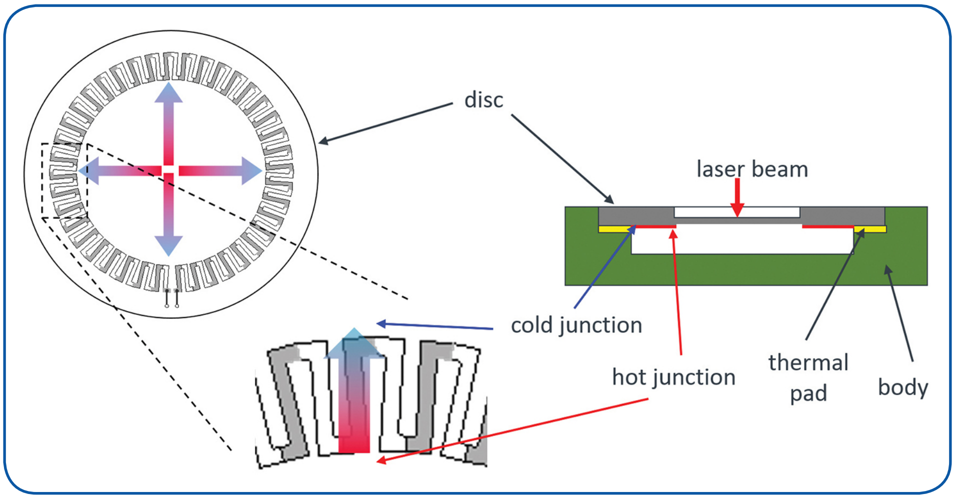


Figure 49. Typical operation of a thermopile sensor.

Pyroelectrics

The configuration of a typical pyroelectric sensor and its operational output can be seen in Figure 50. A pyroelectric material, which is usually crystalline, possesses an electric polarization, even in the absence of an applied voltage. An incident laser pulse heats the crystal, which causes the material to expand and produce a change in the polarization. Charge builds up on opposite surfaces of the crystal which generates a current flow that charges a capacitor. This charged capacitor induces a voltage whose amplitude change is proportional to the original laser pulse energy. Since it is the change in temperature that produces the current, pyroelectric detectors respond only to pulsed or modulated radiation. They respond much more rapidly to variations in radiation than thermopiles and are unaffected by steady background radiation. The response of a pyroelectric detector depends on the thermal time constant (governed by the thermal mass and thermal connections from the element to its surroundings) and the electrical time constant (the effective resistance and capacitance of the detector circuit). Consequently, small detectors with small thermal mass can have extremely rapid response. Adding a black coating to give uniform absorption over a wide spectral range is often used in pyroelectric sensors. However, this coating increases the thermal mass of the detector, which lowers the frequency response. This results in a temporal response vs. sensitivity tradeoff.

Pyroelectric detectors are typically used to measure the energy of pulsed lasers where the pulses may vary in width from fs to ms and in energy from sub µJ to J. As discussed in the previous Section II.A.1 (see temporal response), the thermal and electrical time constants of energy detectors are chosen so that each pulse is integrated. The peak of the output voltage is a measure of the charge produced by the detector and therefore, of the pulse energy. The charge dissipates before the arrival of the next pulse. The integration time, or fall time, imposes limitations on the minimum interval between pulses or the maximum repetition rate which can be accurately measured. It is possible for pyroelectric sensors to be used to measure the powers of low-level *CW* or quasi-*CW* light sources. To accomplish this, the radiation must be modulated or chopped prior to being detected such that the detector produces an AC output signal. Since the frequency response of the detector is much lower than for measuring pulsed radiation, the increased sensitivity typically results in a higher dynamic range in this configuration.

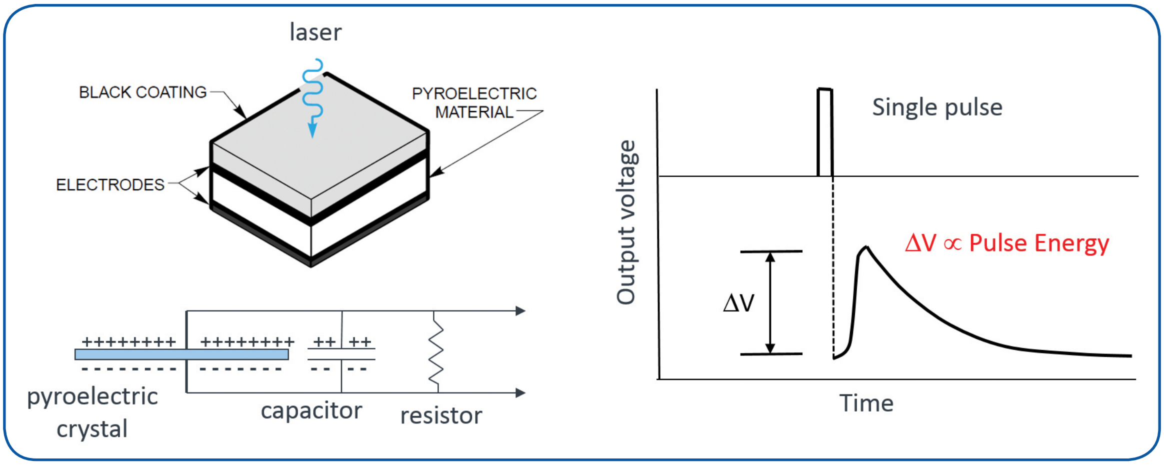


Figure 50. Typical operation of a pyroelectric sensor.

Photodiodes

A photodiode consists of a semiconductor *p-n* junction like the laser diode and LED described in Section I.B.1. However, the fundamental radiation process involved is absorption (see Figure 32). Light falling on the junction causes the formation of electron-hole pairs. In photovoltaic mode, i.e., no applied bias, the electron-hole pairs migrate to opposite sides of the junction, thus producing a voltage (and a current, if the device is connected in a circuit). However, most photodiodes are operated in the photoconductive mode where a reverse bias is applied across the junction. Operating in this mode offers a few significant advantages [1]. As shown in Figure 31, a reverse bias increases the width of the depletion region, which leads to a larger photosensitive area allowing more light collection. Furthermore, the bias produces a strong field in the junction that sweeps the carriers out quickly, making it less likely for recombination to occur. This ensures a large quantum yield or efficient conversion of photons to charge carriers. There are also advantages in terms of response time, which will be discussed in Section II.C.1. In reverse bias photodiodes, the current produced by the bias and charge carriers is proportional to the incident optical intensity over a wide dynamic range.

One critical difference between a semiconductor photon source and a photon detector is that the former requires the use of a direct-gap semiconductor while the latter can utilize an indirect-gap semiconductor (see Figure 29). While the simultaneous requirement for energy and momentum conservation makes photon emission much less likely in indirect-gap semiconductors, this is not the case for absorption [1]. A readily-achievable two-step process occurs where an electron is excited to a high level in the conduction band followed by a relaxation process where its momentum is transferred to phonons. Since this process can be sequential, it is much more likely than an emission process where the two steps must occur simultaneously. A consequence of this is that Group IV elemental semiconductors such as Si and Germanium (Ge) can be efficient photon detectors similar to direct-gap III-V systems like GaAs or InGaAs. The ubiquitous presence of Si in electronic circuits and devices makes it unsurprising that Si photodiodes are the most common detectors of light used in instrumentation (see Figure 51 for a typical device architecture). The spectral responsivity of Si (see Figure 45) covers the UV, VIS, and the NIR. Coverage in other portions of the electromagnetic spectrum are possible with photodiodes of other semiconductor materials. Examples of these detectors are given in Section II.A.3.

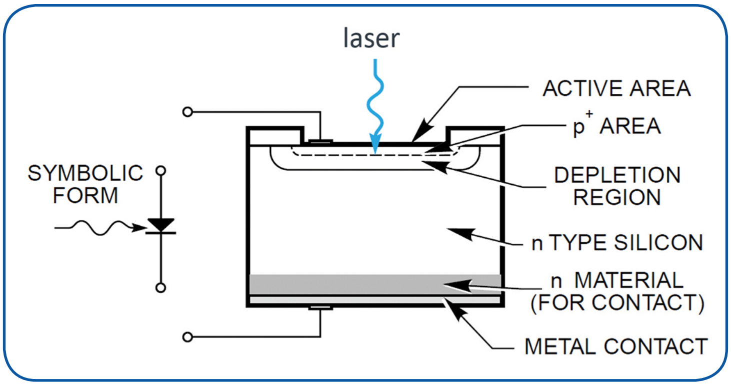


Figure 51. Typical Si photodiode device architecture.

Photodiodes possess several characteristics that set them apart from their thermal counterparts. The conversion of photons to electrons is very rapid and so these detectors have the potential of following fast changing radiation levels. Detectivity can be significantly higher than that of thermal detectors (see Figure 46). The detection mechanism is strongly wavelength-dependent, i.e., there is a peak in responsivity that falls off at short wavelengths due to photon-to-watt conversion and at long wavelengths because of the minimum photon energy required to generate electron-hole pairs. The dynamic range of photodiodes can be extremely large, exceeding values of 10^{10} with a single detector. Due to this large detectivity and dynamic range, photodiodes are often used to measure optical power over a wide range. For *CW* or quasi-*CW* sources, this is straightforward and for pulsed sources the procedure outlined in Figure 48 can be used to estimate pulse energy. Photodiodes can be used as energy sensors as well, provided their temporal response can accommodate the pulse integration. This leads to a reduction in the dynamic range due to decreased detectivity (owing to the faster temporal response) on the low end and due to saturation of the linear response of the detector on the high end. This is because electron-hole pairs begin to recombine instead of flowing through the circuit (see examples in Section II.A.3).

Additional online content discussing the different types of photodetectors can be found in [93, 94, 96-99].

3. MKS Ophir® Power and Energy Sensors

Accurate and reliable laser power and energy measurements are critical for medical applications (e.g., pulsed lasers for cosmetic surgery, corneal reshaping, light microscopy), industrial applications (e.g., material cutting and welding, heat treatment, additive manufacturing, photolithography, process control), as well as in the laboratory (e.g., basic research and applied research in optics, biology, chemistry, and engineering). For over 40 years, MKS Ophir's product group has been a leader in all aspects of laser beam measurement. Critical to this success is the constant innovation that is required as demands for durability and precision increase. Two key aspects of such innovation involve extending the operational limits of devices by improving their resistance to laser damage and improving calibration methods to ensure reliable and accurate instrumentation.

Figure 52. Ophir power and energy sensors.

Sensor Finder

Figure 52 shows the extensive set of both power and energy sensors offered by MKS. The power sensors consist of photodiode sensors and thermal sensors whose functionality and attributes were described in Section II.A.2. The photodiode sensors are typically used for low powers from pW up to 100's of mW but can reach as high as 3 W. The thermal sensors can be used from fractions of a mW up to 1000's of W. These thermal sensors can also measure single shot energies if the pulse rates do not exceed one pulse every 5 s. Pyroelectric energy sensors (also described in Section II.A.2) are capable of measuring pulse rates up to 10's of kHz. To determine the proper measurement device for an application, three important questions must be considered. First, what is the wavelength of the application, or does the application require coverage of a range of wavelengths? This initial question is essential as some sensors are considered broadband while others are highly specific to a narrow spectral region. The second question is: what is the range of powers or energies the sensor needs to measure? This information is critical not only for functionality but also because an incorrect selection can result in a damaged sensor. The third question is: what is the size and shape of the beam to be measured on the sensor? The power density or fluence of the beam is the most destructive element causing sensor failure and so it is important to accurately estimate the size of the beam at the measurement point. MKS provides a sensor finder tool [100] that can help a user determine the proper power or energy sensor. Figure 53 provides a screenshot of this sensor finder.

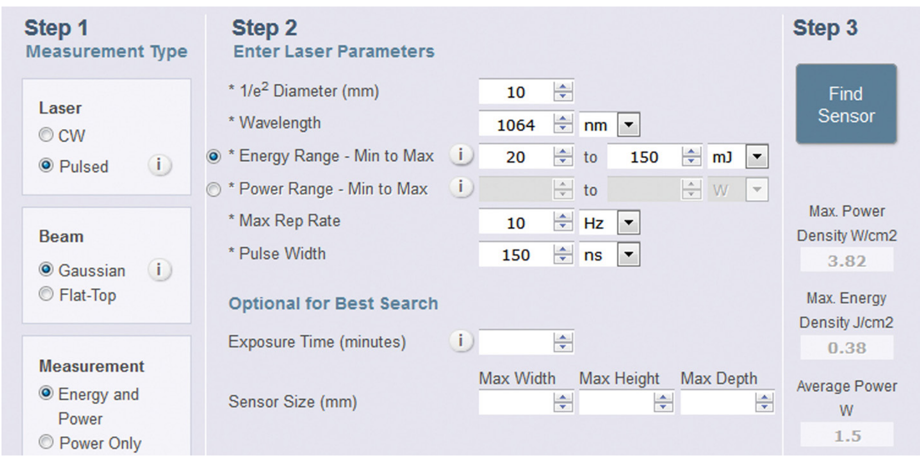


Figure 53. Ophir sensor finder helps determine the proper sensor for an application.

Thermopiles

Thermopile sensors can measure from 10's of mW up to kW, but the thermal range of operation for a single thermopile is limited by the difference in temperature between the hot and cold junctions of the disc. To access different power ranges, discs of different thicknesses and sizes are used in the various thermal sensors with thicker discs used for high powers and thinner ones used for lower powers. Although these thermal power sensors are used primarily to measure power, they can also measure single shot energies as well, where they integrate the power flowing through the disc over time and thus measure energy (see temporal response in Section II.A.1). The typical time it takes for the disc to heat up and cool down is several seconds, so these thermal sensors can only measure one pulse every several seconds at most. There is no limit to how short the measured pulses can be since the measurement is of the heat flowing through the disc after the pulse. However, when measuring short laser pulses with durations of less than 10 µs, volume-absorbing thermopiles are used in place of surface absorbers since the latter can be vaporized by short pulses. As discussed in Section II.A.2, thermopiles have broad spectral response. This is shown in Figure 54, which includes a specification sheet for an MKS general-purpose thermopile. This gives the usable power and energy ranges for the sensor as well as the high damage thresholds which are the result of specially-designed coatings.

Model	30A-BB- 18
Use	General purpose
Absorber Type	Broadband
Spectral Range µm	20 - 0.19
Aperture mm	∅17.5mm
Power Mode	
Power Range	10mW - 30W
Maximum Power Intermittent	N.A.
Power Scales	30W / 5W / 0.5W
Power Noise Level	0.5mW
Maximum Average Power Density kW/cm^2	20 at 30W 28 at 10W
Response Time with Meter (%95-0) typ. s	0.8
Power Accuracy +/-%	3
Linearity with Power +/-%	1
Energy Mode	
Energy Range	6mJ - 30J
Energy Scales	30J / 3J / 300mJ
Minimum Energy mJ	6
Maximum Energy Density J/cm^2	
<100ns	0.3
0.5ms	2
2ms	2
10ms	2
Cooling	convection
Fiber Adapters Available (see page 83)	ST, FC, SMA, SC
Weight kg	0.3

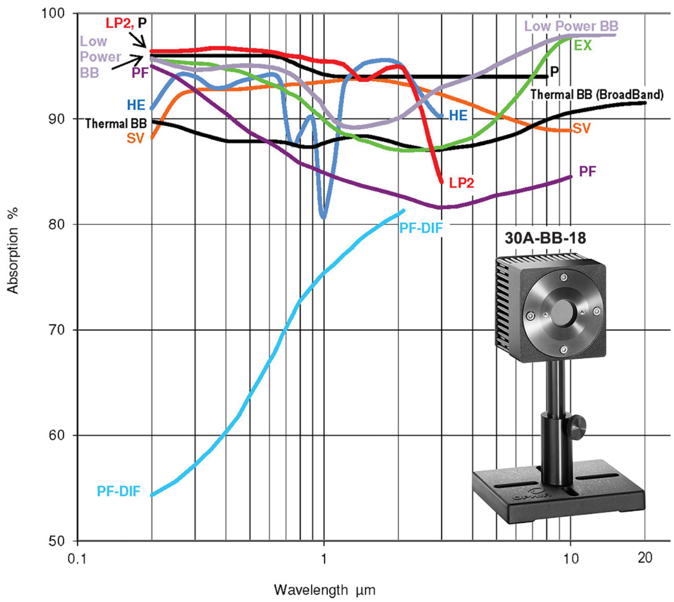


Figure 54. Specification sheet for a general purpose thermopile sensor (left) along with the wavelength-dependence of absorption for Ophir themopile products (right; inset shows an actual thermopile sensor).

Pyroelectrics

Pyroelectric sensors are useful for measuring the energy of repetitively pulsed lasers up to 25 kHz and are sensitive to low energies. They are less durable than thermopile sensors and therefore should not be used when average power measurement is sufficient. This can be seen when comparing the maximum average power density of a typical pyroelectric sensor (Figure 55) to the thermopile sensor above. MKS pyroelectric detectors have a unique and proprietary circuitry that allows them to measure long pulses as well as short pulses and work at a high duty cycle. The duration of the pulse and the associated repetition rate determines the dynamic range of the sensor (see Figure 55). As mentioned in Section II.A.2, pyroelectric detectors can measure average power, and the reduced temporal constraints result in a larger dynamic range. Like thermopiles, these thermal sensors exhibit a very wide spectral response as shown in Figure 55.

Model	PE-50C				
Use	High rep rate				
Aperture mm	∅46				
Absorber Type	metallic				
Spectral Range µm (a)	3 - 0.15				
Surface Reflectivity % approx.	50				
Calibration Accuracy +/-% (a)	3				
Max Pulse Width Setting (d)	2µs	30µs	500µs	1ms	5ms
Energy Scales	10J to 200µJ	10J to 200µJ	10J to 2mJ	10J to 2mJ	10J to 2mJ
Lowest Measurable Energy µJ (c)	10	10	60	80	100
Max Pulse Width ms	0.002	0.03	0.5	1	5
Maximum Pulse Rate pps	10kHz	5kHz	900Hz	450Hz	100Hz
Noise on Lowest Range µJ	0.5	1	6	10	20
Additional Error with Frequency %	%2± to 2kHz %4.5± to 5kHz	%2±	%2± to 750Hz	%2± to 400Hz	%1± to 80Hz
Linearity with Energy for >%7 of full scale (c)	1±.%5				
Damage Threshold J/cm^2 (b)					
<100ns	0.1				
1µs	0.2				
300µs	2				
2ms	6				
Maximum Average Power W	25 ,15 with optional heat sink				
Maximum Average Power Density W/cm^2	20				
Uniformity over surface	%2± over central %50 of aperture				
Fiber Adapters Available (see page 99)	ST, FC, SMA, SC				
Weight kg	0.25				

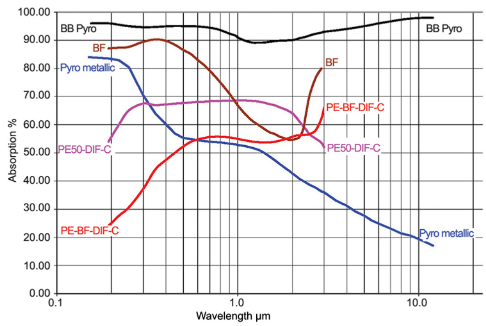


Figure 55. Specification sheet for high repetition rate pyroelectric sensor (left) along with the wavelength-dependence of absorption for Ophir pyroelectric products (right; below an actual pyroelectric sensor is shown).

Photodiodes

As described in Section II.A.1, when a laser is directed at a photodiode detector, a current is generated in proportion to the incident optical power. Many low power lasers have powers ranging from 5 to 30 mW. However, the response of most photodiodes begins to saturate at powers around 2 mW as the electron density in the junction becomes too great. MKS photodiodes are typically constructed with a built-in filter that reduces the power on the sensor so it can accommodate incident laser powers up to 30 mW without saturation. Furthermore, many photodiodes possess an additional removable filter that enables measurements up to 300 mW to 3 W, depending on the model. When used with MKS power meter units (see Figure 57), the photodiode detectors possess a very large dynamic range. The specification sheet for a typical Si photodiode is shown in Figure 56 where the full measurement range goes from 10's of pW to 100's of mW, representing a dynamic range exceeding 10^{10}. The spectral responsivity of photodiodes is wavelength-dependent as shown in Figure 56 for the three different types of semiconductors materials used. Finally, some photodiode sensors can also directly detect pulse energy with an associated reduction in their dynamic range.

Model	PD300-UV/ PD300-UV-193		
Use	Lowest powers from 200-1100nm		
Detector Type	silicon		
Aperture	10x10mm		
Filter mode	Filter out		Filter in
Spectral Range nm	200 -1100		220 -1100
Power Range	20pW to 3mW		2µW to 300mW
Power Scales	3mW to 3nW and dBm		300mW to 300µW and dBm
Resolution nW	0.001		100
Maximum Power vs. Wavelength	nm	mW	mW
	250 - 350	3	300
	400	3	300
	600	3	300
	800 - 950	2.5	150
	1064	3	300
Accuracy (including errors due to temp. variations)			
% error vs Wavelength nm	±6	200-270	±10 220-400
	±3	270-950	±5 400-950
	±5	950-1100	±7 950-1100
Damage Threshold W/cm^2	10		50
Max Pulse Energy µJ	0.4		15
Noise Level for filter out pW	±1		
Response Time with Meter s	0.2		
Beam Position Dependence	±2%		
Fiber Adapters Available	ST, FC, SMA, SC		

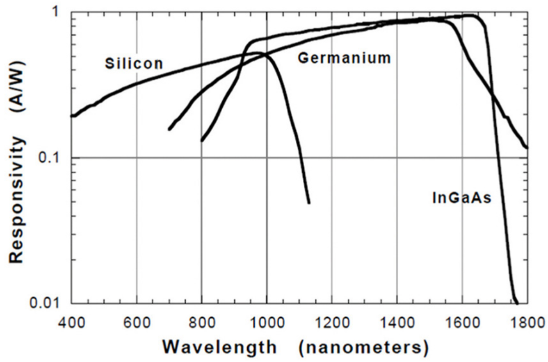


Figure 56. Specification sheet for low power Si photodiode sensor (left) along with the spectral responsivity for Ophir photodiode products (right; below an actual Si photodiode sensor is shown).

Power Meters and Computer Interfaces

In order to monitor, display, and/or store measured values from any sensor, MKS offers a range of power meters and computer interfaces as shown in Figure 57. These devices work on the smart plug principle, which means that they can work with almost any of the optical sensors. The power meters possess superior circuitry that provides low-noise amplification of the signal from the sensor, thereby improving dynamic range. The power meters also have smart keys, allowing for an easy and convenient user interface. Both the power meters and computer interfaces allow for seamless control and data acquisition using either MKS software or a user's own software.

Figure 57. Ophir power meters and computer interfaces, which allow for readout and acquisition of values derived from power and energy sensors.

Calibration

To ensure accurate and precise measurements from the power and energy sensors, MKS has devoted significant time and effort in calibration. As laser absorption varies with wavelength, it is not sufficient to calibrate at a single wavelength. If the wavelength variation is small, then the sensors are calibrated at several discrete laser wavelengths spanning a certain range. If the variation is considerable, the sensor is provided with an absorption correction curve activated by the wavelength of use. These curves are checked at many NIST and PTB traceable wavelengths using a complete line of calibration lasers (both *CW* and pulsed lasers) and corrections are made, if necessary. Furthermore, MKS possesses many sensors calibrated at NIST and PTB, which are used as calibration standards. In addition to wavelength variation, there are other possible sources of calibration error such as nonlinearity, variation with position on the surface and, for pyroelectric sensors, pulse frequency. All of these factors are taken into consideration and are accounted for in a sensor's calibration.

Additional online content discussing MKS' power and energy sensors can be found in [101-105].

4. MKS Ophir® LED Characterization

LEDs are used today in many applications including lighting of homes, streets, and businesses where they are beginning to supplant more traditional light sources. Despite their advantages, measurement techniques of the power, flux (equivalent to power but measured in units of lumens), and spectrum of LEDs is not very different from that of traditional light sources. The main reason for this is that these sources emit incoherent light, which radiates uniformly in all directions (see Sections I.C and I.B.3). These sources produce highly divergent light emission. Accurately measuring the power of such sources using a sensor with a fixed aperture size requires sophisticated collection geometries. The systems used by MKS to characterize the LED output are the subjects of this section.

Integrating Spheres

A power sensor with a finite aperture, such as one of the photodiodes discussed in Section II.A.3, is unable to measure the entire beam of a highly-divergent light source. Furthermore, estimating the fraction of power measured through the aperture is not as straightforward as it is for a laser with a well-defined beam size (see Section II.A.1). Instead, a light-collection device known as an integrating sphere can be used in conjunction with a sensitive photodiode for determining the power of a highly-divergent source such as an LED. As shown in Figure 58, an integrating sphere has its inner surface coated with a highly-reflective surface. When a divergent beam hits the walls of the integrating sphere, the light is reflected and scattered many times until the light hitting any place on the walls of the sphere has the same intensity. A sensor placed in the sphere thus gets the same intensity as anywhere else; the power the sensor detects is proportional to the total incident power independent of the beam divergence. To accomplish this, the sensor must be placed so that it only sees scattered light and not the incident beam. Since the integrating sphere divides the beam uniformly in the sphere, a photodiode can measure the fraction of power through its aperture and then determine the total beam power by knowing the ratio of the area of the aperture to that of the entire sphere. The integrating sphere also lowers the power entering the photodiode, so one can measure much higher input beam powers.

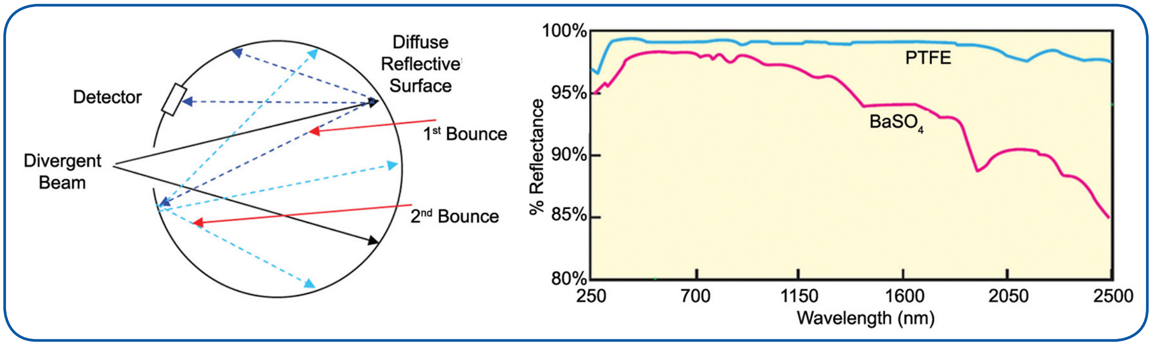


Figure 58. Multiple diffuse reflections of the light from a divergent beam inside an integrating sphere (left). The spectral reflectance of the materials coated on the surface of an integrating sphere (right).

Figure 59 shows MKS' integrating spheres which are designed to be configured in various ways for many applications. With an extensive line of accessories available, a single sphere can perform various integrating sphere tasks such as uniform illumination, light measurement, and reflectance measurement with reasonable accuracy. The spheres have a highly-reflective diffuse white coating (see Figure 58) for high efficiency and readings that are independent of beam size, position, and divergence. The spheres can accommodate light emission in the UV, VIS, and NIR spectral regions while handling input powers up to 30 W. The spheres also come in different sizes since the sphere must scale in size with the input beam aperture to ensure uniform scattering. The larger diameter spheres can be used with either divergent beams, e.g., LEDs, laser diodes, fiber output, or collimated beams from lasers (see Figure 59). This is accomplished by setting the optical geometry such that the input beam never directly hits the sensor and the sensor only sees rays reflected from the wall, as illustrated rightmost in Figure 59. To maintain accuracy, annual integrating sphere sensor calibration is recommended. Finally, any number of the thermopile or photodiode sensors discussed in Section II.A.3 could be suitable choices for LED measurement in combination with an integrating sphere.

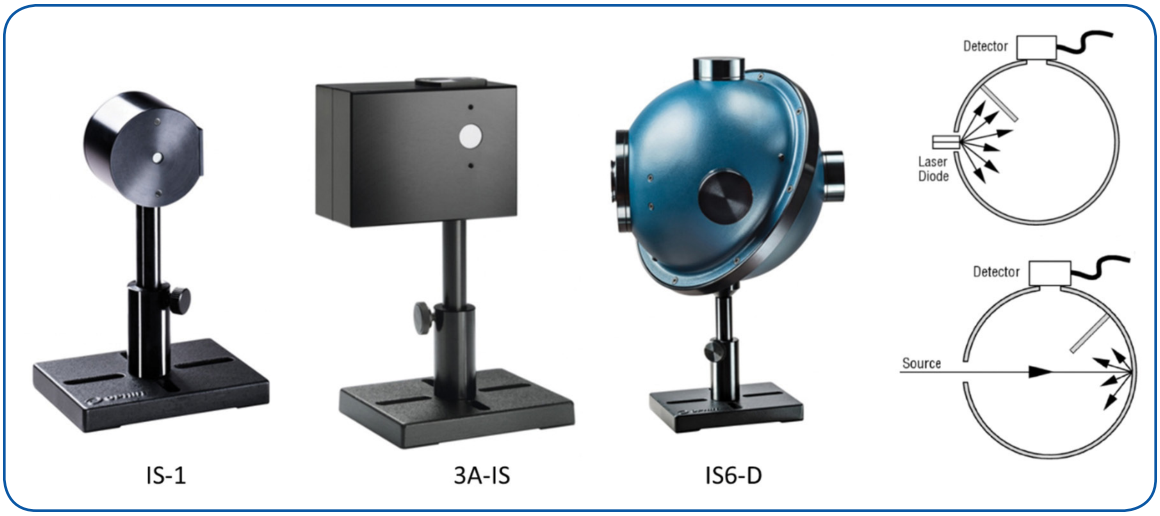


Figure 59. Ophir's integrating spheres for divergent beams with 1", 1.5" and 5.3" inner diameters (going from left to right). The larger diameter integrating spheres can be configured for divergent or collimated beams as shown on the far right.

FluxGage™

The integrating sphere is the standard instrument for measuring power, flux, and color of LED lights. However, the integrating sphere must be at least three times the size of the device under test (DUT). This ensures that the light from the DUT is uniformly scattered within the sphere prior to reaching the sensor and enables an accurate power estimate. For large LED luminaires or lamp assemblies that can approach 2-3 feet, this requires integrating spheres with diameters of 6-10 feet for proper characterization. Additionally, the integrating sphere must be recalibrated for each use to account for the absorption of the DUT itself. MKS developed the FluxGage™ measuring system (see Figure 60) to address these issues using a unique approach: employ detection surfaces which absorb and detect light. By eliminating the need for diffuse reflections, the characterization device can be the size of the DUT, e.g. rectangular, and be calibrated only once. The detection surfaces should be able to efficiently detect light, be independent of incidence angle, and absorb all the light with very little reflection. The FluxGage system's detection surface (see Figure 60) is made of solar panels to detect light, a sheet diffuser to make the panels insensitive to angle, and a dense array of pinholes printed with black ink over the diffuser. This arrangement ensures all three requirements are met.

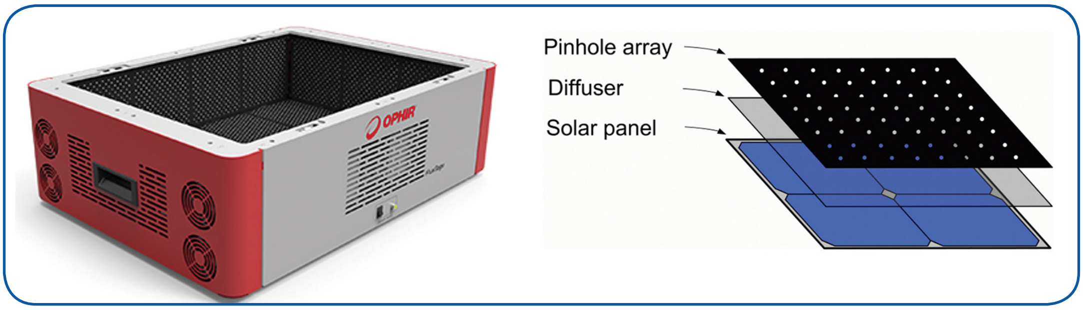


Figure 60. FluxGage measuring system chassis (left) and internal solar panels with diffuser and pinhole array (right).

The FluxGage system is therefore a compact measuring system for LED Luminaires. The system measures total flux or power using solar panels while an onboard spectrometer and fast photodiode complete the system and provide color and flicker (oscillation in power output) measurements. The benefits of using the FluxGage system are small size, robustness, production readiness, ease of use, and less frequent calibration. Increased demand for LED luminaires requires manufacturing operations to adapt to improved quality testing methods that can provide quality measurements while not slowing down the production lines. The value of using the FluxGage system is its ability to take key measurements quickly, but also to allow the manufacturing operation to test the entire luminaire as a fully performing, operational lamp.

Additional online content discussing MKS' LED characterization products can be found in [106-109].

B. Beam Characterization

Most laser-based applications depend on the irradiance or fluence, i.e., power or energy density, of the beam at the point of work. In other words, accurate measurement of the power or pulse energy is necessary for characterizing the laser beam, but it is not sufficient. To quantify the irradiance or fluence, the spatial distribution of the laser must be characterized. The attributes of this distribution include the beam width, its overall shape including amplitude modulation, and whether it is rotationally symmetric or elliptical. Furthermore, applications may utilize the beam at some distance from the output of the laser source or, more commonly, may modify the beam with an optical system. In either case, how the beam distribution changes as it propagates is also important. Such attributes of propagation include the beam divergence and its beam quality factor or M-Squared value. This section discusses the attributes of a laser beam as well as the beam profiling tools that enable its characterization.

1. Laser Beam Spatial Profiles

As discussed in Section I.A.2, the irradiance distribution of a laser beam is determined by the transverse modes that exit the laser cavity. Typically, the lowest-order transverse mode (TEM_{00}) is selected for emission since it propagates with the least beam divergence and can be focused to the tightest spot (see Section I.A.3). The irradiance distribution of this TEM_{00} mode is described by a Gaussian function and therefore much of this section details a Gaussian beam's profile as well as its evolution with distance. Gaussian beam propagation is well understood and even laser beams that do not possess a TEM_{00} mode are often described using a modified Gaussian mode analysis using its M-Squared value. Finally, a brief description of non-Gaussian beams is also given.

Gaussian Beams

The evolution of a Gaussian beam as it propagates along its axial direction (z) is shown in Figure 61. The irradiance has a radial (r) distribution that is circularly symmetric in any plane orthogonal to the axis and the beam's power (φ) is concentrated close to the axis. The functional form of this irradiance distribution (E) is given by:

$$E(r,z) = \frac{2\varphi}{\pi W^2(z)} exp\left[-\frac{2r^2}{W^2(z)}\right]. \tag{11}$$

This illustrates that the beam shape remains Gaussian at any point along the axis and changes only in its width and amplitude. The beam radius W is defined as the radius at which the irradiance decreases to $1/e^2$ or 0.135 of the peak on-axis ($r = 0$) value. This is sometimes referred to as the half-width $1/e^2$ (HW1/e^2) value. As shown in Figure 61, W gradually increases as the distance from the minimum beam radius (known as the beam waist, W_0) gets larger. Since E is the power per unit area, as indicated by Equation (11), the irradiance decreases as one moves away from the beam waist. An online calculator for estimating the power density or fluence of a laser beam can be found in [92]. Integration of the irradiance over the entire radial plane (at any axial position) results in the total optical power. In other words, φ remains constant along the axis. From a practical standpoint, integrating the irradiance within a circle of radius $1.5\times W$ results in 99% of the total power. This is relevant when measuring the optical power of a Gaussian beam as discussed in the aperture portion of Section II.A.1. A convenient calculator for determining the fraction of power through an aperture is given in [91].

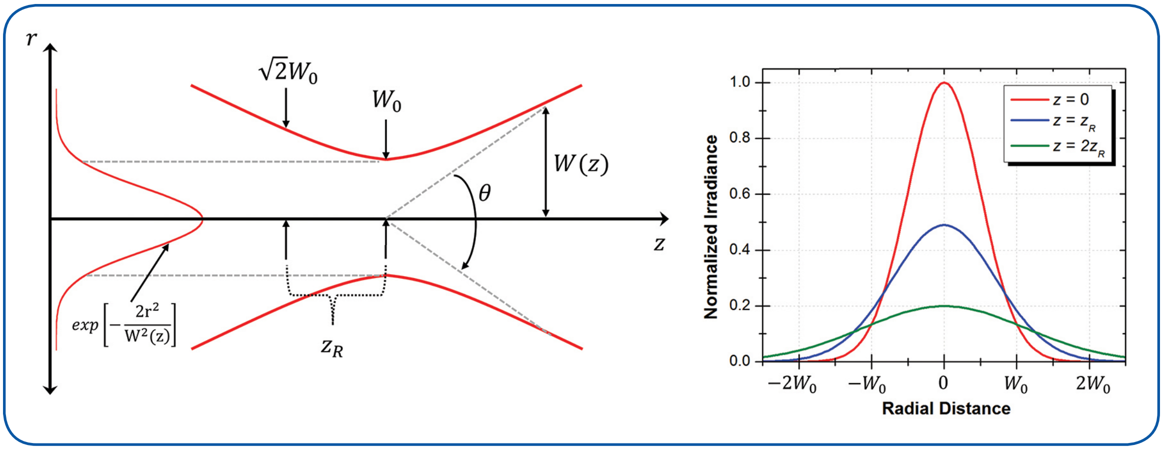


Figure 61. A diagram showing some of the parameters of a Gaussian laser beam (left, [3]). The normalized beam irradiance as a function of the radial distance (right) for different points along the propagation axis.

The equation describing the evolution of the beam radius along the axis where *z* is the distance from the beam waist is given by:

$$W(z) = W_0\sqrt{1+\left(\frac{z}{z_R}\right)^2} \quad where \quad z_R = \frac{\pi W_0^2}{\lambda} . \qquad (12)$$

z_R is known as the Rayleigh range and represents the distance from the waist where the radius increases by a factor of $\sqrt{2}$. If Equation (12) is substituted into Equation (11), then z_R is the distance at which the irradiance has decreased by a factor of ½ from its peak value at the waist. Twice the Rayleigh range is called the confocal parameter (or depth-of-focus) and is a rough estimate of the collimation of a beam. As shown in Equation (12), the beam size increases slowly with axial distance from the waist. As $z >> z_R$, the beam size then increases linearly with *z* with a slope of W_0/z_R. This slope represents the full angular width divergence (θ) of the beam given by:

$$\theta = \frac{4\lambda}{2\pi W_0}. \qquad (13)$$

A clear reciprocal relationship exists between the spot size ($2W_0$) and the divergence. In addition, for a Gaussian beam at a given wavelength, the product of the spot size and the divergence is a constant at $4\lambda/\pi$. This is important in defining how much a beam deviates from a perfect Gaussian beam.

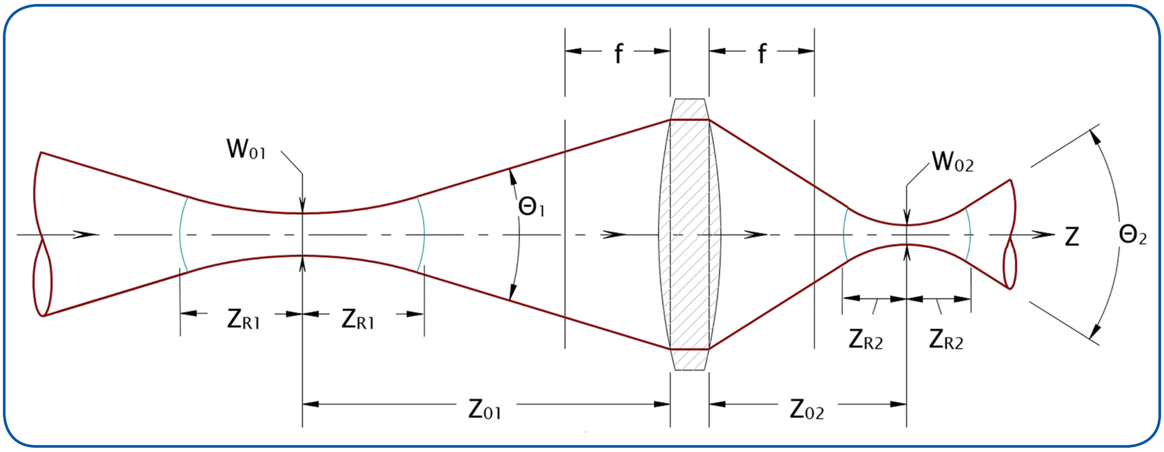


Figure 62. Propagation of a Gaussian beam through a thin lens where z_0 represents the position of the waist from the lens and the subscripts 1 and 2 refer to the original beam parameters and modified beam parameters, respectively.

Laser beams are often propagated through an optical system consisting of lenses, mirrors, dielectric interfaces, or other optical elements. Fortunately, under most conditions, a Gaussian beam remains a Gaussian beam after encountering these elements, ensuring that the propagation equations used above remain valid. An optical element modifies the input beam by changing the position and size of its beam waist. Knowledge of the input beam parameters and the properties of the optical element can be used to determine the new values using a variety of propagation methods [1, 13]. A thin focusing lens (see Section III.A.2 for details) is probably the most important optical element that impacts a laser beam. Figure 62 illustrates beam propagation through a thin lens resulting in a repositioning and resizing of the beam waist (as well as z_R and θ). An online calculator has been developed [110] that can determine these new beam parameters given the input beam parameters and lens focal length (f). When a collimated beam (i.e., $2 \times z_R >> f$) encounters a lens, the resulting beam waist is simply the product of the focal length and original divergence (see Equation (13)) with the waist positioned at the focal point of the lens. These values can also be determined using a simplified online calculator [111].

M-Squared Analysis

The TEM_{00} mode of even a well-designed laser system is not a perfect Gaussian beam. The M^2 ("M-Squared") analysis was developed to characterize the quality of a laser beam. That is, how close it is to an ideal Gaussian beam. M^2 is defined as the product of the spot size ($2W_M$) and divergence (θ_M) of the real beam divided by the spot-size-divergence product of an ideal Gaussian beam:

$$M^2 = \frac{2W_M \theta_M}{4\lambda/\pi}\,. \tag{14}$$

One practical consequence of this definition is that an ideal Gaussian beam (i.e., $M^2 = 1$) can be focused to a minimum spot diameter, whereas beams of higher M^2 values focus to larger spot diameters in proportion to the M^2 value. As a result, M^2 provides meaningful information about lasers, particularly if their application involves small focused spot sizes. While W_M and θ_M are sufficient for determining M^2, these values often cannot be measured directly. By focusing the beam with a lens of known focal length (like Figure 62), the characteristics of the artificially created beam waist and divergence can be measured. To provide an accurate calculation of M^2, the International Organization for Standards (ISO) requires at least 5 measurements in the focused beam waist region and at least 5 measurements in the far fields (two Rayleigh ranges away from the waist area), as shown in Figure 63. These multiple measurements ensure that the minimum beam width is found while a "curve fit" improves the accuracy by minimizing measurement error at any single point.

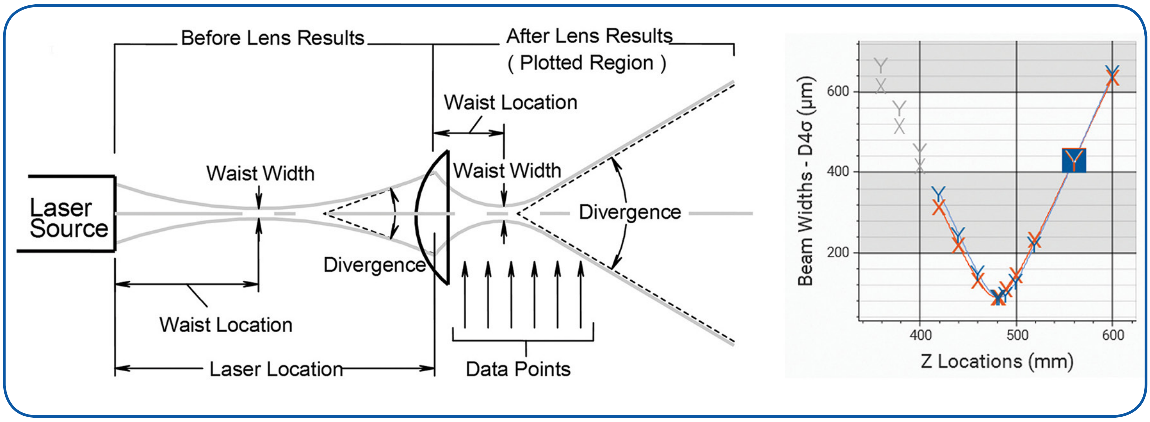


Figure 63. Experimental methodology for determining M-Squared value of a laser beam (left) and acquired data (right).

M^2 values for some laser systems are given in [1]. A commonly-available helium-neon laser emits a near-ideal Gaussian beam with a value of $M^2 < 1.1$. For many solid-state lasers, M^2 is in the range of 1.1–1.3. Collimated laser diodes that emit fundamental TEM_{00} modes possess M^2 of 1.1–1.7, whereas high-energy multimode lasers can generate M^2 factors as large as 10 to 100. Spatial filtering can improve beam quality using Fourier optics. This method will be discussed in Section III.C. Finally, additional beam quality metrics include the Beam Propagation Factor (K) where $K = 1/M^2$ as well as the Beam Parameter Product (BPP), which is defined as $BPP = M^2\lambda/\pi = 2W_M\theta_M$.

Non-Gaussian Beams

While many laser systems operate with near-Gaussian beams, other laser systems possess non-Gaussian beams that propagate differently and exhibit significantly different spatial distributions (see Figure 64 for examples). In some cases, a laser resonator emits a beam with a higher-order TEM_{mn} mode. Depending on the resonator geometry, these modes can be cylindrical in nature and are called Laguerre-Gaussian beams (see Figure 10) or rectangular and are called Hermite-Gaussian beams. The complicated amplitude distributions and axial propagation of these beams are described in [1, 10, 13]. In other cases, a laser beam is modified by an optical system to such an extent that its profile and propagation can no longer be approximated using the Gaussian beam analysis. Flat-top beams are one such example where a beam exhibits a nearly constant irradiance over its beam width (see Figure 64). Given the steep edges of the beam profile, the diameters of these beams are often characterized by their full-width at half-maximum (FWHM) values as opposed to the HW1/e^2 radius values used for Gaussian beams. Such flat-top beams are important for laser-based material processing where a constant irradiance provides more uniform material modification. The propagation of these beams can be quite complicated and is often encountered when a laser beam overfills a focusing objective in order to generate a very small spot size in high-resolution microscopy [18].

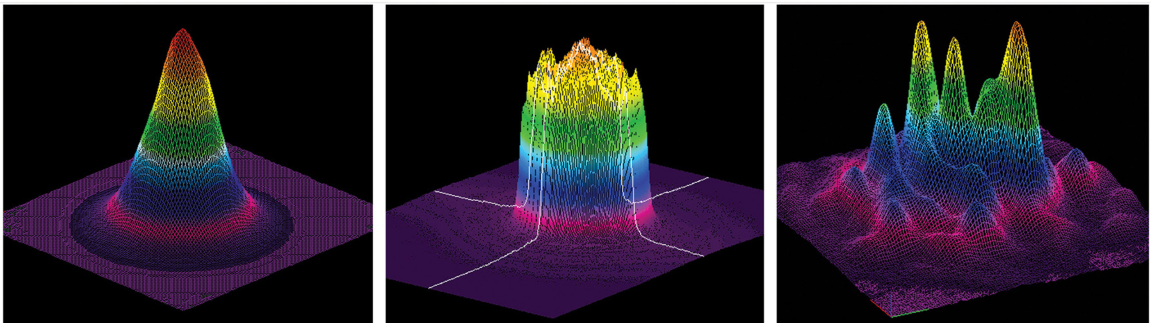


Figure 64. Examples of experimentally measured beam profiles showing a near-Gaussian beam shape (left), a flat-top beam (middle), and a highly-modulated multi-mode beam (right).

Additional online content discussing the characterization of a laser beam's spatial distribution can be found in [112-116].

2. Beam Profilers

Given the different types of laser beam spatial distributions and their evolution as they propagate in free space and through optical elements, measurement of the beam profile is critical to quantify the irradiance at the point of the application. Monitoring the beam profile can also identify laser beam anomalies. This facilitates optimal alignment of a laser cavity or beam delivery optics resulting in significant improvements in laser processing (see Figure 65, for examples). Devices which measure and analyze the spatial distribution of a laser beam are called beam profilers and come in many different types, each with its own advantages and challenges. The basic types include array-based or camera-based profilers, mechanical scanning apertures, knife-edge scanners, and other devices. Since camera-based and scanning aperture profilers are the most common techniques, they are the subjects of this section. Camera-based systems are generally photodiode-based CCD (charge-coupled devices) or CMOS (complementary metal-oxide-semiconductor) devices. However, there are applications using arrays with thermal sensors for detection of longer wavelengths. Scanning aperture systems combine a moving slit and single-element photodetector to reproduce a beam profile by spatially sampling it.

The laser and application dictate the criteria for selecting the best beam profiler [117, 118]. Criteria include laser wavelength, beam size, power, pulsed or *CW*, and how the profiler will be used. The wavelength dictates the type of sensor to be used with Si providing sensitivity in the VIS and NIR, Ge and InGaAs covering the NIR, and pyroelectrics covering the UV and IR. A profiler must have a large enough aperture to accommodate the size of the beam while smaller beams will be limited by the resolution of the profiler, i.e., pixel size for a camera-based profiler and slit size for a scanning profiler. Since every profiler has a sensor, they are susceptible to saturation and optical damage, which limits the allowable incident laser power. The power can be attenuated or reduced prior to reaching the sensor to accommodate higher powers while scanning slit profilers may not require attenuation since they only sample a small portion of the beam's power. Generally, scanning aperture profilers are only effective for pulsed systems above 1-2 kHz due to their mechanical nature while camera profiles can be used for either pulsed or *CW* sources. Camera profilers provide a true two-dimensional image of the beam and will show fine structure and hot spots, which a scanning slit may integrate out. Slit-based profilers will typically provide a more accurate beam size measurement and can measure very small beams directly.

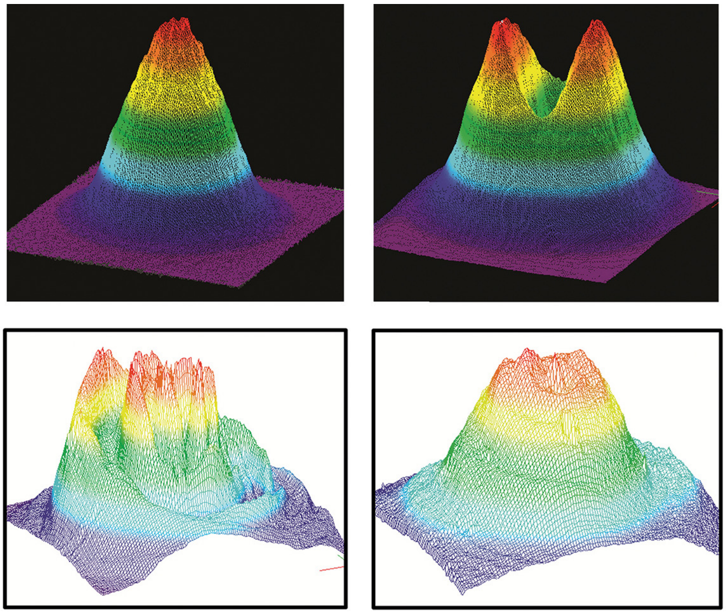


Figure 65. Measured laser beams using a camera-based beam profiler. Top row shows an industrial Nd:YAG laser exhibiting a Gaussian beam (left), which gives rise to a multi-peaked distribution following an increase in power of 1.7 times. Bottom row shows a CO_2 laser profile before (left) and after cavity alignment (right).

Camera-Based Profilers

Figure 66 shows a camera-based beam profiler system consisting of a camera, profiler software, and a beam attenuation accessory. The sensors discussed in Section II.A.2 are single, discrete devices. An array detector contains many sensors arranged in a dense grid pattern and, since all the sensors can respond to incident light simultaneously, these detectors can transform an optical image into an electronic signal to be measured and analyzed. Camera-based profilers use array detectors and the major differences between them lie in the type of sensor used, the sizes of the individual sensor elements or pixels, and the signal handling and readout methods. The sensors are typically semiconductor photodiodes whose spectral sensitivity depends on the type of semiconductor material employed. However, thermal detectors such as pyroelectric sensors and bolometers (a thermometric device which changes its electrical resistance as a function of absorbed radiation) are also used if sensitivity in the IR is desired. The camera consists of an array of photosensors or pixels. The spatial resolution of the camera depends on the pixel pitch, i.e., pixel-to-pixel distance, and this varies depending on the type of sensor. For Si photodiode sensors, the pitch can be below 5 µm while Ge and InGaAs sensors have pitches in the 10's µm range. Pyroelectric arrays have pitches that can reach slightly lower than 100 µm. The electronic signals from the pixels must then be read out. CCDs transfer the charge generated on each pixel to the adjacent pixel and so on until it reaches the edge of the array. The signals are therefore handled in a serial manner by the associated circuitry. By contrast, CMOS detectors read each pixel individually and therefore the signals are handled in parallel. This enables a much faster response for CMOS devices and, because they are based on widely used manufacturing technology, they are less expensive than CCDs. However, CCDs generally possess better linearity and noise characteristics.

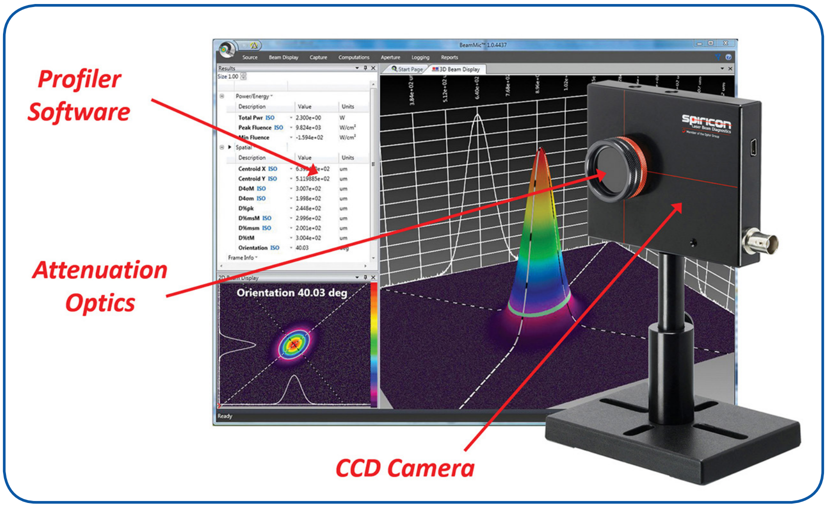


Figure 66. Camera-based beam profiler system consisting of a CCD camera, profiler software, and beam attenuation optics

While the individual photodiodes that make up most cameras have very large dynamic ranges (see Section II.A.2), this is reduced considerably for array detectors mainly due to the signal handling and readout. As a result, detector saturation occurs at much lower incident irradiances, and attenuation optics are often placed in the beam path prior to the detector. Beam profiling software, coupled to the camera, typically provides both image processing and data analysis capabilities. Image processing can include proper baselining, which helps eliminate background noise, electronic gain adjustment to maximize the dynamic range of the detector, and averaging, which can reduce noise. Data analysis should provide information on beam size, position, ellipticity (or deviation from a circularly symmetric beam), and real-time statistics of these values while also allowing for fitting of the beam profile to a Gaussian function. The benefit of a camera-based system is that the beam can be imaged directly and in real time to reveal not only the size and position of the beam but the distribution of the power/energy in the beam.

Scanning Slit Profilers

Figure 67 shows a scanning slit beam profiler in which a narrow slit moves through the beam and the light passing through the slit generates current in a single photodetector. Thus, as the slit scans through the beam, the detector signal is linearly proportional to the spatial beam irradiance profile integrated along the slit. A digital encoder provides accurate slit position. The photo-induced signal is digitized and analyzed to obtain the beam profile in the direction of translation of the slit. Typically, two orthogonal slits are sequentially moved through the beam providing profiles in these perpendicular directions. From these profiles, important spatial information such as beam width, beam position, beam quality, and other characteristics are determined.

A scanning profiler has several advantages over a camera profiler. Since only a single photodetector is employed, sensors with very different spectral sensitivities can be used, enabling profiling from UV to IR. A single detector also has a larger dynamic range compared to an array. Since the slit only allows a fraction of the beam to impinge on the detector, much higher powers can be measured without requiring additional attenuation. Because the slit plane is the measurement plane, it is also possible to measure the tightly focused beams with small spot sizes that cannot be directly measured with a camera. However, the scanning slit profiler only provides measurements along two orthogonal directions of a beam and therefore lacks the two-dimensional information available from a camera. Additionally, because the signal from the slit is integrated in one spatial dimension, the device is most appropriate for Gaussian-like laser beams. Due to the speed associated with the mechanical movement of the slit, the profiler can only measure *CW* and relatively high frequency pulsed beams. Beams with low pulse repetition rates, < 1 kHz, should be measured with a camera.

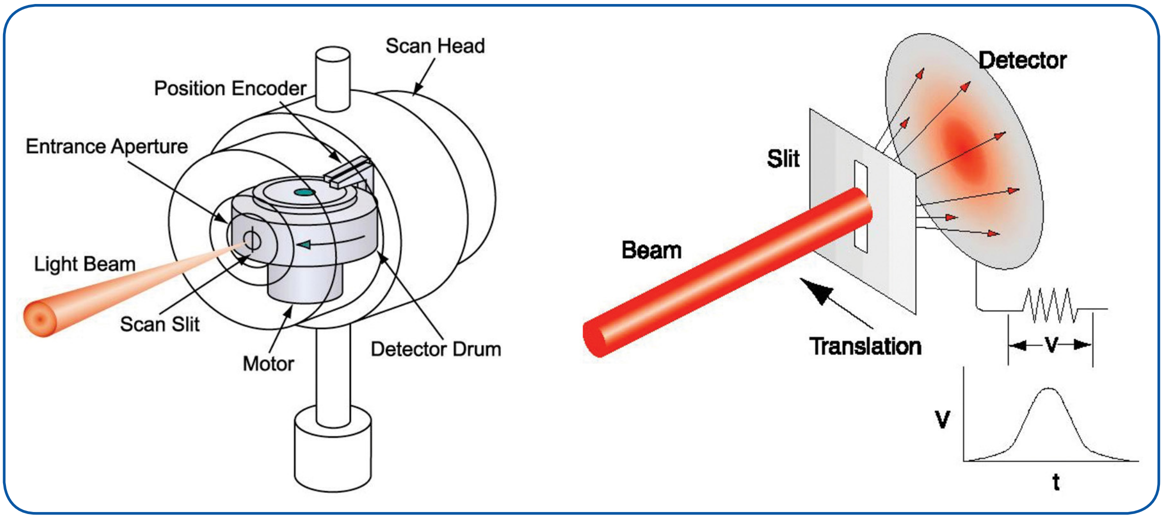


Figure 67. Scan head for scanning slit beam profiler (left) and enlarged view of scanning slit that shows recording of beam profile (right).

Additional online content discussing beam profilers can be found in [119-122].

3. MKS Ophir® Products for Beam Profiling

Accurate knowledge of a laser beam's size, shape, uniformity, divergence, and mode content is critical to the success of any laser-based endeavor. MKS Ophir is a world leader with the largest installed base of laser beam profilers and can provide solutions from the deep UV to the far IR. The beam profiling systems are highly flexible and are used for production environments as well as for rigorous scientific research. In addition to providing the most accurate measurements possible, MKS invests significantly in research and development to ensure its beam profiling products are precise, easy to use, and reliable. Whether choosing a camera or scanning slit system the user must first determine the laser beam measurement environment and what measurements are most important to the success of the application. In addition to knowledgeable product specialists, MKS provides a beam profiler finder [123] that allows a user to determine the profiler that best suits his/her application. A few select beam profiler systems offered by MKS are described next.

Pyrocam™ IIIHR Beam Profiling Camera

Figure 68 shows the Pyrocam camera consisting of a pyroelectric detector array with a very broadband coating enabling operation at essentially all UV and IR laser wavelengths, i.e., 13 nm – 355 nm and 1.06 µm – 3000 µm. The solid-state array has a wide dynamic range (1000:1) and possesses good spatial resolution (80 µm pixel pitch) for accurate profiling. An integrated chopper, i.e., an optical shutter, enables profiling of *CW* beams and thermal imaging. Since the camera operates in *CW* or pulsed mode, it is ideal for analysis of NIR, CO_2, and THz sources. The Pyrocam camera comes bundled with state-of-the-art BeamGage laser beam analysis software for quantitative analysis and image display. The software performs rigorous data acquisition and analysis of laser beam parameters, such as beam size, shape, uniformity, divergence, mode content, and expected power distribution. It runs real-time statistical analysis on all calculated results that can be customized based on the user's needs. The software provides auto-setup and auto-exposure capabilities for fast set-up and optimized accuracy. Finally, the BeamGage software is based on a patented baseline correction algorithm (UltraCal) that helped establish the ISO 11146-3 standard for beam measurement accuracy.

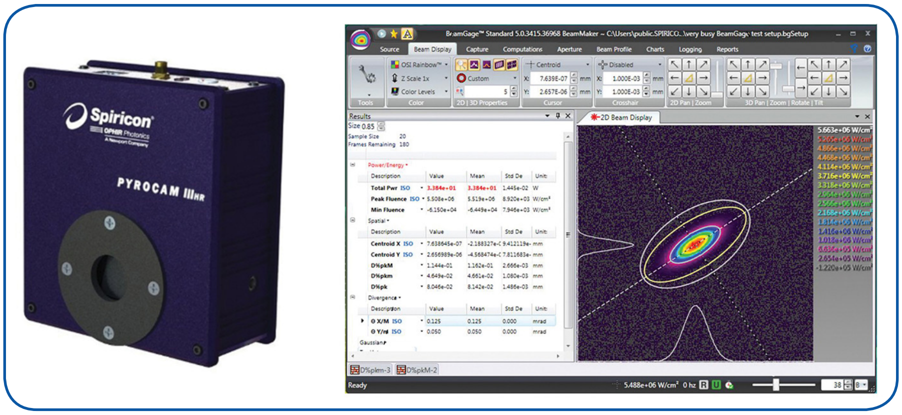


Figure 68. Pyrocam™ IIIHR pyroelectric array camera for beam profiling (left) and state-of-the-art beam profiling software BeamGage® (right).

NanoScan™ 2s Scanning Slit Beam Profiler

Figure 69 shows the NanoScan™ 2s scanning slit beam profiler that provides sub-micron precision for measuring beam position and size. The scan head contains a drum that spins at a 45° angle and possesses two slits on it, one to measure each axis of the laser. At the core of the head there is a single detector element, which measures just the portion of the laser that the slit allows to pass. The entire beam profiling process is done with one drum revolution in less than a second, but the profiler continues to operate (and the drum keeps spinning) to profile the laser beam in real time. The scan head is available with Si or Ge photodiodes, or a pyroelectric detector, which allows it to profile lasers of any wavelength from UV to the far IR over a wide range of laser power levels. The NanoScan 2s slit profiler is the most versatile laser beam profiling instrument available today, providing instantaneous feedback of beam parameters for *CW* and kHz pulsed lasers with measurement update rates to 20 Hz. The natural attenuation provided by the slit allows the measurement of many beams with little or no additional attenuation. The high dynamic range makes it possible to measure beams while adjustments to focus are made without having to adjust the profiler. The NanoScan 2s software includes an extensive set of ISO quantitative measurements, an M^2 Wizard, and the ability to measure laser power. It also allows users to integrate the NanoScan 2s software into OEM systems or to create their own user interface screens.

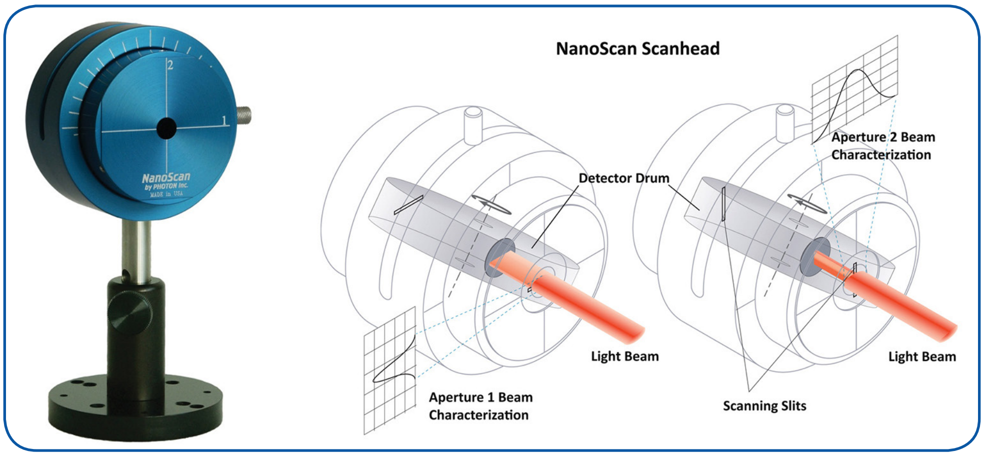


Figure 69. NanoScan beam profiler (left) and internal workings of scan head that enables measurement of orthogonal beam profiles in real time (right).

BeamSquared™

Figure 70 shows the BeamSquared™ system. It is a compact and fully automated tool for measuring the propagation characteristics of *CW* and pulsed laser systems from the UV to NIR. The BeamSquared optical train uses a fixed position lens with movable mirrors and a camera. The mirrors that direct the focused beam into the camera are moved to precise locations, translating the beam through the near field, the waist, and the far field regions. All these measurements and translations, as well as incremental beam attenuation, are automatically controlled by the BeamSquared software. Users can also measure wavelengths above 1.8 μm, including CO_2 and THz in manual mode with a Pyrocam IIIHR camera. The longer optical train and patented Ultracal calibration makes the BeamSquared system the most accurate product on the market and is ISO 11146 compliant. Its operational robustness and reliability ensures continuous use applications in industry, science, research and development. Design improvements in the BeamSquared system have decreased the measurement reporting time, making it possible to report M^2 (as well as *K* and BPP) in under a minute.

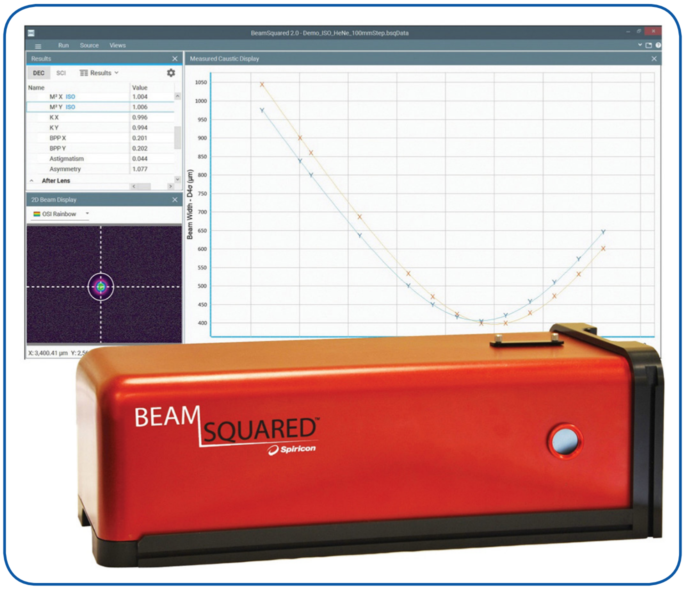


Figure 70. Camera-based M-Squared measurement system called BeamSquared (bottom) and acquired data set consisting of measurements of beam waist versus axial position (top).

Additional online content discussing the MKS' beam profiler products can be found in [124-126].

C. Photoreceivers

Photoreceivers are based on the same types of semiconductor photodiodes used for power sensors (see Section II.A.2) and camera-based beam profiling (see Section II.B.2). Since these applications are mainly focused on steady-state (quasi-*CW*) optical response or, for a pulsed source, the time-integrated response, the actual temporal response of the optical signal is not critical. However, many applications require the monitoring of rapid events and therefore, high-speed optical detection is necessary. The characteristics of such high-speed photoreceivers as well as their associated applications are the subjects of this section. Photoreceivers focused on sensitive detection of optical signals at low-light levels or those that exhibit small differential changes are also addressed.

1. Photoreceiver Characteristics

Since photoreceivers are semiconductor photodiodes, they possess the same intrinsic advantages described in Section II.A.2, e.g., large detectivities and very wide dynamic ranges. They also share many common parameters that must be considered when selecting a detector for any application. These include responsivity, spectral sensitivity, linearity, power handling, bias voltage, and power consumption. However, due to the nature of the specific applications for photoreceivers, there are unique characteristics as well. The attributes discussed in this section include rise time and bandwidth, methods for noise level reduction, as well as detector gain and active area.

Response Time

The conversion of photons to electrons is very rapid in photodiodes, which enables them to follow fast changing radiation levels. To understand this response time, it is instructive to revisit the operation of a reversed bias *p-n* junction (shown in Figure 71) that makes up a photodiode [1]. Electron-hole pairs that are generated in the depletion region experience a strong electric field that causes them to rapidly drift in opposite directions. Charge carriers generated near the depletion layer can end up there following random diffusion and contribute to the external electric current. Carriers generated farther away from the depletion layer are highly likely to recombine and therefore not contribute to the current. The overall response time for current generation depends on the transit time it takes for carriers to drift across the depletion layer as well as the time for carriers to diffuse into this region. Diffusion is a relatively slow process compared to drift and so modifications to the *p-n* junction are often implemented to reduce diffusion (see Section II.C.2 on high-speed photoreceivers).

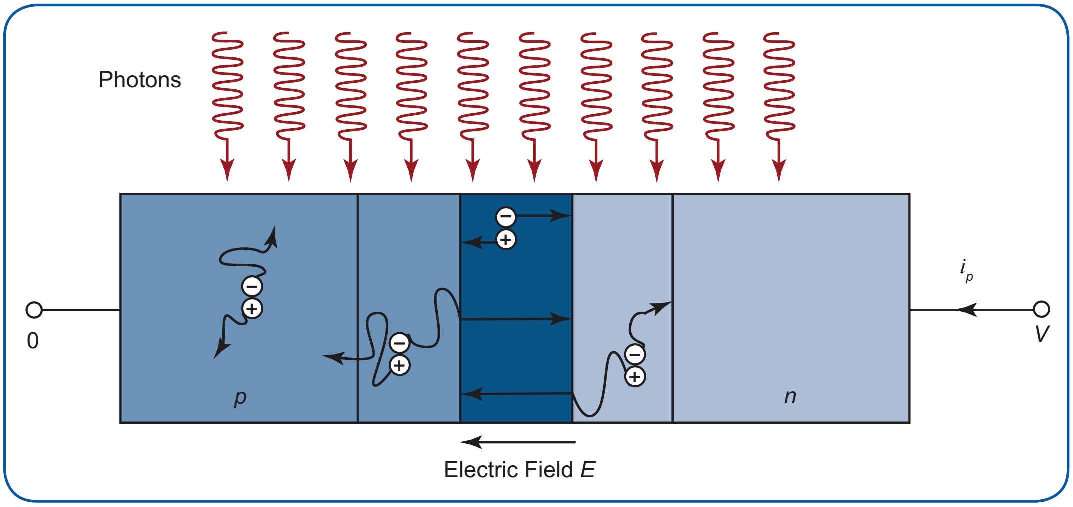


Figure 71. The operation of a reversed bias p-n junction of a photodiode when irradiated by incident photons. Drift of the carriers occurs in the depletion layer (central dark blue region) while diffusion primarily occurs just outside this region [1].

The *p-n* junction of the photodiode is effectively a capacitor with capacitance *C*. When the photodiode is connected to a circuit with a load resistance of *R* (typically 50 Ω), the combination of resistance and capacitance effectively integrates the current. This lengthens the response time by a value known as the *RC* time constant. Therefore, the overall response time of a photodiode depends on the transit and diffusion times as well as the *RC* time constant. In Section II.A.2, the benefits of operating under reverse bias for a photodiode power sensor were stated as greater light collection and a larger quantum yield. However, reverse bias operation also offers advantages in terms of reducing response time [1]. The bias creates a strong electric field in the junction that increases carrier drift velocity leading to a reduced transit time. Also, the increased depletion width reduces the junction capacitance and the associated *RC* time constant. These benefits in terms of reduced response time are why most high-speed photoreceivers operate in reverse bias.

Rise Time and Bandwidth

For measurements in the time domain, the response time of a photoreceiver is quantified by its impulse response. This is essentially the shortest pulse one could expect to see from the detector and is often quantified in terms of the FWHM of the pulse in the time domain. More commonly, the parameter specified for a photoreceiver is the rise time, which is the integral of the impulse response. Figure 72 shows an example of a high-speed photoreceiver's impulse response with a corresponding sub-20 ps rise time. Rise time is the appropriate parameter to consider in digital communications systems where bit streams are comprised of an endless series of rising and falling edges. Generally, the rise time of a detector should be at least three times shorter than the rise time you expect to measure. Sometimes, it is necessary to know the required frequency bandwidth for time-domain measurements. A good rule of thumb is to use a detection 3-dB frequency bandwidth (the frequency at which the power falls off to 50% of the value at DC) greater than $0.44/\tau$, where τ is FWHM of the impulse response.

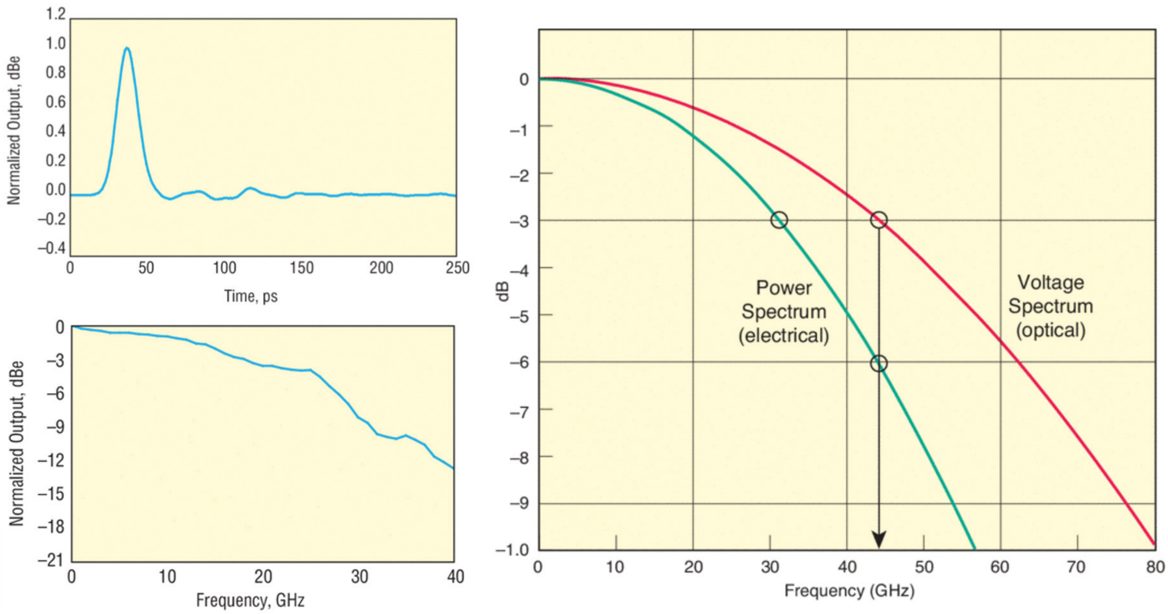


Figure 72. Impulse response of high-speed photoreceiver measured with 18.5 ps rise time (left, top) and a 3-dB power bandwidth of 20 GHz (left, bottom). Comparison between an idealized detector's power bandwidth and voltage bandwidth (right).

For measurements in the frequency domain, such as laser heterodyning or optical modulation experiments, the photoreceiver should have a flat response out to the highest frequency of interest. This is typically quantified in terms of the 3-dB frequency bandwidth. The high-speed photoreceiver discussed above has a frequency bandwidth of 20 GHz as shown in Figure 72. There are many terms used to describe the bandwidth of a photoreceiver, which can lead to confusion when making detector comparisons. The "power bandwidth" (also known as the "electrical bandwidth") is the 3-dB point of the power spectrum and is the most common value reported. The "voltage bandwidth" (sometimes called the "optical bandwidth"), or the 3-dB point of the voltage spectrum, is offset to higher frequencies than the power bandwidth since power is proportional to the square of the voltage. As shown in Figure 72, this implies that the 3-dB voltage bandwidth is always greater than the 3-dB power bandwidth for the same photoreceiver.

Noise Reduction

The sources of noise in photoreceivers and the associated parameters describing noise characteristics were discussed in Section II.A.1. Here, methods to minimize the photoreceiver noise are addressed. These approaches essentially involve choosing the best match of detector response to signal characteristics. The regimes in which detector response can be optimized include time, wavelength, bandwidth, and active area. When measuring pulsed optical signals, gated averagers or integrators are used to improve the SNR. These devices take their names from the gates, boxes, or windows that are used to define the times during which the electronics acquire the signal (see Figure 73). During these gated times, the SNR is improved since noise contributions which would normally accumulate during the non-gated times are absent. If the process is repeated N times, the SNR will improve by a factor of $N^{1/2}$ for most types of noise. The spectral sensitivity of the photoreceiver must cover the wavelength range of the signal. However, if the detector's wavelength sensitivity range exceeds that of the signal, any radiation outside of the signal range will contribute to noise. The use of a bandpass filter (see Section III.A.2) to narrow the range of wavelengths seen by the detector can reduce this undesired noise. The level of noise reduction depends on how optimal the spectral overlap is between the filter and the signal, as shown in Figure 73.

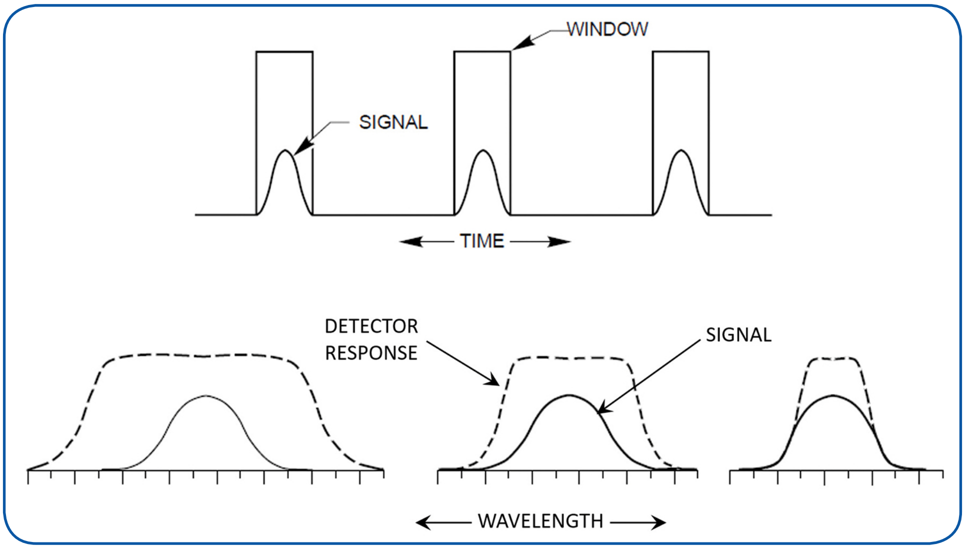


Figure 73. Introduction of gated windows to reduce noise in the time domain (top). Matching of detector responsivity to signal to reduce noise in the spectral domain (bottom) with efficiency increasing from left to right.

Frequency bandwidth can also be adjusted to reduce noise. AC detection methods use electronic and digital filters to match the detection bandwidth to the signal's frequency spectrum. The signal is discriminated from the noise contribution, which will have its power spread over a different and wider band of frequencies. The signal may have its own natural modulation frequency around which the filter can be centered or a known modulation can be applied to the signal. This can be accomplished through optical modulation using a chopper or by electronic modulation. Then, an AC-coupled amplifier with a narrow band filter centered on the modulation frequency helps significantly increase the SNR. More sophisticated approaches involve the use of lock-in techniques where the amplifier actively tracks the modulator frequency, allowing much narrower band filtering to be used.

Active Area

The benefit of a large area photodetector in the context of power measurements was discussed in Section II.A.1. Similarly, alignment of a large laser beam into a photoreceiver is made significantly easier if the detector has an active area ranging from several mm^2 to cm^2 (see Figure 74, for example). However, since noise is proportional to the area of a detector, such large-area photoreceivers are noisy. Reduction of the active area of a photoreceiver is therefore often desired for improved SNR. This is typically accomplished using optical fibers (see Section III.B) to deliver the light directly to the active area (see Figure 74). The active area is essentially matched in size to the core of the fiber, which have diameters of 50-60 µm for multimode fibers and less than 10 µm for single-mode fibers. Reducing the active area also increases response speed. This is a by-product of the reduced junction capacitance, which in turn reduces the response time. Due to the benefits of reduced noise and response time, many high-speed photoreceivers are fiber-coupled.

Figure 74. Examples of large area photoreceiver for free-space applications (left) and small area fiber-coupled photoreceiver for high speed applications (right).

Gain

As discussed in Section II.A.2, a photodiode has a high quantum yield or probability that an incident photon will produce a detectable charge carrier pair. These carriers then produce a single charge in the photoreceiver circuit. In other words, photodiodes do not exhibit intrinsic gain like some other detectors (see description of low-light photoreceivers below). However, a photoreceiver with an electronic amplifier will exhibit gain defined as the average number of circuit electrons generated per photocarrier pair. A more commonly reported parameter for a photoreceiver is the maximum conversion gain which describes the resulting output voltage (V) for a given input optical power (W). Given in V/W, the conversion gain is the product of the photodetector's responsivity, the amplifier's gain, and the input impedance. For an unamplified photodetector, the conversion gain is the product of the photodetector's responsivity and the load impedance. Photoreceivers also possess a fixed gain-bandwidth product such that an increase in gain results in a decrease of the bandwidth, and vice versa. This trade-off between sensitivity and frequency response is associated with the time required for the gain process to take place.

Additional online content discussing photoreceiver characteristics can be found in [127-130].

2. Types of Photoreceivers

The main types of photoreceivers discussed below are those that can detect optical signals with fast temporal responses or those containing high frequency components as well as detectors that are sensitive to low light levels or small differential changes in signals. Depending on the type of semiconductor used, the detectors can have spectral sensitivities anywhere from the UV region of the spectrum to the NIR. Furthermore, depending on the application being targeted, they can possess either free-space or fiber-coupled configurations.

High-Speed

With the advancement of high-transmission-rate systems and short-pulse lasers, many applications now require high time-resolution or equivalently, high frequency-bandwidth optical detection. High-speed photoreceivers are critical for the measurement of the frequency and/or time response of optical systems. In the optical domain, this can include measuring the pulses of mode-locked laser systems, detecting the data stream of a frequency-multiplexed communication system, or providing increased resolution in dynamic, pump-probe spectroscopy. The minimum rise time for high-speed photoreceivers is less than 10 ps. Consequently, for optical signals with faster responses, optical gating techniques are required (see Section II.E). In the frequency domain, applications for high-speed photoreceivers include laser heterodyning experiments and millimeter-wave signal generation. The maximum frequency bandwidths for such detectors can exceed 50 GHz in well-designed devices.

Based on the discussion above, diffusion of carriers to the depletion region is a relatively slow process in reverse biased *p-n* junctions that could serve to limit the response time of a photoreceiver. To minimize this effect, a *p-i-n* photodiode is typically utilized [1, 131] where an un-doped intrinsic layer is sandwiched between the *p* and *n* layers in a *p-n* junction (see Figure 75). This structure effectively widens the depletion layer. This results in a greater proportion of the generated current being carried by the faster drift process instead of diffusion. The increased depletion width also allows for a reduction in the *RC* time constant (via a decreased junction capacitance) and increased area for capturing light. The *p-i-n* device structure is ubiquitous in high-speed photoreceivers and enables fast rise times and large bandwidths. However, the final measured optical signal will be as slow as the slowest component of a detection system even if a sufficiently fast photoreceiver is employed. Therefore, care should be taken when choosing connectors, cables, an oscilloscope, and a spectrum analyzer to measure a fast optical signal.

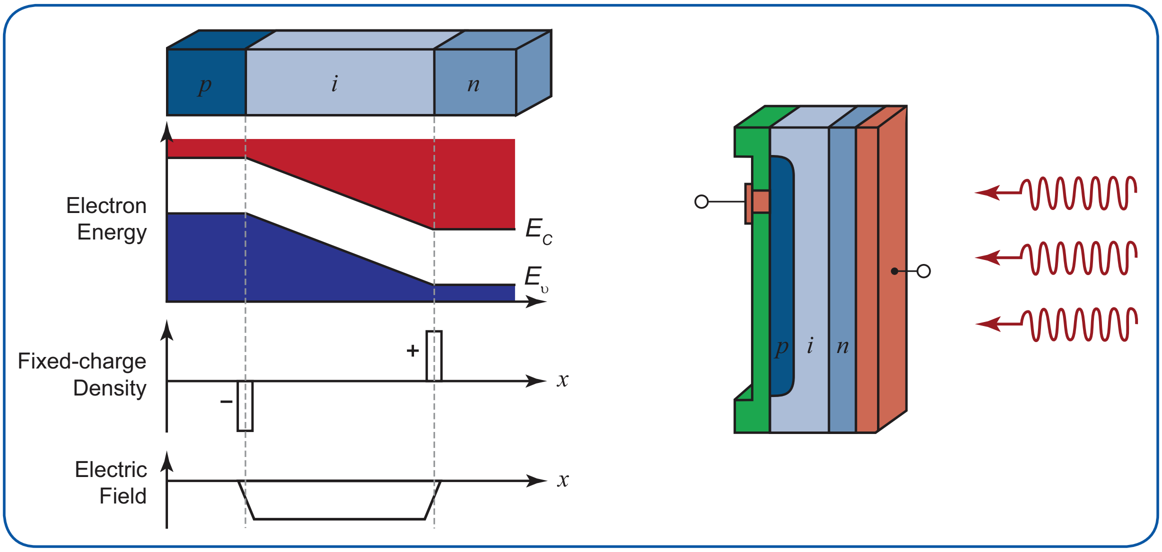


Figure 75. The p-i-n photodiode structure, energy-band diagram, charge distribution, and electric-field distribution (left). The device being illuminated perpendicular to the junction (right) [1].

Balanced Detection

Balanced photodetection is a commonly used detection method to increase the SNR of a signal beyond simple amplification. It is particularly powerful in its ability to cancel laser noise, i.e., common mode noise, and to detect small signal fluctuations on a large DC signal. The simplest balanced detector uses two photoreceivers connected so their photocurrents cancel. The output of the detector pair is zero until there is some difference in the intensity of one of the beams, which causes the pair to become unbalanced and a net signal appears on the output. In practice, this is accomplished by using an auto-balancing circuit that includes a low-frequency feedback loop to maintain automatic DC balance between the signal and reference arms (see Figure 76). This balanced optical receiver permits detection of a small signal with a large background. One application of this balanced detection method is shown in Figure 76 where dynamic changes in a material's optical properties can be measured. Another device that employs balanced detection is a quadrant-cell photoreceiver, which consists of four individual yet identical photodetectors positioned very close to each other. Signals from pairs of these detectors are used to generate differential signals in the horizontal and vertical directions while the sum of all quadrants is provided for normalization purposes. This photoreceiver is ideal for measuring deviations in the position of a laser beam such as that required in beam-stabilization systems.

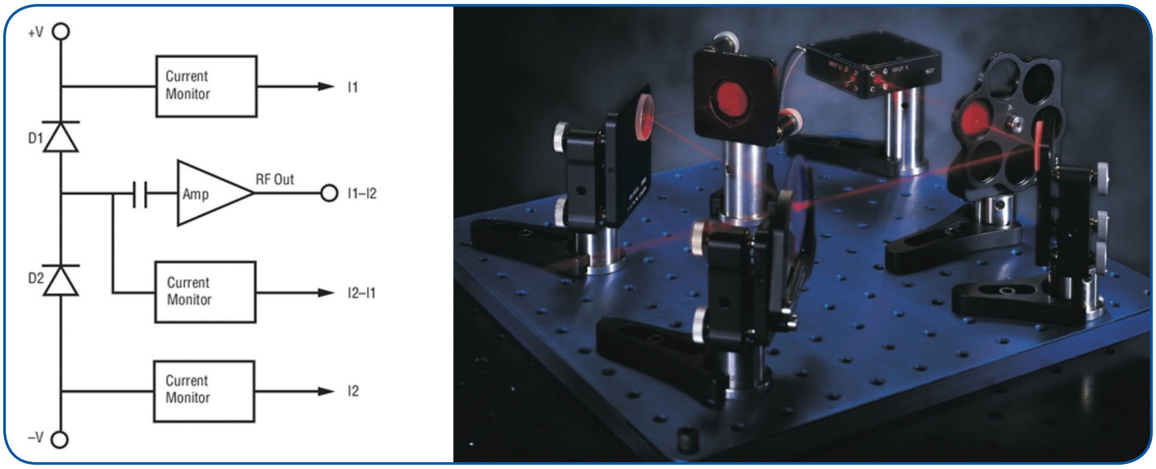


Figure 76. Functional circuit diagram for a balanced photoreceiver (left). Example of a balanced photoreceiver being used to measure absorption properties where a probe beam is focused onto one of the photodetectors while the other detects a reference beam (right).

Low-Light

Certain applications involving spectroscopic or fluorescence measurements require photoreceivers than can detect low-light-level signals. One approach involves utilizing photoreceivers with large conversion gains (up to $2x10^{11}$ V/W). Such large gains are achieved through careful selection of the amplifier-resistor pair, where a large feedback resistor provides the high-gain values while an ultra-quiet amplifier keeps noise to a minimum. By using these photoreceivers in conjunction with an optical chopper and lock-in detection methods, sensitivity levels in the femtowatt range can be achieved. For experiments that require ultra-low light levels, e.g., those requiring only a few photons, a different type of photodiode known as an avalanche photodiode (APD) can be used. This photodiode has a junction so strongly reverse-biased that charge carriers generated in the junction acquire sufficient energy to excite new carriers by a process known as impact ionization. In this way, weak light can generate a current large enough to be detected by the accompanying circuit electronics. Another detector capable of detecting light at the single photon level is a PMT which is not a photoconductive detector like a photodiode, but rather a photoemissive detector. Details about the functional operation of APDs and PMTs can be found in [1, 132].

Additional online content discussing types of photoreceivers can be found in [133-136].

D. Dispersive Spectrometers

A spectrometer measures the spectral distribution of a light source. All light sources emit radiation over a range of wavelengths from a narrowband laser to a high-power LED or broadband incoherent lamp. Spectrometers are critical for the spectral characterization of light sources, as analytical instruments in the field of spectroscopy (for details see Chapter 6, Section III), to enable high-bandwidth telecommunications, and to facilitate extraterrestrial study in both ground- and space-based telescopes. A spectrometer's two principal components are the system that allows for spectral discrimination and a photodetector for sensing the exiting light. Most spectrometers are dispersive, which means they spatially separate the wavelengths of the light prior to detection. The most common dispersive spectrometers are the monochromator and the spectrograph, which are discussed below. The Fourier Transform IR (FTIR) spectrometer, which is an interferometric technique that is used primarily in the MIR, is discussed in Chapter 6, Section III.B.2. The key component of a dispersive spectrometer is the diffraction grating. The underlying physics behind its operation as well as the types of gratings are the initial subjects of this section.

1. Diffraction Grating Physics

When light encounters an obstacle such as an opaque screen with a small opening (or aperture), the intensity distribution behind the screen can look much different than the shape of the aperture that it passed through. Since light is an electromagnetic wave, its wavefront is altered much like a water wave encountering an obstruction. This diffraction phenomenon occurs because of interference (see Section I.A.3 on coherence for details) between different portions of the wavefront. The resulting intensity distribution is called a diffraction pattern. Similarly, when light passes through an opaque screen consisting of multiple elongated apertures (or slits) with a fixed spacing between them, the emerging wavefronts constructively interfere to produce a diffraction pattern with intensities peaked in certain directions as shown in Figure 77. These directions are strongly dependent on both the slit spacing and wavelength of the incident light. Consequently, surfaces with well-defined slit locations can be used to direct light of certain wavelengths into specific directions.

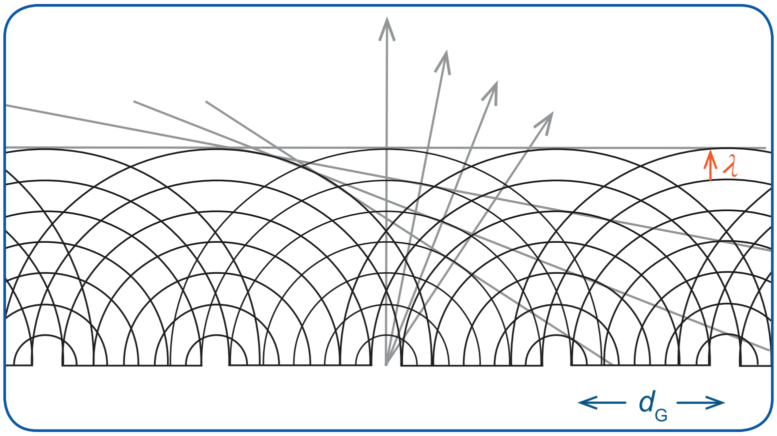


Figure 77. Diffraction of monochromatic light with wavelength λ from a series of apertures with a spacing d_G. The angled lines indicate regions of constant phase while arrows denote directions of intensity peaks in the diffraction pattern.

A diffraction grating is essentially a multi-slit surface. It provides angular dispersion, i.e., the ability to separate wavelengths based on the angle that they emerge from the grating. Gratings can be transmissive, like the multi-slit aperture, but they can also be reflective where the grooved surface is overcoated with a reflecting material such as aluminum. A typical diffraction grating (see Figure 78) consists of a large number of parallel grooves (representing the slits) with a groove spacing (denoted d_G, also called the pitch) on the order of the wavelength of light. This is more commonly reported as the groove density (G), which is the reciprocal of d_G, e.g., typical gratings have G values between 30 and 5000 grooves per mm. The groove spacing determines the angles at which a single wavelength will constructively interfere to form diffracted orders (see below), which are equivalent to the intensity peaks shown in Figure 77. In addition to the spacing of the grooves, the groove profile (see Figure 78) plays a key role in the performance of a grating. When monochromatic light strikes a grating, a fraction of it is diffracted into each order (termed its efficiency). Maximizing the efficiency into a single order, typically the first order, is often desired to ensure increased light collection. To optimize this efficiency for a single wavelength, a procedure known as blazing is performed. This involves modifying the groove profile, including facet angles, shapes and/or depths. The blaze wavelength is the wavelength for which the grating is most efficient.

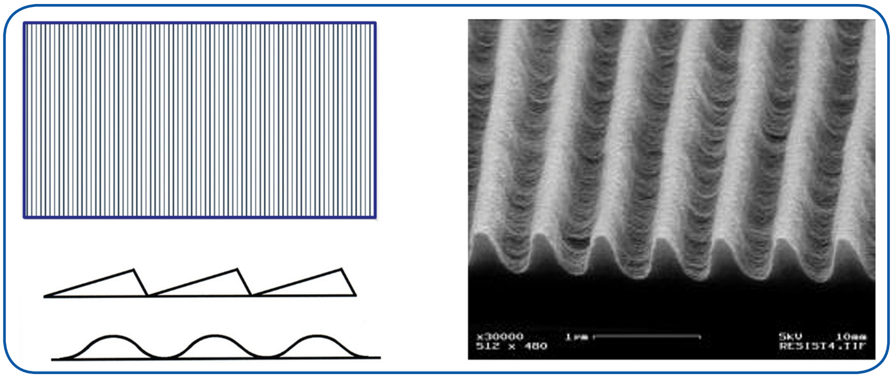


Figure 78. Depictions of top-down view of diffraction grating showing groove pattern (left, top) and side view showing different groove profiles (left, bottom). Scanning electron microscope image of diffraction grating (right).

Grating Equation

The basic grating equation determines the discrete directions into which monochromatic light of wavelength λ is diffracted. The equation is shown below and details of its derivation can be found in [137, 138]:

$$m\lambda = d_G(sin\alpha + sin\beta_m). \quad (15)$$

Figure 79 illustrates this diffraction. Light of wavelength λ is incident at an angle α and diffracted by the grating (with a groove spacing d_G) along a set of angles β_m. These angles are measured from the grating normal, which is shown as the dashed line perpendicular to the grating surface at its center. If β_m is on the opposite side of the grating normal from α, its sign is opposite. In Equation (15), m is the order of diffraction, which is an integer. For the zeroth order (m = 0), α and β_0 are equal and opposite, resulting in the light simply being reflected, i.e., no diffraction. The sign convention for m requires that it is positive if the diffracted ray lies to the left (counter-clockwise side) of the zeroth order and negative if it lies to the right (the clockwise side). When a beam of monochromatic light is incident on a grating, the light is simply diffracted from the grating in directions corresponding to m = -2, -1, 0, 1, 2, 3, etc. When a beam of polychromatic light is incident on a grating, then the light is dispersed so that each wavelength satisfies the grating equation as shown in Figure 79. Usually only the first order, positive or negative, is desired and so higher order wavelengths may need to be blocked. In many monochromators and spectrographs, a constant-deviation mount is used where the wavelength is changed by rotating the grating around an axis while the angle between the incident and diffracted light (or deviation angle) remains unchanged.

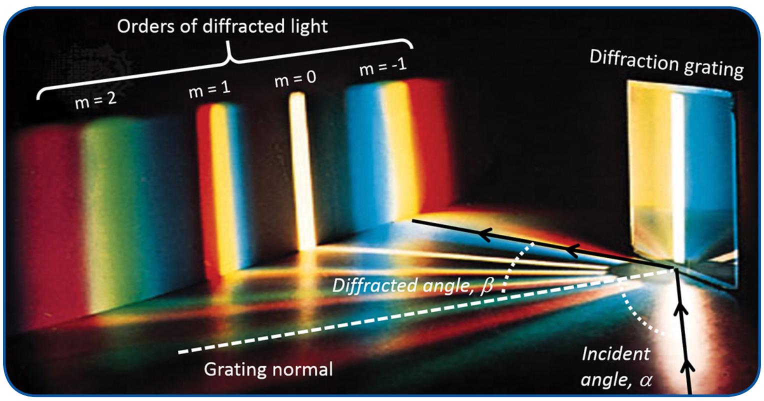


Figure 79. Polychromatic light diffracted from a grating.

Dispersion, Bandpass, and Resolution

By fixing the incidence angle α in the grating equation and differentiating with respect to λ, the angular dispersion (D) or change in diffraction angle per unit wavelength can be determined as:

$$D = \frac{\partial\beta}{\partial\lambda} = \frac{m}{d_G cos\beta_m} = \frac{Gm}{cos\beta_m}. \quad (16)$$

For a given order m, D represents the ability to discriminate between signals at different wavelengths and increases as the groove density (G) increases. Once a grating is incorporated into a spectrometer with an effective focal length (f, see below for details), the linear dispersion of the system is the product of D and f. In practice, the reciprocal linear dispersion (sometimes called the plate factor P), is more often considered, where:

$$P = \frac{1}{Df} = \frac{dcos\beta_m}{mf} = \frac{cos\beta_m}{Gmf}. \quad (17)$$

P is a measure of the change in wavelength (in nm) for a given lateral distance (in mm) and can be used to determine the bandpass and resolution of a spectrometer. The bandpass is the width of the spectrum passed by a spectrometer when illuminated by light with a continuous spectrum. In a monochromator, the bandpass is given by the product of P and the slit width. Reducing the width of the slit until a limiting bandpass is reached gives the resolution of the instrument. In spectral analysis, resolution is a measure of the ability of the instrument to separate two spectral lines that are close together. Figure 80 shows the impact of slit width reduction on the ability to resolve sharp spectral lines from a lamp source. The monochromator resolution is also affected by aberrations in the optical system and by proper grating illumination. Minimizing these contributions ensures that the resolution is determined mainly by P and the slit width. Spectrographs have a bandpass and resolution that are mainly dictated by the detector parameters (see below).

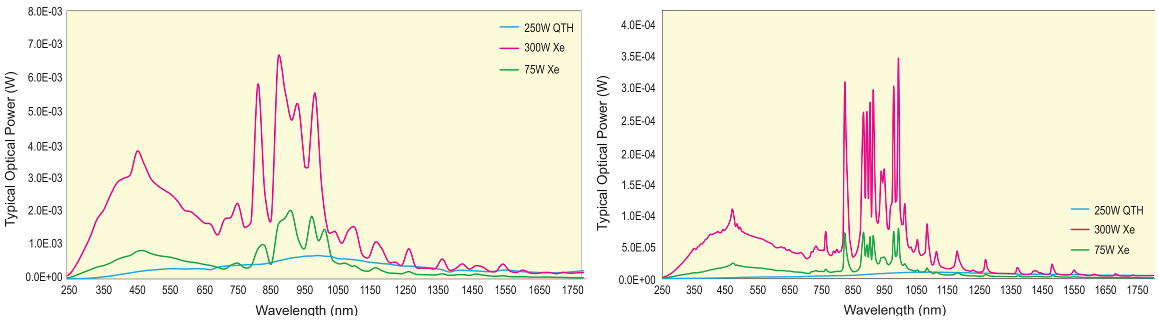


Figure 80. Spectra of an incoherent lamp source when passed through a monochromator with a value of P = 13.2 nm/mm. Reducing the slit width from 760 µm (left) to 120 µm (right) results in an improved spectral resolution from 10.1 nm to 1.6 nm

Types of Gratings

Gratings are produced by two methods, ruling and holography. A high-precision ruling engine creates a master grating by burnishing grooves with a diamond tool against a thin coating of evaporated metal applied to a surface. Replication of the master grating enables the production of ruled gratings, which comprise the majority of diffraction gratings used in dispersive spectrometers. These gratings can be blazed for specific wavelengths, generally have high efficiency, and are often used in systems requiring high resolution. Echelle gratings are a type of ruled grating that are coarse, i.e., low groove density, have high-blaze angles, and use high diffraction orders. The virtue of an echelle grating lies in its ability to provide high dispersion and resolution in a compact system design. Overlapping of diffraction orders is an important limitation of echelle gratings requiring some type of order separation typically provided by a prism or another grating. Holographic gratings are created using a sinusoidal interference pattern which is etched into glass. These gratings have lower scatter than ruled gratings, are designed to minimize aberrations, and can have high efficiency for a single plane of polarization. Gratings can be reflective or transmissive, and the surface of a grating can either be planar or concave. Planar gratings generally give higher resolution over a wide wavelength range while concave gratings can function as both a dispersing and focusing element in a spectrometer.

Additional online content discussing diffraction gratings can be found in [139-143].

2. Monochromators and Spectrographs

Monochromators and spectrographs are the most widely used dispersive spectrometers. They share many common components, including a diffraction grating for wavelength dispersion, imaging optics to relay the light from the entrance aperture to the exit aperture, and a photodetector. They differ primarily in their acquisition and detection of a spectrum. A monochromator typically detects light using a narrow bandpass on a single photodiode and builds up a spectrum by scanning the grating assembly. This results in a spectrum with high-resolution but slow acquisition speed. A spectrograph has a very wide bandpass and essentially detects a whole spectrum using a photodiode array (PDA). This gives rise to fast acquisition but poorer spectral resolution.

Monochromators

A typical configuration for a monochromator, known as the Czerny-Turner configuration, is shown in Figure 81. Diverging light enters through the input/entrance slit and is collimated by a concave mirror prior to reaching the grating. This ensures that parallel light hits the diffraction grating so that it can angularly disperse the light. The focusing mirror then redirects the diffracted light onto the output/exit slit. At any fixed grating setting, only a small spectral bandpass centered at a specific wavelength exits the monochromator and reaches the detector. Rotating the grating scans this central wavelength across the exit slit, essentially building up an entire spectrum. This bandpass is typically chosen to provide the desired spectral resolution for the measurement. As discussed above, the bandpass depends linearly on the slit width and inversely on the focal length of the mirror (for a given grating). Therefore, larger monochromators, i.e., longer focal lengths, and narrower slit widths yield better resolution. However, there is an intrinsic tradeoff between resolution and throughput or the amount of light that can be gathered for detection (see Section III.A.3 for more details). A narrower slit width clearly limits the amount of light exiting the system while a longer focal length restricts the cone of light entering the system. Since the output light is focused through the exit slit, only a single element detector is required. This allows for great flexibility in terms of detection since practically any photodiode or pyroelectric sensor can be used. Finally, in addition to spectral characterization, monochromators can be used to produce continuously tunable sources for spectroscopy provided a broadband input source such as a QTH lamp is used. Details regarding these tunable sources are given in Chapter 6, Section V.

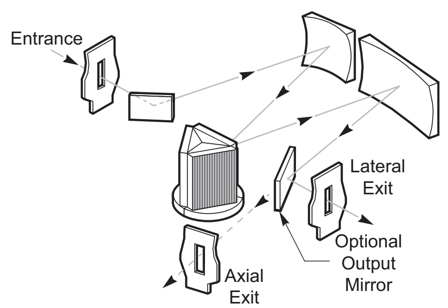


Figure 81. Typical optical configuration of a monochromator (left), which allows for multiple diffraction gratings and two output ports. Monochromator system with optical path shown on top (right).

Spectrographs

A spectrograph is nearly identical to a monochromator except that the grating assembly is fixed and the exit slit is removed. A PDA is placed at the exit focal plane for detection. In this case, the spectral bandpass is no longer limited by the slit but rather by the spatial dimensions of the PDA (see Figure 82). Depending upon the reciprocal linear dispersion of the system, spectral bandpasses of 300 – 600 nm can easily be achieved. A spectrograph is not just a monochromator without an exit slit since additional considerations for the focusing/imaging optics are required to account for this large aperture. However, once this optical design issue is solved, the instrument can possess two exit ports, one with an exit slit, so that it can serve as both a spectrograph and a scanning monochromator. The resolution of a spectrograph depends on the same parameters as a monochromator, except that the role of the exit slit width is replaced by the width of two individual pixels in the recording array. The PDAs used in spectrographs can either be a one-dimensional, i.e., a row of photodiodes, or two-dimensional such as CCDs. A spectrograph configuration that utilizes multiple one-dimensional PDAs is shown in Figure 82. The input to the spectrograph is made up of three optical fibers, each possessing light from a different source; their different spectra can be individually detected on a single PDA.

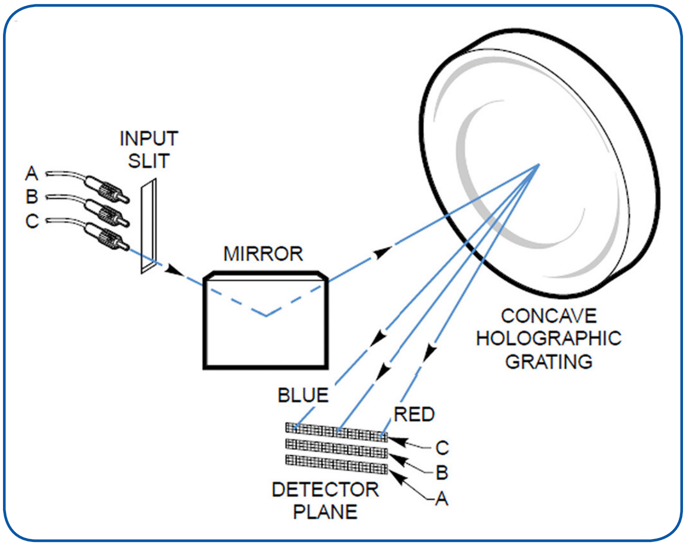


Figure 82. Typical optical configuration of a spectrograph where the detector consists of multiple one-dimensional PDAs, each of which records a spectrum from a different input fiber.

Additional online content discussing dispersive spectrometers can be found in [144-148].

E. Pulse Characterization

Pulsed lasers can achieve large peak powers by squeezing the energy of a pulse into a small duration of time. This enables applications requiring high intensities such as laser processing and nonlinear frequency conversion. Furthermore, ultrashort pulses are critical for probing fast light-matter interactions and for high-speed optical communications. As a result, accurate temporal characterization of pulsed sources is as critical as accurately estimating the pulse energy and spatial and spectral profiles. As discussed in Section I.A.5, methods for pulsed laser generation can lead to pulse durations ranging from μs to fs. Direct measurements of pulses less than 20 ps using photodetectors remain a challenge as the response time of high-speed photoreceivers is currently limited to 10 ps (see Section II.C.2). Consequently, indirect methods for measuring ultrashort pulses using optical gating techniques have been developed and are discussed below. Prior to addressing the characterization techniques, attributes of a pulse laser's temporal profile are detailed.

Laser Pulse Temporal Profile

For a pulsed laser, the output power (φ) is a time-dependent quantity. Like the laser beam spatial profile discussed in Section II.B.1, the temporal profile has both an amplitude, a pulse shape, and a width. One common profile takes the shape of a Gaussian function (as shown in Figure 83) and can be written as follows:

$$\varphi(t) = \varphi_{pk}\, exp\left[-\left(\frac{t}{\tau}\right)^2\right]. \tag{18}$$

The pulse width τ is defined as the radius (HW1/*e*) at which the power decreases to 1/*e* or 0.37 of its peak power (φ_{pk}) value. The temporal width is sometimes reported as its FWHM value, which—for a Gaussian pulse—is larger than τ by a factor of $2(\ln 2)^{1/2}$. Laser systems can produce many pulse shapes, including a Lorentzian, hyperbolic secant, as well as a flat-top (see [149]). Similar to laser beams, laser pulses can undergo modification following propagation which is described in detail in [149]. One important consequence of this modification is that ultrashort pulses, i.e., values of τ less than 35 fs, can experience significant pulse broadening in dispersive media after propagation of only a few centimeters.

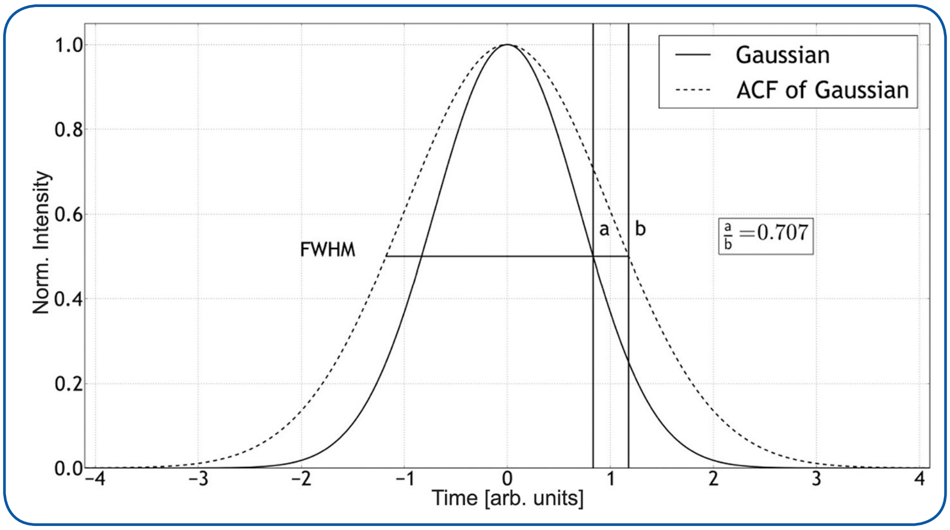


Figure 83. A Gaussian temporal pulse function and its normalized autocorrelation function (ACF) (dotted line).

The value of φ_{pk} is distinctly different from the average power since it gives the instantaneous value of the power at the peak of the pulse. The value of φ_{pk} for a Gaussian pulse can be determined by recalling from Equation (8) that integration of $\varphi(t)$ over the entire pulse gives the pulse energy (Q_e). Therefore, by integrating the Gaussian pulse in Equation (18), one can obtain the peak power, i.e., $\varphi_{pk} = Q_e/\tau\sqrt{\pi}$, where the inverse relationship between peak power and pulse width is clear. Furthermore, by taking this Gaussian form of $\varphi(t)$ and inserting it into the irradiance distribution for a Gaussian beam given in Equation (11), the spatiotemporal irradiance distribution for a pulsed laser is generated. Such a formula is critical for estimating important quantities such as peak irradiance or peak power density. An online calculator for estimating these values for a laser beam can be found in [150].

Autocorrelation

Measurement of a temporal event requires that the response of the detection method be at least as fast as the event being measured. Since response times of high-speed photoreceivers are limited, measurement of a pulse with sub-10 ps duration relies on the optical pulse itself to act as the response function. An autocorrelator is a device (see Figure 84) that utilizes this approach for measuring the pulse shape and duration of ultrafast laser pulses. Autocorrelation is the simplest method for determining pulse widths when phase information of the pulse is not required (see below for a more comprehensive approach). Autocorrelation is based on recording the second order correlation function using a Michelson interferometer (see coherence portion of Section I.A.3 for details about the interferometer). An incoming pulse with electric field $E(t)$ is first split into two replicas by means of a beamsplitter. The replicas are sent down two independent temporal delay lines, one variable and one fixed, to generate a time delay between them. Then, the two replicas, $E(t)$ and $E(t-\tau)$, are recombined in a nonlinear crystal to produce SHG (see Section I.A.6 for details). The total intensity of the SHG signal is proportional to the square of the sum of the fields as shown in Figure 84. There is a component of the signal ($[2E(t)E(t-\tau)]$) that is due solely to the temporal overlap of the two pulses. This component of the signal will only be present when the two pulses are overlapping in time. An iris or aperture allows only this component to be sent to a detector, typically a photodiode. The photodiode squares the signal and integrates it, giving an intensity proportional to the intensity of the two replicas. By recording this signal as a function of the time delay, the intensity autocorrelation of the laser pulse is generated.

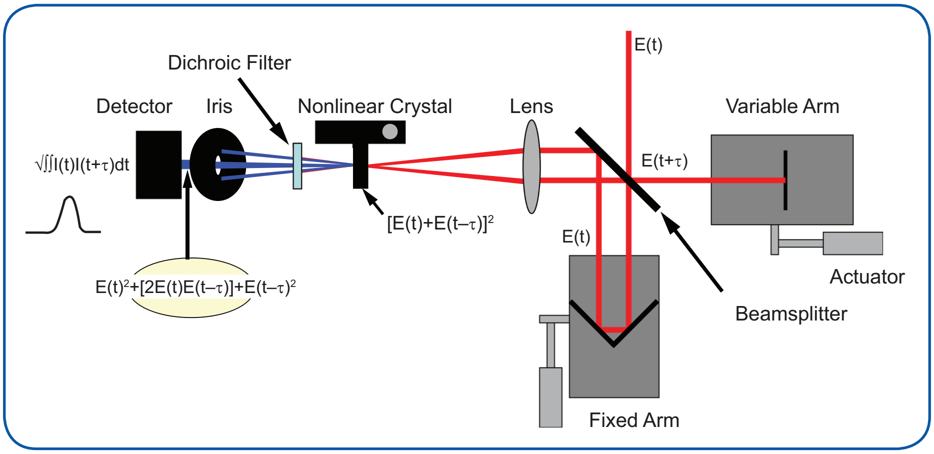


Figure 84. Schematic of a typical autocorrelator.

The equation for the autocorrelation shown on the left side of Figure 84 is a correlation function of the two replicas and not a direct measurement of the pulse. If the two replicas are identical, this equation can be solved analytically and the pulse width can be determined simply by dividing the autocorrelation signal width by a constant factor that depends on the profile of the pulse. Reference [149] lists the autocorrelation factors for common pulse profiles. It should be clear that the autocorrelation width is always greater than the width of the actual pulse. This can be seen for the Gaussian pulse shown in Figure 83, where the autocorrelation width is a factor of $\sqrt{2}$ larger than the pulse width.

Frequency-Resolved Optical Gating

Because of the time-bandwidth uncertainty principle, ultrashort laser pulses carry significant spectral bandwidth (typically from 10 to 100 nm). Autocorrelation can measure the duration of an ultrashort pulse, but it is unable to characterize the phase relationship between the different constituent spectral components. The primary reason for this is that autocorrelators utilize a single element photodetector that effectively integrates over the spectral profile of the pulse. The relationship between the spectral components of light can change upon propagation in a dispersive medium leading to an effect known as chirp, which can cause pulse broadening. Fortunately, techniques exist that monitor the spectral profile of a pulse as a function of time, allowing for complete reconstruction of the electric field. Of these techniques, Frequency-Resolved Optical Gating (FROG) is arguably the most straightforward and easiest to implement. Figure 85 shows the simplest version of this technique (SHG FROG) which is essentially a spectrally-resolved autocorrelator that sends the SHG signal into a spectrometer instead of a single element photodiode. Figure 85 shows typical signals (called traces) generated from a SHG FROG set-up and reveals the impact of pulse broadening following propagation in a dispersive medium. Extracting the electric field from a FROG trace, known as a retrieval, is significantly more complicated than extracting the intensity from an autocorrelation signal. Analysis of FROG traces requires an iterative method incorporating a two-dimensional phase retrieval algorithm. Fortunately, efficient algorithms for phase retrieval exist. There are alternative geometries that can be used for FROG, including the Self-Diffraction FROG (SD FROG). This geometry provides the sign of the spectral phase (not extracted using a single SHG FROG scan) and is more intuitive to analyze, allowing for chirp measurements to be made in real time. The downside is that SD FROG relies on very high peak powers to generate sufficient signals. Details about the advantages and limitations of different FROG geometries, their operation, and the required data analysis can be found in [151].

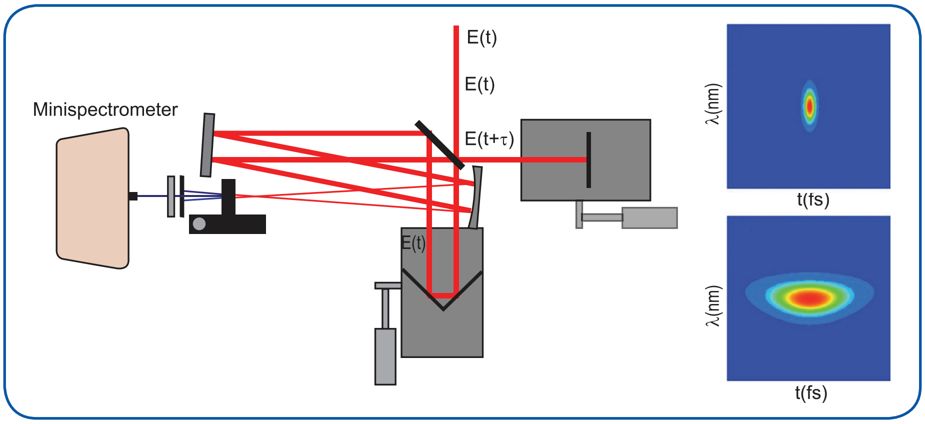


Figure 85. Schematic of a SHG FROG set-up (left) and SHG FROG traces (right) for a 20 fs pulse before (a) and after (b) traveling through a dispersive medium.

Additional online content discussing pulse properties and characterization techniques can be found in [152-156].

III. Components

As detailed in Sections I and II, a well-characterized light source is a critical part of any photonic system. This section discusses the essential photonics components required for managing light. These components provide vital functionalities including the manipulation of the light, positional control of equipment, and isolation from vibrations that can adversely affect performance.

A. Optical Components

Optical components are critical when setting up a benchtop laser experiment, integrating a commercial photonics product, or for developing a space-based telescope or large-scale interferometer. This section discusses the most commonly-used optical components as shown in Figure 86. Optical mirrors and lenses allow light to be routed, collimated, focused, collected, or imaged. The physical phenomena that underlie this manipulation of light are reflection and refraction. These phenomena are usually described by geometrical optics, which depict the propagation of light using light rays. A ray is a fictitious line drawn along the path that light follows in order to understand how it interacts with mirrors and lenses. Other optical components such as filters, beamsplitters, and polarizers can redistribute the incident light into different directions. Their operation is governed by polarization and interference. These phenomena are described by physical optics, which concentrates on the wave nature of light. An illustration of how light can be depicted as both a ray and a wave is given in Figure 77. The operation of these optical components is described in this section as well as their characteristics and applications.

Figure 86. Common optical components.

1. Mirrors

Mirrors are arguably the most commonly-used optical components. They appear in small laboratory experimental set-ups, industrial applications, as well as large-scale optical systems. These components utilize reflection to redirect, focus, and collect light. Optical mirrors consist of metallic or dielectric films deposited directly on a substrate such as glass, differing from common mirrors, which are coated on the back surface of the glass. As a consequence, the reflective surface of an optical mirror may be subject to environmental conditions. This means that durability and damage resistance must also be considered when choosing a mirror as well as how well it reflects light at the wavelength of interest. This section introduces the physical concept of reflection and discusses the important attributes of the mirror as an optical component.

Reflection

Generally, when light reaches a planar interface between two media (see Figure 87), a portion of it is reflected back into the original (incident) medium and a portion is transmitted and refracted into the second medium [1, 138, 157]. Refraction of light is discussed in Section III.A.3. Absorption of the light in either medium is also possible, but non-absorbing media will be assumed here. The direction of the reflected light is governed by two laws. First, the incident ray, reflected ray, and the normal to the interface must lie in the same plane. In this plane of incidence, the angle of incidence (θ_i) is always equal to the angle of reflection (θ_{rfl}). Reflection can occur from smooth surfaces such as those found on mirrors (referred to as specular reflection) or from rough, uneven surfaces (called diffuse reflection or scattering). Although both obey the same laws of reflection, specular reflection leads to rays that reflect as a group at the same angle, whereas diffuse reflection occurs at different angles off randomly oriented surfaces. This enables specular reflection to perform the useful operations of redirecting light.

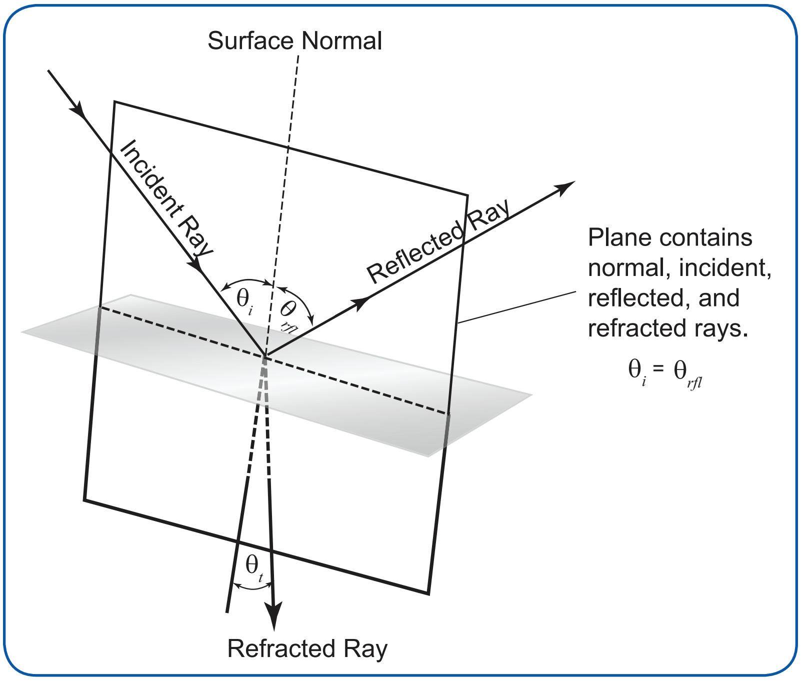


Figure 87. Illustration of the laws of reflection at a planar interface [157].

The fraction of the incident light that is either reflected or transmitted at the interface is described by the Fresnel equations [1, 138] and depends on the angle of incidence as well as the index of refraction of the incident (n_1) and refracting (n_2) media. The fraction of the incident power reflected from the interface is called the reflectance or reflectivity (R), while the fraction refracted in the second medium is the transmittance or transmissivity (T). By assuming that both media are non-absorptive, the sum of R and T must be unity, thus allowing knowledge of one to provide information about the other. Furthermore, different linear polarization components of the incident light (see Section III.A.4 for details on polarization) possess different values of R and T. The Fresnel equations are greatly simplified for light at normal incidence, i.e., $\theta_i = 0$, a situation of significant practical interest. At normal incidence, angular and polarization dependencies are removed from the formula for R (recall that T is complementary), leaving only the dependence on the indices of refraction:

$$R = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 . \tag{19}$$

The index of refraction is a complex value with both real (associated with refraction) and imaginary (related to the transition absorption cross-section) components. Furthermore, there is a wavelength-dependence (or dispersion) associated with the index of refraction. Consequently, R is highly dependent on the materials making up either side of the interface, as discussed in the mirror characteristics section below.

Mirrors made up of planar surfaces, such as that shown in Figure 87, are important components for directing light through the proper path in an optical system. Such mirrors can be combined to form optical components known as retroreflectors or corner cubes. These components consist of three mirror surfaces all perpendicular to one another. Such a geometry enables 180 degrees reflection of the light, regardless of incidence angle, and therefore requires very little alignment compared to a single flat mirror. In addition to stationary mirrors, rapid redirection can be achieved by utilizing rotating planar mirror systems such as those found in scanners or on a smaller scale with micromirrors, which are used for switching in telecommunications and displays. Curved mirror surfaces (also called concave reflectors) can be exploited with the goal of collecting, focusing, and imaging light as illustrated in Figure 88. These mirrors possess an advantage over lenses (see Section III.A.3) in that they perform satisfactorily across a broad-wavelength range without requiring refocusing. The reason for this is that reflection occurs at the surface of these optics, rather than passing through the optic as is the case with a lens (see Section III.A.3), and so the dispersion of the index of refraction does not come into play. Simple spherical reflectors can be used to collect radiation from a source at the focal point (located at half of the radius of curvature of the mirror) and reflect it as a collimated beam parallel to the axis. Since spherical mirrors possess spherical aberration (see Section III.A.3), a parabolic curved surface can be used instead to either collimate light from a focal point or focus light from a collimated beam (see Figure 88). Ellipsoidal surfaces can focus light from one focal point to another (see Figure 88).

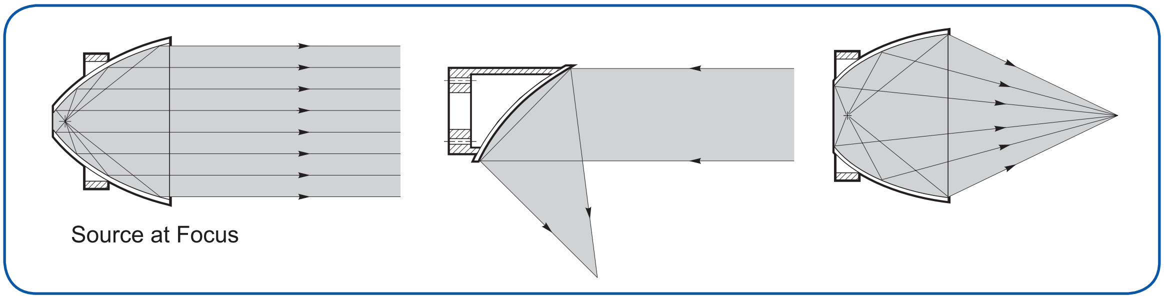


Figure 88. Concave reflectors with different surface shapes allowing for light collection and focusing. A paraboloidal reflector reflects light from the focus into a collimated beam (left). An off-axis paraboloidal reflector refocuses a collimated beam off the mechanical axis (middle). Ellipsoidal reflectors reflect light from one focus to a second focus, usually external (right).

Mirror Characteristics

Selecting the proper mirror for an application requires consideration of a number of factors, including reflectivity, laser damage resistance, coating durability, thermal expansion of the substrate, wavefront distortion, scattered light, and cost. These mirror characteristics depend on the optical coating, the substrate, and the surface quality. The optical coating is the most critical component of a mirror as it dictates its reflectivity and durability. Processes for depositing high-quality optical coatings are discussed in Section III.A.2. Optical mirror coatings are typically made up of either metallic or dielectric materials. A common situation for mirror applications is when light is incident from air (n_1 = 1) onto the optical coating material and so the reflectivity given by Equation (19) is dictated solely by the material's index of refraction (n_2). By virtue of their conductivity, metals have a complex index of refraction with a large imaginary part over a very wide wavelength range. This gives rise to a large reflectivity that is relatively insensitive to wavelength, which gives metallic mirrors their shiny appearance. Metallic coatings are usually made of silver, gold, or aluminum and the resulting mirrors can be used over a very broad spectral range (see Figure 89). Metallic coatings are relatively soft, making them susceptible to damage, and special care must be taken when cleaning. Mirrors with dielectric coatings are more durable, easier to clean, and more resistant to laser damage. However, as a consequence of their dispersive and predominantly real indices of refraction, dielectric mirrors have a narrower spectral reflectivity and are typically used in the VIS and NIR spectral region. There is greater flexibility in the design of dielectric coatings compared to metallic coatings (see Section III.A.2.). When compared with metallic mirrors, a dielectric mirror can offer higher reflectivity over certain spectral ranges and can offer a tailored spectral response (see Figure 89).

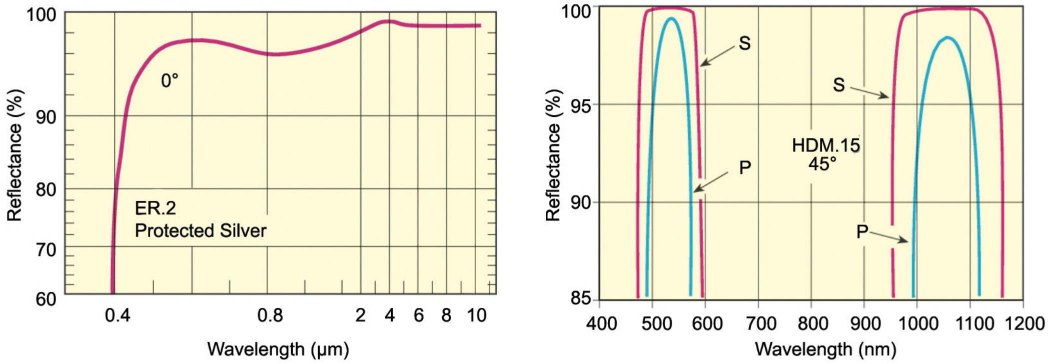


Figure 89. Reflection spectra of silver metallic mirrors showing broadband reflectivity (left) and dielectric laser-line mirror showing two narrow reflection bands (right).

Most substrates upon which the coatings are deposited are dielectric materials [158] and these substrates control the thermal expansion and transmission properties of mirrors. Certain materials have lower thermal expansion coefficients, e.g., PYREX® borosilicate glass or fused silica, than others, e.g. N-BK7 optical glass, but the cost of the material and ease of polishing must also be considered. If light transmitted through the substrate is not required, the backside of the substrate is typically ground to prevent inadvertent transmissions. However, for transmissive mirrors, a substrate material with a homogenous index of refraction is important, e.g. fused silica.

Prior to depositing the optical coating, the substrate's surface must be ground and polished to the proper shape (either planar or curved). The surface quality and flatness determine the fidelity of the mirror performance [159] with the targeted application dictating the requirements for these parameters. Surface flatness is often specified in wavelengths, e.g. $\lambda/10$, over the entire usable area of the mirror. When preservation of the wavefront is critical, a $\lambda/10$ to $\lambda/20$ mirror should be selected, while less demanding applications can tolerate a $\lambda/2$ to $\lambda/5$ mirror with the associated reduction in cost. Surface quality is usually dictated by the severity of random localized defects on the surface. These are often quantified in terms of a "scratch and dig" specification, e.g. 20–10, with a lower value indicating improved quality and therefore lower scattering. For high precision surfaces, such as those found within the cavity of a laser, a scratch-dig specification of 10–5 may be required since it would yield very little scattered light. Surface polishing tolerances in terms of irregularity, surface roughness, and cosmetic imperfections are verified using state-of-the-art metrology equipment. These same parameters and procedures are used to assess the quality and flatness of other optical components such as lenses or windows.

Additional online content discussing mirror types and properties can be found in [160-163].

2. Optical Coatings and Optical Filters

Optical coatings typically consist of thin films made up of single or multiple layers of either metallic or dielectric materials. When properly designed and fabricated, these coatings can dramatically modify the reflection and transmission properties of an optical component. The properties can be controlled from the deep UV to the IR with narrowband, broadband, or multi-band response, and can be polarization sensitive. Optical coatings can be applied directly to the surface of an optical component to tailor its reflectivity, as in the case of an optical mirror or beamsplitter. For other components, such as lenses, the applied coatings may simply improve their overall transmission properties by reducing surface reflectivity. When optical coatings are integrated into a monolithic component for the express purpose of controlling the spectral transmission of light, the component is referred to as an optical filter.

Optical Coatings

The individual layers that make up optical coatings are typically a few tens of nanometers to a few hundred nanometers in thickness, while a single optical coating can be comprised of several hundred layers. Consequently, the techniques used to deposit these layers require a high degree of precision. Generally, the process begins with surface fabrication to minimize surface roughness and sub-surface damage. It continues with surface cleaning and preparation and is followed by deposition of high-performance thin film designs. The deposition technologies include thermal evaporation, electron-beam, ion-assisted deposition, and advanced plasma deposition. The most appropriate coating technology for the intended product design depends on the operating environment, spectral requirements, physical characteristics, application requirements, and economic targets. The optical coating process is completed with comprehensive performance testing using sophisticated metrology tools.

Metal coatings used on optical mirrors typically consist of a single layer approximately 100 nm thick. This ensures that the broadband high reflectivity properties of the metal due to the complex index of refraction are present. In order to provide greater tuning of the reflectivity and over specific wavelengths of interest, dielectric coatings are used [138, 162]. These coatings (sometimes referred to as optical interference coatings) consist of alternating high refractive index (n_H = 1.8 – 4.0) and low refractive index (n_L = 1.3 – 1.7) dielectric layers (see Figure 90). The thickness of each layer is chosen such that the product of the thickness and the index of refraction of the layer is $\lambda/4$. A variation of the formula given by Equation (19) can be used to estimate the maximum reflectivity of the dielectric coating which increases with a greater number of layers but is accompanied by a concomitant reduction in the spectral bandwidth. Dielectric coatings are ubiquitous and their applications are discussed below. Other optical coatings include metal-dielectric hybrid films such as those in cube beamsplitters (see Section III.A.5) and absorptive coatings made of organic materials such as those found in certain optical filters.

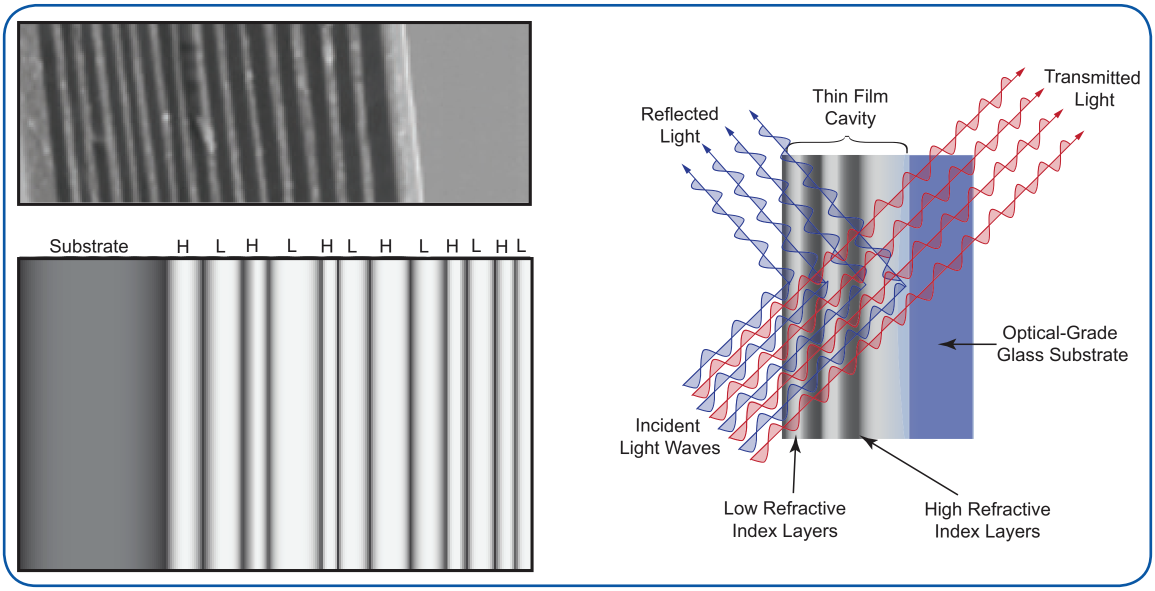


Figure 90. Scanning electron microscope image (top) and schematic (bottom) of an optical interference coating shown on left. Reflection and transmission of light by a filter consisting of an interference coating (right, [164]).

Dielectric optical coatings are used in a myriad of ways. In addition to highly reflective dielectric mirrors (see Figure 89), these coatings are incorporated in broadband beamsplitters (see Section III.A.5) and dichroic filters (color-selective mirrors that transmit and reflect particular wavelengths, see filters section below). When light is incident at an angle to a surface, i.e., not normal incidence, the reflectivity becomes polarization sensitive. This allows dielectric coatings to be polarization selective and such coatings are used in polarizing beamsplitters (see Section III.A.5). In addition to enhancing the reflectivity, dielectric optical coatings can also be used to reduce surface reflections in the form of broadband anti-reflection coatings. These coatings can be applied to any optical component, e.g., lens, prism, beamsplitter, window, to markedly improve its transmission efficiency. Based on Equation (19), the reflection from an air-glass ($n_2 \approx 1.5$) interface gives a reflectivity of 4%, which can be reduced considerably with a broadband anti-reflection coating (see Figure 91). These reflectivities can be reduced even more to improve transmission in laser systems with multiple optical elements, saving valuable laser energy from being lost to surface reflections. This superior performance, however, is achieved at the cost of reduced wavelength range.

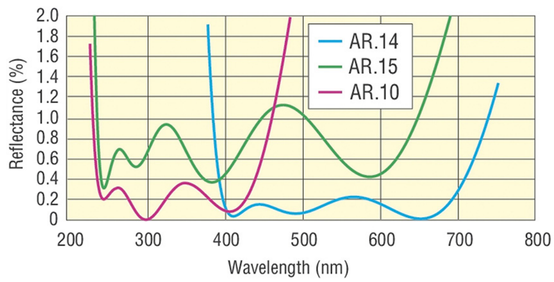


Figure 91. Typical broadband anti-reflection coating in the UV and VIS spectral regions.

<u>Optical Filters</u>

In addition to being deposited on the surface of an optical component, optical coatings can also be incorporated into a single element through deposition on one or many substrates followed by a lamination process. These elements are referred to as optical filters and are used to carefully control or filter the transmission of light. The process of filtering may be accomplished by using coatings that absorb, reflect, or transmit certain portions of the incident radiation. Optical filters can be manufactured in many physical forms and can involve single or multiple elements. An individual filter can be fabricated by applying soft coatings to separate substrate plates that are then laminated together or hard coatings deposited on different components that are then mechanically assembled together. The choice of the method of filtering, the physical form of the filter, the method of assembly of the filter, and the coating technology used to coat the optical filter substrates, will depend upon the constraints imposed by the filter specifications. Some of the most important parameters associated with the specifications of an optical filter are environmental operating conditions, physical dimensions, durability, optical surface quality, and, most importantly, the spectral filter parameters.

Dichroic filters separate a broad spectrum of light into two components: a reflected component and a transmitted component. These filters provide the ability to select different bands from a spectrum and direct those bands to where they can either be used or discarded. Figure 92 shows two examples of dichroic filters: a heat control dichroic filter can be used to remove heat associated with NIR light from an optical system and a dichroic laser beam combiner efficiently combines or separates multiple laser beams at a 45° angle of incidence. Another type of filter is the bandpass filter which possesses a specifically defined region of transmittance, bounded on both sides by regions of spectral rejection. Bandpass filters isolate a band of wavelengths from the total spectrum by providing a band of high transmission and bands of high rejection of spectral energy on the long and/or short wavelength sides of the transmission band. Figure 93 shows two examples of bandpass filters, which show that a very narrow bandpass or a wide passband can be achieved. Bandpass filters can utilize dielectric interference coatings or absorptive coatings. One type of absorptive bandpass filter is a neutral density filter which is a wavelength insensitive filter that is used to suppress the transmission of light by a prescribed amount.

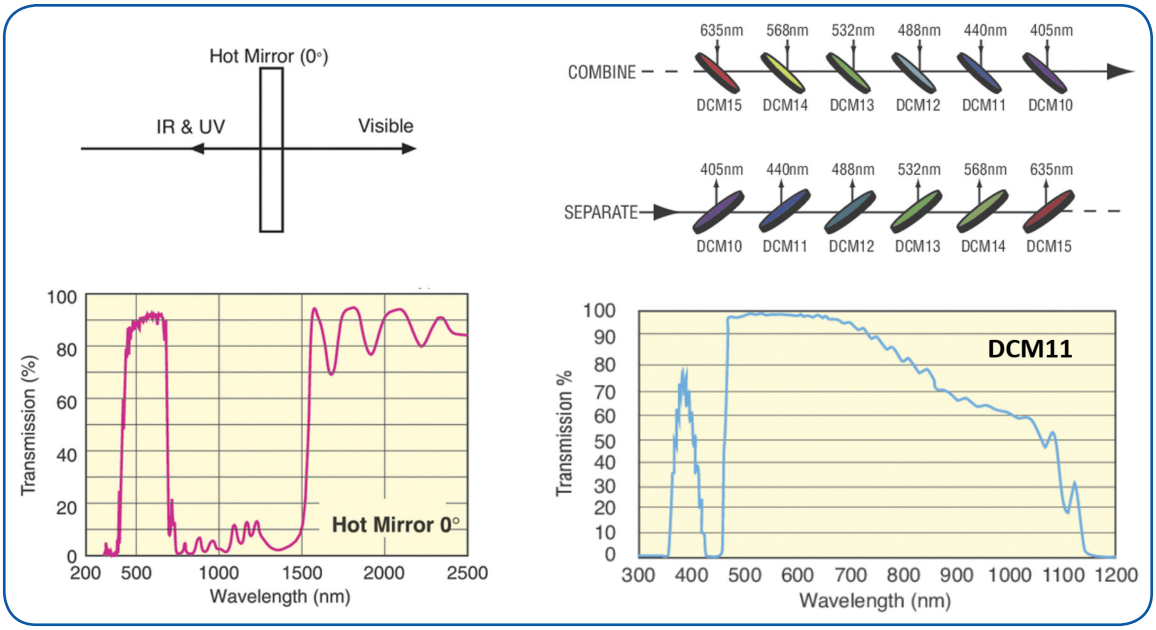


Figure 92. Two applications of dichroic filters: hot mirror reflecting IR and UV wavelengths while transmitting VIS wavelengths (left) and laser beam combining or separating using multiple filters (right). A transmission spectrum associated with a specific filter from the schematic is shown below it.

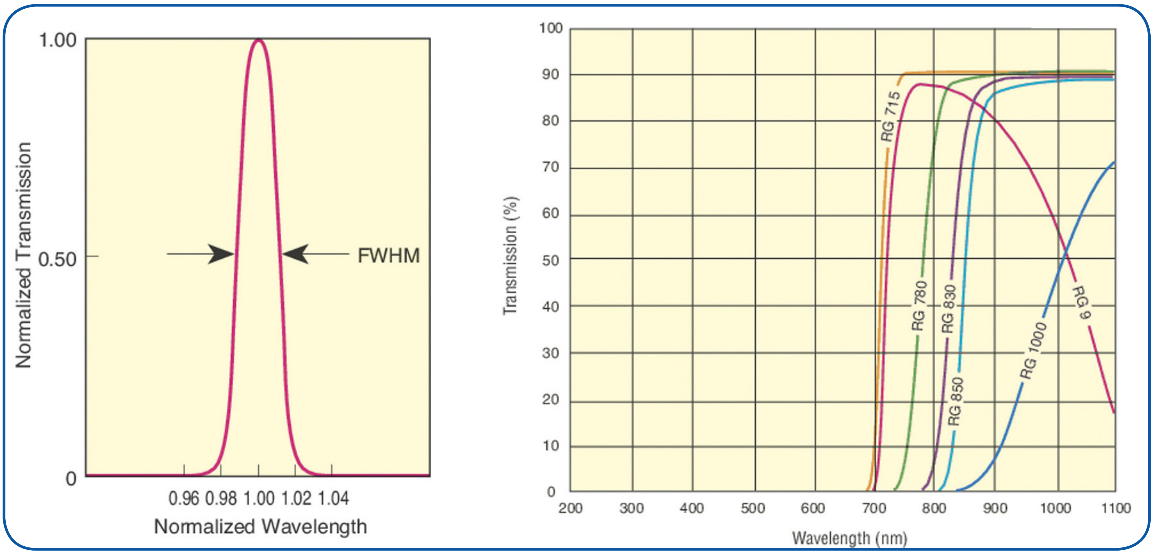


Figure 93. Transmission spectra of two types of bandpass filters: a narrowband laser-line interference filter (left) and an absorptive longpass filter (right).

Additional online content discussing optical coatings and filters can be found in [165-168].

3. Lenses

Lenses are the optical components that form the basic building blocks of many common optical devices, including cameras, binoculars, microscopes, and telescopes. Lenses are essentially light-controlling elements and so are exploited for light gathering and image formation. Curved mirrors (see Section III.A.1) and lenses can accomplish many of same things in terms of light collection and image formation. However, lenses tend to be superior in terms of image formation because they are transparent, which allows light to be transmitted directly along the axis to the detector whereas mirrors require an off-axis geometry. Mirrors are typically preferred in terms of light collection as they can be made significantly more lightweight than lenses and therefore can achieve larger diameters and light collecting ability. This section discusses the mechanism of refraction that underlies the operation of a lens, issues that affect its performance, and the different lens types.

Refraction

In addition to light reflecting off a planar interface between two media, it can also be transmitted and then refracted in the second medium (see Figure 94). Refraction refers to the change in the angle of the incident light when it enters the second medium. Since the speed of light in a medium is inversely proportional to its index of refraction, it will either slow down or speed up when it enters a different medium, resulting in the light changing its direction. Figure 94 shows an example where the index of the second medium (n_2) is greater than the first (n_1), which results in a bending of the light toward the normal to the interface. This phenomenon of refraction is described by Snell's law:

$$n_1 sin\theta_i = n_2 sin\theta_r , \tag{20}$$

where θ_i and θ_r are the angles of incidence and refraction, respectively (see Figure 94). Snell's law enables determination of the refracted ray direction provided the incident ray direction and refractive indices are known. This basic formula governs how lenses control the transmission of light for collection and imaging purposes.

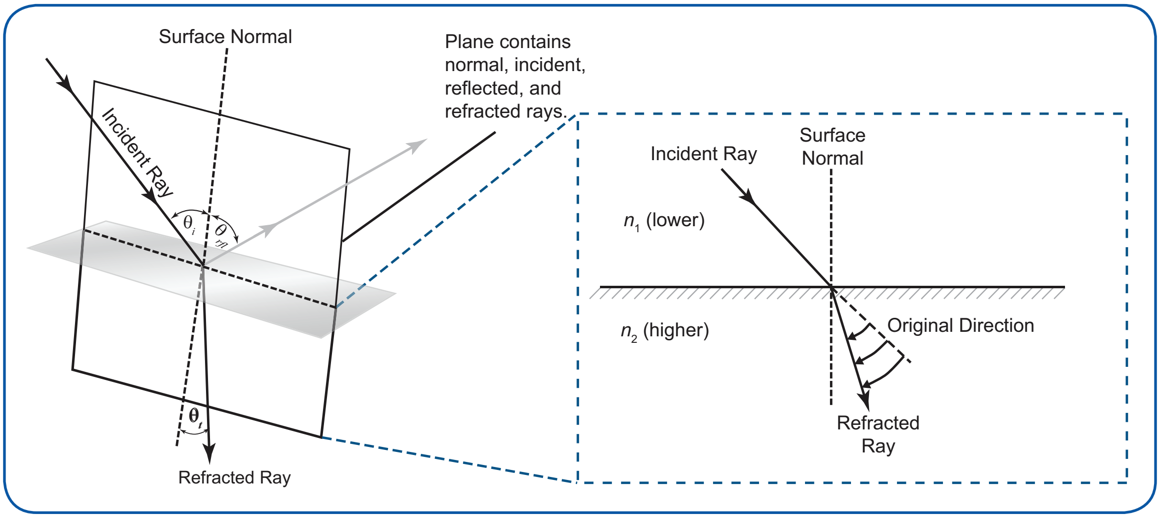


Figure 94. Illustration of Snell's law of refraction at an interface between media of refractive indexes n_1 and n_2 [157].

A lens is typically made up of a transparent dielectric material like fused silica or optical glass with the front and back surfaces having a spherical curvature [1, 157]. Since the surfaces are curved, each ray of light that comes in parallel to the optical axis (as shown in Figure 95) has a different value of θ_i with respect to the surface normal. Each ray then refracts according to Snell's law. For a positive lens, this causes the light to converge toward its focal point on the right side of the lens while light will diverge from the focal point located on the left side of a negative lens. The ramifications of these operations are that lenses can be used for image formation as well as collection and collimation of light (see Figure 95). There are several important aspects to optical imaging with lenses, including the relationship between object and image distances and the resulting magnification as well as the quality of the resulting images. Details about these concepts can be found in [138]. Similarly, important aspects of involving the light gathering ability of lenses including throughput and its relationship to numerical aperture (NA) or f-number (F/#) are described in [169].

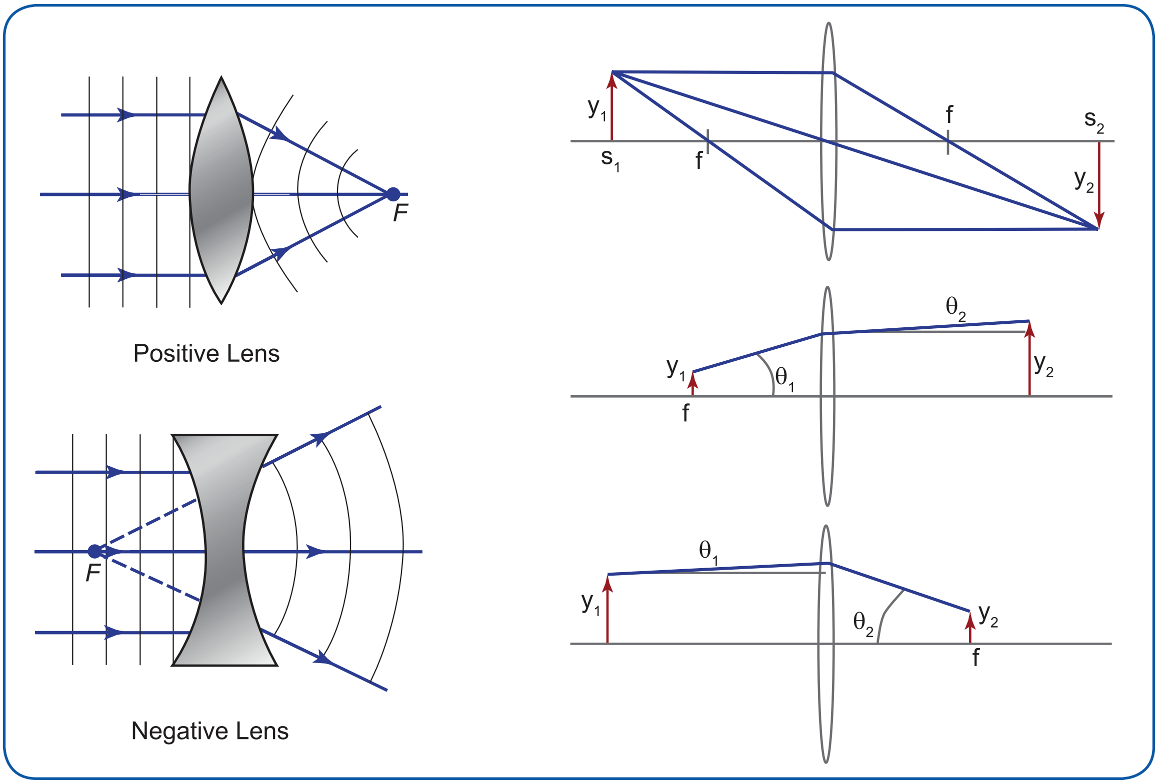


Figure 95. Illustration of how a lens affects incoming parallel light rays (left, [157]). Applications of lenses (right) include creating a magnified image of an object (top), collimating light from a point source (middle), focusing a collimated light source (bottom).

Optical Aberrations and Types of Lenses

Ideal lenses would form perfect images (or exact replicas of the object being imaged) and would be able to focus collimated light to a spot size limited only by diffraction. However, real lenses are not perfect and induce optical aberrations, which cause degradation in the ability to form a high-quality image, collimate a beam, or focus it tightly [138, 170]. Monochromatic aberrations, i.e., no wavelength dependence, are common to both mirrors and lenses and come from the inability of spherical surfaces to focus light properly when it is far from the axis. These aberrations include spherical aberration, coma, and astigmatism. Figure 96 demonstrates the impact of spherical aberration in a lens where rays with smaller angles are effectively collimated while rays with large angles converge instead. Unlike monochromatic aberrations, chromatic aberration only occurs in lenses. Due to dispersion of the index of refraction of the lens material, different wavelengths will refract with different angles according to Snell's law (see Figure 96). This causes degradation in image quality or light gathering ability when broadband light is being used.

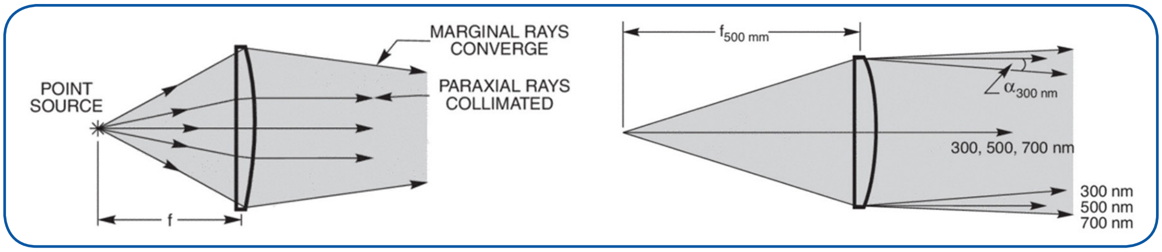


Figure 96. Effects of spherical aberration (left) and chromatic aberration (right) on collimation when a point source is at the focal point.

While spherical lenses do induce aberrations, choosing the proper lens shape can help minimize optical aberrations (see Figure 97). For instance, plano-convex lenses, where only one side is curved, are the best choice for focusing parallel rays of light to a single point. Bi-convex lenses (both sides have curvature that may not be equal to one another) are the best choice for imaging when the object and image are at similar distances from the lens. When a single spherical lens may be unsuitable due to spherical aberration, aspheric lenses may be used. These lenses have surfaces with tailored curvatures that help minimize the impact of aberrations but are typically expensive due to the complexities associated with fabrication. Alternatively, multiple spherical lenses can be used where one lens can cancel the aberration caused by another, as shown in Figure 97. In addition to correcting for monochromatic aberrations, an achromatic doublet can be used to minimize chromatic aberrations by choosing the dispersion of the materials in the two lenses to produce a focal length that is independent of wavelength. Microscope objectives are multi-element lens systems that can significantly reduce the impacts of aberrations but are more expensive due to the complexity of the design. All the aforementioned lenses are rotationally symmetric, that is, light focuses the same regardless of which transverse axis it passes through. On the other hand, cylindrical lenses focus or expand light in one axis only and so are ideal for modifying an asymmetric beam like the output of a laser diode (see Figure 35) into a symmetric one. There are many features to take into consideration when choosing a lens, including focal length, lens shape, F/#, lens material, transmission properties, wavefront distortion, scattered light, types of coating, and cost.

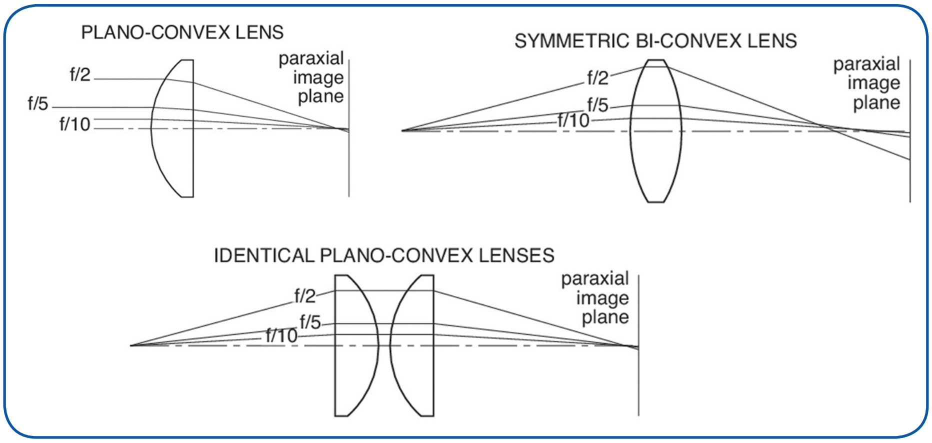


Figure 97. Using single and multiple lens systems to minimize optical aberrations for a specific imaging application.

Additional online content discussing refraction, image formation and aberrations, as well as different types of lenses, can be found in [171-175].

4. Polarization Optics

Precise control of polarization behavior is necessary to obtain optimal performance from optical components and systems. Characteristics such as reflectivity, insertion loss, and beamsplitter ratios will be different for different polarizations. Polarization is also important because it can be used to transmit signals and make sensitive measurements. Even though the light intensity may be constant, valuable information can be conveyed in the polarization state of an optical beam. Deciphering its polarization can reveal how the beam has been modified by numerous material interactions (magnetic, chemical, mechanical, etc.). Sensors and measurement equipment can be designed to operate on such polarization changes. For these reasons, optical components capable of filtering, modifying, and characterizing a light source's polarization are valuable. Such polarization control can be accomplished by exploiting the reflection, absorption, and transmission properties of materials used in these components. The physical phenomena that enable polarization control, as well as the key components that exploit them, are discussed below.

Polarization

The electric field of a light wave vibrates perpendicularly to the direction of propagation as shown in Figure 98. Since the electric field is a vector quantity, it can be represented by an arrow that has both a magnitude (or length) and a direction of orientation. This orientation direction is the polarization of the light [1, 138, 176]. There are basically three polarization states: linear, circular, and elliptical. These terms describe the path traced out by the tip of the electric field vector as it propagates in space. Figure 98 shows a snapshot in time of a linearly polarized wave. Although the electric field alternates direction (or sign), it stays confined to a single plane. Therefore, sitting at a fixed point in z as time passes, the arrow tip would oscillate up and down along a line. The angle (θ) of this line with respect to some reference set of axes completely specifies this linear polarization state. For circular polarization, the electric field vector tip forms a helix or corkscrew shape. For a fixed point in z, the vector would rotate in time, like the second hand on a watch. Circularly polarized light can be either left-handed or right-handed, depending on the clockwise or counterclockwise nature of the rotation. Elliptical polarization is the most general case of polarization. It is the same as circular polarization but with unequal major and minor axes (for circular polarization, these are equal).

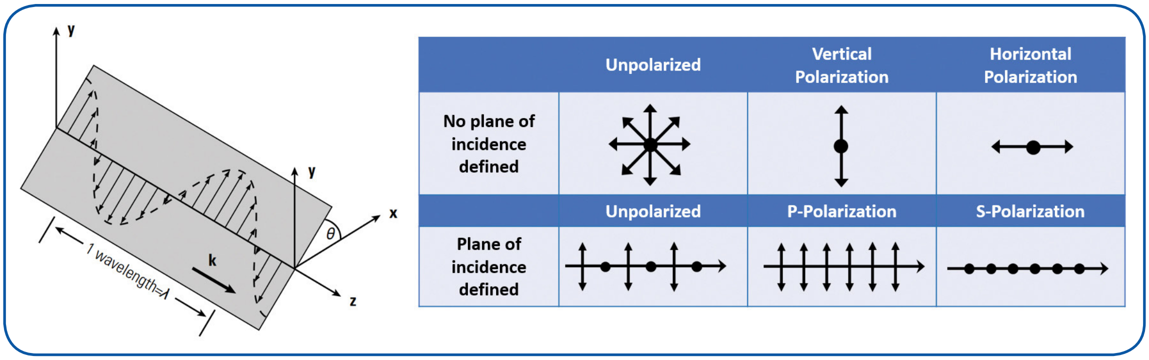


Figure 98. Depiction of a linearly polarized wave (left) and standard symbols for linearly polarized light (right, [176]).

Incoherent sources such as lamps, LEDs, or the sun typically emit unpolarized light, which is a random superposition of all possible polarization states. On the other hand, the output light from a laser is typically highly polarized, that is, it consists almost entirely of one linear polarization. Analyzing laser polarization is easier if it is decomposed into two linear components in orthogonal directions. In this way, depicting the polarization can be done using the standard symbols shown in Figure 98. The upper part of the table lists the symbols generally used for unpolarized, vertically polarized, and horizontally polarized light. For the graphic shown in the figure, the vertical direction would be along the y-axis while the horizontal direction would lie along the x-axis. When a plane of incidence is specified (see lower part of table in Figure 98), the polarization components acquire specific designations. S-polarization refers to the component perpendicular to the plane while P-polarization refers to the component in the plane. Examples of the depictions of linearly polarized light are illustrated in the remaining figures of the section.

The way in which polarized light interacts with an optical material can enable selective filtering of the polarization (similar to the spectral filtering discussed in Section III.A.2) or conversion of the incident polarization state to a different one. This polarization control relies on a material's optical properties to respond differently depending on the polarization of the incident light. A material that exhibits birefringence, or different refractive indices for different input polarizations, is said to be anisotropic. This anisotropy affects the transmission and absorption properties of light and is the primary mechanism used in polarizers and waveplates as discussed below. However, even isotropic materials (same index for different polarizations) can enable polarization selection via reflection. As discussed in Section III.A.1, the Fresnel equations describe the change in reflectivity as a function of angle of incidence. For a linearly polarized beam, both S- and P-polarizations exhibit different changes in reflectivity versus incident angle. There is an incident angle known as Brewster's angle (θ_B) at which P-polarized light is transmitted without loss, or exhibits zero reflectance, while S-polarized light is partially reflected. This angle can be determined from Snell's law to be $\theta_B = arctan(n_2/n_1)$. Figure 99 shows this response when light is incident from air onto a dielectric material where $\theta_B \approx 56°$. This polarization-selective reflectivity is exploited in laser cavities to produce strongly polarized light (see Figure 15) and for fine tuning of the output laser wavelength (see Figure 25).

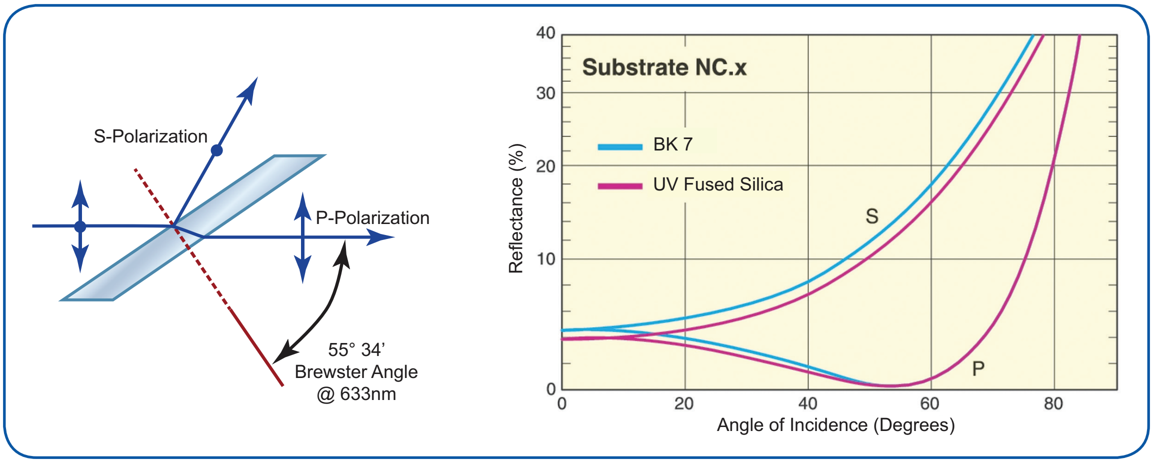


Figure 99. Polarization of incident light when encountering a plate oriented at Brewster's angle (left), reflectance of the plate as a function of incident angle showing the minimum in the P-polarization at the Brewster angle (right).

<u>Polarizers and Waveplates</u>

A polarizer is an optical component whose transmission depends strongly on the incident polarization of the light [1, 138]. Polarizers typically filter linear polarization, so an ideal polarizer would transmit 100% of one polarization component while rejecting all of the orthogonal component (see Figure 100). In practice, a portion of the undesired polarization will be transmitted. The transmittances of the target polarization and the undesired polarization through the polarizer are measured (by simply rotating the polarizer by 90 degrees) and the extinction ratio is defined as the ratio of these transmittances. Extinction ratios for a variety of different polarization optics are discussed in Section III.A.5 with larger values representing a higher-degree of polarization purity of the transmitted light. The difference between a polarizer and a Brewster plate is that the former results in strong polarization-dependent transmission while the latter does not (only the reflection is highly polarized).

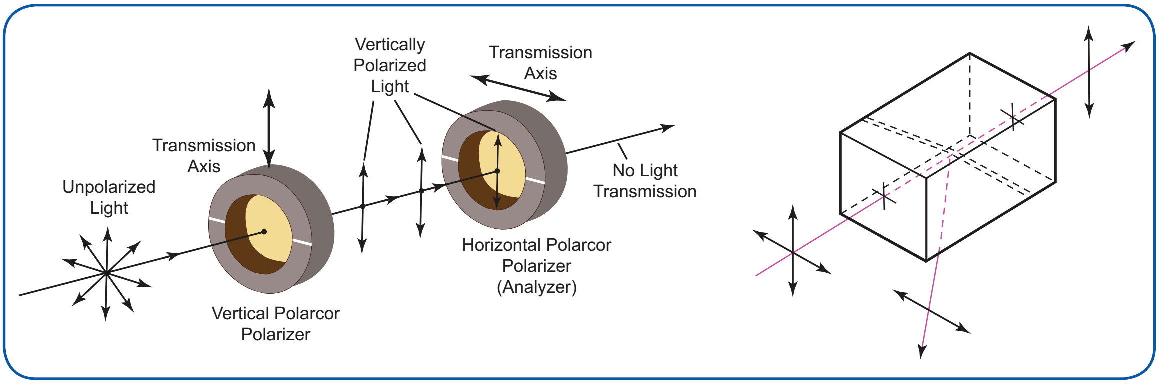
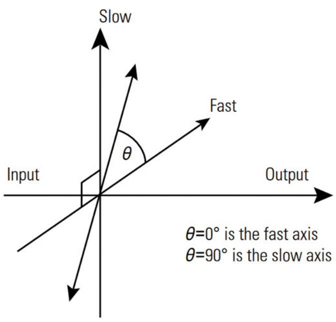


Figure 100. Impact of a dichroic polarizer on unpolarized light (left, [176]) and the separation of P- and S-polarizations following passage through a Glan-Laser calcite polarizer (right).

Polarizers rely on birefringent materials and, since the index of refraction is complex, these materials can exhibit a polarization-dependent absorption and refraction. The first polarizers were based on selective absorption of incident light and are usually denoted as dichroic polarizers. Typical materials used for this anisotropic absorption are stretched polymers or elongated silver crystals; their operation is shown in Figure 100. The strongly absorbing axis of the material is placed perpendicular to the desired output polarization such that the undesired polarization is strongly absorbed. A different type of polarizer is based on the anisotropic refractive indices of a birefringent crystal such as calcite. As described in the nonlinear frequency conversion portion of Section I.A.6, a birefringent crystal will produce an *o*-wave or *e*-wave depending on the axis of the crystal to which the polarization component is aligned. These waves experience different refractive indices and will possess different critical angles for TIR (see Section III.D), resulting in one polarization component being reflected while the other is transmitted (see Figure 100). By placing two calcite prisms back-to-back to form a rectangular optic, the transmitted beam will follow the same direction as the incident beam. The gap between these prisms can either be air or an optically transparent cement, depending on whether a high damage threshold or large acceptance angle, respectively, is desired.

Polarizers are used to filter the input polarization, increase its purity, or separate orthogonal components of a linearly polarized beam. However, a polarizer cannot convert the polarization state of the input light into a different polarization state. For this type of modification, an optical component known as a waveplate or retarder is required. To understand its operation, it is important to know that any polarization state (not just linear) can be decomposed into orthogonal components. The difference between the polarization states then results from the phase difference between the orthogonal components. Linear polarization possesses components that are in-phase, i.e., no phase difference, but have different amplitudes depending on their angle θ. Circular and elliptical polarization components possess a phase difference of $\pi/2$ or a quarter of a wavelength (circular polarization has the same amplitudes for the different components while elliptical has different amplitudes). Consequently, in order to convert one polarization to another, the phase difference between the two components must be controlled. This can be accomplished by sending a polarized beam into a birefringent crystal such that the *o*-wave or *e*-wave each experience a different phase delay. The operation of a waveplate and a summary of how quarter and half waveplates convert one polarization state to another are shown in Figure 101. An important case of polarization conversion is shown on the right side of Figure 101. A half waveplate can rotate the angle of a linearly polarized beam to any other angle, which can be used for rotating a vertically polarized laser beam to obtain horizontal polarization. Furthermore, proper combinations of waveplates and polarizers can be used to form optical systems that allow for variable attenuation of a laser beam (see Section III.C for details) or for isolating a laser cavity from spurious reflections [1, 138].



Input	Output
Quarter-wave	
Linear, θ=45°	Right circular
Linear, θ=-45°	Left circular
Right circular	Linear, θ=-45°
Left circular	Linear, θ=45°
Linear, any $\theta \neq$45°	Elliptical
Half-wave	
Linear, angle θ	Linear, angle -θ
Left circular	Right circular
Right circular	Left circular
Any wave plate	
Linear, θ=0° or 90°	Unchanged

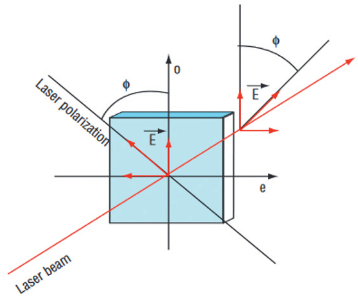


Figure 101. General polarization conversions using waveplates (left) and common application of a half waveplate used to rotate linear polarization to twice the input angle (right).

Additional online content discussing polarization and its associated optical components can be found in [177-181].

5. Beamsplitters and Windows

A beamsplitter is an optical component that is used to split a beam of light into two beams with one beam transmitted through the component and the other reflected at a 90 degree angle to the beam. Beamsplitters can take a variety of different shapes including cubes (see Figure 102), circular disks, and plates. They are made from high grade glass materials with very good surface flatness and quality and have a tight tolerance on the splitting ratio, i.e., ratio of power in transmitted beam to that of the reflected beam. The ability to split the beam is facilitated via application of a dielectric or hybrid metal-dielectric thin film coating similar to those described in Section III.A.2. Consequently, beamsplitters can be made for broadband light or narrowband laser-line applications. They can also be fabricated for non-polarizing and polarizing beamsplitting applications (see Figure 102). The former components evenly split the input laser power and are essentially insensitive to input polarization. Polarizing beamsplitters, particularly beamsplitting cubes, bear resemblance to the calcite polarizers discussed in Section III.A.4. The main differences (beyond the use of glass versus calcite for the optical materials) are that polarizing beamsplitters generally offer better polarization purity for the reflected beam and this beam is directed orthogonal to the input beam for convenience. However, the extinction ratio of the transmitted beam is significantly lower than for the calcite polarizers (see Figure 102).

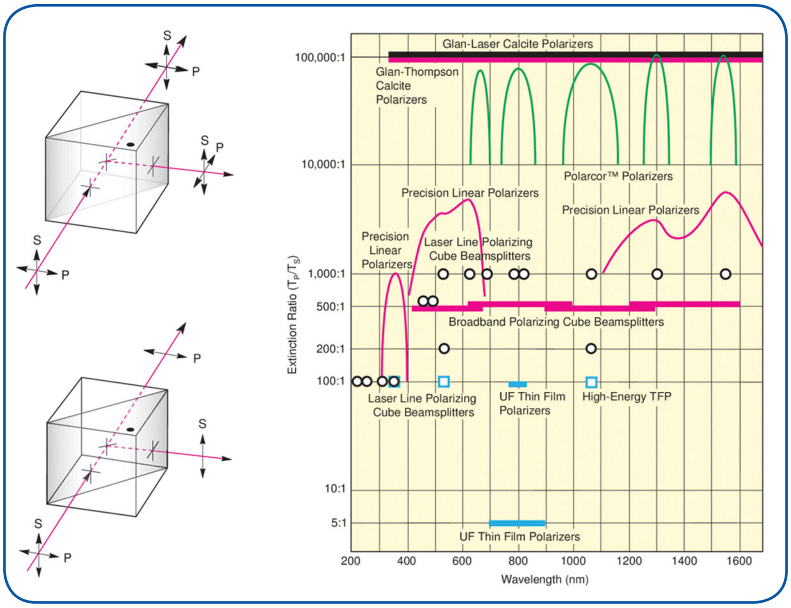


Figure 102. Operation of non-polarizing and polarizing cube beamsplitters (upper left and lower left, respectively). Extinction ratios of various polarizers and polarizing beamsplitters (right).

An optical window is a plate with good surface flatness made of an optically transparent material. The materials that make up optical windows are chosen based on the required transmission properties at the wavelength of interest and are typically dielectrics like glass, fused silica, sapphire, or semiconductors such as germanium or zinc selenide. Parallel or wedged optical windows are typically available with laser grade surface quality and surface flatness for common, or very demanding, interferometer measurements. Uncoated optical windows can serve as Brewster windows (see Figure 15) or as broadband beam samplers that split off or sample 1-10% of a beam's energy via Fresnel reflection from a single surface (see Figure 99, for example). This enables beam monitoring with minimal transmitted power loss and typically requires the optic's back surface be slightly wedged and anti-reflection coated.

Additional online content discussing beamsplitters and optical windows can be found in [182-185].

B. Opto-mechanical Components

For any optical system, no matter how simple or complex, each optical component must be mounted in some way. Furthermore, system performance depends not only on the precision of the optics but also their mounts and their positioning accuracy. Opto-mechanical components act as the interface between the optical components and the actual work surface. This section discusses the most common opto-mechanical components and their primary functionalities: optical mounts that hold, position and align optical components, posts and bases that couple the mounts to the work surface, and beam routing and rail systems, which can help form optical systems.

Type of Optic	Primary Function	Main Adjustments	Degrees of Freedom
Mirror	Changes beam direction/length (routing, path length change)	Angle (ϕ_Y, ϕ_Z), Focusing (X)	2 or 3 
Lens	Changes beam cross-section (focus, expand, collimate, shape)	Focusing (X), Centering (Y,Z), Angle (ϕ_Y, ϕ_Z)	2, 3, or 5 
Polarizer	Changes wave orientation	Rotation about optical axis (ϕ_X)	1 
Filter	Attenuates beam power	None	0 
Plate Beamsplitter	Splits beam into two beams	Angle (ϕ_Y, ϕ_Z), Focusing (X)	2 or 3 
Cube Beamsplitter	Splits beam into two beams	Angle (ϕ_X, ϕ_Y, ϕ_Z)	2 or 3 
Prism	Separates beam wavelengths	Angle (ϕ_X, ϕ_Y, ϕ_Z)	3 
Diffraction Grating	Separates beam wavelengths	Angle (ϕ_X, ϕ_Y, ϕ_Z)	2 or 3 

Table 7. Listing of various optical components, their functions, the adjustments that must be made to align them, the degrees of freedom for alignment, and examples of optical mounts used.

Optical Mounts

Optical mounts represent the most important opto-mechanical component since they perform the critical operations of holding, positioning, and aligning the optics. Typically, the most important parameters in determining the performance of mounts are adjustment range, resolution, repeatability, orthogonality of motion, stability, thermal drift and cost. Understanding these parameters and individual mount mechanisms can help a designer select the right mounts for a given application. Range is the total angle or linear distance over which the mount can be adjusted. Increasing the range typically increases the mount cost, so low-cost mounts may permit only a small amount of “tweaking” around the nominal position. Resolution is the minimum movement that the mount can make. For instance, if a mirror is to reflect a beam of light onto a point on a surface, it must have adequate angular resolution to position the beam. Repeatability measures how closely the mount returns to the same position for a given adjustment. Repeatability is required most often in laboratory setups, where a set of conditions must be accurately replicated several times during data collection. Orthogonality describes the independence of the degrees of motion. For example, it would address whether angular adjustments also produce small translational movements. Stability specifies how well a mount remains in a set position over time. The adjustment mechanisms and mount materials influence mount stability. Finally, thermal drift measures the shift in mount position as the temperature changes. Mount design and materials influence thermal drift. Table 7 lists the major optical components as well as examples and properties of their associated optical mounts.

Since mirrors are the most commonly-used optical components, optical mirror mounts are ubiquitous and warrant a more detailed description. The mirror mount is a device that securely holds a mirror in place while allowing for precision tip and tilt adjustment. It generally consists of a movable front plate that holds the mirror, and a fixed back plate with adjustment screws. Adjustment screws drive the front plate about the axes of rotation in the pitch (vertical) and yaw (horizontal) directions. An optional third actuator often enables z-axis translation. Kinematic mirror mounts (see Figure 103) are the most common type, owing to excellent stability and relatively low cost. The kinematic mechanism for mirror adjustment is the best for providing the required performance for the vast majority of experiments performed in labs today. It does, however, have drawbacks, such as cross-coupled adjustment, beam translation, and limited angular travel. Other adjustment mechanisms for mirror mounts include flexure, which offers a more compact design with limited range, and gimbal, which allows for a large range of motion but with increased cost. Mirror mounts can be adjusted by hand with a micrometer head or adjustment screw or can be motorized for automation by using a linear actuator. The choice of the adjuster also determines the resolution of the mount.

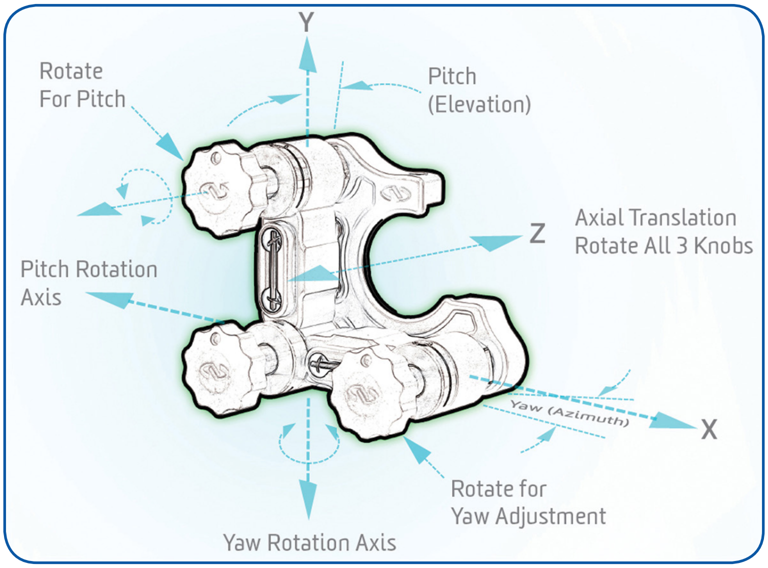


Figure 103. Kinematic mirror mount showing adjustment using two angular degrees of freedom (pitch and yaw) and one translational degree of freedom (axial).

Optical Posts and Bases

Once an optical mount is chosen, the optic and mount must be attached to the working surface, which can be an optical breadboard or optical table (see Section III.G.2) or a specially-designed platform. When the surface is a breadboard or table, the typical method for coupling between the mount and surface is by first attaching a post system to the mount and then attaching this post system either directly to the surface or to a base, which is attached to the surface (see Figure 104, for example). The two most common post systems are optical posts with post holders and pedestal risers. Posts with post holders are the more conventional alternative and are convenient because they offer continuous height adjustment. Pedestals, on the other hand, offer significantly more stability because of their wider diameters, but are available only in fixed heights. These two post systems react differently to forces such as vibrations or the force applied to the optical mount when making adjustments, which can affect the application. This motion becomes apparent during alignment of an optical system. The amount of motion is determined by the stiffness of the post or pedestal. The post system (particularly pedestals) can be attached directly to the working surface. This provides increased stability for the entire system but sacrifices flexibility in terms of modifying an optical system. Alternatively, post systems can be attached to kinematic bases (see Figure 104), which utilize high-strength magnets to securely couple the top plate (attached to the post system) to the bottom plate (attached to the surface). These bases are useful in setups where optical elements must frequently be interchanged, or when reference elements must occasionally be inserted in an optical path. Generally, the way the optical mount is coupled to the surface can make the difference between a stable setup and one that suffers from deleterious vibrations. When designing such a setup, it is important to try and minimize beam height, i.e., the height of the optic and mount above the working surface, use fewer parts (for instance, using a single post or pedestal rather than stacking them), and use parts made of thick, stiff material such as steel.

Figure 104. Opto-mechanical assembly for coupling an optical mount to an optical table including a post, post holder, and a kinematic base, which allows for mounting, removing, and replacing the assembly.

Beam Routing and Optical Rails

When a number of optical and opto-mechanical components are combined together to form an optical system (see Section III.C for examples), beam routing and optical rail systems allow for sensible grouping of these components on a common platform. A beam routing system provides a modular and versatile means for routing, shielding, and enclosing a laser or similar light beam above an optical table. The system is composed of a hollow cube to hold and enclose optics mounts, a stout post holder and clamping fork combination for mounting the cube to an optical table, and a large selection of beam tubes in various lengths to route the optical beam into and out of the cube (see Figure 105). Enclosed beams are therefore protected from air currents, dust, and stray light which can adversely affect an application. Enclosures can also help prevent accidental misalignment of a system. Two examples of optical systems that could benefit from a beam routing assembly are a variable laser attenuator (see Section III.C) or a fixed height periscope for vertical beam routing. For systems that may require translational adjustment with respect to one another, a rail system may be appropriate. Optical rail systems make repeatable linear alignment of optical components quick and easy. These systems typically consist of an optical rail that attaches directly to the working surface. The rail has a dovetail design which accepts rail carriers on top of which sit the optical component and its mounting hardware (see Figure 105). There is a linear scale engraved on the side of the rail for determining absolute and relative positions of the rail carriers. The beam expander and spatial filter optical systems described in Section III.C can easily be integrated onto an optical rail system.

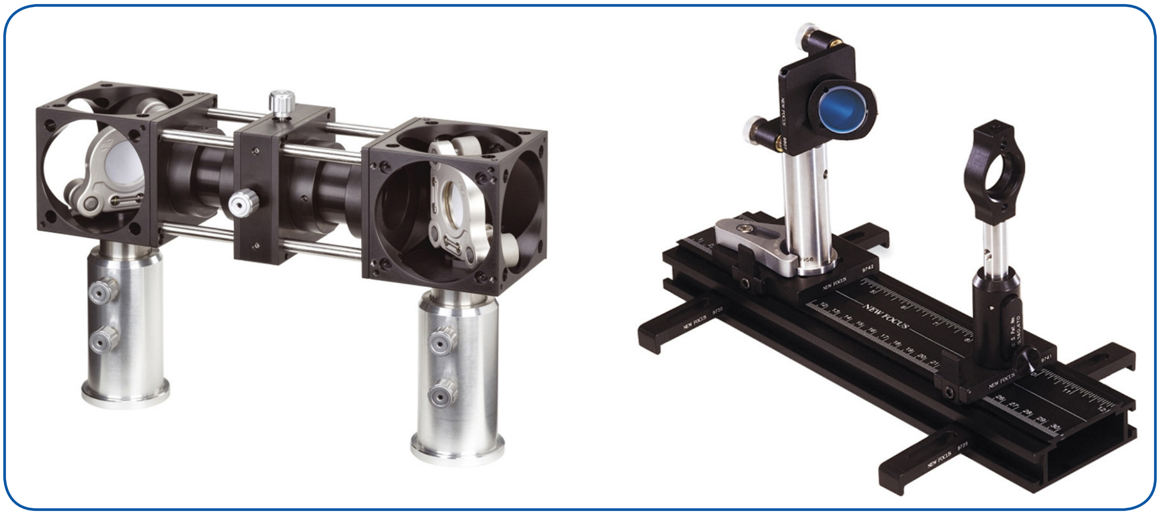


Figure 105. Examples of a beam routing (left) and rail carrier (right) system which allow for integration of multiple opto-mechanical components in a single assembly.

Additional online content discussing opto-mechanical components can be found in [186-190].

C. Optical Systems

By combining multiple optical and opto-mechanical components together, an optical system can be produced. Such a system can have significantly different functionalities from the individual components with which it is made. This section describes a few examples of such systems whose functionalities include modifying the input size of a light beam, improving the quality of the spatial beam profile of a laser, and continuous power control of a polarized laser beam.

Beam Expander

Beam expansion or reduction is a common requirement in many applications using lasers or other light sources. While beam expansion is discussed here, the optical system can simply be reversed to serve as a beam reducer. Simple beam expanders are essentially telescopes which, in their most basic forms, consist of two lenses. The input beam is assumed to be collimated. The first lens must have a diameter larger than the diameter of the input beam to avoid clipping the beam (see apertures topic in Section II.A.1). Similarly, the diameter of the output lens should be larger than the expected output beam diameter. The size of the output beam is the product of the magnification of the system (or expansion ratio) and the input beam diameter. The magnification of a two lens system is equal to the ratio of the focal length of the second lens to that of the first lens. The spacing between the two lenses is equal to the sum of the focal lengths of the lenses. One common beam expander geometry is based on a Galilean telescope (see Figure 106) where the first lens has a negative focal length and the second lens has a positive focal length. Because of the difference in signs of the focal lengths there is no focal point between the lenses and the distance between the lenses is the difference between the magnitudes of these focal lengths. This can be contrasted with another expander geometry based on a Keplerian telescope, which uses two positive focal length lenses [191]. In this case, the design is necessarily longer due to the addition of the focal lengths. Furthermore, the presence of a focal point within the system means that care should be taken when using lasers with high pulse energies, as energy densities can be large enough to ionize air. For the Galilean system, it is best to use a plano-concave lens for the negative lens and a plano-convex lens for the positive lens with the plano surfaces facing each other. To further reduce aberrations, only the central portion of the lens should be illuminated, so choosing oversized lenses is often recommended. An example of a Galilean-based beam expander on an optical rail system is shown in Figure 106.

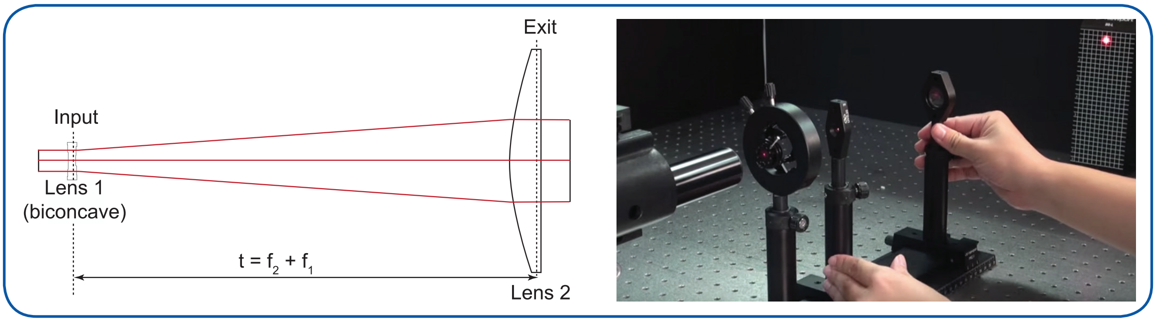


Figure 106. A Galilean telescope geometry for expanding a laser beam (left) and a Galilean beam expander system resulting in a 12 times expansion of the initial laser beam diameter (right).

Spatial Filter

Spatial filters provide a convenient way to remove random fluctuations from the intensity profile of a laser beam, which can be critical for applications like holography and optical data processing. Laser beams pick up intensity variations from scattering by optical defects and particles in the air. This is easily visualized by expanding a laser beam onto a screen where the patterns, holes, and rings superimposed on the ideal pattern of uniform speckle are due to scattering. Spatial filtering is conceptually simple (see Figure 107): an ideal coherent, collimated laser beam behaves as if generated by a distant point source. Spatial filtering involves focusing the beam and producing an image of the “source” with all its scattering imperfections defocused in an annulus about the axis. A pinhole centered on the axis can block the unwanted noise annulus while passing most of the laser’s energy. The proper pinhole size depends on the focal length of the focusing lens or objective and the diameter of the original beam. Interestingly, the Keplerian telescope geometry discussed in the beam expander section is ideal for use in a spatial filter set-up, as the intermediate focal point is an appropriate location to place the pinhole. It is necessary to make certain that the energy density of the beam at the focal point does not exceed the damage threshold limitation of the pinhole material. A spatial filter can be set up on an optical rail system. However, compact spatial filter assemblies are also typically available (see Figure 107). Proper positioning of the pinhole with respect to the focused beam is made using a positioner that allows for three degrees of translation.

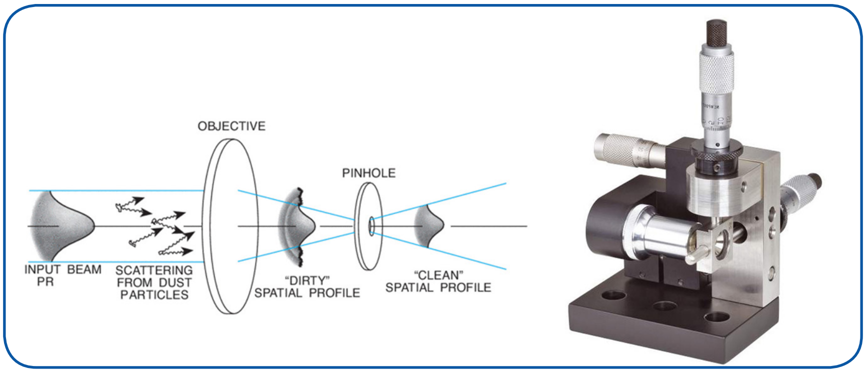


Figure 107. Illustration of the operation of a spatial filter in removing intensity fluctuation from a laser beam profile (left) as well as a compact spatial filter assembly consisting of a three-axis stage, a pinhole, and a focusing objective (right).

<u>Variable Laser Attenuator</u>

To adjust the laser power (or intensity) down to the level required for a specific application, neutral density filters or other devices are often used to reduce the pump power. These techniques typically do not allow for continuous control of the laser power. On the other hand, variable laser attenuators consist of essentially two optical components, a half waveplate, and a polarizer with a good extinction ratio, and allow for continuous power control for linearly polarized lasers. The principle of operation of a laser attenuator is based on Malus's law [138]. In this case, the intensity *I* of the polarized light that passes through the polarizer is given by:

$$I = I_0(cos\theta)^2 \ , \tag{21}$$

where I_0 is the input intensity, and θ is the angle between the beam's initial direction of polarization and the axis of the polarizer. Rotation of a half waveplate allows for continuous rotation of the beam's polarization direction (see Section III.A.4). When following a half waveplate with a polarizer, the total system provides variable attenuation according to Equation (21). Figure 108 shows the operation of a laser attenuator as well as its associated optical layout. Such an optical system would be ideal for a fully-enclosed beam routing system since the optical components are in fixed locations and the rejected light from the polarizer would be contained.

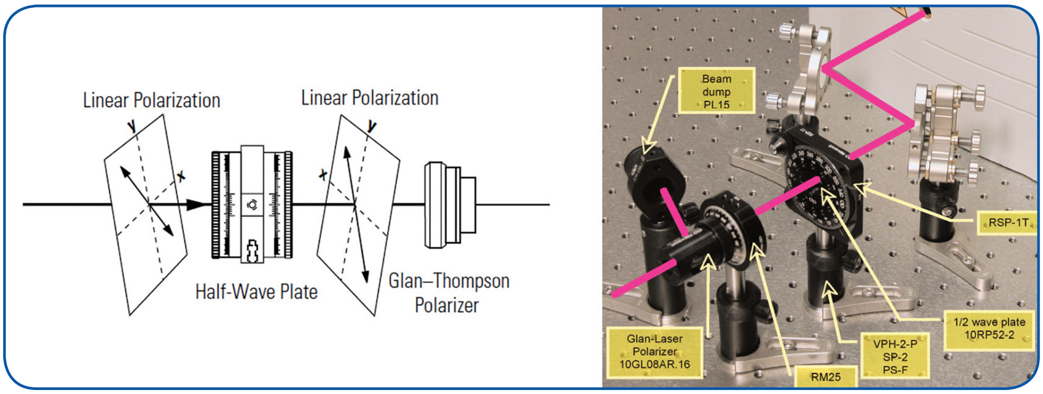


Figure 108. Operation of a variable attenuator for a laser (left) and example optical system which includes the two required polarization optics: a half waveplate and calcite polarizer (right).

Additional online content discussing the aforementioned optical systems can be found in [192-195].

D. Fiber Optics

Optical fiber and fiber optic cables are used as a means to transport optical energy and information over short or long distances. In combination with semiconductor laser diodes and photoreceivers, optical fibers have enabled the rapid development and proliferation of fiber optic telecommunication systems over the past thirty years [196]. Optical fibers are circular cross section dielectric waveguides consisting of a central core surrounded by a concentric cladding with a slightly lower (by ≈ 1%) refractive index. They are typically made of silica with index-modifying dopants such as GeO_2. A fiber optical cable is an optical fiber enclosed in a protective coating. This allows for ease of handling, reduces cross talk between adjacent fibers, and suppresses loss that occurs when fibers are pressed against rough surfaces. In addition to the benefits of light transmission, the confinement of light to a small area within the core of an optical fiber has enabled the development of fiber lasers (see Section I.A.4) and photonic crystal fibers. This section discusses the fundamental physics of optical fibers, their practical implementation, and the various types of optical fibers.

Basics of Optical Fibers

Figure 109 illustrates the directions of incident rays of light when they encounter an interface with a rarer medium, i.e., $n_2 < n_1$ as when light goes from glass to air [157, 197]. Rays 1 and 2 refract according to Snell's law where, unlike Figure 94 which shows an interface with a denser medium, rays bend away from the normal after passing through the interface. At a specific incident angle known as the critical angle θ_c, the angle of refraction is 90° (Ray 3), causing the light to travel along the interface between the two media. For any angle greater than θ_c, there is no refracted ray and the light experiences total internal reflection (TIR), causing it to simply obey the laws of reflection (as exemplified by Ray 4). When light is incident from a medium of higher index, determination of θ_c is accomplished by using Snell's law: $\theta_C = arcsin(n_2/n_1)$. As shown below, it is the phenomenon of TIR that enables propagation of light in optical fibers.

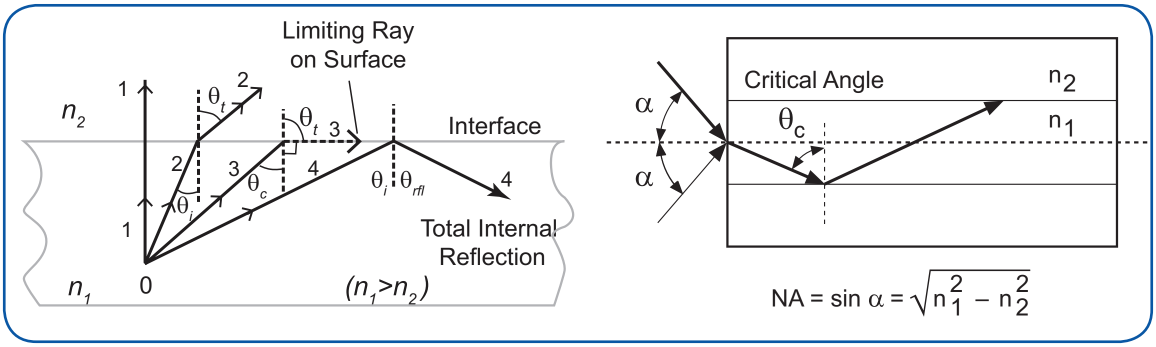


Figure 109. Illustration of critical angle and TIR (left, [157]). Light rays impinging on the core-cladding interface at an angle greater than the critical angle are trapped inside the core of the fiber (right). The relationship between the acceptance angle (α), the NA, and the refractive indices is also shown.

An optical fiber is a circular dielectric waveguide whose core has a higher index of refraction than the cladding. As shown in Figure 109, light will be confined to the core if the angular condition for TIR is met. The NA of a fiber is defined as the sine of the largest angle (α) an incident ray can have for TIR in the core [157, 197]. Qualitatively, NA is a measure of the light gathering ability of a fiber. It also indicates how easy it is to couple light into a fiber. The fiber geometry and composition determine the discrete set of electromagnetic fields, or fiber modes, which can propagate in the fiber. There are two broad classifications of modes: radiation modes and guided modes. Rays launched outside the angle specified by a fiber's NA will excite radiation modes. These modes carry energy out of the core where it is quickly dissipated. Rays launched within a fiber's NA typically give rise to guided modes which are confined to the core. These modes propagate energy along the fiber, transporting information and power. If the fiber core is large enough, it can support many simultaneously guided modes, i.e., multimode propagation. When light is launched into a fiber, the modes are excited to varying degrees depending on the conditions of the launch, e.g., input cone angle, spot size, axial centration, and can take on a variety of spatial distributions [1]. Much like for the transverse modes of a laser (see Section I.A.3), the lowest-order mode of an optical fiber has a near-Gaussian spatial distribution and therefore possesses many of the same benefits. This is why single single-mode operation is often desired in a fiber. The normalized frequency parameter of a fiber, also called the V number, is a useful specification in this regard. It describes the number of modes at a given wavelength based on the fiber's NA and its core radius [1, 197].

Light power propagating in a fiber decays exponentially with length due to absorption and scattering losses (see Figure 110). Attenuation is the single most important factor in fiber optic telecommunication systems, as it directly impacts acceptable signal levels. In the NIR and VIS regions, the small absorption losses of pure silica are due to tails of absorption bands in the FIR and UV. Impurities, notably water in the form of hydroxyl ions, are much more dominant causes of absorption in commercial fibers. Recent improvements in fiber purity have reduced attenuation losses to the order of 0.1 dB/km. Scattering losses also contribute to attenuation in the form of small-scale index fluctuations in the fiber when it solidifies and irregularities in the core diameter and geometry.

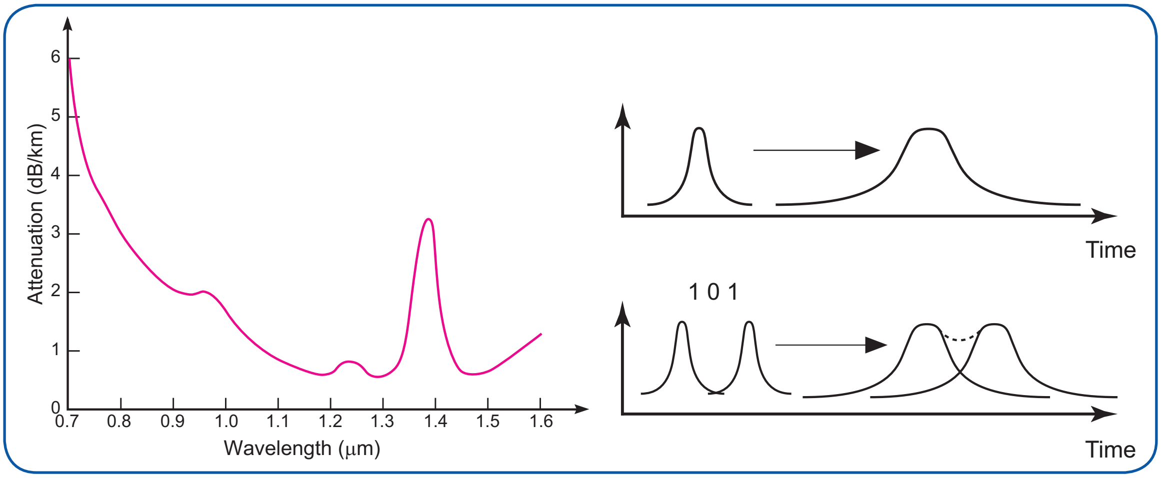


Figure 110. Typical spectral attenuation in a silica fiber (left). Dispersion causes spreading of a single optical pulse in time as it propagates down a fiber (upper right). Multiple pulses representing a bit stream of information become unrecognizable after propagation due to dispersion (lower right).

The bandwidth of an optical fiber determines its data rate [196, 197]. The mechanism that limits a fiber's bandwidth is known as dispersion. Dispersion is the spreading of the optical pulses as they travel down the fiber. The result is that pulses begin to spread into one another and the symbols become indistinguishable (see Figure 110). Dispersion limits both the bandwidth and the distance that information can be transported. There are two main categories of dispersion: intramodal and intermodal. There are two distinct types of intramodal dispersion: chromatic dispersion and polarization-mode dispersion. Chromatic dispersion is simply the result of the variation in the material's refractive index with wavelength. Polarization mode dispersion is due to orthogonal polarization modes in a fiber traveling at different speeds due to birefringence. Intermodal dispersion occurs because different propagating modes travel with different velocities. Therefore, this category of dispersion only applies to multimode fibers.

A single-mode fiber supports a mode which consists of two orthogonal polarization modes. This is the result of the asymmetry in the fiber core cross-section. Normally, external stresses are randomized and the resulting induced birefringence helps to scramble or randomize the polarization. A specialty fiber known as a polarization-maintaining fiber intentionally creates a consistent birefringence pattern along its length. This is accomplished by modifying the geometry of the fiber and the materials used to create a large amount of stress in one direction. This large induced birefringence dominates the random birefringence, allowing polarization to be maintained during propagation within the fiber. Controlling the polarization state in an optical fiber is similar to the free space control using waveplates via phase changes in the two orthogonal states of polarization (see Section III.A.4). This is accomplished by applying stress-induced birefringence to a fiber. This induces a retardation enabling the creation of a waveguide-based waveplate. Figure 111 shows one such polarization device which consists of a fiber squeezer that rotates around the optical fiber. Applying a pressure to the fiber produces a linear birefringence, effectively creating a fiber wave plate whose retardation varies with the pressure.

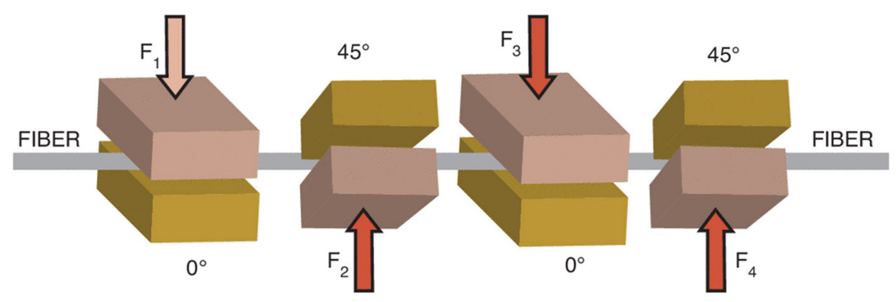


Figure 111. Polarization control in an optical fiber initiated by squeezing fiber from various directions.

Fiber Coupling and Fiber Types

The characteristics of the focused beam (typically a laser beam) must match the fiber parameters for good coupling efficiency [198, 199]. The general guidelines are (1) the focused spot should be comparable to the core size, (2) the focused beam should be centered on the fiber core, and (3) the incident cone angle should not exceed the NA of the fiber. Conditions (1) and (2) are illustrated on the left side of Figure 112 and condition (3) is illustrated on the right side of the figure. The first two conditions are easy to accommodate for multimode fibers owing to their large core diameters. Consequently, good coupling efficiency is achieved in a multimode fiber by matching the coupling lens to the fiber NA. Coupling into single-mode fibers is a fundamentally more difficult problem. Single-mode fibers have small core diameters requiring more opto-mechanical components that enable coupling of the focused beam with sub-micron positioning resolution. Furthermore, the mode of the incident laser light must match the mode of the fiber. In other words, the coupling efficiency depends upon the overlap integral of the Gaussian mode of the input laser beam and the nearly Gaussian fundamental mode of the fiber.

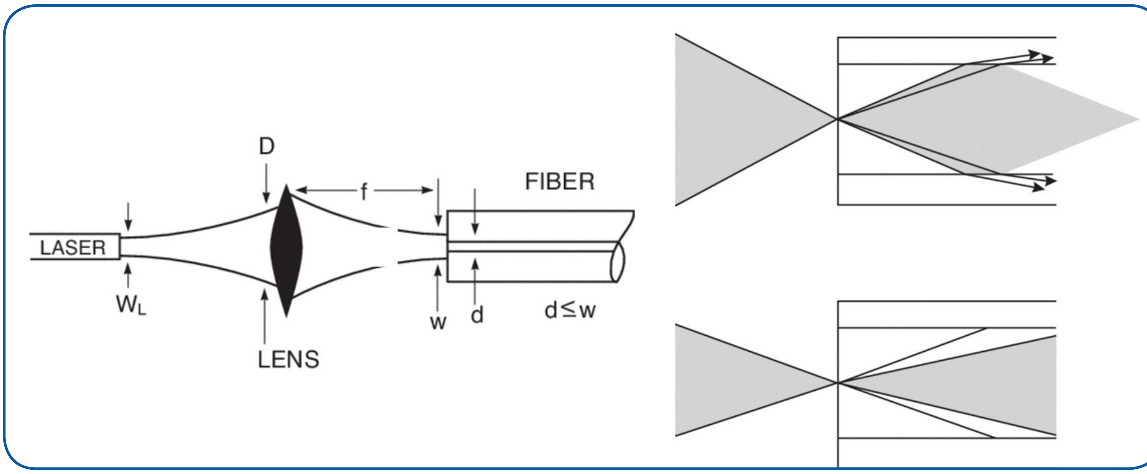


Figure 112. A schematic of coupling of light into a multimode or single-mode optical fiber (left). Launching conditions in a multimode optical fiber resulting in an overfilled (upper right) and underfilled (lower right) fiber.

There are a large variety of optical fibers available that can exhibit very different geometries as shown in Figure 113. Standard single-mode fibers for telecommunications applications have small core diameters (< 10 μm) whereas multimode fibers have core diameters between sixty and several hundred microns. The latter can also have index profiles that are either graded or stepped. Specialty fibers are also common, including polarization maintaining fiber, high power delivery fiber optic cables, bend-insensitive optical fiber, and fibers for extreme temperature conditions. Due to their ubiquity and utility, two specific types of specialty fibers are described in more detail below: rare earth doped fibers and photonic crystal fibers (PCFs).

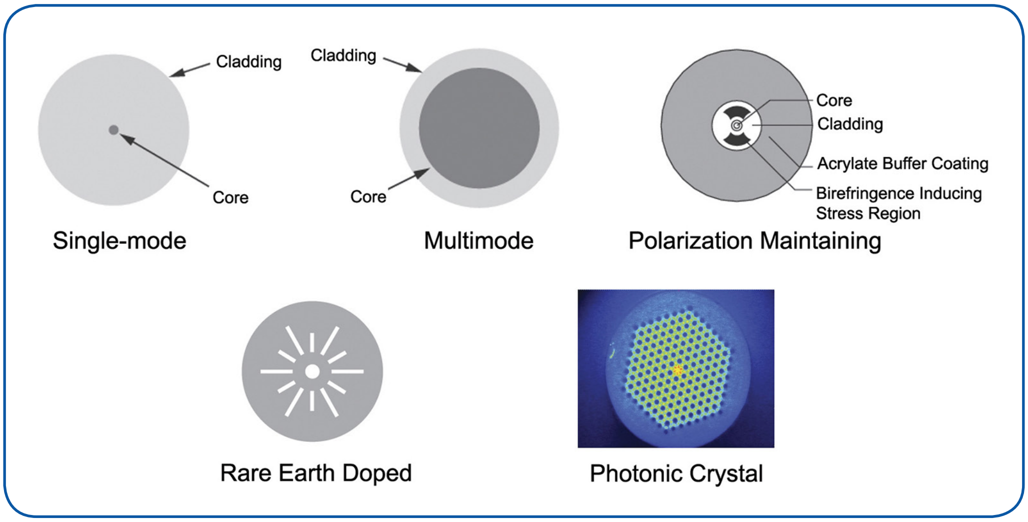


Figure 113. Various types of fibers.

Rare earth doped fibers are particularly important for fiber lasers since these dopants, e.g. Nd, Yb, and Er-Yb co-doped, can act as laser gain media (see Section I.A.4). As shown in Figure 20, the use of double-clad rare earth doped fibers can allow for efficient matching of the pump beam, whether it is delivered by free-space focusing or via another optical fiber. These doped fibers can also be used as photosensitive fibers for fabricating FBGs. A Bragg grating is a periodic modulation in a material's index of refraction that enables reflection of light with a wavelength of twice the grating period. High quality FBGs can be constructed by exposing photosensitive optical fibers to periodic patterns of UV light (rare earth dopants absorb strongly in the UV). The grating forms when the fiber is exposed to a periodic pattern of UV light, typically generated with a phase mask. Such fabrication methods are clearly attractive from a production point of view since they allow for rapid and reliable manufacturing. FBGs enable high reflectivity (up to 99%) over a narrow wavelength band (see Figure 114), which is useful for producing a cavity mirror in a fiber laser (see Figure 20) or as spectral filters in fiber optic telecommunications systems.

A photonic crystal is a microstructured material in which there is a periodic variation in the index of refraction as a function of position. In PCFs, this periodic variation is achieved through a regular pattern of voids, or air holes that run parallel to its axis (see Figure 113). Unlike traditional fibers, both the core and cladding are made from the same material. All the waveguiding properties in a PCF thus derive from the presence of the voids. PCFs are generally divided into two main categories: index guiding fibers that have a solid core, and photonic bandgap fibers that have periodic microstructured elements and a core of low index material, e.g. hollow core. PCFs provide characteristics that ordinary optical fibers cannot, such as single-mode operation from the UV to IR with large mode-field diameters, exceptionally high nonlinearity, NA values ranging from very low to about 0.9, and optimized dispersion properties [1, 200]. Applications of PCFs are found in a wide range of research fields like spectroscopy, metrology, biomedicine, imaging, telecommunication, industrial machining, and defense.

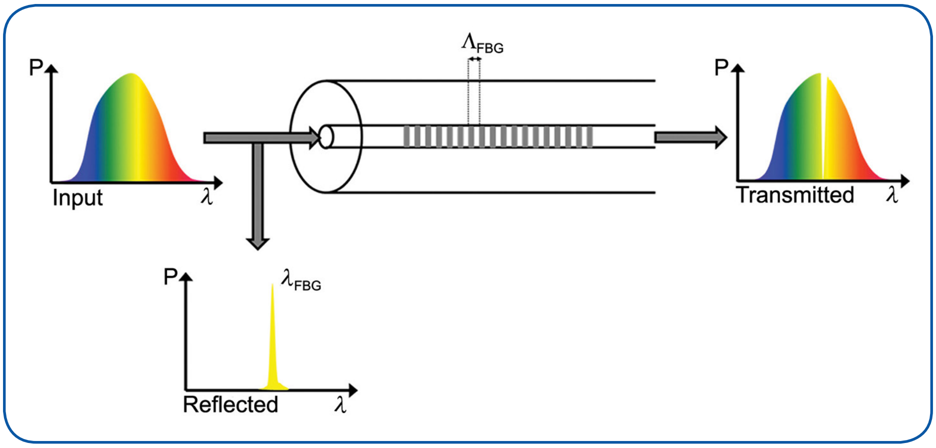


Figure 114. Schematic of an FBG and representative transmission and reflection spectra.

Additional online content discussing fiber optic basics and applications can be found in [201-206].

E. Optical Modulators

Electro-optic amplitude and phase modulators control an optical beam's amplitude, phase, and polarization state electrically. Such modulators differ from acousto-optic modulators where acoustic waves are used to control light. Details about acousto-optic modulators can be found in [1] while this section describes electro-optic modulators. In communications systems, electro-optic modulators impress information onto an optical frequency carrier. Unlike direct modulation of the laser itself, external modulators do not cause any degrading effects to laser linewidth and stability. In measurement systems, amplitude modulators can be used as actuators to hold laser beam intensity constant, or as optical choppers to produce a pulse stream from a *CW* laser beam. Phase modulators are used to stabilize the frequency of a laser beam, or to mode-lock a laser. This section discusses phase and amplitude modulators and the underlying phenomenon that enables them, the electro-optic effect.

Electro-optic Effect

The linear electro-optic effect or Pockels effect, is the modification in the index of refraction in proportion to an externally applied electric field [1, 207]. The effect of an applied electric field on the index of refraction, as seen by an optical beam polarized in an arbitrary direction in a crystal, is described by a third-rank tensor. This change in the refractive index (Δn) of a crystal has the form:

$$\Delta n = \frac{1}{2} r n^3 E_{app} \,, \tag{22}$$

where n is the unperturbed index of refraction, r is the appropriate element in the electro-optic tensor (known as the electro-optic coefficient), and E_{app} is the applied electric field. Crystals with large electro-optic coefficients are chosen to minimize the required fields. Such crystals include those made from BBO, potassium dideuterium phosphate (KD*P), $LiNbO_3$, magnesium-oxide-doped $LiNbO_3$, and KTP. Even for crystals with large values of r, the electro-optic effect is small. For example, an electric field of 10^6 V/m applied to a lithium niobate crystal will produce a fractional index change of roughly 0.01%. Consequently, relatively large fields must be applied to these crystals when being used in electro-optic modulators.

Phase and Amplitude Modulators

There are two types of modulators: bulk and integrated-optic. Bulk modulators are made out of discrete pieces of optical crystals and are typically used on a lab bench or an optical table. They feature very low insertion losses, and high power-handling capability. Integrated-optic modulators, which are not discussed here, use waveguide technology to lower the required drive voltage and are wavelength specific. Unlike bulk modulators, these modulators are fiber pigtailed and compact. When choosing a modulator, it is important to understand the application requirements, i.e., phase or amplitude modulation, broadband versus resonant (or single-frequency) operation, and the operational wavelength range (which dictates the anti-reflection coating to be used on the crystal). A typical phase modulation set-up is shown in Figure 115.

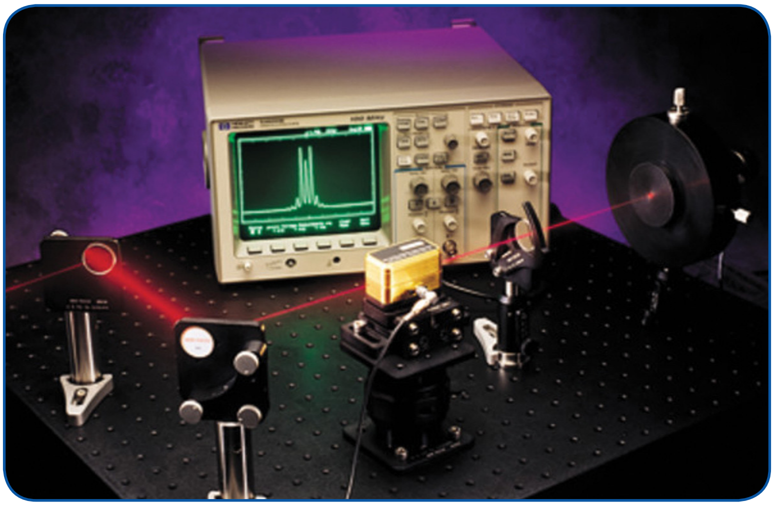


Figure 115. A typical set-up involving a phase modulator: a laser beam is sent through a resonant phase modulator and the beam is focused into an optical spectrum analyzer. The laser's phase-modulated spectrum, with its characteristic frequency sidebands, is observed on an oscilloscope.

Phase modulators are used to vary the phase of an optical beam and are the simplest electro-optic modulators. Here, an electric field is applied along one of the crystal's principal axes. Light polarized along the principal axis experiences an index of refraction change, hence an optical path length change, that is proportional to the applied electric field (see Equation (22). The phase of the optical field exiting from the crystal therefore depends on the applied electric field. The most common bulk phase modulator is the transverse modulator, as shown in Figure 116, which consists of an electro-optic crystal between parallel electrodes. These modulators develop large electric fields between the electrodes while simultaneously providing a long interaction length in which to accumulate phase shift. The optical phase shift is obtained by applying a voltage between the electrodes (or E_{app}) and a commonly used figure-of-merit for electro-optic modulators is the half-wave voltage, V_π. It is defined as the voltage required to produce an electro-optic phase shift of π (or 180°). When driven sinusoidally, phase modulators can generate frequency sidebands at multiples of Ω on a *CW* optical beam with a central optical frequency of ω. The electric field of an optical beam after passing through a phase modulator can be seen in Figure 116. This frequency-modulating property makes phase modulators useful in laser mode-locking (see Section I.A.5).

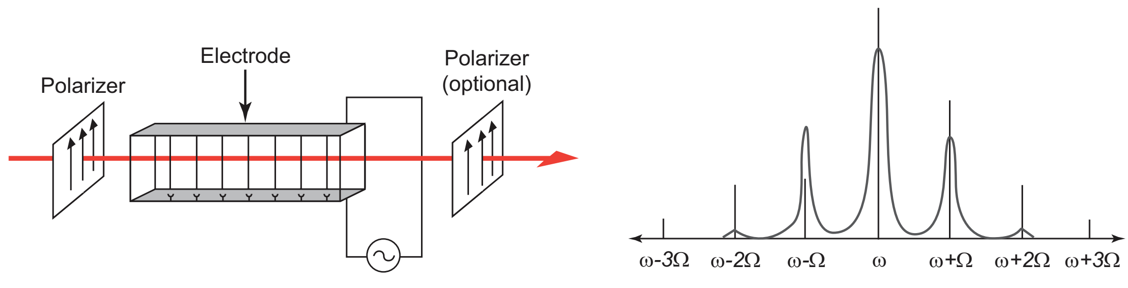


Figure 116. Typical operation of a phase modulator (left) and the resulting phase-modulated electric field (right). The optical intensity of each sideband is proportional to the square of the electric field amplitude.

Bulk electro-optic amplitude modulators consist of a voltage-tunable wave plate followed by a polarizer. The geometry of a simple amplitude modulator, as shown in Figure 117, consists of a polarizer, an electro-optic crystal, and an analyzer. The input polarizer guarantees that the optical beam is polarized at 45° to the crystal's principal axes. When an external electric field is applied, the electro-optic effect changes the indices of refraction along the two crystal directions to a different degree, thereby changing the retardation of the effective waveplate. The crystal acts as a variable waveplate, changing the exit polarization as the applied voltage is increased. The analyzer transmits only the component of the exit polarization that has been rotated. The relationship between the transmission and applied field has the same functional form as that of laser beam attenuator (see Equation (21)). To obtain linear amplitude modulation, these modulators are often biased at 50% transmission and only operated with small applied voltages (see Figure 117). Two ways to bias the modulators are by adding a DC voltage through a bias tee, or adding a quarter-wave plate before the analyzer. In order to suppress birefringence variations due to temperature changes, two matched crystals can be arranged in series with their applied electric fields oriented at 90° relative to each other (Figure 117).

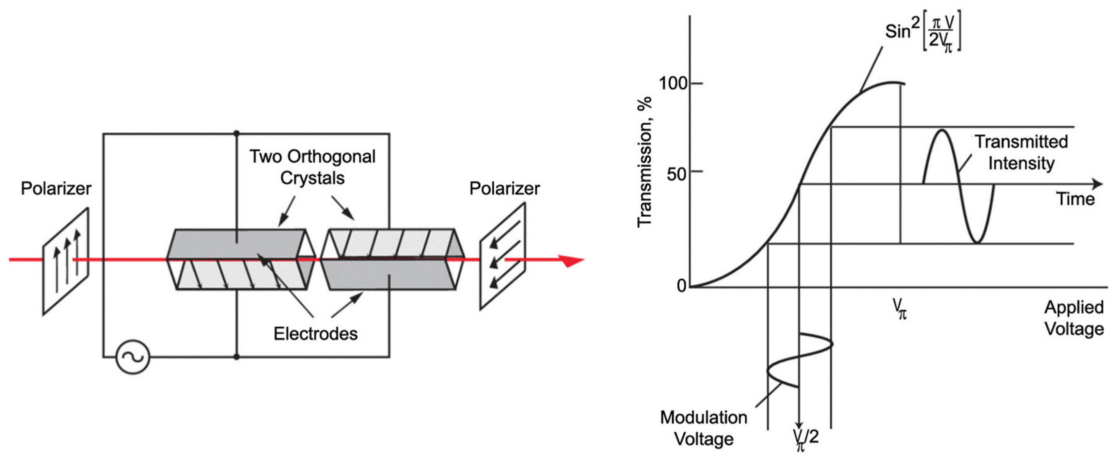


Figure 117. Typical operation of an amplitude modulator (left) and the transfer function of an amplitude modulator between crossed polarizers biased at the 50% transmission point (right).

Additional online content discussing the electro-optic effect and modulators can be found in [208-212].

F. Motion Control

Reliable and repeatable positioning of equipment is critical in laser research, fiber optic communications, semiconductor wafer manufacturing, metrology, bio-medical research, defense and security, and industrial manufacturing. A motion control system provides the accuracy and precision necessary to properly position a piece of equipment for a given application. Generally, such a system consists of a controller, a driver, and a motion device. A very simple example is an optical mirror mount, which typically allows for rotational alignment of the mirror (see Section III.B). Alignment is often performed manually and so the user is both the controller and driver of the system while the mount is the motion device. Due to inherent limitations in manual adjustment sensitivity and the complexity associated with multidimensional positioning, automated motion control systems are utilized for many applications. In this section, the basics of both manual and automated motion devices, such as stages and actuators, are detailed. Then the primary components of an automated system are discussed, including the electronic controller and the motor driver, which control the dynamics of the motion device.

1. Motion Fundamentals

The function of a motion device, such as a stage, is to generate a desired motion in an ideal trajectory. Furthermore, the device must be able to reliably and repeatably reach a specified target position. Factors like friction between moving parts, quality of guides or bending induced by moving loads, often deviate the motion from this ideal trajectory or result in differences between the actual and target position. Details regarding the various factors that impact the motion of these devices and how to characterize their positioning precision and accuracy are the subjects of this section.

Positioning Basics

Any positioning stage is considered to have six degrees of freedom: three linear translational, along the x, y, and z-axes and three rotational about those same axes. The purpose of a stage is to produce motion along an ideal trajectory, which requires constraining motion to certain degrees of freedom. Any motion in non-constrained directions will contribute to deviation from the ideal trajectory and/or position. For instance, the function of a linear translation stage is to constrain motion along an ideal straight line. Runout of a linear stage is the linear (versus angular) portion of off-axis error and represents the departure from desired, ideal straight-line motion. Runout consists of two orthogonal components referred to as flatness and straightness. Figure 118 shows a linear stage whose ideal straight line motion is confined to the x-axis. In this case, flatness deviation is displacement along the z-axis while straightness deviation is displacement along the y-axis. Similarly, angular runout (or tilt) of a linear stage is the rotation of the moving functional point or the point where a measurement and/or process is occurring. It has three orthogonal components commonly referred to as pitch, roll, and yaw (see Figure 118) and can be a complex combination of the three components. Furthermore, if multiple stages are connected to one another to create a multi-axis system, cross-coupling can occur which means that a change in one axis can result in an unwanted change to another axis.

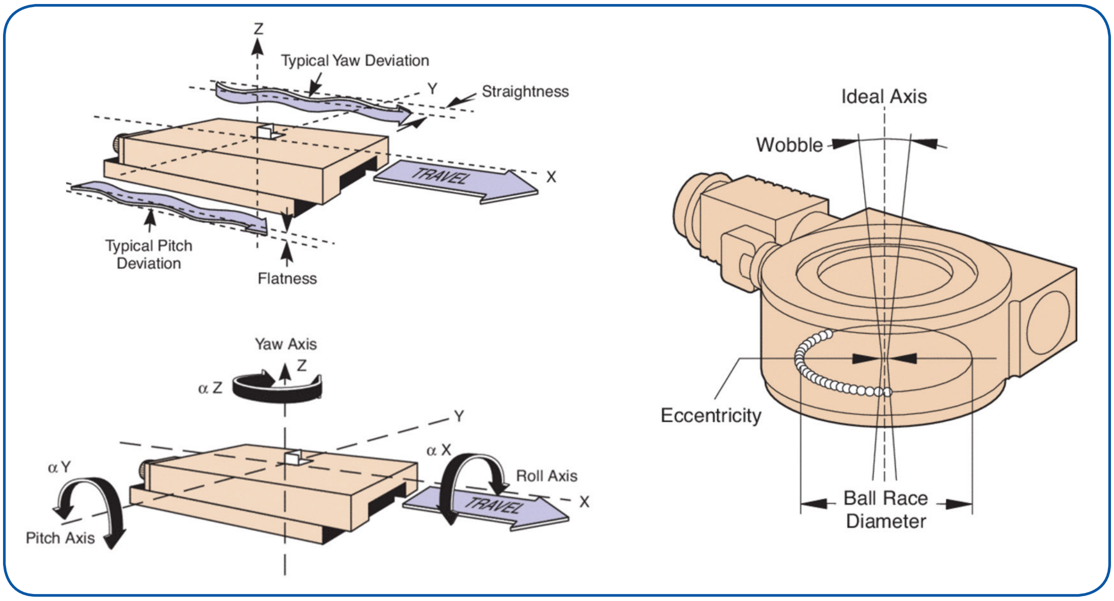


Figure 118. Flatness and straightness runout of a linear stage (upper left). Roll, pitch and yaw angular runout of a linear stage (lower left). Off-axis deviations in a rotary stage (right).

For a rotational stage, eccentricity is the radial (perpendicular to the axis of rotation) deviation of the center of rotation from its mean position as the stage rotates through one revolution (see Figure 118). It is also referred to as radial runout. A perfectly centered stage with perfect bearings would have no eccentricity. Wobble of a rotary stage is the tilt of the axis of rotation relative to the ideal axis over one revolution. It is most easily observed as a cyclic tilting of the rotating surface or table top of a stage. Like eccentricity, it is generally the result of imperfect bearings.

In addition to traveling an ideal trajectory, a stage should reliably reach and maintain a specified target position. Reversal error is the distance between the actual positions reached for a given target position, when approached from opposite directions (see Figure 119). This value is a combination of backlash and hysteresis. Backlash is the result of relative movement between interacting mechanical parts of a drive system that does not produce output motion. Contributing factors include clearance between mechanical parts such as gear teeth and mechanical deformation. Not all systems have backlash but, when they do, it mainly affects bi-directional repeatability (see below). Backlash can be compensated by motion controllers due to its repeatable nature. Hysteresis is a component of reversal error that is dependent on the recent history of the system. It is the result of elastic forces in the various components and is observed when the forces acting on a system reverse direction. Hysteresis affects both bi-directional repeatability and accuracy (see below). Unlike backlash, hysteresis is present in all mechanical systems although its value may be low. While reversal error relates to a stage's ability to reach the target position, position stability is the ability to maintain a position within a specified position range over a specified time interval. It is the sum of drift and vibrations. Drift is the slow deviation from a stable position. It mainly depends on the migration of lubricants and thermal variations. Vibrations are fast alternative motions of small amplitude generated by the environment, e.g. noise from the flow, air fans, and electronics, e.g. a motor driver.

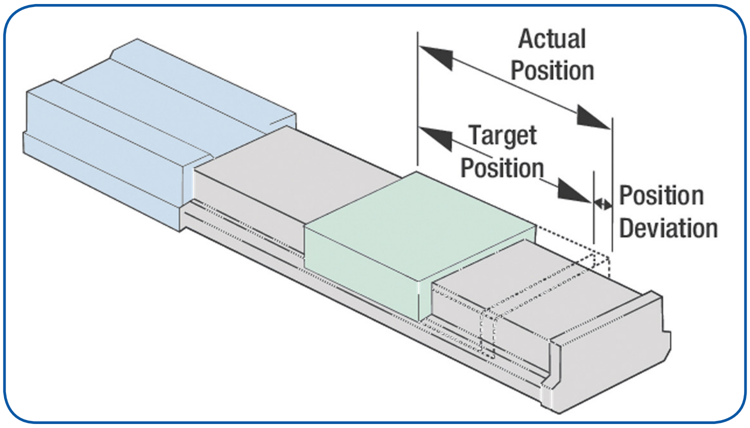


Figure 119. Illustration of a position deviation for a linear stage, which is the actual or measured position reached by the functional point minus the target position.

In addition to position, the rate of change of position, or speed, is an important factor in motion systems. The maximum speed specification is provided at the stage's normal load capacity (see below). Higher speeds are possible for lower loads or larger motor drivers. Minimum speeds are highly dependent on a motion system's speed stability. Speed stability is a measure of the ability of a motion system to maintain a constant speed within specified limits. It is usually specified as a percentage of the desired speed. Acceleration is the rate of change of speed, which is often set to achieve the maximum speed in a set amount of time. Friction can play a significant role in the speed, speed stability, and acceleration of a stage. Friction is defined as the resistance to motion between surfaces in contact. Elements contributing to friction may be in the form of drag, sliding friction, depleted lubrication, system wear, or lubricant viscosity. Stiction is the static friction that must be overcome to impart motion to a body at rest. Since static friction is generally greater than moving friction, the force which must be applied to impart motion is greater than the force required to keep the body in motion. As a result, when a force is initially applied, the body will begin to move with a "jump" that results in position and/or speed overshoot.

Load capacity is the maximum allowable force that can be applied to a stage in a specified direction while meeting stage specifications. This maximum force includes static (mass times gravity) and dynamic forces (mass times acceleration). Dynamic forces must include any external forces such as vibrations acting upon the stage. The amount of acceleration a stage can impart to a mass is limited to the accelerating force it can produce without exceeding the load capacity. In particular, the centered normal load capacity is the maximum load (centered on the carriage and in a direction perpendicular to the axis of motion) that can be applied to a linear stage (see Figure 120). For rotary stages, it is the maximum load along the axis of rotation. Transverse load capacity, also called side load capacity, is the maximum load that can be applied perpendicular to the axis of motion and along the carriage surface. This is typically smaller than the normal load capacity. Axial load capacity is the maximum load along the direction of the drive train. For linear stages mounted vertically, the specified vertical load capacity is usually limited by the axial load capacity. The maximum load capacity of a stage is diminished when the load is not centered. Inertia is the measure of load's resistance to change in speed. The larger the inertia, the greater the force required to accelerate or decelerate the load. If there is a constraint on the amount of force available, then the allowable acceleration and deceleration must be adjusted to an acceptable value. Inertia is a product of mass elements and the square of their distance from the axis of rotation. The maximum inertia specified for rotary stage is a value based on available torque.

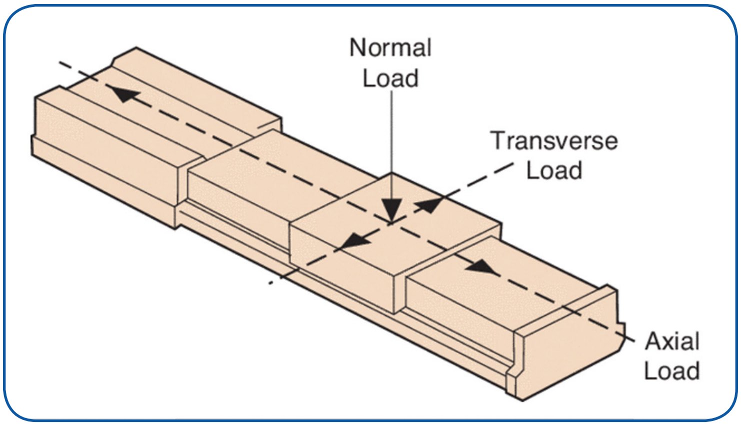


Figure 120. A stage's different load capacities.

Motion Control Specifications

When selecting the appropriate positioner for an application, it is common to evaluate them with regard to the product specifications, in particular, repeatability and accuracy. Repeatability is often confused with accuracy but, as illustrated in Figure 121, a system may be very repeatable yet lack accuracy. Accuracy is a measure of the degree to which a given displacement conforms to an agreed upon standard. For instance, runout is also known as straight-line accuracy. The accuracy of a motion system can also be highly influenced by the test set up, environmental conditions, and the procedure used to measure displacement. With the majority of modern controllers, linear error compensation can be easily accomplished by entering a compensation factor into the controller. Therefore, accuracy after compensation is often specified for a particular stage.

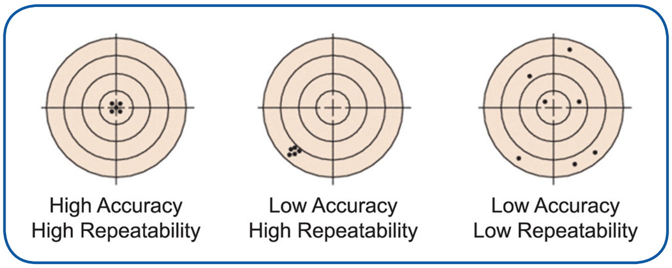


Figure 121. A depiction of the differences between accuracy and repeatability.

Repeatability is a measure of the positioning system's ability to sequentially position. It can be unidirectional (when approaching the target position always from the same direction) or bidirectional (when approaching the target position from either direction). In a number of applications, the repeatability of a motion system is more important than the accuracy. Systematic errors can be taken into account and compensated, but the repeatability is the ultimate limit that is reached after all compensation. The general idea of repeatability is the measure of the ability of a system to achieve a commanded position over many attempts when approached from either the same or different directions (see Figure 122). Reversal error and position stability essentially determine a system's repeatability. Repeatability is essential to ensuring quality in many different areas. Without proper repeatability in a manufacturing environment, there can be no reliable process to assure that products are manufactured the same way and meet the same specifications all the time.

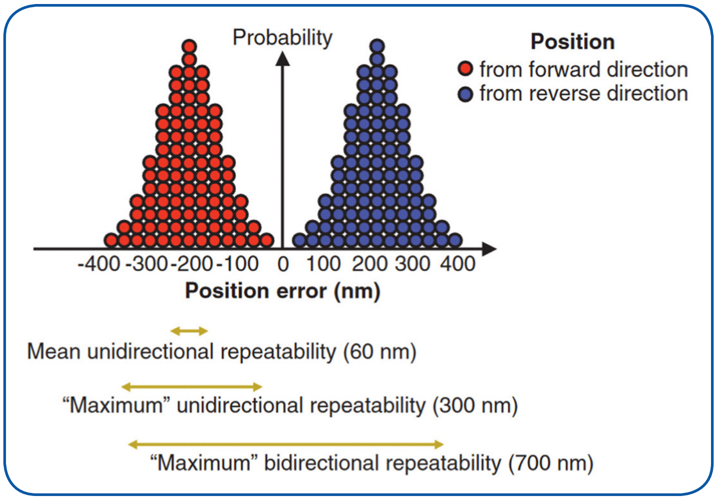


Figure 122. Measurements that yield a distribution of position errors define a motion control system's repeatability.

Resolution is the smallest increment that a motion system can be commanded to move and/or detect. A system may or may not be able to consistently make incremental moves equal to the resolution. Minimum incremental motion (MIM) is the smallest increment of motion a device is capable of consistently and reliably delivering. Factors that can affect this motion output include friction, load, external forces, system dynamics, controller, vibrations, and inertia. The MIM should not be confused with resolution, which is typically based on the smallest controller display value or smallest encoder increment. Resolution can be significantly smaller than the smallest actual motion output, a key distinction, but not always well understood. Sometimes MIM is referred to as "practical resolution."

Additional online content discussing the motion fundamentals can be found in [213-216].

2. Motion Control Systems

An automated motion control system consist of three main components: a motion controller, a motor driver or amplifier, and a motion device. The primary purpose of a motion controller is to control the dynamics of the motion device. The motor driver converts the command signals from the motion controller into power signals required to move the motor. The motion device is any mechanical device that provides motion and is actuated by a motor. Such motion devices typically contain feedback devices to provide information such as position and velocity to the motion controller. In this section, motion devices are discussed first and mainly in the context of manual positioning. This discussion is equally applicable to automated positioning and motorized drivers and electronic controllers are then detailed.

Motion Devices

Motion devices are mechanical positioning devices such as linear translation stages, rotation stages, and actuators. While the specifications of a stage or actuator are important selection criteria, they may not be exhaustive enough or directly applicable for each application. For this reason, it is important to have sufficient understanding of the inherent abilities of the components that make up a stage. This section provides a brief discussion of the most common components used in high precision positioning equipment with their pros and cons. The main components of a motion device are the materials used for the body construction, the mechanism that enables translation or rotation, and the drive mechanism.

Each material used for mechanical components in motion control has its own unique set of advantages and disadvantages. Table 8 provides a summary of the properties for the most commonly used materials in motion mechanics. Stiffness is a measure of the amount of force required to cause a given amount of deflection. Young's modulus is a material-dependent constant that quantifies the stiffness with large values indicating greater stiffness. Thermal expansion is the change in size or shape of an object, such as a stage, due to a change (increase or decrease) in temperature. When temperature change across a component is non-uniform, such as when a heat source like a laser diode is present, a material which does not dissipate heat may be susceptible to distortions caused by thermal gradients. In this case, the relative thermal distortion, i.e., ratio of the coefficients of thermal expansion to thermal conductivity, becomes important with lower values being preferred. Aluminum is a lightweight material, with good stiffness-to-weight ratio, and has low thermal distortion. It is also fast-machining, cost-effective, and does not rust. However, anodized surfaces are highly porous, making them unsuitable for use in high vacuum. Steel has very good stiffness, good material stability, low thermal expansion, and is well suited to high vacuum applications. Machining of steel is much slower than aluminum, making steel components considerably more expensive. Corrosion of steel is a serious problem, but stainless steel alloys can minimize these problems. Brass is a dense material and fast machining. The main use of brass is for wear reduction where it can be used to avoid self-welding effects with steel lead-screws or shafts. Brass has a less desirable stiffness-to-weight ratio and does not have ideal thermal expansion or thermal conductivity properties. Granite is an extremely hard material allowing polishing to very flat surfaces, which is beneficial in positioning accuracy and repeatability of a total system. Granite also has a very low thermal expansion coefficient. However, for large structures and table surfaces, the mass of a granite structure can become impractically large.

Parameter	Steel	Aluminum	Brass	Granite
Young's Modulus (stiffness), E, Mpsi (GPa)	28 (193)	10.5 (72)	14 (96)	7 (48)
Thermal Expansion, a (μin/in/°F)	5.6	12.4	11.4	4
Thermal Conduction, c (BTU/hr-ft-°F)	15.6	104	67	2
Specific Stiffness, E/ρ	101 (25.4)	108 (27.7)	45.6 (11.3)	70 (17.8)
Relative Thermal Distortion, a/c	0.36	0.12	0.17	2
Density, ρ, lb/in^3 (gm/cc)	0.277 (7.6)	0.097 (2.6)	0.307 (8.5)	0.1 (2.7)

Table 8. Properties for common stage materials.

The load and trajectory performance of a translation or rotation stage is primarily determined by the type of bearing or flexure used. Bearings are the preferred mechanism since they provide smooth low-friction rotary or linear movement between two surfaces. They are the primary elements that determine the runout errors of a stage, define the stiffness, and the static load capacity of a stage. Bearings employ either a sliding (dovetail) or rolling action (ball or crossed-roller) as shown in Figure 123. In both cases, the bearing surfaces must be separated by a film of oil or other lubricant for proper performance. Dovetail slides are primarily used for manual positioning and consist of two flat surfaces sliding against each other. They can provide long travel, and have relatively high stiffness and load capacity. However, they do possess high stiction, and the friction varies with translation speed, which makes precise control difficult and limits sensitivity. Ball bearing slides reduce friction by replacing sliding motion with rolling motion. Balls are constrained by vee-ways or hardened steel rods and the friction is very low, resulting in extremely smooth travel. Since the contact area available to transmit loads is smaller in vee-groove bearing ways, ball bearings have a lower load capacity than crossed-roller or other bearings. In order to carry the same sized load, the balls would need to be larger in diameter or be greater in quantity. Crossed-roller bearings offer all of the advantages of ball bearings but with higher load capacity and higher stiffness. This is a result of replacing the point contact of a spherical ball with the line contact of a cylindrical roller. Due to the averaging characteristic of line contacts, angular and linear deviations are generally lower than those found in ball bearings. However, crossed-roller bearings require more care during manufacture and assembly resulting in higher costs. A flexure mechanism uses the elastic deformation of a material (typically a high-strength steel spring) to provide translation. This mechanism requires no lubrication and is virtually free of the stiction normally associated with bearings. However, when used in a translation stage, travel range is limited to just a few millimeters. Also, care must be taken so that permanent deformation does not occur, causing reduced functionality. In addition to these mechanical bearings, air bearings can also be used which provide a low-friction interface via a thin film of pressurized gas. Details regarding these types of bearing are discussed in Chapter 2, Section II.B.5.

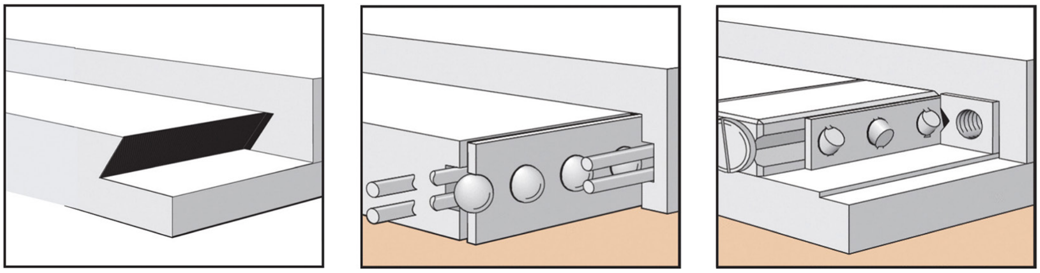


Figure 123. Different types of bearing mechanisms allowing for stage motion: dovetail (left), ball bearing slide (middle), and cross-roller bearing (right).

A stage can be driven directly by a motor (see below) or indirectly based on different mechanical systems (see Figure 124). A popular technique for moving loads is to use the axial translation of a nut riding along a rotating screw. Lead screws use sliding contact, so their wear rate is directly proportional to usage. The advantages of lead screws include self-locking capability, low-noise motion, low initial costs, ease of manufacture, and a wide choice of materials. In order to eliminate possible backlash between the screw and the nut, the nut needs to be preloaded to the screw via an external spring, gravitational forces (applicable only to vertical use), or by a double nut with a spring in between. Recirculating ball screws are essentially lead screws with a train of ball bearings riding and rolling between the screw and the nut in a track. The primary advantage of ball screws is less screw heating, which can impact the stage's repeatability and accuracy. Also, because of the reduced friction, most ball screw stages can run at higher speeds and can perform smaller incremental motions compared to lead screw-driven stages. The large number of mating parts makes tolerances critical, thus increasing manufacturing costs. Also, ball screws generate more noise than lead screws due to the recirculating balls in the nut. The worm gear system transforms rotary motion from one plane into another plane by meshing a screw (worm) with a gear (worm wheel). As the screw is turned, the worm threads mesh with the gear, causing it to rotate. Worm drives are commonly used as a drive system for rotation stages and allow very low-profile design. In order to eliminate backlash, the worm and the worm wheel need to be in perfect contact with each other, which requires a sophisticated worm preloading system with high transversal stiffness.

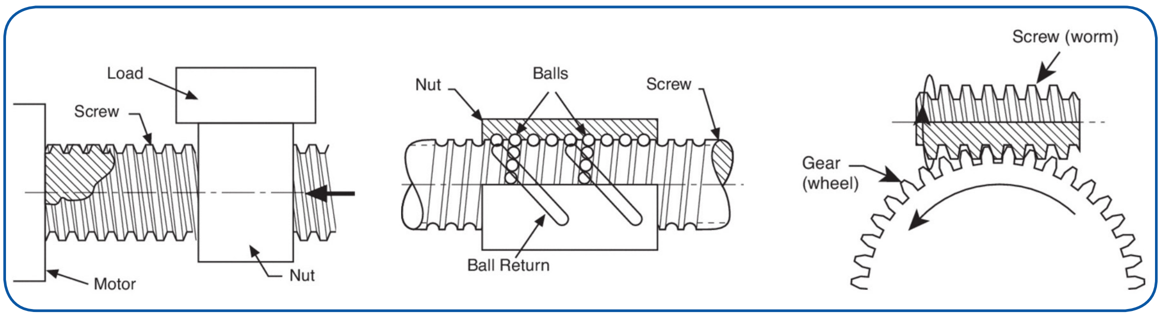


Figure 124. Different mechanisms for indirectly driving a stage: lead screw (left), ball screw (middle), worm drive (right).

Actuators (see Figure 125) also allow for indirectly driving a stage but are typically externally coupled and therefore, provide flexibility in terms of matching a particular stage with the desired drive mechanism. Manual actuators are simple, low-cost options for positioning and can be described as a high sensitivity lead screw with a knurled knob. Unlike the lead screw system described above, the nut of the screw is fixed to the stage body, and the screw itself moves back and forth. Springs press the carriage against the screw tip to make good contact and to preload the screw and eliminate backlash. Micrometer heads are the adjustment mechanism of choice if accurate position read-out or repeatable positioning is needed. Standard metric micrometer heads feature a scale in units of 10 µm but, with an additional vernier, can reach a resolution of 1 µm. When resolution of much less than one micron is needed, a differential screw is recommended. These devices use the difference between two screws of nearly the same pitch to produce very fine motion. Motorized linear actuators provide the ability to motorize manual linear translation stages for remote and/or computer control. Such actuators can either use the lead screw mechanism described above or can utilize the piezoelectric effect, which exploits interactions in certain crystalline materials to produce mechanical movement when an electric field is applied. These piezo actuators can achieve resolution of a few tens of nanometers and are sometimes referred to as nanopositioners. This increased resolution typically comes at a cost of reduced speed and/or travel range.

Figure 125. Various types of actuators including a manual actuator (left), a micrometer (middle), and motorized actuators (right).

Motorized Drivers

A motion device can be electronically controlled through either a direct or indirect motorized drive system. Common indirect drive systems for linear and rotary stages are based on the lead screws, ball screws, and worm drives discussed in the previous section. Shaft couplings, transmission belts, and gearboxes are often located between the drive system and the driving motor. These components affect system dynamics such as speed and torque capacity, but can also introduce backlash and hysteresis. The two most-common motors used for indirect drive systems are brushed DC motors and stepper motors. A brushed DC motor consists of a rotor placed in a magnetic field, which causes rotation when current is applied to the motor windings. The rotational speed is proportional to the applied voltage, while the torque is proportional to the current. DC motors are best characterized by their smooth motion and high speeds. A stepper motor operates using the basic principle of magnetic attraction and repulsion. Steppers convert digital pulses into mechanical shaft rotation. The amount of rotation is directly proportional to the number of input pulses generated, and speed is proportional to pulse frequency. One difference between a DC and a stepper motor is that when a voltage is applied to a DC servo motor, it will develop both torque and rotation. However, when a voltage is applied to a stepper motor, it will develop only torque. For the stepper motor to rotate, the current applied must be commutated or switched. Stepper motors are often used in open-loop control systems, a low-cost alternative to closed-loop DC servo systems. The pulse count is a good indicator of position, and stepper motors work reliably when used within their specified torque and speed range. However, the motion of a stepper motor becomes unpredictable outside of its specified range and skipping steps, extra steps or motor stalling can result. Stepper motors typically develop torque almost instantaneously, faster than with a DC brush motor. Hence, stepper motor-driven stages can deal with mechanical stiction better than DC motor-driven stages that often generate position overshoots when the motor torque exceeds the stiction.

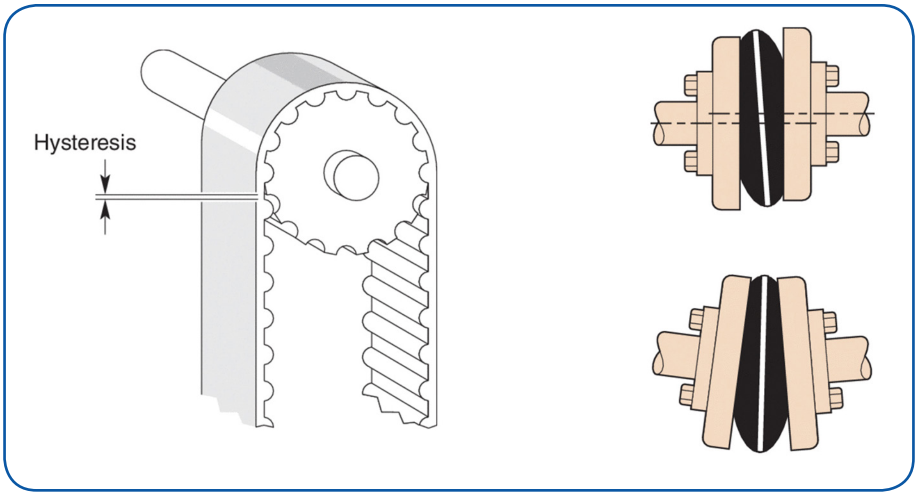


Figure 126. Components for connecting the drive system and the driving motor, including a belt drive (left) and a flexible shaft coupling (right).

In direct drive systems, the motor is directly coupled to the motion, with no screw or transmission system in between. The most common high-precision direct drive systems feature either a brushless linear motor for linear stages or a brushless torque motor for rotation stages. A linear motor consists of a permanent magnet assembly which establishes a magnetic flux, and a coil assembly which generates a force proportional to coil current. Linear motors have become very important components of precision positioning systems, with numerous advantages over traditional mechanical actuators such as ball screws. These systems typically provide higher quality, frictionless motion, and higher speeds and acceleration.

A motor driver receives input signals from a controller and converts them to power to drive a motor. A motor driver can be a simple amplifier or it can be an intelligent device that can be configured through software for varying operational parameters. Different motor drivers support the different types of motors used in motion control. The stepper motor driver receives input signals from the motion controller commanding it to step the motor to a commanded position. The driver then applies current to the stepper motor windings in order to move the stepping motor to the next step or increment. Drivers for DC motors simply convert a –10 V to +10 V analog control signal from the motion controller to a usable current to drive the motor. Most brushless DC motor drivers are simple amplifiers that convert control signals from the motion controller to a usable current to drive the motor, with the motion controller providing the motor commutation.

Electronic Controllers

In a motion system, the controller is used to manipulate stages and actuators so that they move or stop in a desired manner. Common motion systems use three types of control methods: position control, velocity control, and torque control. Each control method is based on a feedback device whose basic function is to transform a physical parameter, e.g., a scale reading, into an electrical signal for use by the controller. Most motion systems use the position control method. The purpose of the motion controller in these systems is to command a motor so that the actual position of the moving mechanism tracks the desired position specified by a preplanned trajectory. In this case, the primary feedback device is an encoder which directly monitors the position and provides the motion controller with actual displacement information. Typically, this is done optically by detecting light passing through a series of accurately-spaced slits in a metal or glass disc (see Figure 127). Velocity control is used in applications where velocity regulation is of primary importance, such as in spindles or conveyor belts. In these applications, the primary feedback device is a tachometer. Torque control is used in applications such as robotics where the torque applied by end-effectors must be controlled accurately in order to grasp or release objects. In these applications, the primary feedback device is a torque/force sensor such as a strain gauge.

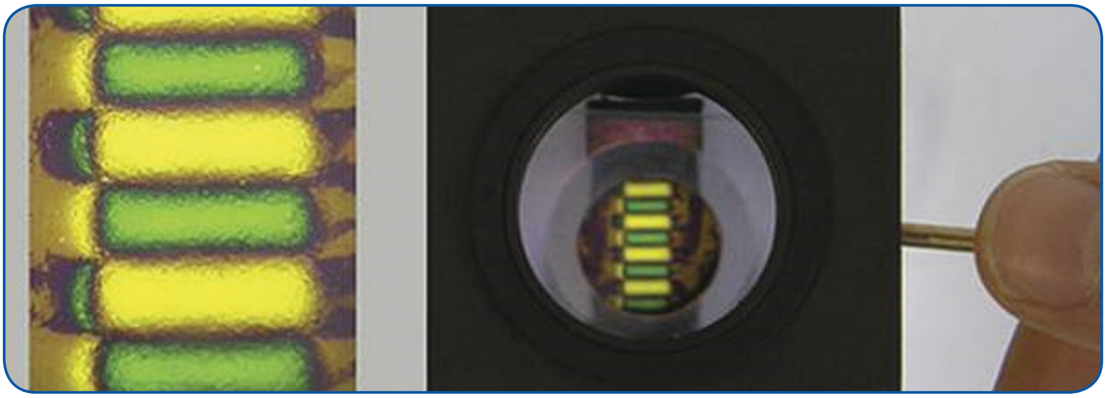


Figure 127. Example of a linear steel scale encoder that ensures positioning with accuracy of ± 1 μm.

While the main objective of a motion controller is to control a motion device, many advanced motion controllers provide additional functions such as:

- Trajectory generation for moving devices from one point to another or for coordinating the motion of multiple devices
- An interface to let users configure and command the motion system to perform various tasks
- Monitoring end-of-travel limits, amplifier faults, feedback errors, etc. for safety of the system
- Digital input/output lines to synchronize external events to motion or vice versa
- Memory for storing and running on-board motion programs

Furthermore, the output of the motion controller can be configured depending upon the type of motor used to move a motion device.

Additional online content discussing motion control systems can be found in [217-222].

G. Vibration Control

Laboratories and industrial facilities contain a variety of vibration sensitive equipment used for research and production in fields such as laser-based research, high-resolution imaging, semiconductor manufacturing, and biotechnology. Inserting electrical probes into nuclei of living cells, etching sub-micron lines in nanostructures, or taking SEM images requires laboratory environments to have vibration levels well below human perception thresholds. Floor vibrations can cause imaging components, specimens, lasers, or substrates to move relative to each other, causing blurry images, low yields, and erroneous results. The degree to which this relative motion affects results depends on the amplitude and frequency of the environmental vibrations as well as the sensitivity of the experiment. Common sources of noise and vibration are shown in Figure 128.

Vibration control systems that include vibration isolators and optical tables, are intended to minimize the impact of environmental vibration. The optical table serves as a common base for the whole opto-mechanical assembly. Opto-mechanical components such as posts, rods, and mounts, as well as positioning stages, are made to anchor optical elements in place so that the optical paths will be undisturbed by environmental impacts such as vibration. This section describes the fundamentals of vibration and how these concepts relate to measurable parameters such as compliance and maximum relative motion. These concepts are then used to understand how various systems implement isolation and damping to minimize the impacts of vibration.

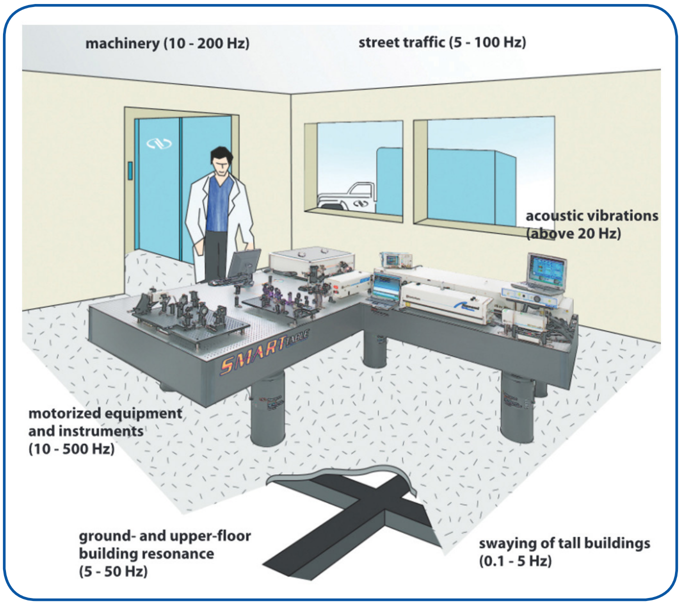


Figure 128. Sources of mechanical noise or vibration in a typical laboratory setting.

1. Vibration Fundamentals

Vibration is essentially a periodic force. In order to understand its potential impact on an optical component, it is instructive to start with a simple model that assumes a rigid body (the optical table to which the component is attached) is coupled via a spring to an applied force (the vibration). The resulting interaction can provide a lot of information about how the body moves in response to the force, including the magnitude of the movement and its frequency response. Introducing a mechanism that allows for this movement to be deadened or reduced, i.e., damping, provides for a more realistic system. While this damped harmonic oscillator model is instructive, real structures, such as an optical table, are much more complicated since they can undergo complex deformations when they experience vibrations. The compliance of a system describes how such a structure moves as a result of an external force. While this parameter is both descriptive and quantitative, other parameters such as the relative motion can provide more germane information about how an optical component attached to a table will respond to vibration. These concepts describing responses to vibration are detailed below.

Damped Simple Harmonic Motion

Vibration and vibration isolation are both intimately connected with the phenomenon of resonance and simple harmonic motion. One example of harmonic motion is a mass connected to a flexible cantilevered beam (see Figure 129). An external force, either from a one-time impulse or from a periodic force such as vibration, will cause the system to resonate as the spring alternately stores and imparts energy to the moving mass. The frequency at which the system resonates is called the natural frequency. In the example of the mass and beam, the natural frequency is determined by two factors: the amount of mass, and the stiffness of the beam, which acts as a spring. As illustrated in Figure 129, a lower mass and/or a stiffer beam increase the natural frequency while a higher mass and/or a more compliant (softer) beam lower the natural frequency. This example represents an undamped system in which there is no mechanism to dissipate mechanical energy. With a small amount of damping, the system will vibrate for quite a long period of time before coming to rest. Damping dissipates mechanical energy from the system and attenuates vibrations more quickly. For example, when a finger touches the resonating mass-beam system lightly, this damping action rapidly dissipates the vibrational energy.

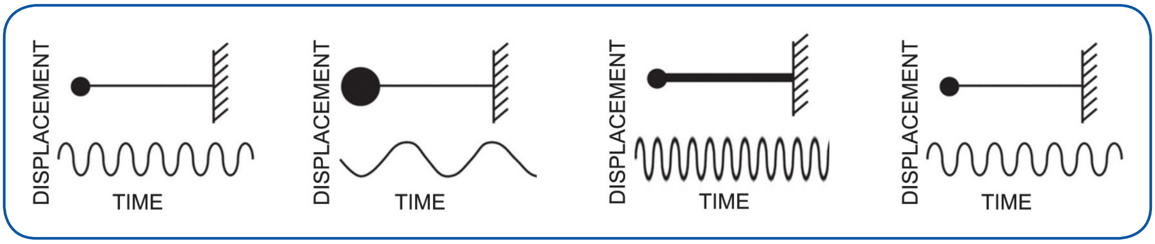


Figure 129. Illustration of how the mass (left two figures) and spring (right two figures) affect the natural frequency of a system undergoing simple harmonic motion.

More formally, as shown in Figure 130, the simple harmonic oscillator consists of a rigid mass M connected to an ideal linear spring [223]. The spring has a static compliance C, such that the change in length of the spring (Δx) that occurs in response to a force F is: $\Delta x = CF$. The compliance of the spring is its susceptibility to move and is the inverse of the spring stiffness (k). When the system oscillates, its natural frequency (ω_0) is given by: $\omega_0 = \sqrt{1/CM}$. The natural frequency of the system is determined solely by the mass and the spring compliance. It decreases for a larger mass or a more compliant (softer) spring. The ability to control this resonance frequency is critical in the proper design of optical tables when the goal is to reduce the impact of vibrations (see Section III.G.2). If the spring-mass system is driven by a sinusoidal displacement, e.g., vibration, with frequency ω and peak amplitude $|u|$, it will produce a sinusoidal displacement of the mass M with peak amplitude $|x|$. The steady-state ratio of the amplitude of the mass motion $|x|$ to the spring end motion $|u|$ is called the transmissibility T. This transmissibility is a function of the driving frequency ω as shown in Figure 130. There are three key features of this system. Well below the resonance frequency ($\omega << \omega_0$), $T = 1$ and so the motion of the mass is the same as the motion at the other end of the spring. Near resonance ($\omega \approx \omega_0$), the motion of the spring end is amplified, and the motion of the mass $|x|$ is greater than that of $|u|$. Above resonance ($\omega >> \omega_0$), the displacement $|x|$ decreases in proportion to $1/\omega^2$. In this case, the displacement $|u|$ applied to the system is not transmitted to the mass. The spring acts as an isolator. Controlling the transmissibility curve is the key to designing effective isolators as shown in Section III.G.2.

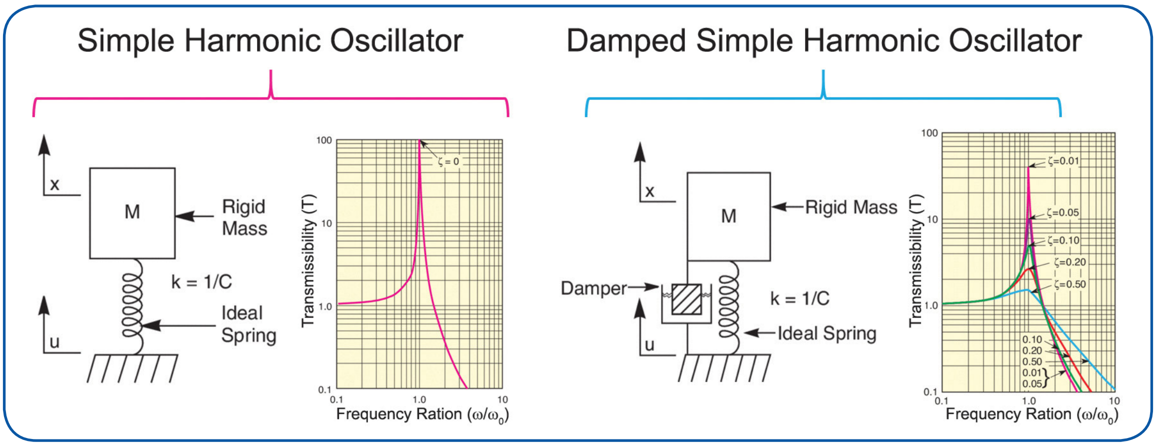


Figure 130. Depiction of a simple harmonic oscillator and its transmissibility curve (left) and a damped simple harmonic oscillator and its transmissibility curve for various values of the damping coefficient (ζ) (right).

A more realistic system includes a mechanism to dissipate mechanical energy from the system (very often as heat). A damped simple harmonic oscillator is shown schematically in Figure 130. A rigidly connected damper is expressed mathematically by adding a damping term proportional to the velocity of the mass to the differential equation describing the motion [223]. Similar to the undamped system, an external force results in a displacement amplitude of the spring, which is related to the mass displacement by the transmissibility. Plots of T versus ω are shown in Figure 130 for various values of the damping coefficient (ζ). When ζ approaches zero, the curve becomes exactly the same as for the undamped system. However, as the damping increases, the amplitude at resonance decreases and the “roll-off” at higher frequencies decreases, i.e., the transmissibility declines more slowly as damping increases. It is the reduction in amplitude at resonance that explains why the use of damping is a key strategy in both the design of optical tables as well as isolators (see Section III.G.2).

Compliance

While the discussion of a rigidly connected mass-spring system is informative, no actual structure is a perfectly rigid body. All structures vibrate by flexing and twisting. The response of structures to random vibrations can be quite complicated because they vibrate with complex deformations and have more than one resonant frequency. The resonant frequencies of an object are the natural frequencies of vibration determined by the physical parameters of the vibrating object. The compliance curve, the classic method of measuring dynamic rigidity, is a useful tool for evaluating the basic dynamics of a vibrating structure [224]. Compliance is a measure of the susceptibility of a structure to move as a result of an external force. The greater the compliance, i.e., the lower the stiffness, the more easily the structure moves as a result of an applied force, e.g., vibration. Compliance curves show the displacement amplitude of a point on a body per unit force applied ($C = |x| / |F|$), as a function of frequency. The units of compliance are typically given in mm/N or in/lb. The dynamic performance of a table top is usually characterized with a compliance curve, a log–log plot of the table's dynamic response to random vibration (see Figure 131). The compliance of a rigid body is proportional to $1/\omega^2$ and is graphed as a straight line with slope of -2. This line, which is called the ideal rigid body (IRB) line, represents the dynamic performance of a theoretically perfect rigid table. For non-rigid bodies, a compliance curve shows the structure's resonant frequencies and its maximum amplification at resonance.

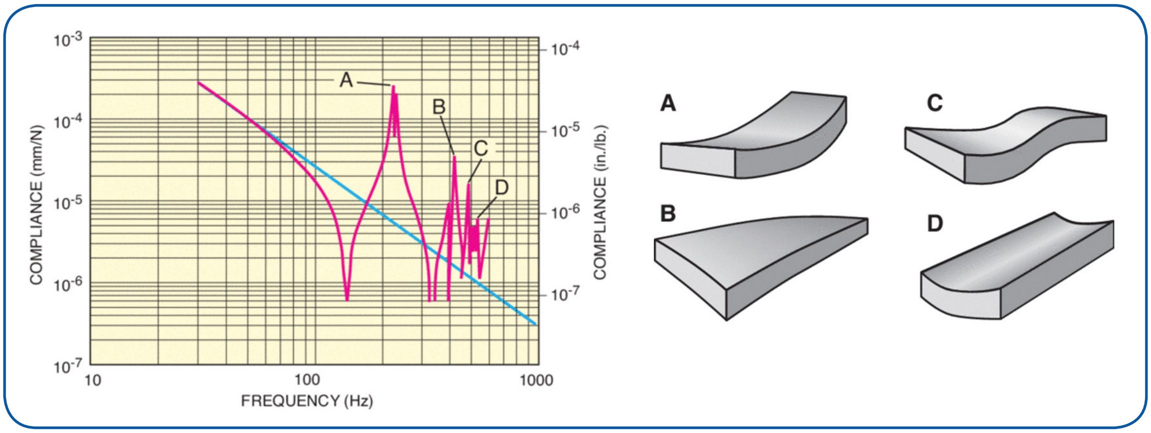
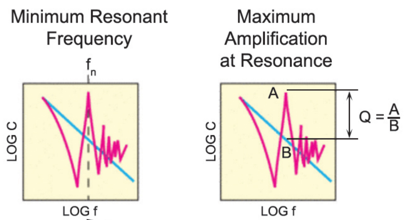


Figure 131. Typical compliance curve of an undamped table top with the IRB in blue (left) and vibrational modes of the table top indicated in the compliance curve (right).

Figure 131 shows the relationship between an undamped table top's vibration modes and the peaks on its compliance curve. Each peak in the curve, marked A through D, corresponds to a fundamental vibration mode. A table top's response to vibration depends on the frequency range. For low frequencies, the compliance decreases inversely proportional to ω^2. In other words, the structure is behaving as an ideal rigid body. For frequencies greater than 80 Hz, the compliance curve begins to deviate from this line, structural vibrational modes are excited, and the table begins to deform. The rigid body compliance falls off rapidly as frequency increases, so the largest displacements are generally caused by low-frequency resonances. The first peak on the left (~ 220 Hz in Figure 131) usually has the highest amplitude and dominates the table's response to vibration. The highest possible minimum resonant frequency is desirable, because the amplitude of table displacements is much smaller at higher frequencies, providing greater stability. Damping of table top resonance modes is critical for maximum stability. Effective table top damping reduces compliance, i.e., reduces the height of resonance peaks. The goal is to design a table top whose compliance curve deviates as little as possible from its theoretical IRB line. With other information, compliance curves can also furnish a reliable estimate of how a particular system will perform in a particular application. Maximum amplification at resonance, or *Q*, is a measure of how much the compliance curve deviates from the IRB line. In exact terms, it is defined as the maximum compliance value of the highest peak above the IRB line (usually, but not always, the first peak on the left) divided by the IRB response at the same frequency (see Figure 132). The lower the *Q* a structure has, the better it is damped and the more stable the structure will be. A structure that is more stable is less likely to vibrate in response to disturbances. The compliance curve supplies information on the two key parameters that govern dynamic performance — minimum resonant frequency and maximum amplification at resonance. These values can be used to calculate more meaningful table top specifications such as the dynamic deflection coefficient and relative motion.

Maximum Relative Motion

The dynamic rigidity of a table top, i.e., the resistance of the top surface to movement to vibration, is the single most important measure of vibration control performance. But compliance curves, the classic method of measuring dynamic rigidity, do not go far enough in providing a quantitative measure of table top vibration control capabilities. The dynamic deflection coefficient, a figure of merit that can be derived from any compliance curve, enables one to compare dynamic performance directly. When the ambient vibration level is known, the dynamic deflection coefficient can be used to calculate the relative motion value, which can then be used in selecting the most appropriate table for an application. The equation for the maximum relative motion is shown in Figure 132 (details on its derivation are given in [225]). This equation allows for calculation of the worst-case relative motion between two points on a table at the natural frequency (f_n). Each term can be understood in terms of the physical mechanisms described in the previous two sections. The dynamic deflection coefficient (second term in the equation) is derived from the table top's minimum resonant frequency and damping efficiency, which together quantify the table top's dynamic performance. The third term, the applied power spectral density (PSD), is the contribution of the applied vibration intensity level, which can be measured directly or estimated using a table of typical values. Isolator transmissibility, the fourth term, accounts for the attenuation of ground vibrations at the frequency range of interest through the support structure. Values of the dynamic deflection coefficient and relative motion for a typical lab environment are often specified for an optical table.



Typical PSD Values*

Relatively Quiet Lab	$<10^{-10}$ g²/Hz
Lab Next to Street	$<10^{-9}$ g²/Hz
Light Manufacturing	$<10^{-8}$ g²/Hz

*Assumes random vibration as found in real world environments

$$\text{Maximum Relative Motion at } f_n = \underbrace{g\left(\frac{1}{32\pi^3}\right)^{\frac{1}{2}}}_{\text{Constant}} \bullet \underbrace{\left(\frac{Q}{f_n^{\;3}}\right)^{\frac{1}{2}}}_{\text{Dynamic Deflection Coefficient}} \bullet \underbrace{(\mathrm{PSD})^{\frac{1}{2}}}_{\text{Floor Vibration Level}} \bullet \underbrace{T}_{\text{Isolator Transmissibility}} \bullet \underbrace{2}_{\text{Multiplier}}$$

Constant [g = 386 in/sec² (9.8 m/sec²)]; Dynamic Deflection Coefficient; Floor Vibration Level (g²/Hz); Isolator Trans-missibility (<0.01 Above 10 Hz Typical); Multiplier For Worst-Case Motion Between Two Points

Figure 132. Formula for determining the maximum relative motion between two points on an isolated table top from any compliance curve (g is the acceleration due to gravity).

One important practical example of using the maximum relative motion is to approximate worst-case beam deflection. The simple setup illustrated in Figure 133 (a mirror mount attached to a table top) indicates how a laser beam reflected by the mirror will be affected by two types of vibrational modes: axial deflection from translational table top motion and angular deflection caused by table top bending motion. The relative contributions of each mode to the beam deflection are estimated in [226]. Essentially, displacement in translation does not cause significant deflection in most experiments or applications. However, table top bending can have serious consequences for optical performance for two reasons: when a mirror is rotated, the angle of the reflected beam is twice the tilt angle and the "doubled" error also increases linearly with reflected spot distance from the mirror. This effect, multiplied by the number of mirrors in an optical system and the total path length, can have a significant effect on experimental results if the table does not provide an adequate level of vibration-control performance.

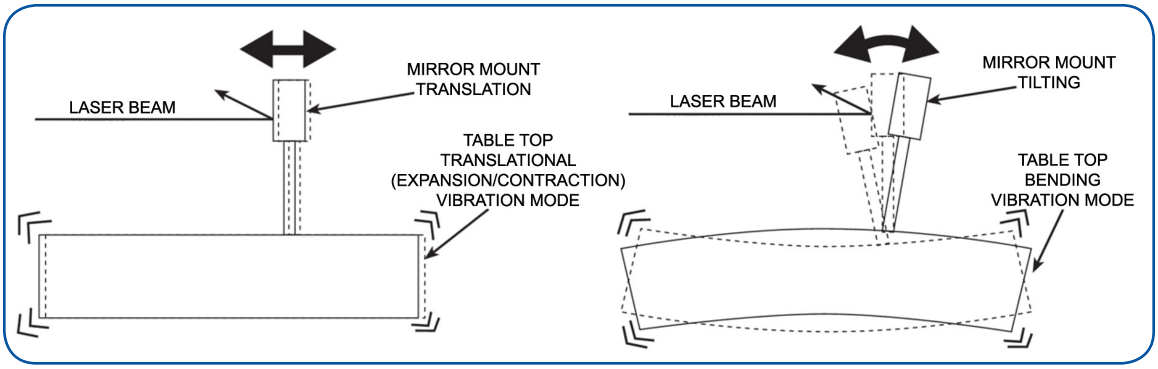


Figure 133. Examples where translational motion of the table top (left) and table top bending (right) can contribute to the maximum relative motion leading to unwanted beam deflection. Table top bending is the most significant cause of beam deflection.

Additional online content discussing the fundamentals of vibration can be found in [227-230].

2. Vibration Control Technologies

In optical tables, "noise" refers to three types of vibration: seismic, acoustic, and tabletop forces. Seismic vibrations come from the ground underneath the table and can be caused by foot and vehicle traffic, wind, and building vibrations. Acoustic noise comes from sound waves that travel through the air and walls. Tabletop forces are caused by vibrations on the working surface, such as a moving positioning stage or vacuum-system tubing. The term optical table system typically refers to a vibration-isolation system, that is, the tabletop plus the support system (see Figure 134). Such tables generally come in several grades, determined by the level of vibration damping they provide. Basic optical tables for general-purpose use are suitable for quiet environments and provide little or no vibration damping mechanisms, but because of their honeycomb composite construction, they provide a rigid platform for general optical experiments that are relatively insensitive to vibration such as spectroscopy, velocimetry, and non-phase-dependent applications. An intermediate-grade table typically incorporates broadband damping or a moderate level of tuned mass dampers to reduce the compliance of the table to disturbances. It is typically preferred for experiments like bioimaging, Raman spectroscopy, micropositioning, and machining. For more demanding applications like interferometery, nanopositioning, and imaging, or for labs with excessive vibration and acoustic noise, a higher-performance table should be used that provides the maximum damping and consequently the quietest table surface for the application. Tables in this category incorporate more tuned mass dampers or active dampers. The various components that make up such vibration control systems are the subjects of this section.

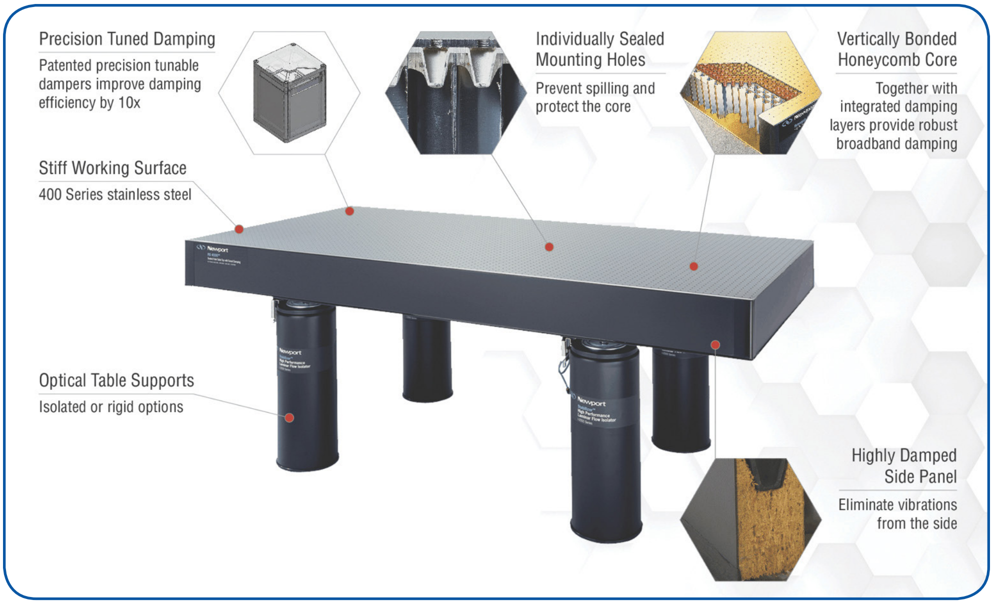


Figure 134. A vibration-isolation system which includes an optical table and the support system.

Optical Table Design

Since one cannot completely eliminate the sources of vibrational disturbances, the goal is to reduce relative motion between different elements by connecting them with a structure that is as rigid as possible. In a perfectly rigid body, which exists only in theory, the distance between any two points remains constant in time. In other words, the size and the shape of the body do not change while it is undergoing force inputs from vibrations, static forces, or temperature changes. If all of the elements are mounted together to form an ideal rigid body, the different elements will not move relative to each other and system performance will not be impaired. Since it is impossible to create a perfectly rigid structure, an effective vibration isolation system must take into consideration dynamic forces (vibration), static forces, and temperature effects. Dynamic forces cause structural deformations that vary with the frequency of the driving force. Structural resonance can amplify the relative motion between optical components. There are two main methods to mitigate these effects: connect all of the critical elements together in a dynamically rigid structure that is designed to eliminate (damp) structural resonances and isolate the system from vibration with mechanical filters or active cancellation technology. Static forces cause deformations that are constant in time. However, the addition or movement of equipment in the system will change the static forces and cause misalignment of system elements. To combat this, it is necessary to build a statically rigid structure that deforms as little as possible under the application of external forces. Non-uniform temperature changes usually cause a slow bending of the structure, with time constants of one hour or more. The key techniques for reducing thermal effects are controlling the environment to reduce temperature variation and designing structures to be as insensitive to temperature as possible.

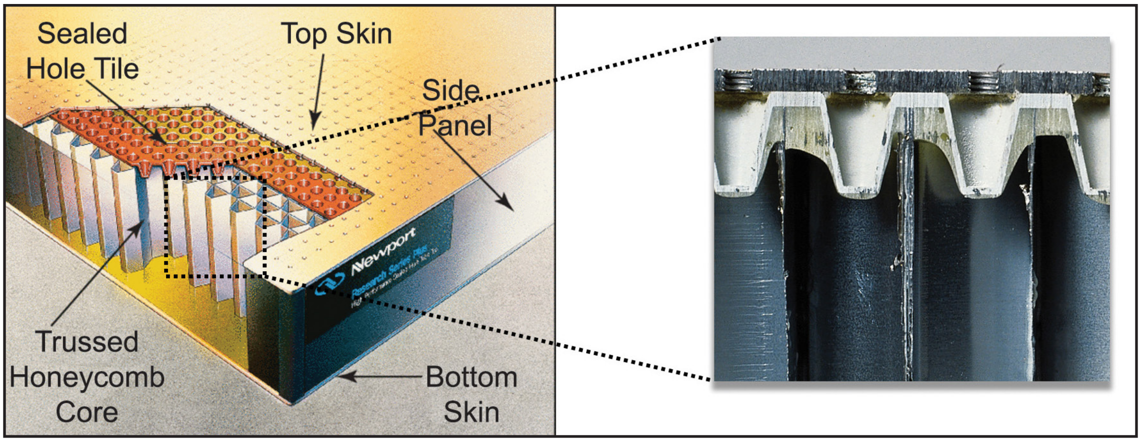


Figure 135. Depiction of an optical table revealing the internal honeycomb core construction (left). Cross-sectional view of the top portion of an optical table (right).

Optical tables provide a rigid platform for high-precision optical experiments and systems. They are designed to eliminate errors caused by relative motion between optical components in the beam path. Rigidity is the primary consideration in optical table design. Table rigidity can be quantified in terms of static or dynamic rigidity. Static rigidity describes the ability of an optical table to resist deflection when the static or quasi-static load distribution is changed. It defines the table performance when stages move across the table or equipment is relocated, added, or removed. Dynamic rigidity describes the ability of an optical table to resist deflection in response to mechanical excitation. It describes the table performance in response to floor vibration, acoustic noise, and mechanical sources on top of the table. A honeycomb design is commonly used to produce very low weight, highly rigid structures. Reduced weight dramatically improves the dynamic rigidity of the structure by moving structural resonance modes to higher, less detrimental frequencies. The structural resonance modes are the frequencies at which the platform deflects, thus causing relative motion across the optics mounting surface. For a given input vibration force, the deflection will be reduced as the mode frequency increases. This is the primary reason why steel honeycomb (see Figure 135) has replaced granite in most high-end optical applications. Since granite is relatively heavy, the resonance modes occur at lower frequencies and therefore, produce higher amplitudes of surface deflection. There are additional improvements that can increase the efficacy of the honeycomb design. A trussed core design uses an additional steel member to bridge across the center of the honeycomb cell. This extra mechanical component significantly stiffens the cell with very little increase in weight, resulting in improved static and dynamic rigidity. By vertically bonding along the height of the honeycomb core, the table's rigidity-to-weight ratio can be maximized. The trussed core design possesses a triple core interface at each honeycomb cell. By bonding the three sheets over the full table height, greater stiffness is produced.

Passive and Active Damping

Disturbances caused by relative motion between optical components generally occur at the structurally dominant bending or torsional modes. In addition to pushing these natural modes to higher, less detrimental frequencies, a major advantage of honeycomb over granite is the high level of damping present in the honeycomb structure. Damping attenuates the amplitude of the natural modes and reduces the relative motion across the table surface. Generally, the two types of passive damping used for optical tables include narrowband tuned damping and broadband damping. Tuned damping techniques use individual mode selected vibration absorbers to both eliminate a particular "narrow" mode and mode harmonics across the broader band. On the other hand, broadband damping techniques indiscriminately absorb moderate amounts of vibration over the broadband. Narrowband tuned damping is the most effective means for eliminating structural resonances, which can be seen by comparing their respective compliance curves in Figure 136. These narrowband dampers selectively cancel vibration modes to minimize the dynamic deflection coefficient and make the table behave more like an ideal rigid structure.

While some narrowband damper designs use oil, others use a mass-spring mechanism, which can improve performance and allow for tuning to the exact frequency needed to damp the resonance. While broadband damping is less effective than narrowband techniques, it can still improve table performance. One method for broadband damping involves the use of constrained layer structures, which usually consist of two or more metallic sheets separated by a compliant material. For instance, the three metallic sheets in the trussed honeycomb core are each separated by an adhesive. Although the adhesive is rigid, its damping factor is much higher than that of steel and introduces substantial damping into the core. Other broadband damping techniques include introducing a polymeric material that serves to seal the honeycomb core and dampen the working surface. The polymer experiences the same bending and shear stresses as the stainless steel top but, since the damping factor of this material is much higher than that of steel, considerable damping is introduced into the work surface. Damped table sides can also improve vibration sensitivity. In this case, the sides of the table are made of a highly damped, epoxy sealed wood composite. Compared to metal sides, the composite wood sides offer significant damping to the structure and eliminate another source of resonance.

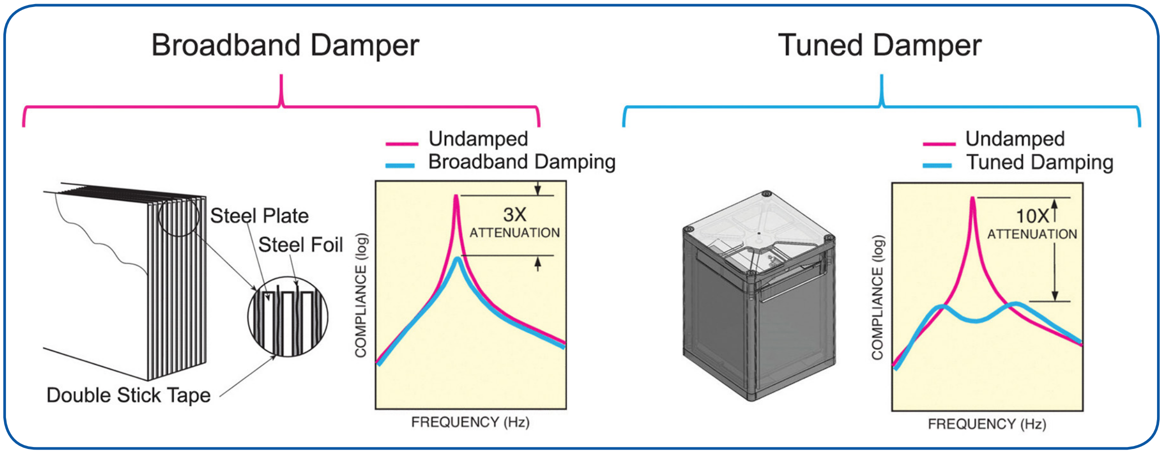


Figure 136. Broadband damper technology and corresponding compliance curve illustrating moderate damping over a wide range of frequencies (left). Tuned damper technology and associated compliance curve showing concentrated damping at the frequencies of dominant resonance modes (right).

Tuned narrowband dampers are effective at suppressing flexural resonance vibration of the table. However, these dampers are typically tuned to the table's particular resonant frequencies and cannot be adjusted for significant changes to table loading (which also change the resonances). Methods of active vibration control can provide high efficiency without the restrictions of passive methods. Active vibration control involves monitoring vibrations of a structure and utilizing the vibration signal to generate a force with the proper phase and amplitude to attenuate the vibration. An additional advantage of an active approach is the ability to supply a vibration signal that can be used independently for monitoring the vibration environment. Two sensor-actuator assemblies are typically integrated into the structure of the optical table at two corners. The design ensures rigid coupling of the sensor and the actuator, and includes a stiff tubular structure coupling the damper to both top and bottom facesheets. The damping performance is on par with that of high-quality passive tuned absorbers. However, if the table is loaded by a weight comparable to the weight of the table, the passive vibration absorbers can become "mistuned," whereas an active damper will work equally well after re-tuning.

Isolation

Pneumatic isolators filter vibration before the mechanical noise can reach the optical bench work surface. Improved vibration isolation reduces errors caused by relative motion between optical components in the beam path. Pneumatic isolators combine with the optical table and payload to form a mass/spring/damper system. Pneumatic systems are used instead of mechanical springs since they offer self-leveling and minimize the effect of varying mass on isolation. Conventional isolators use a compliance chamber to act as an air spring and a damping chamber to increase system stability (see Figure 137). The compliance chamber is sealed with a flexible diaphragm to form a piston and support the optical table on compressed air. If the piston is pushed further into the compliance chamber, the pressure of the gas increases and provides a restoring force—somewhat like a soft spring. The isolation performance is primarily related to the volume of the compliance chamber. After the compliance chamber, air is pumped to the damping chamber through a flow restricter - usually a thin tube or orifice. The restricter dissipates energy in the air and essentially damps the system. The design of both chambers and the restricter must be optimized to minimize the natural frequency/damping trade-off. The performance of the isolator is defined primarily by its natural frequency and damping characteristics as summarized in the transmissibility plot in Figure 137. The pneumatic isolator is essentially a simple harmonic oscillator that uses the “fast roll-off” at higher frequencies to act as a low pass mechanical filter. Below the natural frequency of the harmonic oscillator, the isolator is essentially rigid and vibration is passed directly to the platform. At the natural frequency, vibration is actually amplified. Therefore, a primary goal is to lower the natural frequency since this improves low frequency isolation and overall isolation bandwidth. Another primary goal is to damp the harmonic oscillator amplitude at resonance. This lowers the magnification of vibration at low frequencies and improves system stability. Unfortunately, there is a compromise between the natural frequency and damping. As damping is increased, the isolator natural frequency moves slightly higher, and higher frequency isolation is decreased.

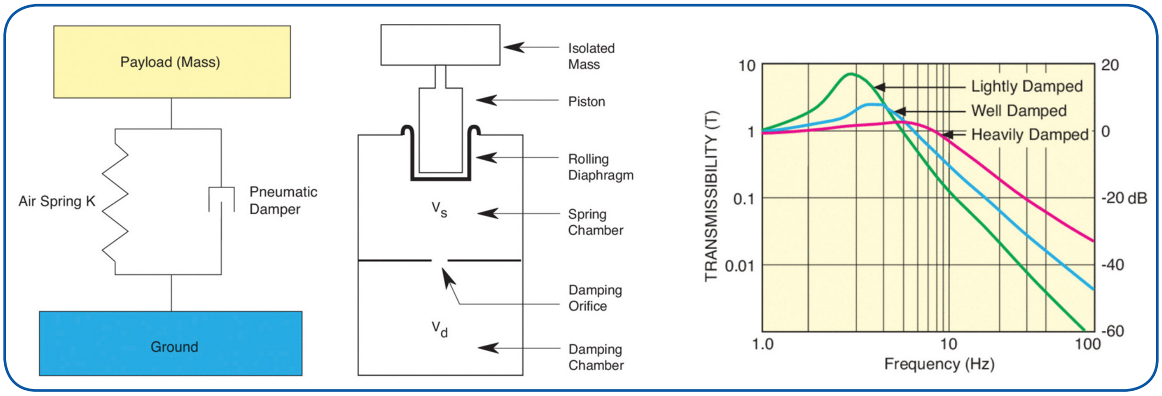


Figure 137. Depiction of a pneumatic isolator with damping (left) and a typical transmissibility plot showing damping effects (right).

Additional online content discussing vibration control systems can be found in [231-237].

This page intentionally left blank.

Chapter 2

Photonics in Semiconductor Manufacturing

This page intentionally left blank.

I. Introduction

Photonics technology is widely used in semiconductor device manufacturing processes. While processes that use advanced photonics technology, such as Deep UV (DUV) and Extreme UV (EUV) lithography, are very much current topics of conversation within the industry, optical methods that produce, manage, and measure light are ubiquitous within device fabrication plants (fabs). In addition to lithography processing, semiconductor applications of photonics include critical dimension (CD) measurements, process overlay measurements, mask inspection, process sensing, and many other technologies. Without the use of photonics technology, device manufacturers would be unable to create the precisely overlaid and layered structures that constitute an integrated circuit (IC). This chapter reviews many applications of photonics technology in semiconductor device manufacturing.

A. Semiconductor Manufacturing Overview

A detailed discussion of semiconductor device fabrication processes is available in [5]. This section will provide a brief overview of semiconductor manufacturing and describe the various optical techniques used to define, inspect, and measure feature sizes and layer properties throughout the manufacturing process. For information on basic semiconductor physics, see Chapter 1, Section I.B.1.

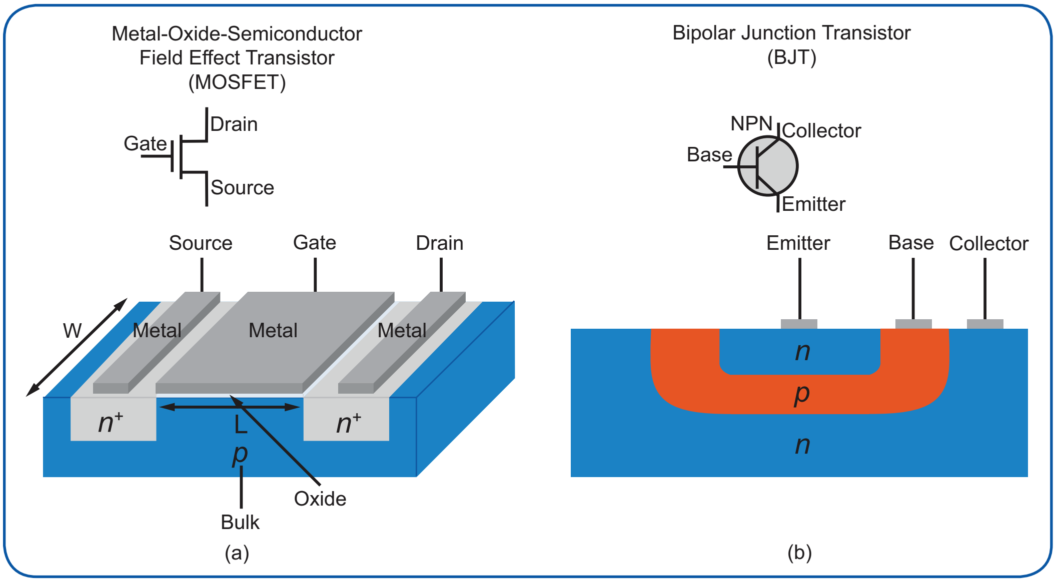


Figure 138. (a) Metal-oxide-semiconductor field effect transistor (MOSFET) symbol and structure; (b) bipolar junction transistor (BJT) symbol and structure [5].

ICs are composed of layers of materials with different electrical properties, i.e., semiconductors, insulators, and metals, [238]. These materials are patterned on a substrate surface in such a way that they form microscopic electronic devices (transistors, resistors, capacitors, etc.) that, when connected by patterned metal lines, produce an IC. Transistors constitute the heart of an IC and their function and design generally fall into one of two categories (see Figure 138): metal–oxide–semiconductor field-effect transistors (MOSFETs) or bipolar junction transistors (BJT). Such transistors are used in many circuit applications, including logic, microprocessors, dynamic random access memory (DRAM), and flash memory (also denoted as NAND). Figure 139 shows a specific example of a CMOS digital IC that illustrates the backend metal connections, and the transistors formed directly on the surface of a silicon substrate. Such CMOS circuits contain complementary and symmetrical pairs of *p*-type and *n*-type MOSFETs for logic functions.

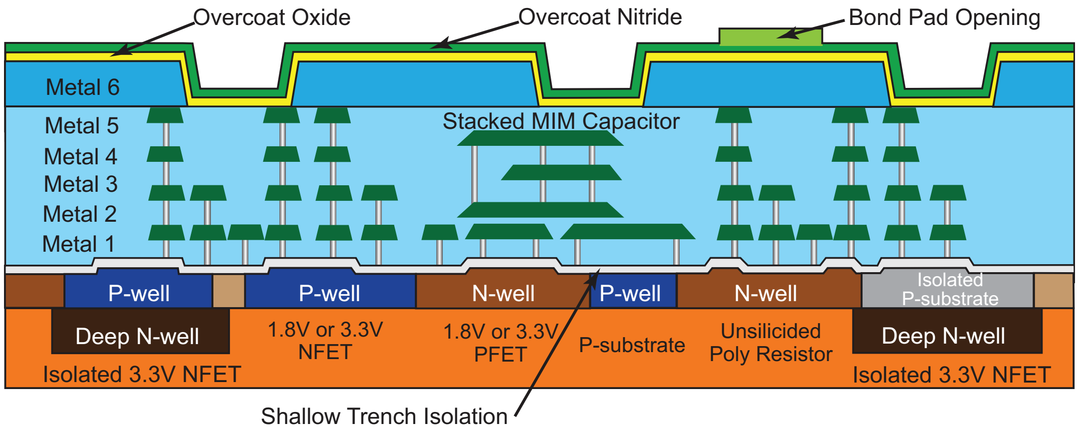


Figure 139. Cross-sectional schematic of CMOS device [238].

New IC architectures are continually being developed that achieve higher integration densities, functionalities, and performance. Figure 140 shows some examples of the different approaches used to improve cost and function. 3D NAND architectures stack memory cells like a skyscraper to reduce the overall footprint; FinFet logic transisters use a 3D approach to transistor design that reduces electrical leakage; and DRAM is scaled to ever smaller sizes. The rate of increase in functionality and performance has followed the famous Moore's law curve that predicts a halving of cost every two years [239]. However, as device design rules move into the few nm regime, the basic laws of physics have acted to slow the rate of improvement in cost and function, preventing new IC architectures from maintaining Moore's law. However, new heterogenous integration schemes, sometimes referred to as "more than Moore," are being used to continue the scaling.

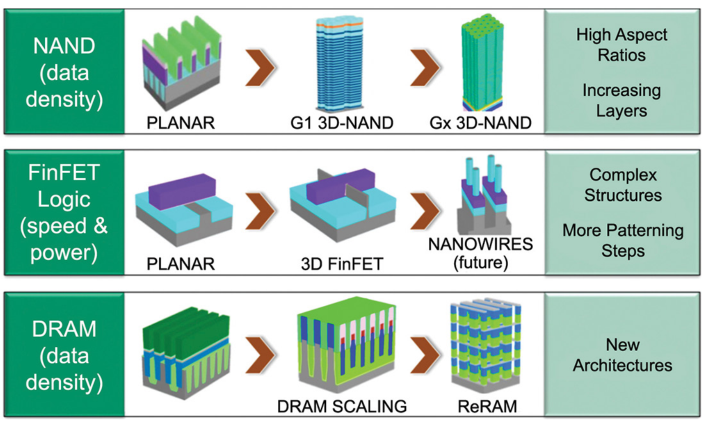


Figure 140. 3D device road map (figure is reprinted with permission of Nanometrics [240]).

Figure 141 shows the relative number of processing steps required for past technology nodes, with 500 needed for the 45 nm node. More recent nodes, such as at 7 nm, require more than 1000 processing steps. The number of processing steps required in a device fabrication scheme can be seen to scale with increasing complexity in the device. More importantly, each additional step adds time and cost to the production of an IC and increases the likelihood for critical defect formation. The need for multiple patterning steps, i.e., litho-etch-litho-etch, to reduce the minimum feature size may be reduced to some extent through the use of advanced EUV lithography (discussed below).

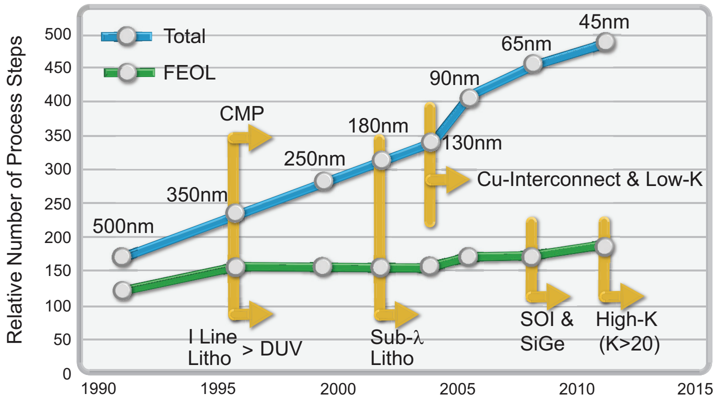


Figure 141. Evolution of process complexity and inflections [241].

Semiconductor device manufacturing requires hundreds of sequential processing steps to produce a functional, cost-effective device and each step must be performed with near perfection to avoid the creation of a device-killing defect [5]. The MKS Handbook of Semiconductor Devices and Process Technology [5] provides a detailed discussion of device fabrication and therefore this section will provide only a brief outline of the process, identifying those steps where precision photonics enable inspection and metrology. The process steps are illustrated in Figure 142 and the sequence is described in more detail in Table 9.

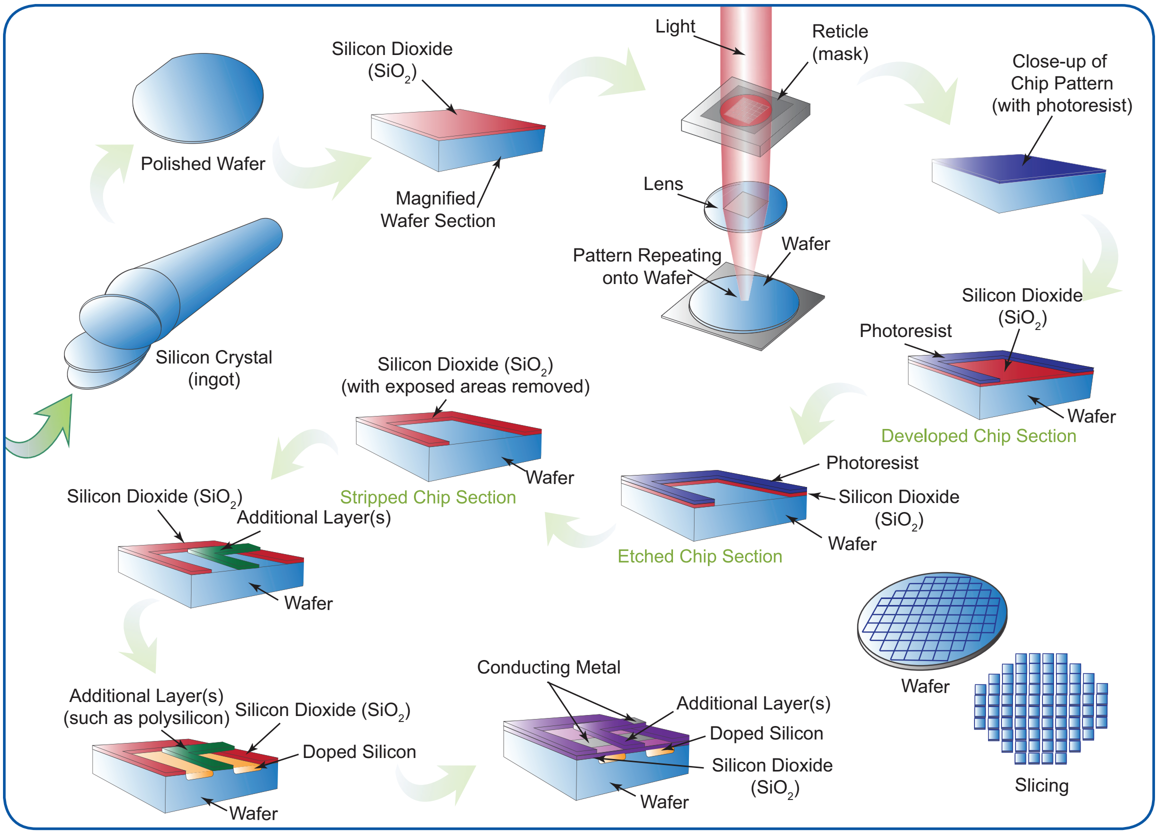


Figure 142. Overview of the semiconductor manufacturing process [242].

Step	Title	Description	Inspection (•) Metrology (Δ)
1	Silicon Boule Growth	Electronics grade single crystal silicon boules are grown via the Czochralski process.	
2	Wafer Fabrication	The boule is ground to the desired diameter, flats are added, and the boule is cut into thin slices (wafers) with a diamond encrusted blade or steel wire. The wafer is ground and lapped to µm-scale surface flatness, while damaged surface layers are etched away. The wafer undergoes chemical mechanical polishing to produce a polished wafer. Epitaxial layers may be deposited.	• Δ
3	Active Area Isolation	Thermal oxidation, nitride deposition, and lithography are used to define the active device areas on the wafer surface. Shallow isolation trenches are etched into the substrate around the active area and filled with deposited silicon dioxide.	
4	N- and P-Well Formation	The wafer is patterned by coating with photoresist, which is developed selectively using specific wavelengths of intense light shone through a reticle containing the pattern. The light is focused to achieve the final dimensions. N- and P-well regions are formed using ion implantation.	•
5	Transistor Formation	Thermal oxidation, polysilicon deposition, patterning, and etch are used to produce transistors on the wafer surface.	• Δ
6	Contact Formation	Silicon dioxide deposition, patterning, and etch define contact holes that are then filled with deposited metal.	• Δ
7	First Level Metal Formation	Silicon dioxide or other insulating film is deposited, patterned using lithography and then etched. Interconnects are then patterned in dielectric using lithography and etched. Metal film is then deposited in the openings in the oxide to form the via and metal interconnects. This is called a dual damascene process. Excess material is then planarized using a chemical mechanical polishing step.	• Δ
8	Step 7 is repeated to produce back-end connections	Step 7 is are repeated to create a multilevel metal structure such as that shown in Figure 139.	• Δ
9	Dicing	The chips are tested and separated for packaging.	• Δ

Table 9. Process sequence in a typical IC device fabrication.

B. Impact of Photonics on Semiconductor Manufacturing

Metrology and inspection equipment identify issues associated with defects, contamination, misalignments, and undesirable topographies at every critical step in semiconductor manufacturing, thereby improving yield, down-time, and cost. Figure 143 shows a plot of the known points of error generation in semiconductor manufacturing. Precision optical components are incorporated into tools for imaging, scatterometry, shape metrology, reticle inspection, alignment, and lithography. Advanced photonics technology, e.g. motion, positioning, light sources, and optics, enables higher resolution lithography to increase circuit density, reduces the number of processing steps, increases yields, and reduces the time to market for semiconductor devices.

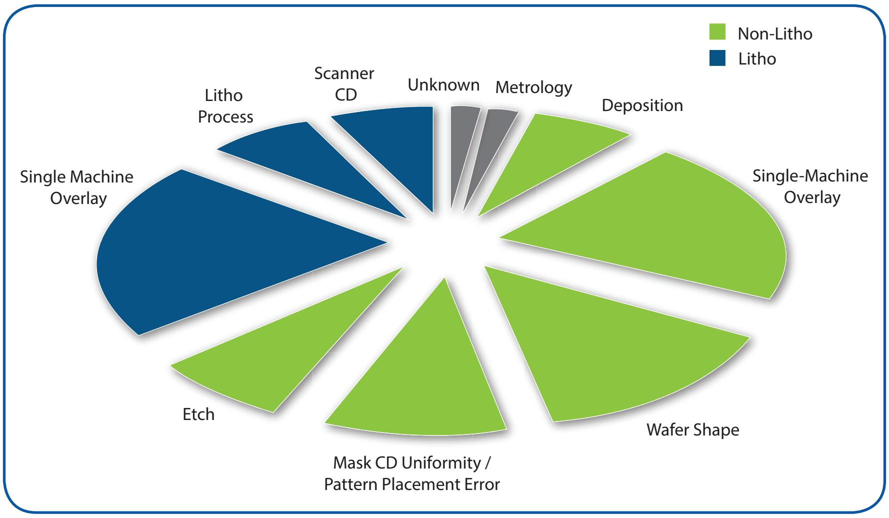


Figure 143. Breakdown of sources of CD error in semiconductor manufacturing [243].

Section II below will cover the inspection and metrology solutions described in the introduction. Photolithography is detailed in Section III thereafter. Semiconductor wafer scribing and dicing using lasers for die singulation is covered in Chapter 3 – Photonics in Laser Machining.

II. Process Diagnostics: Defect Inspection and Metrology

The progression to ever smaller feature sizes on both substrates and masks/reticles has resulted in increasingly lower tolerance to defects in both starting materials and finished devices. Near-zero tolerances are being approached for known defect types, e.g., particles, crystal defects. In addition, manufacturers continue to discover device sensitivities to entirely new types of nanoscale defects. These issues require the available metrologies for defect detection to sense and quantify defects at very near the noise level of their operating principles and therefore, new approaches to defect detection are continually being developed. Wafer and reticle inspection techniques are described below, along with some of the characteristics of current inspection tools.

A. Bare Wafer Inspection

1. Scanning Defect Inspection

This section highlights the kinds of defects present at the bare wafer stage and the techniques used to identify them. Prior to starting production, bare wafers are qualified at the wafer manufacturer and again upon receipt by the semiconductor fab. These qualifications locate, map, and differentiate pre-existing defects from those arising in the IC manufacturing process. Only the most defect-free wafers are used in production, and their pre-production defect maps allow manufacturers to track regions that are likely to result in poorly functioning chips. Bare or non-patterned wafers are also measured before and after being subjected to a passive or active process environment to determine a baseline for particle contributions from a given process tool.

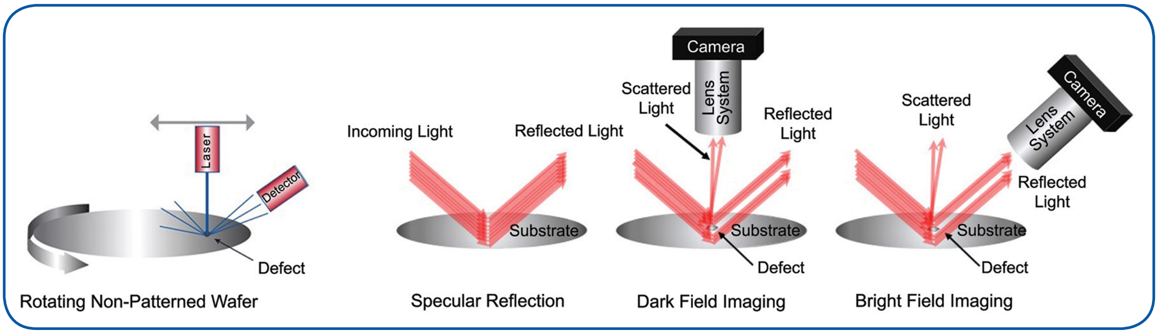


Figure 144. Defect detection on rotating non-patterned wafers (left) and the use of specular reflection in dark-field and bright-field image illumination (right) [244].

Device manufacturers use optical detection systems to inspect wafers and masks for particles and other types of defects and to determine the position of these defects in an X-Y grid on the wafer. The basic principle used for defect detection on non-patterned wafers is relatively simple. A laser beam is radially scanned over the surface of a rotating wafer to ensure that the beam is projected onto all parts of the wafer surface. The laser light is reflected from the surface as it would be from a mirror, as is shown in Figure 144. This type of reflection is referred to as specular reflection. When the laser beam encounters a particle or other defect on the surface of the wafer, the defect scatters a portion of the laser light. Depending on the illumination arrangement, the scattered light can be detected directly (dark-field illumination) or as a loss in intensity in the reflected light beam (bright-field illumination). The rotational position of the wafer and radial position of the beam define the position of the defect on the wafer surface. In wafer inspection tools, the light intensity is electronically recorded using a PMT or CCD and a map of the scattered or reflected intensity over the wafer surface is generated, as shown in Figure 145. This map provides information on defect size and location and on the condition of the wafer surface due to issues such as particle contamination. This method requires highly accurate and repeatable rotary and linear motion control of the wafer stage and optical components.

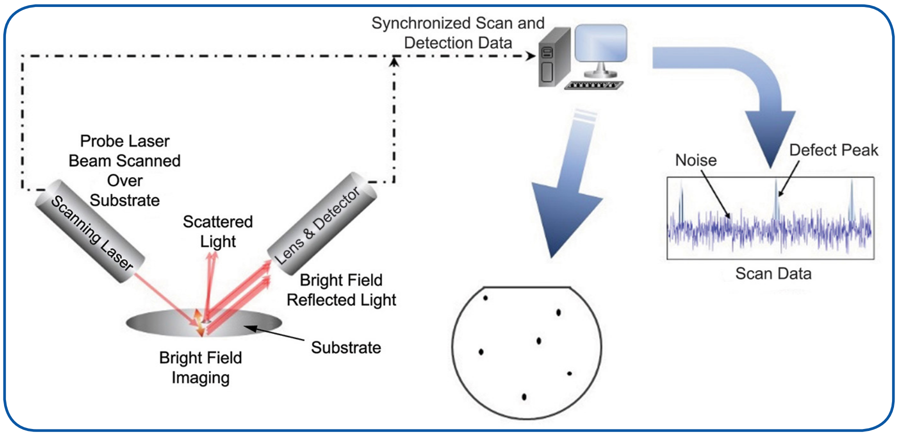


Figure 145. Light collection, processing and wafer mapping in an inspection tool [245].

In general, dark-field inspection is preferred for non-patterned wafer inspection since high rastering speeds are possible and this enables high wafer throughput. Patterned wafer inspection is a much slower process. It may use bright-field and/or dark-field imaging, depending on the application. Note that the complexity of scattering from patterned surfaces reduces the overall photon flux to the detector, resulting in longer integration periods for wafer inspection.

Sub-100 nm inspection tools are currently used in manufacturing environments to provide quality assurance on incoming wafers and for process tool monitoring and qualification for high volume manufacturing. These tools employ the same basic operating principles as tools designed for larger scale defect detection but use DUV illumination-enhanced optical systems. Sophisticated image analysis algorithms are claimed by some manufacturers to achieve sub-20 nm sensitivity. As might be expected, a high degree of precision and accuracy is required in the motion control of the wafer stage and the optical components in systems used in these applications.

As inspection tools are required to sense and quantify ever smaller particles, the impact of factors such as surface microroughness (haze) begins to influence the detectability of small particles owing to reductions in the SNR of the scattered light signal. Sub-100 nm inspection for non-patterned wafers is complicated by issues of scale with SNR being a critical parameter in determining an inspection system's detection limits for particles and other defects on wafer surfaces. Surface chemical contamination from sources such as ambient humidity can also contribute to reductions in SNR. To help counter such effects, inspection tools for sub-100 nm defect detection employ highly sophisticated optical spatial filtering, analysis of polarization of the scattered signal, and specialized signal processing algorithms to detect defects in the presence of surface haze.

2. Topography Inspection

Bare wafer topography is measured for many reasons. For example, a wafer can bend or the chuck (either electrostatic or pneumatic) that supports a wafer can produce indentations in the wafer at points of contact. Such deformations can affect pattern imaging at nm scales. Extremely precise interferometric tools have been developed to measure such variations in wafer shape prior to processing.

The basic design used to measure topography on bare wafers is similar to the Fizeau interferometer shown in Figure 146 [246]. A description of interferometry is given in Chapter 1, Section I.A.3. This interferometric technique compares the wafer to a reference wedge (or reference flat) of very high quality and flatness. The wedge angle ensures that reflections from the first surface of the flat do not contribute to the interferometric signal. Light reflected from the second surface serves as the reference while a portion of the light passes through the flat to interrogate the wafer (test flat). Light reflecting from the wafer and test flat is directed to the imaging system by the beamsplitter. The interference pattern is analyzed, and measurements are stitched together using software to form complete wafer maps with nm-scale resolution. In practice, interferometric tools that measure bare wafer topography are extremely sophisticated and utilize motion solutions, large optics and illumination sources that help expand the boundaries of design for manufacturability.

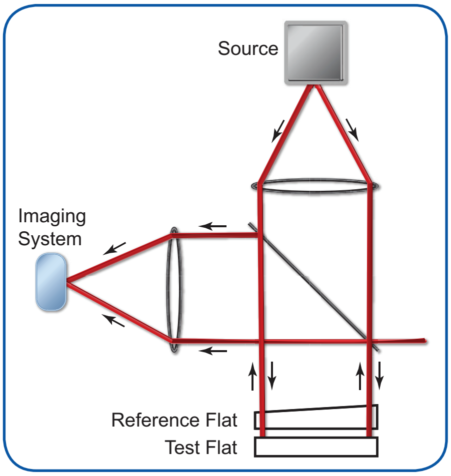


Figure 146. Fizeau Interferometer [246].

B. Patterned Wafer Inspection

1. Differential Image Detection

As previously noted, the optical inspection of patterned wafers can employ bright-field illumination, dark-field illumination, or a combination of both for defect detection. Patterned wafer inspection systems compare the image of a test die on the wafer with that of an adjacent die (or of a "golden" die known to be defect free). Image processing software subtracts one image from the other. Any random defect in one of the dies will not zero out in the subtraction process, showing up clearly in the subtracted image (Figure 147). The positions of the defects allow a defect map to be generated over the wafer, similar to the maps generated for non-patterned wafers. As with the inspection of non-patterned wafers, patterned wafer inspection requires precise and repeatable motion control of both the wafer stage and the optical components of the inspection system since they are moved simultaneously.

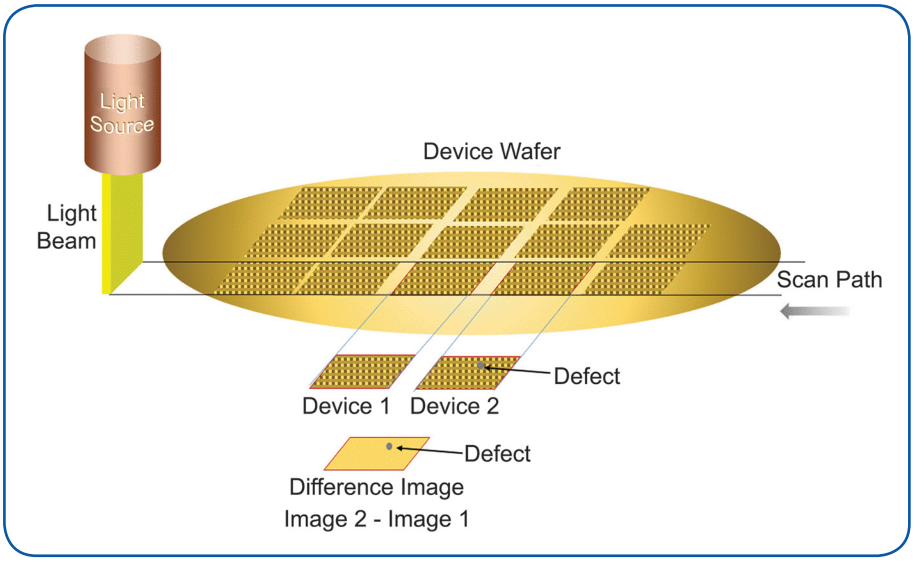


Figure 147. Patterned wafer inspection procedure [247].

2. DUV Inspection for Sub-100 nm Features

Defect inspection for sub-100 nm patterned wafers presents greater challenges than for non-patterned wafer inspection. DUV-based optical inspection for patterned wafer applications uses the same image comparison principle as older VIS and UV light inspection systems. However, DUV-based methods require a much greater degree of sophistication in the optics, motion control, and image analysis algorithm.

DUV inspection tools have become the industry standard for patterned wafer inspection down to 65 nm feature sizes; inspection rates of up to several wafers per hour make these systems suitable for production applications. DUV inspection tools have demonstrated high sensitivities for the detection of defects such as shallow trench isolation voids, contact etch defects, and photoresist micro bridging at sub-100 nm geometries [243]. Using broadband DUV/UV/VIS illumination, modern bright field patterned wafer inspection systems currently achieve the required sensitivity for defect inspection of all layers on DRAM and flash devices down to 55 nm feature sizes.

While their well-understood characteristics coupled with relatively low cost and high throughput make the continued use of DUV optical inspection systems attractive, some manufacturers have reported that DUV inspection systems do not have the needed precision and sensitivity for geometries below 65 nm. One study claims that the limiting defect sensitivity for DUV dark-field optical pattern inspection systems is about 75 nm in memory technology (e.g. SRAM) and much larger on logic areas [248]. DUV bright-field systems have somewhat better limiting sensitivities, ~50 nm in SRAM and, as with dark-field, larger in logic. Furthermore, the use of DUV lasers to illuminate very small and consequently fragile structures on patterned wafers has produced some unusual problems such as laser ablation of surface material. The solution to these problems may lie in the use of either broadband plasma illumination for optical inspection systems (existing DUV systems employ 266 nm wavelength and are moving to 193 nm illumination) or in the use of production-capable electron beam inspection tools. Recently-introduced inspection tools based on plasma-generated broadband illumination are available for use in production environments. Sub-10 nm resolution is claimed for these systems, because shorter wavelengths provide more accurate inspection at this smaller scale.

3. E-beam Inspection

Electron beam (EB) imaging is also used for defect inspection, especially at smaller geometries where optical imaging is less effective. EB inspection can provide material contrast with a dynamic range of resolution much larger than optical inspection systems. However, EB applications are limited by the slow speed of measurement, making it useful primarily in R&D environments and in process development for the qualification of a new technology. New EB tools are available for defect inspection applications at 10 nm and lower nodes, and multi-EB tools are being developed with up to 100 columns, or channels of measurement.

C. Reticle Inspection

Reticles are transmissive or reflective projection masks patterned with fine features, typically 4x-5x larger than the desired pattern scale on the wafer. They are used with optical illumination systems that image and de-magnify the patterned light to selectively develop photoresists as part of the wafer patterning process. Arguably, reticle inspection is far more important than either non-patterned or patterned wafer inspection. This is due to the fact that, while a single defect on a bare or patterned wafer has the potential to “kill” one device, a single, undetected defect on a reticle can destroy thousands of devices since the defect is replicated on every wafer processed using that reticle. With EUV this problem is further complicated by the finer resolution of the pattern, the presence of a thin protective membrane, and the reflective design of the reticle.

Reticle inspection systems work on the same principles and have similar physical requirements as wafer inspection tools, with the exception that reticles are normally inspected using transmitted rather than reflected light. Transmitted light is used to locate UV-opaque stains and other transmission defects. Reticle inspection tools employ high resolution imaging optics and either VIS or UV illumination, depending on defect tolerances and/or feature sizes, to find defects on a reticle blank or a patterned reticle. Inspections are routinely performed during the reticle manufacturing process and throughout the reticle usage. Reticle inspection tools employ sophisticated image analysis software and motion control systems similar to those used in wafer inspection tools. The use of conventional optics in reticle inspection systems has been extended down to 90 nm feature sizes by using UV illumination. Reticle inspection at smaller feature sizes is possible using EB since lower throughputs can be tolerated in comparison to patterned wafer inspection.

As with wafer inspection, the reticle inspection tools used in sub-100 nm applications (both blank and patterned reticle inspection) employ DUV illumination, typically using a single wavelength at 266 nm or 193 nm. Figure 148 shows a block diagram of a reticle inspection platform. Note that, in addition to the objective optical system, motorized stage, and light source, the platform employs various controllers and data analysis modules. Reticle inspection systems can be configured to employ either transmitted light through the reticle or reflected light from the reticle surface in the inspection process. As with other inspection systems, this reticle inspection tool requires highly accurate and precise motion control for the optical components and the air bearing reticle stage.

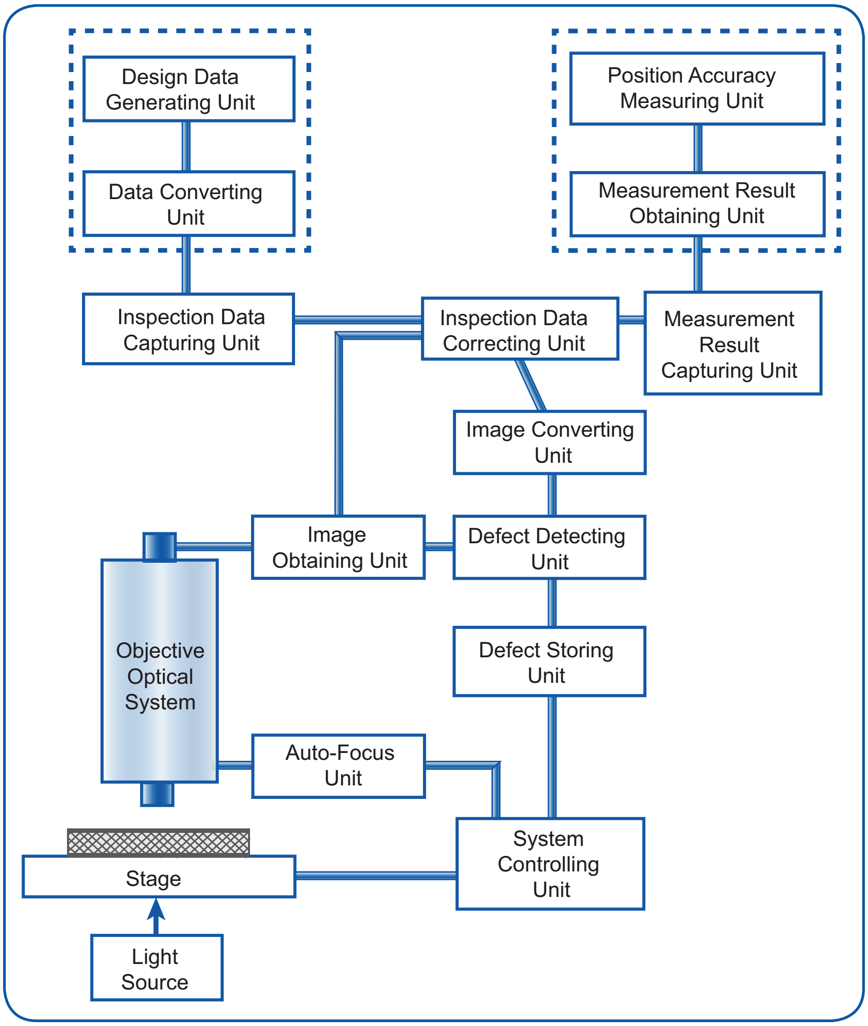


Figure 148. Block diagram of the components in a reticle inspection system. Figure is reprinted with permission from Fujitsu Ltd. [248].

D. Spacing and Overlay Alignment Inspection

Each layer of an IC must align correctly with the previous layers. This requires specialized tools that optically measure overlay alignment and precisely position the wafer. Specialized light sources and detectors are used to detect features on the wafers and allow for its positioning on the stage to nm accuracies. This becomes more important when feature sizes shrink, as the overlay budget is reduced [249]. Multiple patterning steps are required to produce smaller sizes and each additional step results in additional cumulative overlay error.

E. Lasers and Photonics for Semiconductor Manufacturing

MKS is a major global supplier of standard optical components, opto-mechanical assemblies, laser light sources, optical measuring instruments, vibration control equipment, and motion control systems. Details regarding the fundamental operation of these components can be found in Chapter 1 – Introduction to Photonics. Popular MKS products include polarizers, beamsplitters, gratings, optical filters, optical mounts, motorized optical stages, optical tables, vibration isolators, laser systems, laser beam profilers, and other optical instruments. These products are frequently built into commercial OEM products owing to their high quality, competitive pricing, and ready availability.

1. Custom Optical and Opto-Mechanical Solutions

Many semiconductor optical and opto-mechanical systems require non-standard products. These can range from custom-designed lenses or lens assemblies, to simple modifications of standard products. For example, semiconductor manufacturers use both free-space optics and specialized fiber optic patch cables, compatible collimators, and detectors for beam routing. Motorized fiber alignment solutions also enable dynamic fiber coupling of free-space light. Often such custom solutions require customer-specified testing and documentation. MKS supports this technology with technical depth, advanced manufacturing and test equipment, and quick turnaround.

MKS created its Integrated Solutions Business (ISB) to service larger projects having optical, opto-mechanical, and electronic control elements. ISB projects range in scope from electro-optical sub-assemblies used in OEM equipment to entire instruments or systems in a custom-designed housing. A typical ISB project starts with exploratory meetings and in-depth discussions between the customer and ISB's business development staff. These produce a specification document that is formally reviewed by ISB and iteratively modified by both parties as needed to define a viable program. Projects include electromechanical assemblies for semiconductor manufacturing equipment, photovoltaic device manufacturing equipment, biomedical instruments, and military/space optical assemblies, and include reviews of technical performance, product improvements, schedules, and shipments. Highly regulated optical manufacturing procedures and operational control are required to achieve the technical specifications demanded for these applications.

2. Motion Solutions

The throughput of wafer inspection and metrology is dependent on the wafer handling capabilities of motion stages. Ideally, these stages need to accelerate, scan rapidly, stop, and settle quickly, all while isolating the optical inspection equipment from vibrations. Accuracy and repeatability in the motion and positioning are critical to ensuring tool-to-tool consistency and accurate test measurements. Many of the motion stages used in semiconductor manufacturing meet these requirements with air bearings that provide a low-friction interface via a thin film of pressurized gas along with specialized control electronics. The materials used in motion stages are carefully selected to minimize contamination and maximize flatness. MKS motion solutions designed to meet these product requirements are described in detail in Section II.B.5.

3. Lasers and Other Light Sources

Light sources are used in many of the steps in semiconductor inspection and manufacturing. Tunable lasers are used in interferometry owing to their narrow linewidths and frequency scanning capabilities. Pulsed lasers are used for wafer scribing and dicing. Laser applications in micromachining are discussed in Chapter 3 – Photonics in Laser Machining.

4. Laser Power Measurement and Beam Profiling

The optical techniques used in semiconductor manufacturing require accurate measurements of laser power or energy to achieve repeatable results that ensure tool-to-tool consistency independent of geographical location. Applications requiring laser power measurements range from logging emitted power level from light sources to confirming the optical throughput of optical assemblies. Detectors use wavelengths ranging from 190 nm for DUV systems to 12 μm for IR systems and have power sensitivities ranging from nW to kW. Measurement accuracy is a concern for tool-to-tool consistency and NIST-traceable calibration of optical components is a necessity. Figure 149 shows the wide range of wavelength, power, and energy ranges for detectors used for semiconductor manufacturing.

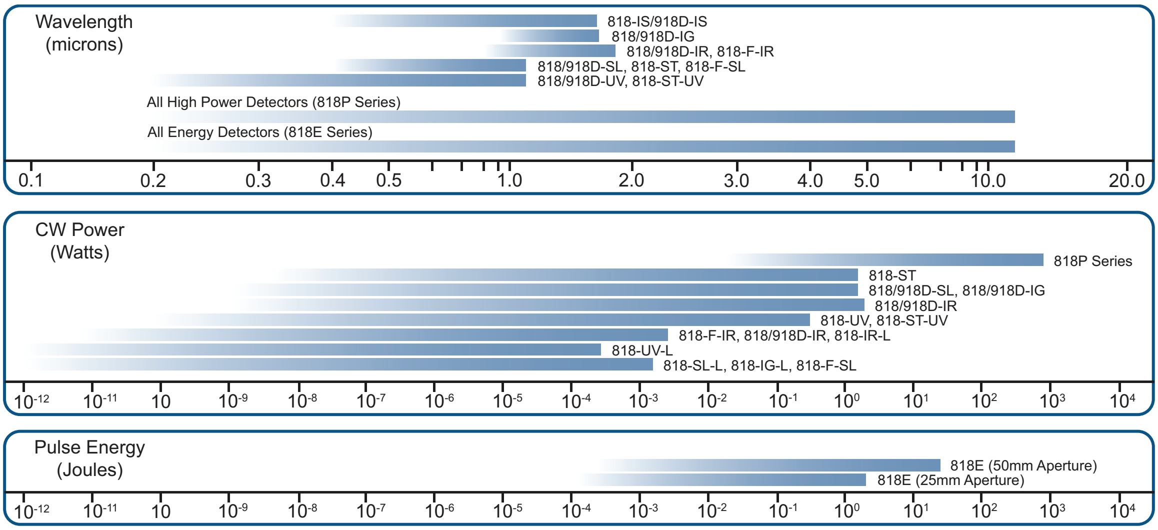


Figure 149. MKS detector performance charts. Note: The gradient shaded regions of the detector power and energy operating ranges reflect the dependence of the measurement noise floor on the amplifying electronics used with the detector. Newport brand optical meters are designed to minimize the noise floor.

Laser beam quality is another vital parameter for optical systems in semiconductor manufacturing since issues such as beam clipping can reduce throughput and greatly increase measurement times or decrease accuracy. Advances in beam profile measurement instrumentation and software have enabled laser manufacturers to greatly improve laser beam quality. These beam profilers are used for rapid troubleshooting at the tool level and during R&D for tool development.

5. Semiconductor Applications Using MKS Products

MKS Precision Optics and Custom Optical Solutions

MKS specializes in optical sub-system design and manufacturing and has designed and manufactured assemblies for wafer and reticle inspection tool manufacturers that feature optics from the MKS precision optics fab. ISB has designed and manufactured optical sub-systems for lithography, wafer inspection, excimer and EUV light sources, metrology and mask writing applications, among others. ISB has produced custom optical assemblies for applications ranging from illuminators for water inspection tools, as shown in Figure 150.

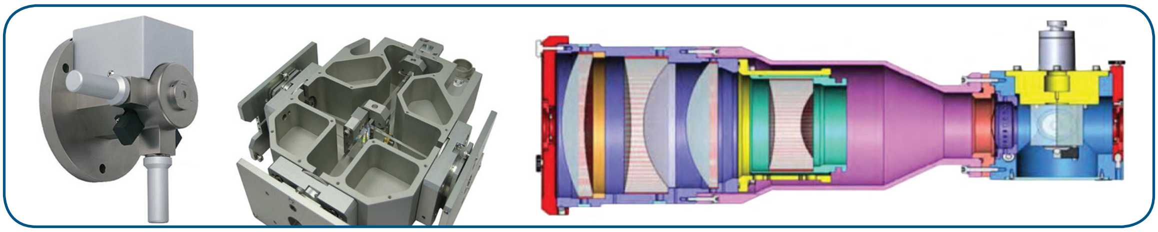


Figure 150. 157 nm objective lens for VUV photolithography (left); high power LED illuminator for use in inspection tools (middle); custom optics (right).

MKS precision optics fabrication manufactures advanced optics for OEM customers. Products include lenses, mirrors, windows, beamsplitters, waveplates, filters, and polarizers. These feature high precision surfaces polished to tight irregularity, surface roughness, and cosmetic specifications which are verified using state-of-the-art metrology equipment. Thin film coatings are engineered over a broad spectral range (DUV to VIS to IR). They also employ multiple deposition methods to produce broadband, narrowband, and multi-band anti-reflective, reflective, filter, and polarization coatings that are verified with sophisticated coating metrology tools (see Sections III.E.1 and III.E.2 for details).

MKS Motion Control for Wafer Inspection and Reticle Inspection

MKS offers a variety of high-performance air bearing stages suitable for use in wafer inspection tools and other motion control applications. Experienced MKS application engineers collaborate with OEM customers to provide automated motion solutions specific to the semiconductor manufacturing process being developed. This section describes the technologies used in these systems to increase accuracy and dynamic performance. Table 10 illustrates some typical accuracy requirements for motion components.

MKS Motion Solutions	Type	Required Accuracy
Stack of IDL stages	Mechanical bearings	0.5 µm
Integrated XY CeraMech™	Mechanical bearings	0.25 µm
HybrYX™	Hybrid mechanical and air bearings	0.15 µm
DynamYX® (ZTT)	Air bearings	0.1 µm
DynamYX® DATUM®	Air bearings	0.05 µm

Table 10. MKS Motion Solutions typical accuracy.

The IDL Industrial Linear Stage Series (Figure 151) was developed for high dynamic performance; it is available in standard versions that facilitate fast response time in building XY systems. The stage has an aluminum structure with a design optimized for stiffness that combines with mechanical bearings, linear motors, linear encoders and integrated cable management to provide reliable and accurate positioning. The stack XY IDL's key attributes for automated optical inspection are apparent based on its specifications, which vary depending on the travel range.

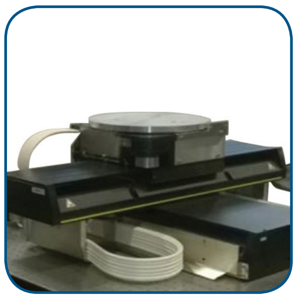


*Figure 151.
IDL Industrial Linear Stage.*

CeraMech^{TM} (Figure 152) is a cost-effective XY stage that employs ceramic construction and mechanical bearings to provide exceptional performances without the use of air bearing guide ways. The stiffness and stability of the ceramic stage is three times higher than that of an aluminum stage. MKS has developed advanced manufacturing capabilities for flat

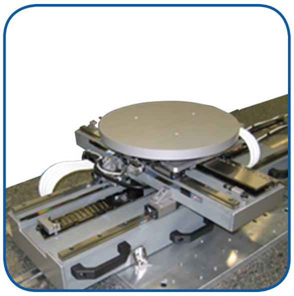


Figure 152. CeraMech™ XY Stage.

Figure 153. HybrYX™ XY Hybrid Air Bearing XY Stage.

Figure 154. DynamYX® DATUM® Ultra-High-Performance Stage.

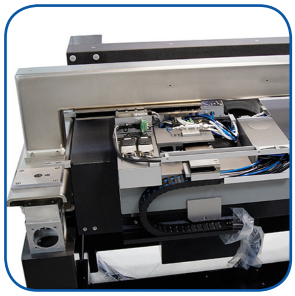


Figure 155. DynamYX® RS Stage.

and stable reference surfaces that enable this technology to achieve high dynamic performance for both scanning and step-and-settle applications.

The HybrYX™ XY Hybrid Air Bearing-Mechanical Bearing XY Stage (Figure 153) is a relatively low-cost, single plane, air bearing stage well-suited for semiconductor wafer or panel inspection systems as well as many other scanning applications that require ultra-low velocity ripple and dynamic following error, i.e., the difference between the current position and the desired or theoretical position setpoint at any point in time. The Z Tip Tilt Theta option for the stage allows fine alignment of wafers or panels for serpentine XY motion profiles. HybrYX™ systems provide reliable, long-life operation that is ideal for high duty cycle environments such as wafer inspection applications. The HybrYX™ 300, for example, is a cost-effective hybrid ball bearing and air bearing stage optimized for wafer inspection having scanning speeds of up to 1 m/s, with stability < 0.1%, flatness of ± 250 µm over 300 mm, and XY position accuracy ± 250 nm over 300 mm.

The DynamYX® family of stages (Figure 154) provides the highest level of commercially available positioning performance. Specifically designed for 300 mm wafer processing and inspection applications, DynamYX® stages make extensive use of ceramic materials to provide exceptionally rigid structural stability. These stages are designed with a low profile to facilitate their use in OEM applications.

The DynamYX® Reticle Positioning Air Bearing Stage (Figure 155) is designed for use in reticle inspection and repair applications. The DynamYX RS stage has a much smaller footprint than traditional open-frame solutions and a full-open aperture that accommodates flexible optical component integration and ease of service access. It features a stiff and stable ceramic holder in the front side, that ensures high stability and cleanliness during reticle inspection and/or repair processes.

The DATUM® series stages are also suitable for use in wafer inspection applications. They provide ISO 2 level cleanliness at the wafer level and scan up to 1 m/s with the following specifications: Dynamic XYZ jitter ± 30 nm, flatness of ± 150 µm over 300 mm, and XY position accuracy of ± 100 nm over 300 mm.

Motion products are also used for beam routing if the laser source is located separately from the tool. For example, MKS' fast steering mirrors (FSMs) use a feedback loop and motorized mirror mounts to control the positioning of a laser beam. In R&D applications this can be useful for beam stabilization between two independently floating optical tables. In fabrication applications, FSMs are used in a periscope configuration when the laser source is separated from the tool by a long distance. A mirror on the light source (for angular correction), and one on the tool (for positional correction), provide control of all degrees of freedom for the beam entering the tool.

Figure 156.
TLB-8800 (left) and TLM-8700 (right) Venturi ECDL Tunable lasers.

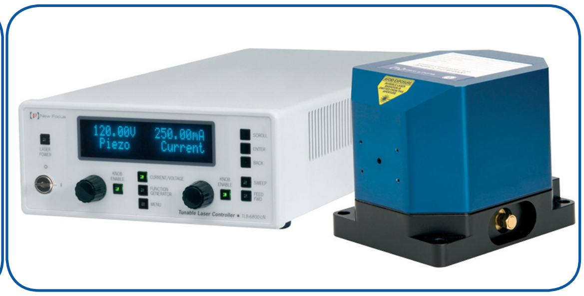


Figure 157.
Vortex Plus / Stablewave™.

MKS Lasers for Reflectometry and Interferometry

MKS Spectra-Physics® brand provides a broad portfolio of high-performance, reliable laser systems backed by a team of applications engineers and a responsive global support organization. Spectra-Physics lasers are used in a wide variety of industrial manufacturing applications, including stereolithography, rapid prototyping, and laser shock processing.

Figure 158.
SWL-7500

Spectra-Physics External Cavity Diode Lasers (ECDL) are tunable frequency light sources with a very narrow linewidth (~100 kHz) for application in laser reflectometry and interferometric analysis of semiconductor wafers [5]. MKS provides OEM engineering capabilities for customizing these sources to customer specifications, e.g. mechanical and optical re-designs, form factor modification, fiber coupling, optical isolator addition, extended mode-hop-free tuning range, electronics integration, and tuning speed or tuning linearity customization. Figure 156 illustrates ECDL models designed for semiconductor manufacturing. The TLB-8800 tunable laser uses a Littman-Metcalf design optimized for laser scanning and reflectometry metrology. It offers ultrafast piezo tuning (2000 nm/s and option for 20,000 nm/s) for rapid analysis at high repetition, e.g. 30 Hz at 10,000 nm/s, and wide (100+ nm) mode-hop-free tuning. The TLM-8700 is the OEM version of the TLB-8800, which offers a compact form factor.

The TLB-8800 (Figure 157) produces tunable red light for wafer inspection, metrology, and lithography positioning. The SWL-7500 (Figure 158) is a compact, single wavelength source with narrow linewidth (< 200 kHz) and high coherence length (> 100 m). The 632 nm model is suitable for replacing HeNe lasers in interferometry.

MKS Calibrated Power Detection

MKS Newport brand power meters and photo detectors offer ISO-9001 or ISO-17025 certified NIST-traceable calibration for the most accurate power and energy measurements possible (Figure 159). These are custom-tailored for OEM applications and stocked as catalog products for R&D. Photodiode (818 and 918D series), thermopile (919P), and pyroelectric (919E) detectors convert incoming photons to a measurable electrical signal that is read by compatible power meters, corrected for detector responsivity, and displayed as power. The fundamental operation principles for the detectors are described in detail in Chapter 1, Sections II.A.1 and II.A.2. Various form factors are available to accommodate most applications, including the customer-inspired 818-ST2 wand-style detector with 5.4 mm clearance. Integration sphere sensors (918D-IS, 819C, or 819D) that are compatible with most fiber optic connector types reduce measurement uncertainty further by homogenizing light over the sensor active area.

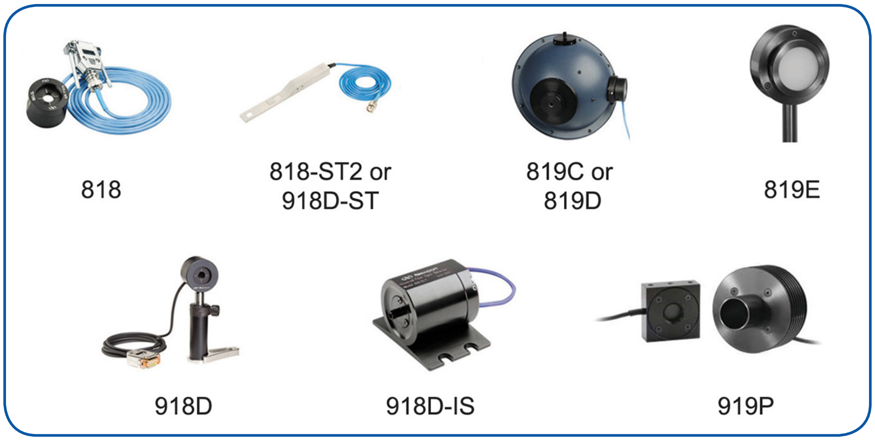


Figure 159. Newport photo detectors.

Available MKS Newport brand power meters (Figure 160) range from a low cost, hand-held meter (843-R) to the most advanced dual channel benchtop power meter currently on the market (2936-R). The 1936-R/2936-R series features state-of-the-art analog boards with a 250 kHz sampling rate and femtowatt level resolution. MKS' Ophir brand offers an extended range of high-power measurement options for laser machining. For further details, see Chapter 3, Section III.B.3.

Figure 160. Newport power meters.

<u>MKS Laser Beam Profilers for Beam Quality</u>

MKS laser beam profilers offer a wide range of beam analysis equipment and several classes of software (Newport LBP2™, software, or Ophir BeamMic™, BeamGage®, and NanoScan™). CCD camera beam sampling systems are used for beam profiling of lasers with powers from 10 mW to 400 W. A UV imager extends the detectable wavelength range from 190 nm – 1100 nm for VIS models; an anti-Stokes phosphor filter extends the detectable wavelength range from 1440 nm – 1605 nm for IR models (Figure 161).

Figure 161. LBP2™ series laser beam profiler.

Figure 162 is a 3D graphical image generated by MKS software. It shows the power intensity distribution of a pulsed Nd:YAG laser that is used to trim active circuits. Further laser beam profiler applications are covered in Chapter 1, Section II.B.3 and in Chapter 3, Section III.B.4.

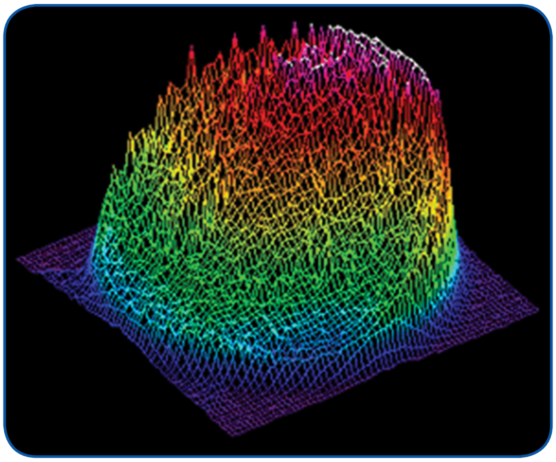


Figure 162. 3D beam profile of a laser used for trimming active circuits. [250].

MKS Cleanroom Compatible Vibration Isolation

MKS optical vibration isolation solutions are customized for each application and can be built for vacuum compatibility or cleanroom compatibility. Our cleanroom optical tables combine a rigid trussed honeycomb core with innovative tuned damping. During production, the tapped hole array is sealed in a protective membrane throughout the assembly process to prevent contaminant penetration. This cleanliness and stability make the difference between success and failure in leading edge applications. Performance specifications and basic construction are the same as the RS4000 Series tuned damped optical table tops, which use six tuned mass dampers to selectively eliminate torsional and bending table nodes. MKS also offers cleanroom S-2000AC isolators for vibration isolation with pneumatic re-leveling (Figure 163).

Figure 163. UCS cleanroom compatible optical table top and S-2000AC isolators.

F. Future Directions for Photonics in Metrology and Inspection

The latest-generation broadband plasma inspection tools have higher sensitivity and speed because of the use of DUV light from the plasma source, smaller inspection pixels, advanced algorithms, and optical apertures [251]. These systems enable full-wafer inspection in about an hour. The trend is to use increasingly shorter UV wavelengths and higher power to reduce scan time and increase SNR. DUV optics for these systems are manufactured using rigorous anti-photocontamination protocols. These include using only UV-compatible materials, specialized fabrication processes, engineered coatings, advanced packaging, cleanroom handling, and extensive metrology for alignment. These protocols will continue to evolve as more semiconductor manufacturers adopt DUV. Cleanliness and contaminant reduction also matter for precision motion solutions. The move to shorter wavelengths is driving a need for fast, accurate motion stages with high cleanliness and materials with high flatness.

III. Photolithography

A. Brief Overview of Applications

Lithography is defined as “a method of printing from a flat surface (such as a smooth stone or a metal plate) that has been prepared so that the ink will only stick to the design that will be printed” [252]. In semiconductor device manufacturing, the stone is the silicon wafer while the ink is the combined effect of the deposition, lithography and etch processes that create the desired feature. Since lithography for device fabrication involves the use of optical exposure to create the pattern, semiconductor lithography is commonly called “photolithography.” As with the inspection and metrology techniques already discussed, photolithography is the technique of choice for patterning because it is optical, and thus enables small features and high wafer throughput. This contrasts with other techniques such as direct writing and imprint.

B. Basic Principles

This section will outline the key steps for photolithography. Full details on the process can be found in [5]. Figure 164 illustrates a typical photolithography process used to define shallow trench isolation features. Such a process consists of the following steps:

1. Substrate cleaning and preparation
2. Form layers of thermal oxide and deposit a layer of silicon nitride on the clean substrate
3. Deposit a carbon hard mask followed by a layer of anti-reflective material
4. Deposit a layer of photoresist
5. Pre-bake the photoresist
6. Align the substrate/photoresist and reticle and expose the photoresist using UV radiation and 4x-5x imaging. Step and scan to repeat
7. Post exposure bake
8. Develop the pattern in the photoresist and hard bake to remove remaining solvent
9. Perform etch to open dielectric anti-reflective coating (DARC) and hard mask pattern and remove photoresist and DARC
10. Perform etch to open trenches in substrate and remove hard mask
11. Clean surface

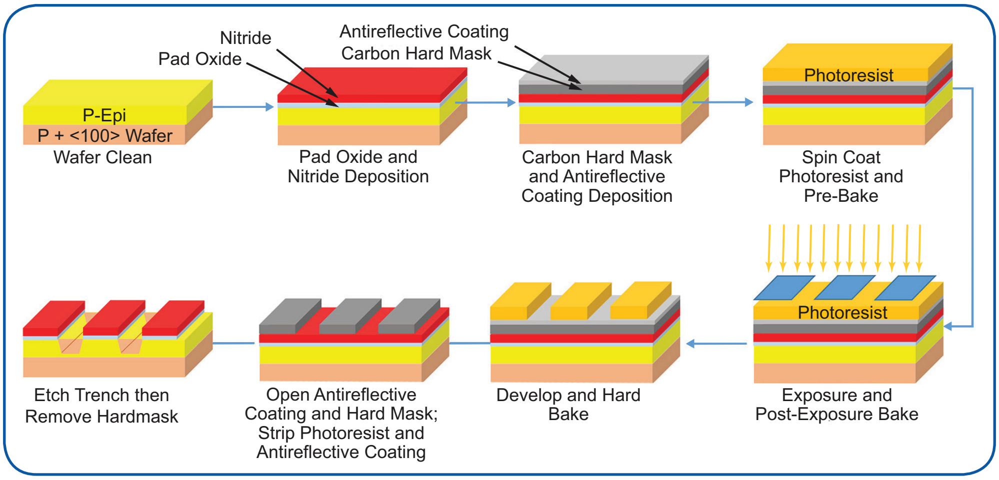


Figure 164. A schematic representation of a semiconductor device patterning process.

C. Deep UV Photolithography

DUV technology for photolithography is exclusively based on projection optics since the pattern on the photomask is much larger than the final pattern developed on the photoresist. The optical system in a 193 nm photolithography tool is known as a catadioptric system. The term means that it uses both lens (refractive) and mirror (reflective) elements for directing and conditioning the light from the laser. Details regarding the operation of such elements can be found in Chapter 1, Sections III.A.1 and III.A.3. The advantage of this type of system is that it accommodates a broad bandwidth of the source laser light while limiting chromatic aberration. Refractive elements in the optical system are fabricated from either synthetic fused silica or calcium fluoride, materials that have low absorption of 193 nm light. Photomasks (or reticles) in these systems are typically made from fused silica with chrome patterns. Figure 165 shows a schematic of the exposure mechanism and relative motions of the reticle and wafer in a step-and-scan exposure system along with how water immersion is maintained between the objective lens and the wafer. In the step-and-scan process, a slit of light is scanned across one or more dies patterned on the reticle. The light reproduces the part of the pattern on the reticle that is illuminated on the wafer, albeit at much reduced feature size owing to passage through the reduction lens. Simultaneous (and highly precise, accurate, and repeatable) movement of both the reticle and the wafer is used to produce the full image of the die on the wafer. Once a die has been patterned, the process “steps” over to the next die area to be patterned.

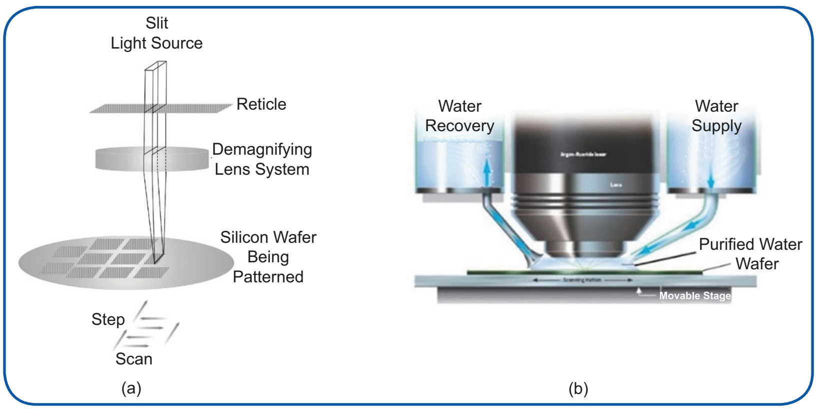


Figure 165. (a) Configuration and relative motions in a step-and-scan exposure tool [253]; (b) water immersion arrangement (figure reprinted with permission from Lachina [254]).

The catadioptric projection optical approach is discussed below with other semiconductor photolithography technologies described in available monographs [255, 256]. The physics of light are a strong determinant for the ultimate resolution (or minimum feature size) achievable in a given photolithography process (along with other factors related to substrate and resist properties and design methodologies). The smallest linewidth (W) that can be printed is determined by the wavelength (λ) of the exposing light and by the NA of the projection optics according to the Rayleigh equation:

$$W = k_1 \frac{\lambda}{NA}. \tag{23}$$

k_1 is a factor that accounts for the processing characteristics such as the quality of the resist and the use of resolution enhancement techniques like off-axis illumination. k_1 has a theoretical minimum of 0.25, although values below 0.3 are considered too difficult or expensive for common use. The NA is a measure of the optical system's ability to collect and focus the light from the source. Figure 166 shows the relation between the NA of a lens system and other relevant parameters in the system. In terms of smaller feature sizes, a larger NA is desirable since it reduces the minimum feature size achievable in the photolithography system. The maximum NA of a lens operating with air as the imaging medium is 1.0, although values greater than 0.95 are not normally found. Typical linewidths for single exposure patterning with λ = 193 nm immersion scanners are around 40 nm. This can be reduced with design methods and additional multiple patterning techniques to produce 22 nm feature sizes and smaller (see below).

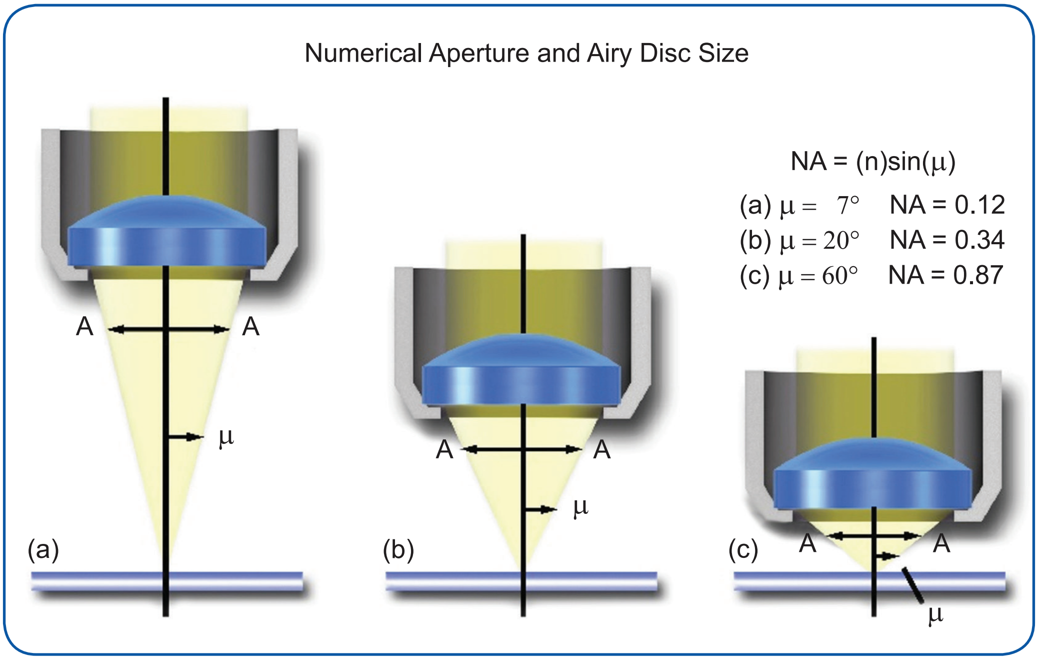


Figure 166. The relationship between NA, the half-angle of the light cone, and the refractive index of the imaging medium between the lens and the substrate (reprinted with permission of Molecular Expressions.com at Florida State University) [257].

The aim of photolithography is to produce accurate 2D images, but the optical imaging technique is inherently 3D because the aerial image is projected into air and then onto the photoresist. A consequence of this is that the expected sharp contrast between light intensity in the bright and dark areas of an image is reduced because a gradient is present in the light intensity (see Figure 167). This can reduce the quality of a lithographically-patterned line. The Normalized Image Log-Slope (NILS) method is used to quantify the aerial image quality [258]. Minimum acceptable NILS values can be calculated using empirically-determined constants [259].

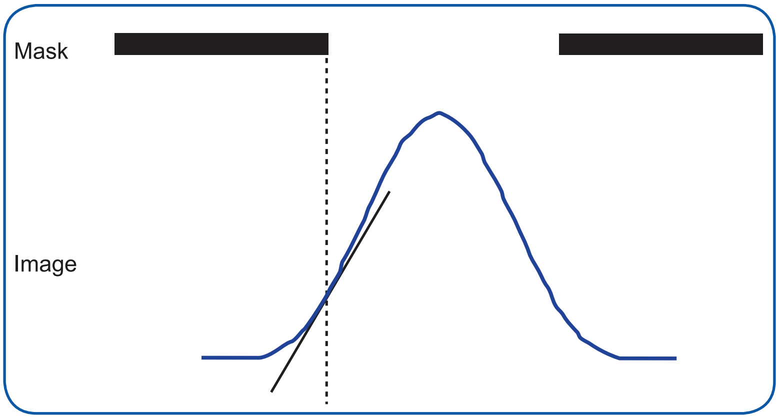


Figure 167. The NILS method used to assess image quality in photolithography [258].

The depth of focus (DOF) is the vertical distance over which the image remains in focus. A sufficiently high DOF is required for the entire resist layer to be developed during photolithography. The DOF is also determined by λ and NA according to:

$$DOF = k_2 \frac{\lambda}{(NA)^2}, \quad (24)$$

where k_2 is related to k_1 and has a minimum value of 0.5 with conventional resist technologies. Using this representative value for k_2 and a maximum NA of 0.95 shows that resist thicknesses are limited to about 100 nm for conventional 193 nm applications. The use of systems with higher k_2 values, e.g., systems employing advanced resists such as polymethyl methacrylate and other system modifications, allows thicker resists to be used.

These rules highlight the key parameters that have been adjusted as photolithography technology has progressed to smaller and smaller feature sizes. The Rayleigh equation (Equation (23)) shows that, to reduce the feature size, one must either reduce the wavelength of the exposing light or increase the NA of the projection optics. The DOF (Equation (24)) must be sufficient to ensure accuracy and precision in the feature size through the entire thickness of a resist. Simplistically, the feature size that is achievable in a lithography process depends on the Rayleigh equation, while the process yield is dependent on the projection system's DOF.

Figure 168 shows the historical progression of IC feature sizes and the wavelength of the photolithography light source required to achieve these feature sizes. Until recently, photolithography equipment designers have focused primarily on wavelength reduction to achieve smaller feature sizes:

- Arc lamp systems developed in the 1970's and early 1980's (λ = 436 nm) are useful down to feature sizes of about 450 nm
- Mercury lamp I-line systems developed in the mid-1980's (λ = 365 nm) are useful down to feature sizes of about 380 nm
- KrF excimer laser-based systems developed in the 1990's (λ = 248 nm) are useful down to feature sizes of about 250 nm
- ArF excimer laser-based systems (λ = 193 nm) are useful down to feature sizes of about 65 nm

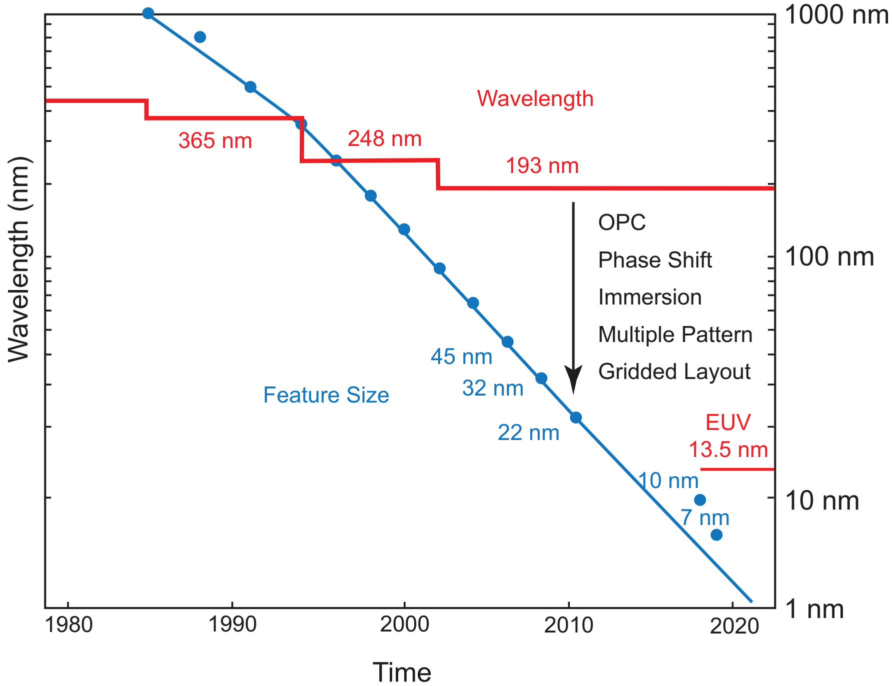


Figure 168. Historical progression of IC feature size and photolithography technologies.

Modern DUV photolithography systems employ ArF excimer lasers at 193 nm and would be conventionally limited to feature sizes ≥ 65 nm or larger when air (NA = 1) is the medium between the optical system and the substrate. Since light sources below 193 nm were not readily available at the time (see below for EUV sources), advanced techniques were required to achieve the needed reductions in feature size for the 45, 32, and 22 nm technology nodes. By creative use of different combinations of optical proximity correction (OPC), phase shift, immersion lithography, and multiple patterning, manufacturers have extended 193 nm lithography to produce feature sizes significantly below conventional expectations. Specific details of these techniques are addressed below. Dry 193 nm photolithography in combination with double patterning has been successfully employed in 45 nm patterning technology. This technology has been adapted to include immersion lithography that has achieved 32 nm (and below) patterning technology.

OPC techniques compensate for image distortions that occur when printing feature sizes smaller than the wavelength of the exposing light. Typically, these distortions result in shortening of line ends, corner rounding or changes to linewidths. OPC corrections are made at the mask level and involve changes to the mask image such as the addition of serifs at design corners and augmentation to linewidths. Figure 169 shows a simple example of how OPC corrections are made at the mask level.

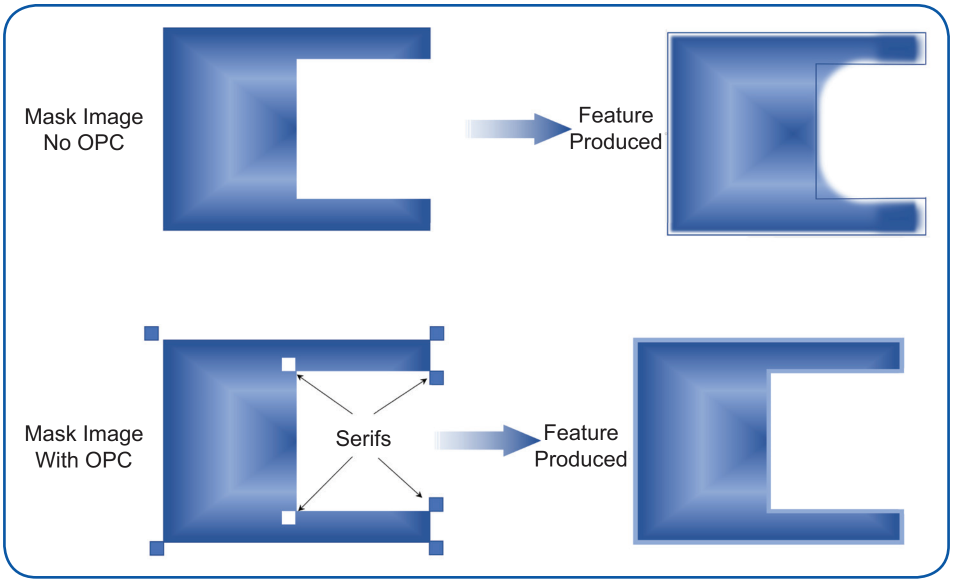


Figure 169. Representative OPCs to a photomask [260].

Phase shift [261] is a technique that enhances edge contrast in the image being patterned, removing defects that occur due to diffraction limitations at sub-wavelength patterning. Phase shift masks do this by changing the thickness of different sections of the pattern on the mask, which changes the phase of the transmitted light passing through (Figure 170). A useful description of how phase shifting works is provided in [255].

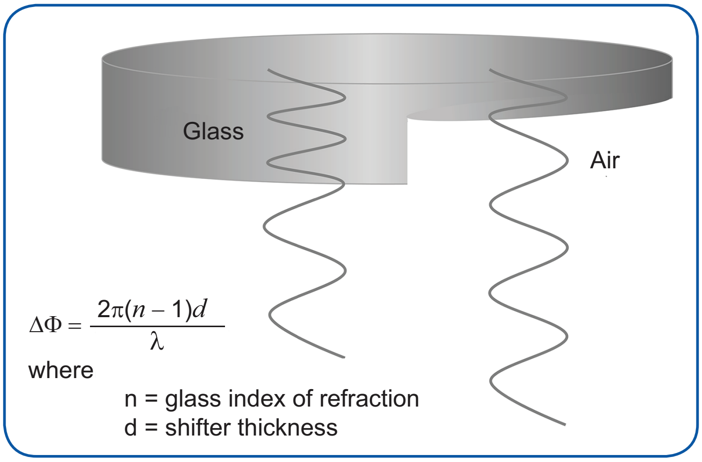


Figure 170. Shifting the phase of light by using different mask thicknesses [262].

Figure 171 depicts immersion lithography, which bypasses the feature size limitations of dry lithography by changing the medium between the optical system and the substrate from air to water. Since water has a refractive index of 1.44, this increases the value of NA beyond 1.0, leading to a reduction in minimum single-exposure feature size to about 40 nm when using 193 nm light. Water has come into standard use as the medium in immersion lithography systems since immersion techniques both increase the amount of light that can reach the resist (increasing the resolution) and change the phase of the light so that it improves DOF. For these and other reasons, single exposure immersion lithography is the dominant approach for patterning in advanced device fabrication processes at design nodes down to 45 nm.

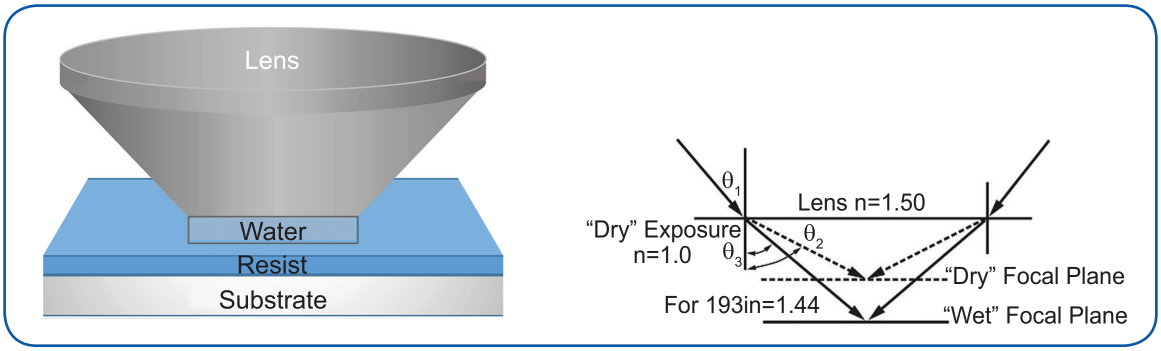


Figure 171. Layout and optical characteristics of immersion lithography [263].

Below the 45 nm node, the combination of 193 nm immersion lithography with enhanced techniques such as multiple, i.e., double, triple, quadruple, patterning provided the smallest possible feature sizes until the advent of cost-effective EUV lithography. Quadruple patterning using multiple, shifted exposures, such as the process shown in Figure 172, effectively lowered the feature size limits. Quadruple patterning has provided a solution for patterning features as small as 5 nm.

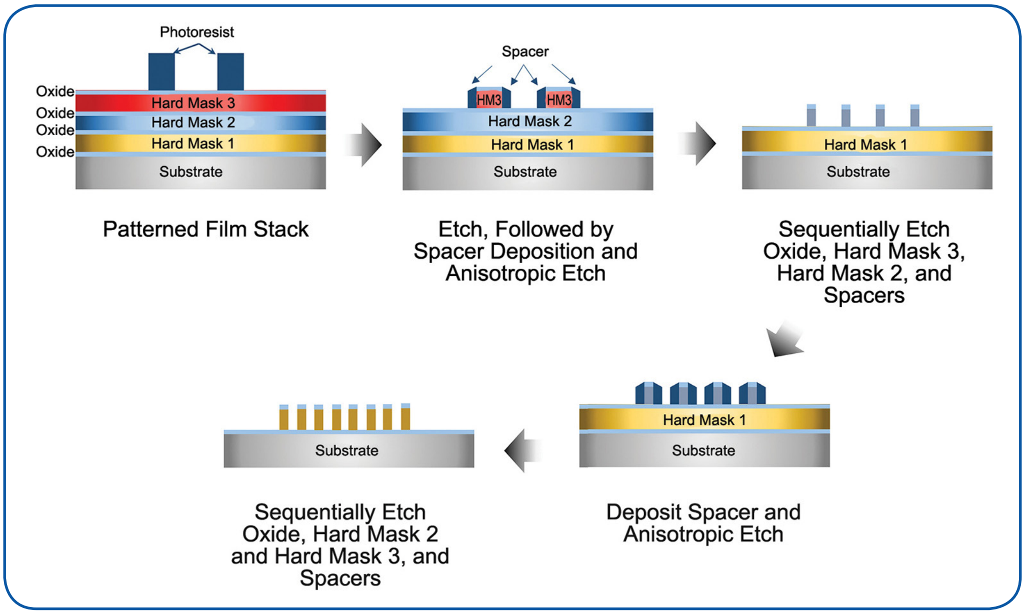


Figure 172. Multiple patterning technique [264].

Table 11 provides examples of successful application of 193 nm lithography in IC manufacturing. 193i+ refers to extended 193 nm photolithography aimed at avoiding the use of EUV photolithography. Modifications include increased refractive indices of the immersion fluid, lenses, and resists [265].

k$_1$		NA								
		193 nm			193 nm water immersion				193i$^+$	
		0.75	0.85	0.93	1.07	1.2	1.3	1.35	1.55	1.80
Half-Pitch Size	130 nm	.505	.573	.620						
	90 nm	.350	.396	.429	.499	.560				
	65 nm		.286	.310	.360	.404	.438			
	45 nm					.280	.303	.315		
	32 nm								.257	.298

Table 11. k1 values for specific NA, and feature sizes for representative 193 nm patterning processes.

D. Extreme UV (13.5 nm)

EUV lithography [266] is being developed to fulfill single-exposure patterning requirements at feature sizes below 22 nm (Figure 173). Unique to this technology is the nature of the light source. There is no readily available conventional light source having a wavelength below 157 nm which is the wavelength of light from a F_2 excimer laser. The F_2 excimer laser source has not gained broad use for lithography at small feature sizes. This is probably due to factors including difficulties with the required CaF_2 optics, e.g. increased mask registration errors due to birefringence, and the success of the extension of 193 nm patterning technology that allowed it to function down to the single-exposure limit of 157 nm lithography, i.e., 32 nm. Instead, lithography equipment developers have turned to an entirely new way of generating light at EUV wavelengths for use in lithography at sub-32 nm feature sizes.

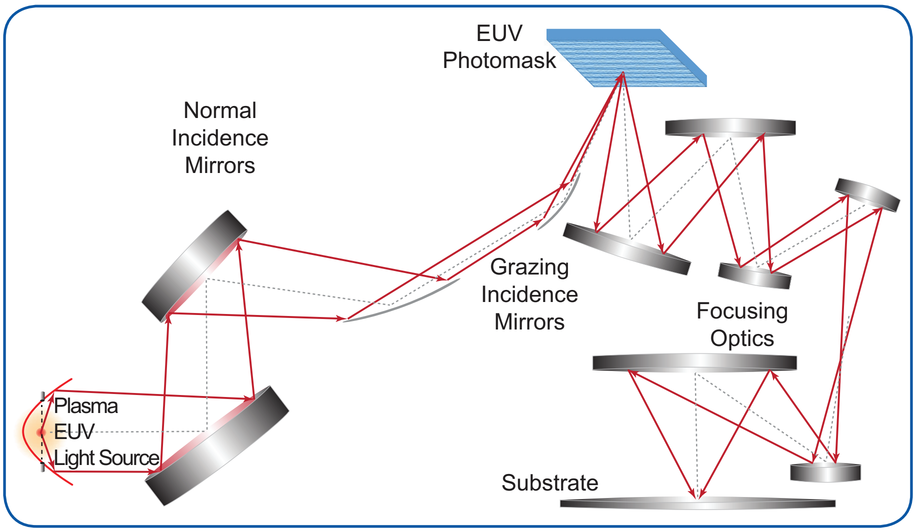


Figure 173. EUV lithography schematic [266].

13.5 nm light sources of consist of a high-power CO_2 laser, laser beam transport and focusing optics, and a light source vessel. Within this system, 13.5 nm light is generated using laser pulsed plasma excitation of tin molecules by a CO_2 laser. The light is generated inside a vacuum vessel where droplets of tin pass through the focal point of pulses of high intensity light from the CO_2 laser [267]. As the small (~30 μm diameter) droplets of molten tin encounter the pulse of high intensity laser light, they are vaporized, and the tin atoms undergo electronic excitation and ionization, creating a hot plasma with electron temperatures of 10's of electron volts (1 eV = 11,605 K). Electron-ion recombination and ion deexcitation within the plasma emit photons at 13.5 nm. Conversion efficiencies (CEs) for this mode of light generation are relatively low, with CEs of 3-5% reported for some tin-based EUV sources. The light emitted by the plasma is collected using an ellipsoidal collector that reflects it to an intermediate focus at the front of the light source. The mirror is unique in that it is not a conventional polished optical mirror, but rather a Bragg reflector (Chapter 1, Section I.B.3 discusses Bragg reflectors) made up of thin layers of molybdenum (Mo) and silicon (Si) designed to reflect the greatest amount of 13.5 nm light. From the intermediate focus, the light is transported to the optical train in the photolithography system using solely reflective optical components. The entire optical system is maintained under high vacuum since 13.5 nm light is strongly absorbed by all solids, liquids, and gases. The mirrors are also Bragg reflectors and can reflect up to 70% of the incident light. They are critical system components that must have extremely low surface roughness (a few atoms) and highly precise flatness and curvature. Since reflection is not 100%, the number of reflectors between the source and the substrate has a significant impact on the energy of the EUV beam at the substrate surface. Since the optics are all reflective, the masks used in EUV scanners must also be reflective, since no optical materials are transparent to EUV. In theory, transmission masks could be used if the substrate was less than 100 nm thick, but this is not a practical solution. EUV masks are fabricated on very low thermal expansion substrates using multilayer Bragg reflector technology. The multilayer reflector has 40-50 pairs of Mo and Si thin film layers each and a total thickness of about 300 nm topped by a capping layer of about 11 nm of Si. A buffer layer, typically SiO_2, is deposited on top of the stack, followed by an absorber material with a thickness of about 100 nm. This absorber layer and the buffer layer are patterned to produce the reflective mask.

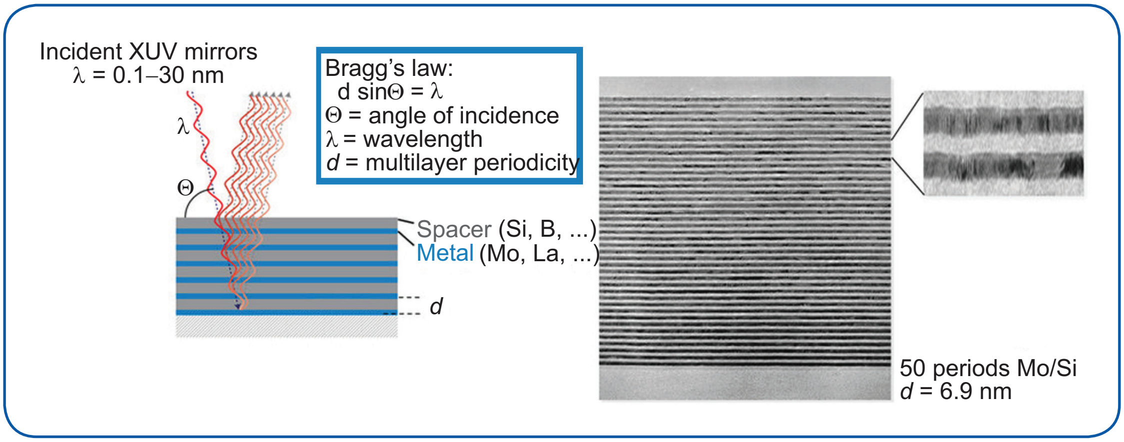


Figure 174. The principle employed in a Bragg reflector (left); cross-sectional transmission electron microscopy image of multilayer EUV mirror grown by e-beam evaporation and ion beam sputter deposition (right). Figure reprinted with permission from F. Bijkerk, University of Twente [268].

Figure 174 shows a transmission electron microscopy image of the multilayer stack in a EUV mask while Figure 175 provides a simplified process flow for EUV mask fabrication. Since there are no optical materials that are transparent at 13.5 nm, any pellicles employed with EUV masks must be extremely thin (pellicles are the enclosures used to protect the photomask). 50 nm thick pellicle membranes supplied by ASML can transmit 85% of the EUV light (useful for lower power EUV) [269]. Great care is required in maintaining these masks in a defect-free state. Transport of masks to and from a lithography tool is performed using specialized dual pod containers to ensure mask integrity. There are currently no actinic inspection techniques for EUV masks, i.e., inspection using a wavelength of light that is similar to the feature sizes on the mask.

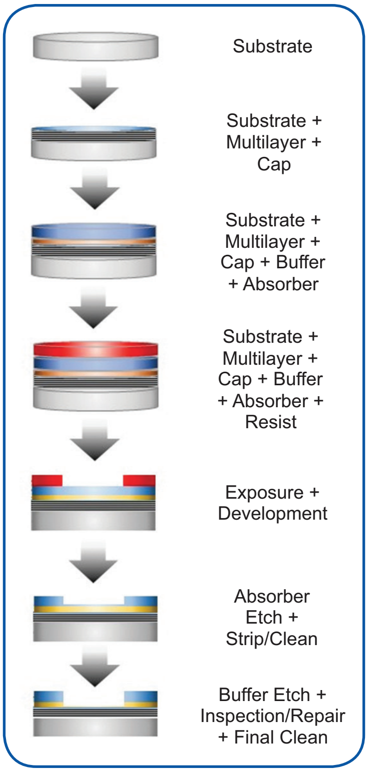


Figure 175. EUV mask fabrication step.

Photoresist materials for EUV continue to be developed and improved. It is expected that photoresist formulations will have some unique requirements compared with conventional resist technology. Chemically amplified resists that served the industry well since the introduction of 248 nm exposure wavelengths appear to have certain inherent resolution/sensitivity trade-offs that will force the development of more advanced resist formulations for EUV technology. Currently, hybrid approaches such as those that employ a sacrificial spin-on carbon film and a resist layer containing organometallic molecular linkages are receiving considerable attention [270].

E. Lasers and Photonics for Photolithography

1. Precision DUV Optics

Photo-molecular contamination is a significant issue for reliability and lifetime in semiconductor photolithography systems. Consequently, a great deal of care must be taken to prevent the exposure of the optics in these systems to adhesives, lubricants, and any other organic carbon as well as to siloxanes, phosphonates, or sulfates. The organics are absorptive in the DUV range and can react to form a variety of damaging contaminants after adsorption on the surface of optics and subsequent DUV illumination. Trace contamination by volatile organic compounds, condensable Si-O and inorganic compounds can lead to loss of light either in transmission or reflection and can cause wavefront distortion, significantly affecting performance and causing unplanned downtime. It can cause Strehl reduction (or optical imaging quality reduction), polarization changes and even de-tuning, i.e., shifts in the optical wavelength, to negatively affect the performance of a photolithography system. The scope of photo-molecular contamination covers DUV wavelengths (190 – 355 nm) and EUV wavelengths (sub-190 nm, typically 13.5 nm). Photo-molecular contamination mechanisms are complex and are highly dependent on power levels as well as type and concentration of chemical compounds. There are multiple aspects to consider and specific solutions must be understood and proven in specific applications. As wavelengths continue to shorten and power increases, prevention of this form of chemical contamination becomes increasingly important.

Optics fabricators have methodologies for eliminating or limiting the effects of photo-molecular contamination that can help achieve high performance, long lifetime optics and systems. These include proprietary optical materials and compounds, proprietary polishing, cleaning and coating processes, and cleanroom handling equipment and processes. This section describes the capabilities that set DUV optics apart from standard catalog optics, including:

- Extensive R&D in materials science
- Reliability and lifetime testing
- Initial design of the system
- System design for long life and minimal preventative maintenance cycles
- Sub-tier supply chain management and control
- Internal cleanroom environments and production control
- Preservation of cleanliness and packaging

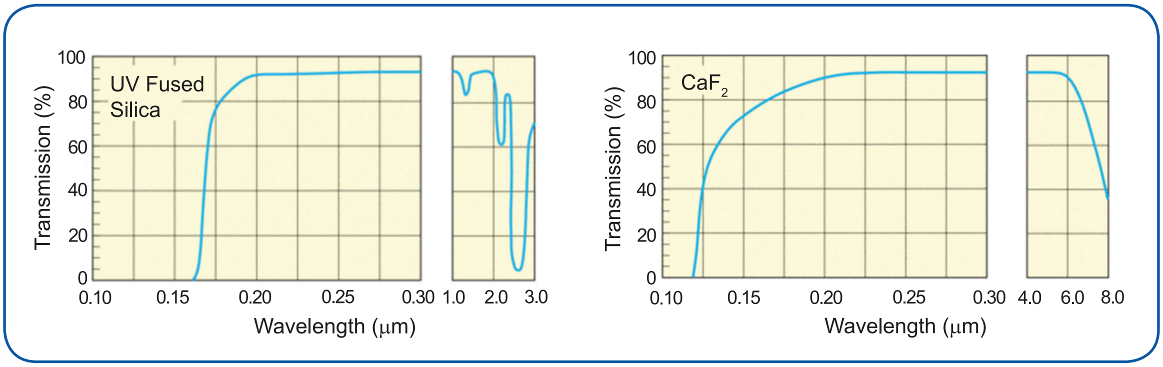


Figure 176. DUV optical material transmission plot [158].

Many optical materials have low transmission below 200 nm, and so UV fused silica or calcium fluoride (CaF_2) are favored for DUV transmissive optical substrates. Figure 176 shows the typical transmission for these materials, which extends below 200 nm and then steeply drops off. However, CaF_2 optics are subject to defects and slip-planes if not coated using specially optimized processes. Extensive research and testing have gone into the selection of polishing compounds and processes that are compatible with DUV wavelengths. Some polishing materials/compounds will absorb UV/DUV light and this can affect reliability and lifetime of the optical component. Others may contain chemical compounds that directly react with DUV light to cause damage and failures. The surface polishing requirements are tighter for precision optics, and Table 12 displays the capabilities of the MKS precision optics fab. Optics fabrication is completed using processes such as computer numerical control (CNC), magneto rheological figuring (MRF), pitch lapping (PL), and single-point diamond turning (SPDT).

Custom Optic Type	Fabrication Technique	Spectral Region		Material Type	Finished	Radius Of	Specification					
							Surface Irregularity		Cosmetic Quality (Scratch-Dig)		Wedge	Surface
							Typical	Best	Typical	Best		
Spherical Lenses	CNC, MRF, PL	VUV	120 nm - 220 nm	CaF_2, FS	120 mm - 250 mm	15 mm - infinity	Lambda/10	Lambda/20	10:05	05:01	30"	< 3A
		DUV	193 nm - 220 nm									
		UV	220 nm - 400 nm									
		VISIBLE	400 nm - 700 nm	ALL					40:20	10:05		< 5A
		NIR	700 nm - 1.5 microns									
	SPDT, CNC, MRF, PL	MWIR	5 - 3 microns	Ge, ZnSe, FLIR ZnS, multi-spectral					80:50	40:20		< 8A
		LWIR	12 - 5 microns									
Aspheric Lenses	CNC, MRF	VUV	120 nm - 220 nm	CaF_2, FS					20:10	10:05		
		DUV	193 nm - 220 nm	FS								
		UV	220 nm - 400 nm									
		VISIBLE	400 nm - 700 nm	ALL					40:20	20:10		
		NIR	700 nm - 1.5 microns									< 40A
	SPDT	MWIR	5 - 3 microns	Ge, ZnSe, FLIR ZnS, multi-spectral ZnS, Si					80:50	40:20		
		LWIR	12 - 5 microns									
Mirrors	MRF, PL	VUV	120 nm - 220 nm	CaF_2, FS	10 mm - 250 mm	15 mm - infinity	Lambda/20 Surface Flatness	Lambda/40 Surface Flatness	10:05	05:01	10"	< 3A
		DUV	193 nm - 220 nm	FS								
		UV	220 nm - 400 nm									
		VISIBLE	400 nm - 700 nm	ALL					40:20	10:05		
		NIR	700 nm - 1.5 microns									
	SPDT, PL	MWIR	5 - 3 microns	Si, Ni-plated Al, Ni-plated Be, Al					80:50	40:20	30"	< 10A
	SPDT, PL	LWIR	12 - 5 microns									
Cylindrical Lenses	CNC, PL	VUV	120 nm - 220 nm	CaF_2, FS	10 mm - 200 mm	10 mm - infinity	Lambda/4	Lambda/10	40:20	20:10	TBA	< 5A
		DUV	193 nm - 220 nm									
		UV	220 nm - 400 nm									
		VISIBLE	400 nm - 700 nm	ALL					60:40			
		NIR	700 nm - 1.5 microns									
	SPDT	MWIR	5 - 3 microns	Ge, ZnSe, FLIR ZnS, multi-spectral ZnS, Si	12 mm - 250 mm	15 mm - infinity	Lambda/2	Lambda/4	80:50	40:20	30"	< 40A
	SPDT	LWIR	12 - 5 microns									

Table 12. MKS precision optics custom fabrication capabilities matrix.

Coatings, equipment, and processes are tailor-made to increase LIDT (see Chapter 1, Section I.A.7 for details) values for high fluence applications. MKS pioneered 193 nm and 157 nm coating development, having qualified the first 193 nm coatings over 15 years ago. This investment in developing and qualifying coating materials in the DUV has continued to achieve enhanced coating processes and coating equipment. MKS optical components support extended lifetimes in high power, excimer grade laser systems and are used by the top excimer laser manufacturers in the world. Typical coating types include:

- Anti-reflection coatings for lenses to reduce ghost images and increase transmission
- Filters: bandpass, notch, high pass, low pass
- Dielectric mirrors
- Metallic mirrors – porous or densified coatings
- Partial reflectors

Details regarding optical coatings can be found in Chapter 1, Section III.A.2.

In parallel with the development of coating processes and capabilities, optical cleanliness enables high LIDT, high reliability and lifetime, and prevents the risk of photo-contamination. Along with controlling the materials and processes during the fabrication process, MKS has developed proprietary cleaning processes that enable high reliability optical components and systems. All optical components are final cleaned, coated, tested, and packaged in particulate/molecular contamination-controlled cleanrooms with special secondary containment equipment to protect and prolong the cleanliness of the optical components. MKS' DUV protocol employs material restrictions, cleanroom environments, minimal handling, and specialized packaging to control photo-molecular contamination. All optics are packaged with proven materials and methodologies to ensure clean, high precision and long lifetime DUV optical components.

2. Opto-Mechanical and Electro-optical Assemblies and Systems

MKS specializes in optical sub-system design and manufacturing. Many photolithography applications require extremely high uniformity of the focused beam and very low image mapping distortion over large fields of view. Achieving these levels of performance starts in the design process and emphasizes concurrent engineering between the optical designers and the manufacturing engineers in the earliest stages of the design phase. Developing fully engineered optical solutions requires a thorough understanding of the opto-mechanical interactions between the optics and the structures in their intended final environment as well as process considerations in the manufacture and final alignment of the optical components. These high-performance optical systems often require very precise control of optical element positions, tilts, and air spaces, typically using advanced metrology methods and highly skilled technicians during assembly and integration in contamination-controlled environments.

While careful attention to stray light is generally important in optics, achieving very low surface roughness and the reduction of mid-spatial frequency surface errors reaches critical importance in the UV wavelength regime. All DUV systems are designed and qualified to ensure long lifetime and reliability. The design process includes the proper selection of materials for mechanical components, surface treatment, mechanical cleanliness, bonding methodologies, optical placement, purge flow, motion systems, electronics selection and placement, and chemical compatibilities. All materials and processes must consider the materials of manufacture and how they are processed, contaminant exposures during manufacturing, and contaminant cleaning processes. MKS has a proven and qualified supply chain with proprietary methods and requirements to ensure clean materials are utilized in the manufacturing of optics products.

As wavelengths get shorter and power increases, the list of usable materials for optical systems becomes shorter, necessitating the development of new techniques. For example, certain metals have inherent chemical properties that can induce photo-contamination effects, making them incompatible with DUV wavelengths. Surface treatments on some metals are available to ensure compatibility. MKS has fully qualified a set of materials that are compatible in the DUV.

When metals are machined, they are exposed to chemicals that help the machining process but are not compatible to DUV wavelengths since, trace levels of these chemicals can induce photo-contamination effects even after a thorough cleaning of the worked metal part. A qualified list of chemicals and cleaning processes for optics fabrication ensures that only metal parts that do not have a negative effect on lifetime and reliability are incorporated into DUV systems. In addition, MKS employs qualified bonding methodologies and assembly techniques that mitigate the risks of outgassing hydrocarbons and inorganic elements.

Cleanrooms are designed and manufactured to exact specifications for particulate control and molecular contamination. MKS cleanrooms are designed, assembled, tested and qualified to extreme levels of cleanliness for ambient air quality with organic/inorganic airborne molecular contamination controlled and monitored to ensure acceptable levels of cleanliness. All packaging materials and processes have been proven to preserve the cleanliness of the product as well as to protect it during shipment and storage.

3. Accurate Motion Stages

DUV lithography requires high velocity motion stages for wafer handling that have high accuracy and stability and fast step-and-settling times. Overlay (the relative position of one patterning layer to another), CD size, and throughput drive these requirements in the reticle and wafer stages, with typical overlay tolerances of 15% of the CD in 193 nm technologies. Throughput requirements (up to 200 wafers/hour) limit the maximum processing time to less than 20 s per wafer. This means that relatively high velocities and accelerations occur in the reticle and wafer translation operations. Motion control systems in these lithography tools must be able to achieve these velocities and accelerations with no impact on the vibration levels of the reticle or wafer, since this would impact the achievable CDs. Rapid step-and-settling requires active vibration isolation to minimize oscillation of the optics column and subsequent delay in illumination.

In addition to the higher velocities, throughput is also maximized by increasing the die size so that fewer dies are processed per wafer [271]. However, this approach increases the requirements for positioning accuracy. Lithography applications require motion stage calibration to ensure repeatability in the positioning of many different stages in the fab.

MKS Newport offers precision motion control products with a broad array of general duty and customized motion solutions applicable in wafer patterning and other applications. MKS has developed an extensive catalog of manual positioning and motion control standards and custom products for applications ranging from industrial to nanopositioning. For further information on motion control systems, see Chapter 1, Section III.F.2.

4. Vacuum Control

EUV places strict requirements on specifications for optical assemblies and vacuum control. EUV light sources require hard vacuum because all gases absorb 13.5 nm light. Process control equipment like MKS flow controllers, valves, and pressure gauges are used. Examples are included in [5].

5. Photolithography Applications Using MKS Products

MKS Precision Optics

MKS offers the highest quality optics, including lenses, mirrors, filters, beamsplitters, polarizers, optical systems, windows, and many other components and systems. These high-performance optic components cover UV, VIS, NIR, and IR wavelengths and are suitable for photolithography applications. For further information on optical components, see Chapter 1, Section III.A.

MKS Optical Sub-Systems

MKS specializes in optical sub-system design and manufacturing for excimer and EUV light sources. MKS has designed and manufactured optical assemblies for a diverse range of UV applications. Some examples of previous optical assemblies that have been designed and manufactured at MKS include:

- All refractive and reflective, catadioptric, and hybrid diffractive assemblies
- Objective lens assemblies
- Afocal telescopes, expanders, and collimators
- Anamorphic beam expanders

- Zoom systems
- Relay projection lenses
- Bright field and dark field illuminators
- Excimer laser pulse stretchers
- Low distortion scan lenses for ultrashort pulse lasers
- Refractive top hat beam shapers
- Variable optical attenuators
- Passively athermal systems
- Metrology test optics
- Fizeau interferometers for wafer inspection

MKS Vacuum Control

MKS vacuum solutions include pressure gauges, vacuum switches, mass flow controllers, mass spectrometers, valves, and heating jackets. The combined breadth of products enables exciting new product design synergies for OEM customers. Full details on this technology are available in [5].

Figure 177. Baratron® Capacitance Manometer.

MicroPirani™/Piezo gauges contain two gauges in one package: a MEMS MicroPirani sensor and a Piezo sensor. They are designed for load locks, measuring pressures that range from atmospheric pressure down to medium vacuum (1000 to 1×10^{-5} Torr). This wide range allows the gauge to be used in vacuum chamber applications requiring absolute vacuum/pressure switching capabilities. The full line of sensors and transducers based on MEMS MicroPirani™ technology is detailed on the MKS website for Granville-Phillips® indirect vacuum gauges.

Figure 178. 226A Differential Capacitance Manometer.

Baratron capcitance manometers (20 mT to 1000 T) provide accurate, repeatable measurements of pressure which are stable over long periods of time. They are an ideal choice for measuring pressure and vacuum in many semiconductor and critical thin film applications, including photolithography for semiconductor device manufacturing and optical coating. MKS Baratron® manometers are constructed from corrosion resistant materials and are insensitive to the types of aggressive process gases, such as halogens, typically used in semiconductor etch processes. Baratron® manometers for typical semiconductor applications have full scale pressure ranges from 20 mTorr up to 1000 Torr. Standard Baratron® manometers such as the one shown in Figure 177 measure absolute pressure and are widely used for stand-alone pressure sensing as well as being integrated into process control equipment. Baratron® manometers come in a variety of application-specific configurations, including both unheated and heated versions. MKS also supplies capacitance manometers that measure differential pressure (Figure 178), as well as ancillary equipment for the use of these manometers in process settings, including power supplies, readouts, and other accessories. MKS also provides calibration subsystems and services.

MKS Motion Control

MKS offers a series of high-performance air bearing stage solutions (Figure 179) suitable for use in semiconductor photolithography applications and customized tools for automated manufacturing and process control. These extremely rigid structures can accommodate wafers with diameters up to 300 mm. Very high accelerations (up to 5G in some models) and velocities (400 mm/sec to 1000 mm/sec) are achievable while simultaneously retaining high repeatability (25 – 50 nm) and accuracy (0.2 – 0.4 μm). MKS performs high-accuracy calibrations on all stages that achieve operational stability and high positioning accuracy and enable the same level of accuracy from tool to tool.

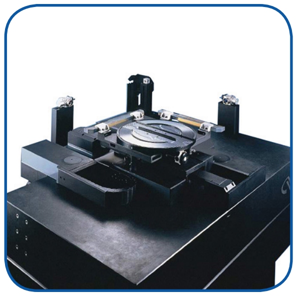


Figure 179. DynamYX 300 Wafer Positioning Stage.

The MKS Z Tip Tilt and theta stage provides high accuracy positioning for large dies and features and an autofocus capability to dynamically maintain the Z position of the wafer [271], enhancing the throughput of lithography systems. For example, the DynamYX® GT stage provides six degrees of freedom for wafer positioning (XYZ, tip, tilt, theta) with XY position accuracy of ± 150 nm over 300 mm. It has flatness of ± 200 μm over this range and step settle for 25 mm (± 50 nm) within 200 ms. The DynamYX® DATUM® GT stage supports 600 mm panel sizes for panel processing, increasing process throughput in these large-scale applications.

MKS Lasers and Light Sources

Many precision positioning applications employ lasers and other light sources for measurement and calibration purposes in the application. MKS Spectra-Physics® offers a broad selection of light sources suitable for these and other applications. For further information on light sources, see Chapter 1, Section I.

F. Future Directions for Photonics in Lithography

While EUV is the semiconductor industry's focus for faster and more energy efficient precision ICs [272], it is important to note that, in parallel, DUV remains vital for particular technologies and continues to undergo development [273]. DUV is expected to evolve to use higher power lasers and other general enhancements with the device focus on memory fabrication. Indeed, older technologies and processing tools predating DUV are also being used to produce lower cost ICs a few generations behind the cutting edge. These lower-tech devices are in demand for less rigorous applications, and markets continue to exist for manufacturing equipment that is being refurbished and for slightly improved technology used in applications requiring inexpensive production costs.

EUV lithography is proving extremely expensive owing to the complexity of the technique and the massive scale of the equipment [272]. From a systems engineering perspective, EUV illumination sources are some of the most complicated machines ever produced and the uptime requirements are unparalleled. In addition to design specifications that optimize X-ray control in vacuum, EUV optics and components have extremely strict and unforgiving handling and cleanliness requirements. In addition, EUV installations require a high degree of redundancy, serviceability, and modularity.

There are currently projections for tens of EUV systems being produced over the next few years. Efficiencies that help mitigate the cost for EUV systems are being vigorously sought out. Green technology is being proposed to alleviate the huge electricity requirements and there is a focus on recycling of resources like nitrogen and other gases. Efficiency will also be boosted by increased power and availability in CO_2 lasers used in EUV sources that can lead to decreased exposure times and improved wafer throughput. Note that advanced laser safety precautions will be required for lasers with powers in the multi-kW regime.

Reticle defects continue to be a top risk factor in the semiconductor industry since single defects can print on thousands of chips. This is driving a strong need for actinic reticle inspection. 13.5 nm light will be used to find more defects than optical inspection [274], however, the use of this wavelength is complicated by the presence of a pellicle, a thin membrane that protects the mask from particles. Pellicle materials are being developed that can withstand operational temperatures over 1000°C in anticipation of higher EUV photon flux. Higher optical power will increase throughput and reduce Line Edge Roughness (LER). LER represents degradation in the printed quality of lines due to shot noise that is present when low intensity EUV light interacts with photoresist [275]. The photoresist materials may also be redesigned in synergy with this effort.

As with DUV, 3D geometry impacts the patterning achievable using a reflective EUV reticle [276]. Reflective reticles are constructed as a single highly flat reflector with tantalum-based blocking material on the surface that has a finite height of around 60 nm. EUV systems illuminate the reticle at a 6° angle of incidence (chief-ray incidence angle), and the blocking material clips the light, impacting the reflected light anisotropically. This 3D mask effect causes shadowing. Future reticle designs must account for the resulting size-dependent focus and pattern placement shifts until thinner absorbers are developed and the optical path is re-designed with high NA optics.

The NA of an optic determines the solid angle over which it can collect light. In general, higher NA systems collect light from a wider range of angles and condense it down to resolve finer features. In the case of EUV lithography, high NA technology may enable sub-8 nm resolution by 2020 [277]. These systems have higher accuracy and G-force requirements for wafer placement and thus the motion control is more stringent and will continue to evolve. In addition, magnification of patterns in high NA systems are different in X and Y axis; because of divergence in the optical path. Thus, the print speed differs for each dimension, requiring advanced motion control programming capabilities.

Chapter 3

Photonics in Laser Machining

This page intentionally left blank.

I. Introduction

Drilling, sawing, milling, welding, and other mechanical techniques have been fundamental processes for manufacturing in a wide range of industries. These mechanical approaches, although effective in delivering high throughput at low cost, have certain limitations. As processes that require contact between the tool and the workpiece, they are dependent on the quality of the tool, which degrades with wear, resulting in inconsistent machining results and a need for frequent tool replacement. With these conventional techniques, fine features and complex geometries can be difficult or impossible to realize, depending on the required dimensions and shapes. When high machining quality is required, a secondary polishing process is commonly required to achieve the final desired quality. Additionally, certain materials, such as low melting point plastics and brittle or hard materials including glass, ceramics, and silicon carbide, can also be challenging to machine. Mechanical techniques also tend to generate significant acoustic noise and produce undesirable waste byproducts. For example, some techniques generate contaminated cooling waste water, which requires remediation and disposal.

An increasingly popular solution to overcome these limitations is to use laser machining. Chapter 1, Section I.A provides a detailed description of the operation of a laser and its unique characteristics as a light source. Table 13 shows a high-level comparison of laser and mechanical processes. Laser machining is a non-contact process that can achieve higher precision and quality, smaller features, improved consistency with no tool wear, and possesses the flexibility to machine different materials. Simultaneously achieving the necessary combination of machining quality, throughput, and cost with laser machining has been the key challenge in supplanting traditional mechanical processing techniques. Continuous improvements in laser technology, including semiconductor laser diode power and cost, have enabled dramatic advances in laser power, performance, and costs. As a result, laser machining is rapidly displacing conventional mechanical processes and enabling new processes not previously possible.

	Mechanical	Laser
Fast Throughput	✓✓	✓
High Machining Quality	✓	✓
Machining Fine Features	X	✓✓
Machining Challenging Materials e.g. low temperature plastics, glass, ceramics	X	✓✓
No Post Processing Required	X	✓
Consistency and High Yield (no tool wear)	X	✓✓
Noisy (sound)	X	✓
Environmentally Friendly	X	✓
Upfront Cost	✓✓	✓
Low Cost of Ownership	✓	✓

Table 13. High-level comparison of mechanical machining processes (e.g., drilling, sawing, milling, sandblasting) to laser-machining processes.

The benefits of laser machining enable a wide range of end user applications. In Figure 180, some of these key applications are listed based on their typical operational laser power and pulse width. Laser machining applications are typically grouped into two primary categories, microprocessing and macroprocessing. Microprocessing applications generally require lower output power levels (< few hundred W) and typically involve machining fine features (< tens of μm) on thinner parts (thicknesses < few mm). Macroprocessing applications require higher output power levels (> few hundred W) and typically machine larger features (> one mm) on thicker parts (> one mm). Macroprocessing relies on thermal mechanisms (heating, melting) to perform the machining while microprocessing utilizes a combination of thermal and non-thermal, e.g., instantaneous vaporization, mechanisms to machine finer features.

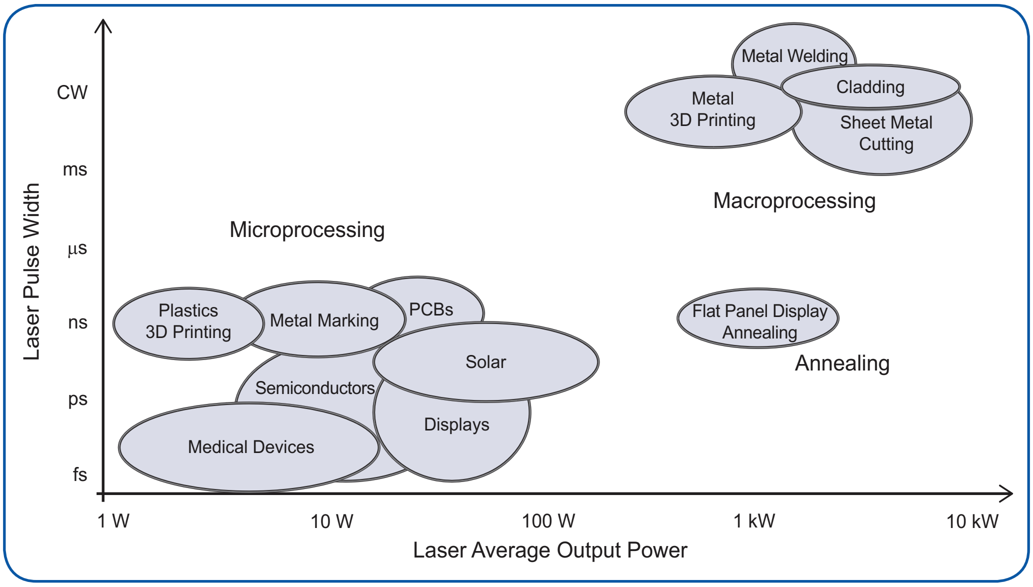


Figure 180. Laser machining applications along with their typical operational laser power and pulse width.

A third application category in Figure 180 is laser annealing, which is often utilized for processing flat panel displays. This process is used to crystallize amorphous silicon on glass or plastic sheets with the goal of enhancing the performance of thin film transistors in this multi-crystalline layer [278]. The laser processes used to crystallize silicon are highly specialized and rely on different laser output properties than those used in microprocessing and macroprocessing and will not be described here. This chapter first covers the applications, requirements, and examples of microprocessing before discussing the details of macroprocessing.

II. Microprocessing

Micromachining of fine features has become an integral part of high-volume manufacturing in markets as diverse as consumer electronics, medical devices, and automotive. Tiny holes, fine cuts, and narrow scribes are created with precision drills, saws, routers and, with increasing frequency, lasers. The smartphone, for instance, contains thousands of components, incorporates millions of drilled holes and precisely cut parts, and is manufactured in quantities of billions of units per year. Adding more functionality to smartphones has forced manufacturers to fabricate printed circuit boards (PCBs) with smaller, more densely-packed holes to accommodate the higher density wiring associated with an increasing number of smaller electronic components. New design features and slim profiles have also created a need for precisely rounded edges and closed features on the display. Similarly, medical devices such as stents, catheters, and other implantable devices are becoming smaller and more complex as functionality is added to these devices. Micromachining of intricate geometries, including small holes and accurate cuts in different materials, is required to manufacture these devices. For the automotive industry, the evolution of self-driven, smart electric vehicles has required microelectronics devices, precise sensors, and electric batteries to become integral components of vehicle performance and safety. Again, micromachining of precise, accurate and small features in different materials is required for this new generation of automotive components.

For consumer electronics, laser micromachining is widely used for high volume manufacturing to achieve the desired precision, quality, throughput, and cost per machined part. For example, in devices such as smartphones, tablets or wearables, laser micromachining is used in the fabrication of the semiconductor chips and their packaging, electronic components, and the circuitry connecting these components in the form of PCBs. The manufacture of high-resolution touchscreen displays also relies on many laser micromachining processes. The camera windows, sensors, and even the external enclosure utilize laser micromachining to achieve the required machining results. Figure 181 depicts the range of laser micromachining processes used in the manufacturing of these devices.

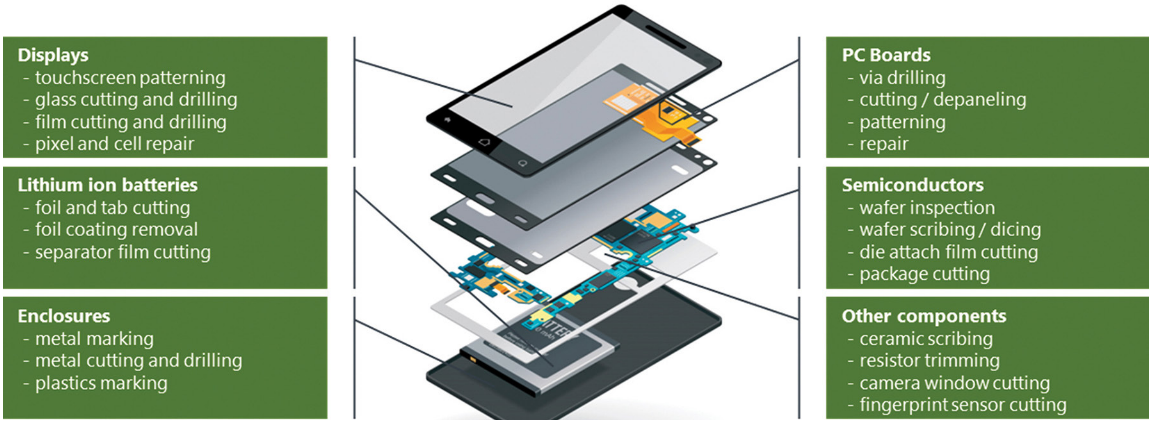


Figure 181. Laser micromachining processes in mobile device manufacturing.

Clean energy is another important application for laser micromachining. There are three main areas of focus for investment and innovation to meet future energy demands while addressing concerns about pollution and climate change: renewable energy generation, improving energy efficiency, and energy storage. Figure 182 outlines some of the laser micromachining processes exploited for clean energy. The manufacturing of photovoltaics for solar power utilizes laser micromachining to pattern large areas with fast throughput while realizing fine features. Lithium ion batteries used for energy storage are also manufactured using laser micromachining in combination with mechanical techniques. For energy efficiency, laser micromachining is used for the manufacturing of high efficiency LEDs for lighting and is being evaluated for manufacturing lightweight materials such as carbon fiber reinforced plastics used in automotive and aerospace applications.

Figure 182. Laser micromachining processes used in clean energy applications.

Implantable medical devices such as stents, intraocular lenses, prosthetics and catheters are becoming increasingly complex. The feature sizes of these devices are shrinking to both address new usages and improve patient outcomes. In the case of metal stents, including drug-eluting metal stents, shrinking their dimensions allows for deployment in smaller coronary, peripheral and neurovascular blood vessels – important areas being targeted for treatment. Some studies have suggested a correlation between clinical outcomes and the amount of metal deployed within the vessel. Stent makers have thus been motivated to work with thinner-walled and smaller-diameter tubes with more complex stent features. A recent report forecasted that the next generation of coronary stents will have features – including struts and links – that are nearly half the size of current stent features [279]. Another trend is to add a controlled surface texture or geometry to stents and prosthetics to improve bio-compatibility, for example, to reduce the risk of restenosis (the recurrence of arterial narrowing after treatment). Feature size reduction and the incorporation of new materials are proving challenging for the fabrication process. Not only does the quality of the cut edge become more critical as stent strut dimensions shrink, it becomes more difficult to achieve. For smaller stents it is difficult to remove damaged material by post processing since damage from heating, or the heat-affected zone (HAZ), starts to become a limiting factor [280]. Fortunately, advances in laser micromachining are making strides towards addressing these issues.

A. Laser and Photonics Requirements and Implications

One of the challenges for laser micromachining is removing only the desired material, usually through localized heating, while at the same time minimizing the extent of the HAZ to any of the remaining material. Delivering laser irradiation with near-perfect beam quality precisely to the target region is a necessary step to achieving this desired result. Shorter wavelengths and shorter pulse widths are advantageous to achieving higher-quality results. Wavelengths in the VIS and UV spectral regions, for example, are more readily absorbed in most materials, resulting in shallow absorption depths and a significantly reduced HAZ. UV lasers also can be focused into tighter spots for smaller, more precisely machined features, and they have longer depth-of-focus for improved process yields. Ultrashort pulse widths in the ps and fs regime yield intense peak powers that result in nonlinear absorption at the sample for instantaneous material vaporization, very minimal heat transfer into the material, and a negligible HAZ [281, 282]. This process, commonly referred to as a “cold ablation” process, is depicted in Figure 183.

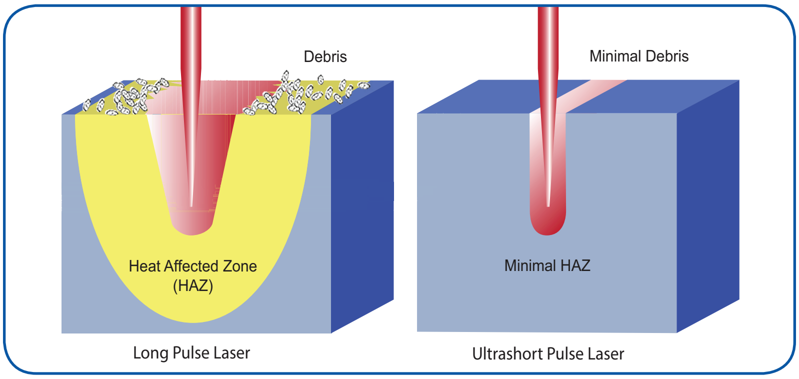


Figure 183. The impact of laser pulse width on machining quality for a long pulse ms laser (left) versus an ultrashort pulse laser (right) [283].

A second challenge for laser micromachining is achieving high machining throughput. Increased average output power can translate into higher ablation rates but with certain limitations. Laser fluences (or energy densities, typically given in J/cm^2) outside an optimal range result in decreased material removal efficiency. Excess fluence is partly deposited as heat into the material, causing a reduction in throughput and quality, while insufficient fluence results in reduced ablation rates. This is particularly true for ultrashort pulses where delivery of the optimal fluence is required to stay within the desired operating regime for cold ablation. Thus, to achieve high average output powers required for increased throughput while maintaining optimal fluences for machining quality, high repetition rates with sufficiently high pulse energies are necessary. The relationships between average power, fluence, repetition rate, pulse energy, and other radiometric quantities are described in Chapter 1, Section II.A.1. Higher repetition rates also tend to improve the quality and efficiency of machining. However, other components in the machine tool can place limits on usable laser repetition rates. Faster scanners and improved motion control techniques are necessary to take advantage of the higher repetition rate. Examples of MKS Newport's advanced motion control solutions are discussed below. In some cases, spatial beam splitting and shaping techniques can be utilized to maintain optimal fluence without scaling up repetition rate. However, these techniques require additional system complexity.

One useful approach to improving both throughput and quality is tailoring of the pulse sequence with pulse bursts and pulse shapes [284]. With this approach, the temporal profile of the laser energy deposition can be tailored and optimized for a particular material. In this manner, the incident energy can be applied almost entirely to material removal rather than excess heating. As such, pulse tailoring is an additional "knob" for process optimization and can result in large gains in process throughput and quality [285, 286]. The Spectra-Physics IceFyre® industrial ps laser uses a hybrid fiber laser architecture that combines a flexible fiber laser front-end with a DPSS power amplifier to deliver high power, ultrashort pulses with programmable pulse capability as shown in Figure 184.

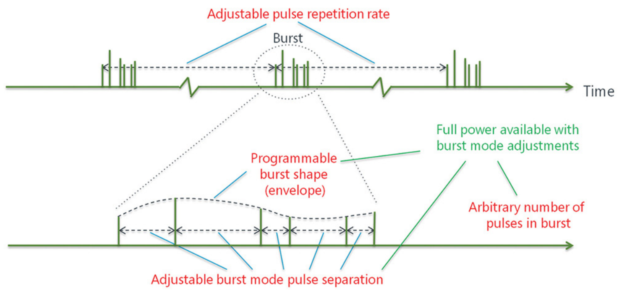


Figure 184. Diagram of the programmable pulse shaping capability (top) for a Spectra-Physics IceFyre® industrial ps laser (bottom).

The spatial shape of the laser beam delivered to the workpiece directly affects the spatial profile of the material removed by each individual pulse. Both the laser beam profile and the performance of the beam delivery optics are of critical importance for laser micromachining applications. It is necessary for the laser output beam profile to be a single spatial mode, i.e., TEM_{00}, with a near-Gaussian profile in both axes, a circular shape, and a low astigmatism to realize a tight, round focal spot at the sample. Details regarding the parameters used to quantify the shape and quality of a laser's spatial beam profile are given in Chapter 1, Section II.B.1. For the beam delivery, optics are used to transport the beam to the target position at the sample with the desired beam shape and focus. However, each optical element between the laser and the target workpiece will incrementally distort the beam quality. Therefore, both the number and quality of each optical element should be optimized. For example, optics with poor surface quality would distort the beam shape, and optical elements with an insufficient clear aperture would crop the beam and introduce undesirable diffractive effects. Distortions of the beam shape can result in undesirable machining quality and unintended feature geometries.

Beam delivery optics can significantly affect other characteristics of the delivered laser beam as well. In particular, each optical element will incrementally introduce optical loss (reducing power at the sample) and dispersion (modifying the delivered pulse width). For ultrafast fs pulses, minimizing undesired dispersion is critical to prevent pulse broadening. Optics specifically designed for handling ultrafast pulses by introducing minimal dispersion, such as Newport ultrafast mirrors and optics, as shown in Figure 185, should be used to ensure proper beam delivery of fs pulses.

Figure 185. Newport ultrafast optics, designed to deliver fs pulses from the laser to the workpiece.

A key consideration is how to maintain stability of the laser beam at the sample over a range of timescales and environmental conditions since instability can introduce errors and yield loss in the laser micromachining process. Vibration control systems including passive isolation tables are necessary to minimize instability of the beam delivery optical system over short timescales (see Chapter 1, Section III.G.2 for details). For longer timescales up to many hours, thermal changes due to fluctuations in the environment or changes in the process can cause drift in the laser beam delivery to the workpiece. Due to the number of optical elements in a typical beam delivery system, optical mounts can particularly impact the thermal stability. MKS has developed opto-mechanical components such as the Newport Suprema® ZeroDrift™ thermally compensated mirror mounts, as shown in Figure 186, to minimize temperature-induced alignment drift. For long-term stability of the system, reliability of the beam delivery system is required. Minimizing residual stress in the optics and optical mounts as well as minimizing potential sources of contamination that degrade optical performance (photo-contamination) are key factors in maintaining long-term stability. Photo-contamination can limit long-term reliability in the UV, and utilizing optics optimized for handling UV wavelengths, along with clean components such as MKS Newport's vacuum-compatible optomechanics, can deliver favorable results. Active beam pointing and position control, as provided by Newport Picomotor™ piezo mirror mounts (see Figure 186), can also be used to compensate for thermal and long-term drift. As for the stability of the laser source itself, MKS Spectra-Physics lasers are developed and manufactured to address all the above considerations by providing exceptional stability and reliability, including small pulse-to-pulse fluctuations, low thermal dependence, environmental robustness, and long-term reliability.

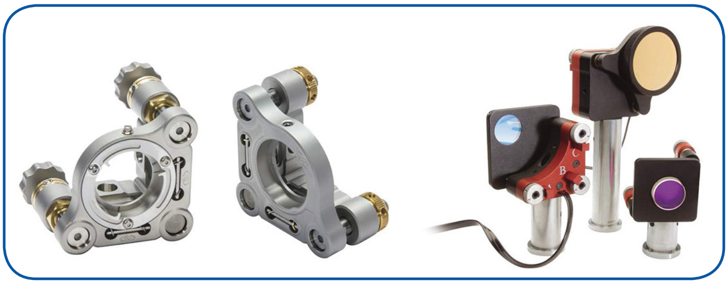


Figure 186. Suprema® ZeroDrift™ mirror mounts (left) are resistant to temperature-induced alignment drift. Picomotor™ piezo mirror mounts (right) are used for active beam pointing and position control.

To monitor the laser power and beam quality delivered to the sample and ultimately to the location of the laser micromachining process, laser power meters and beam profilers are often used as diagnostic tools either off-line or in-line. MKS Ophir beam measurement and analysis tools are the market leading solution for such diagnostics, and examples are shown in Figure 187. Detailed descriptions of the operation of power and energy meters and the wide variety of products provided by MKS Ophir are given in Chapter 1, Section II.A. Furthermore, Chapter 1, Section II.B discusses the key attributes of beam profiling tools as well as specific systems offered by MKS Ophir.

Figure 187. Ophir power meters and beam profilers are used for laser micromachining system process monitoring.

Figure 188. Newport DynamYX® air-bearing stages precisely move the workpiece at high speeds.

In order to laser micromachine shapes and features on a workpiece, precision motion of either the laser beam and/or the sample are required. Galvanometers or galvo scanners, in combination with a large area focusing lens, called an f-theta lens, are commonly used to move the laser beam in a programmed pattern. Due to the very light weight of the galvo mirrors, these devices can precisely scan the beam relative to the workpiece at rapid speeds (several m/s) in arbitrary shapes, including around tight corners. However, galvo scanners are limited in their field-of-view (FOV, the area over which the beam can be scanned) to a few hundred mm. In addition, the focal spot size is limited to tens of µm. Furthermore, since the FOV and focal spot size limitations are inversely correlated, a tighter focal spot size results in a smaller FOV. Movement of the workpiece itself using precision motion stages can help circumvent some of these issues. This approach allows one to use a fixed focusing objective to achieve tight focal spot sizes down to several µm without making sacrifices in the area to be fabricated. However, due to the inertia associated with the larger mass of the stage, this approach has some limitations on acceleration both for linear travel and around corners. MKS Newport motion solutions provide optimized performance for precision micromachining. A high-performance example is the Newport DynamYX® air-bearing stage (shown in Figure 188), that can achieve tens of nanometers precision with up to 5G's of acceleration and can reach speeds over 2 m/s. A third approach to expand laser machining capabilities is to combine galvo scanners with motion stages using a step-and-scan approach. In this manner, motion stages provide the step-wise motion necessary to stitch together separate galvo FOVs to create a larger effective FOV than what can be achieved by the galvo itself.

Finally, cost is a key criterion for micromachining. The cost differential between laser processing and mechanical processing for each manufactured part may ultimately be the most important figure. This metric includes amortization of the upfront laser and system cost, cost of operation, lost productivity from downtime, and process yield. Thus, in addition to the upfront cost, laser system reliability is of critical importance, as are serviceability and cost of service. Process yield losses also directly contribute to the operating cost associated with the system. These costs can be minimized if the system has high process stability over time and if the system provides consistent reliable performance.

B. Microprocessing Applications Using MKS Products

1. PCBs and Electronic Components

One large and growing market for laser processing is PCB manufacturing. The market has traditionally used CO_2 and excimer lasers but is increasingly moving towards UV DPSS lasers. The migration to UV DPSS laser technology is driven by several factors, including the need to machine smaller features with higher precision and density and the ability of UV wavelengths to process many types of materials with good quality. Furthermore, UV DPSS products available on the marketplace have experienced reduced cost and cost of ownership.

The Spectra-Physics Talon® UV product family (as shown in Figure 189) consists of industrial UV lasers with ns pulse widths. The Talon product family is highly versatile with a wide range of configurations in terms of power, energy, and pulse repetition frequency (PRF). For example, the 15 W UV Talon offers a high pulse energy of 300 µJ at a 50 kHz PRF for larger feature machining in thicker materials such as FR4-based rigid PCBs. On the other hand, the 20 W UV Talon has lower maximum pulse energy but maintains an elevated output power level at very high PRFs. These parameters are ideal for processing thinner materials with a tight beam focus and can enable high-speed, high-resolution machining.

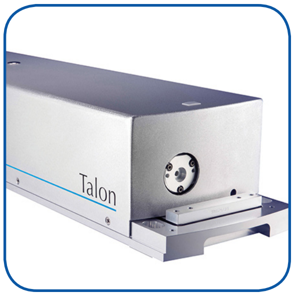


Figure 189. Spectra-Physics Talon® UV ns Laser.

As electronic devices shrink and improve in performance, the need for compact and thin flex-PCBs to which the electronic devices are mounted is growing rapidly. Flex-PCBs are typically made using layers of materials that have several to tens of µm thicknesses. Materials used include copper (Cu) foils, polyimide (PI) sheets, and adhesives that enable the creation of various laminates. A typical flex-PCB material is comprised of a 12 µm thick PI layer laminated between two Cu foils of similar thickness. In flex-PCB fabrication, laser processes may include blind- and thru-via drilling, straight-line and contoured cutting, as well as 2D patterning. Given the high power and high PRFs available from the Talon UV laser, MKS has been able to optimize various high-speed processes for use with flex PCB materials.

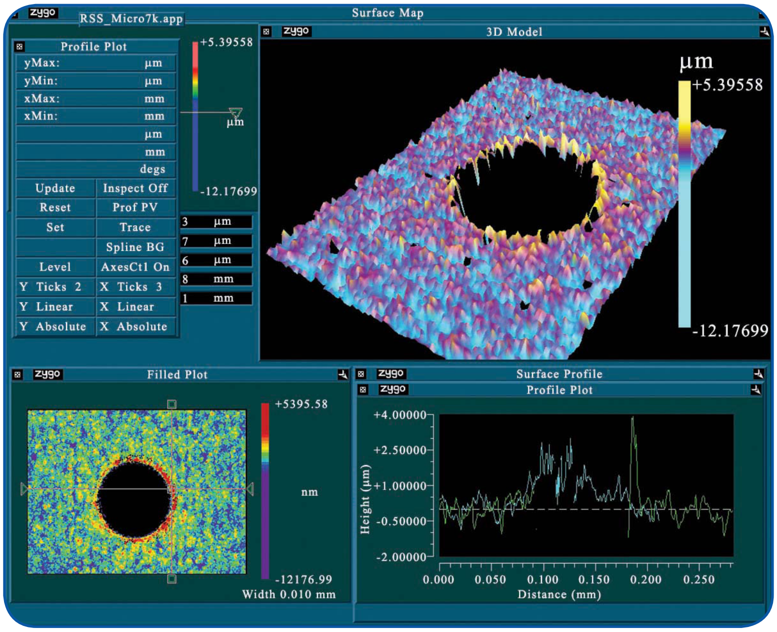


Figure 190. Profile of a via hole drilled in flex-PCB using a Spectra-Physics Talon UV ns laser. The flex-PCB layer is a Cu/PI/Cu laminate. The via hole was trepan cut with an 80 µm diameter. The resulting average burr height was < 2 µm.

It is possible to achieve high quality cutting and drilling performance at high speeds by optimizing the scanning speed and PRF to achieve the proper spot overlap in the material. The Talon laser's high average power at high PRFs provides sufficient pulse energy to ablate Cu material up to several hundred kHz. Cu/PI/Cu laminates can be cut through with the Talon 355-20 laser at speeds exceeding 450 mm/sec while operating at 500 kHz. As shown in Figure 190, exceptional cutting quality is achieved with minimal burring and a small HAZ [287]. The Talon UV laser has also been tested for high-speed percussion drilling of blind vias in a similar Cu/PI/Cu laminate material. In this case, the middle PI layer was twice as thick, at 25 µm.

High-quality vias with sub-30 μm diameter openings were processed at very high PRFs to minimize the drilling time. With the Talon laser operating at 300 kHz, just twenty pulses were required for drilling each hole, which equates to a throughput of 15,000 holes per second. Thin PI films are also used extensively in flex-PCB manufacturing as coverlay materials to protect circuitry from harsh environments, similar to solder masks in thick PCB manufacturing. UV wavelengths are very effective for processing PI with high quality and precision due to the strong optical absorption by the material and subsequent photo-ablation. With commonly-used thicknesses of 12-25 μm, optimal cutting of these materials is achieved with a high PRF, lower energy laser source. The Talon laser is ideal for such requirements and a wide range of cutting speeds are possible with the various product offerings.

2. Flat Panel Displays

As the flat panel display market moves from liquid crystal displays (LCDs) to organic light emitting diode (OLED) display technology, laser patterning and cutting of heterogeneous materials (comprised of multiple organic material films) with high quality and accuracy is needed. Within OLED technology, substrate materials are moving from glass to flexible plastics to create lighter, thinner and more durable displays. Flexible OLED structures consist of a multilayer stack of functional materials deposited on heat sensitive plastic films such as PI and polyethylene terephthalate (PET) as shown in Figure 191. The sensitive thermal and optical properties of these films make these stacks difficult to machine using conventional laser sources.

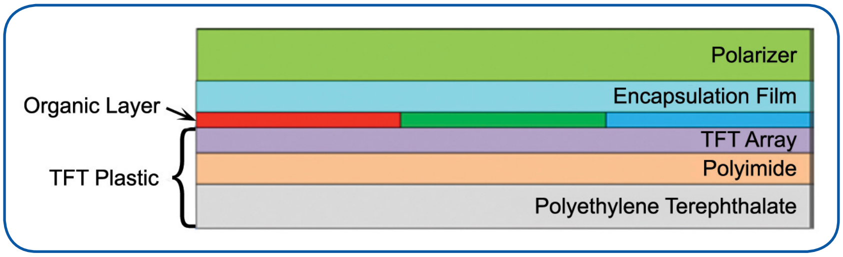


Figure 191. Schematic representation of a flexible OLED structure. TFT stands for thin-film transistor.

Fs lasers are capable of machining plastic materials with a minimal HAZ and precise control of the material removal. Although the processing quality achieved meets industrial demands, processing speeds need to be improved to satisfy the industrial user. To process parts quickly and cost-effectively, a fs laser system with high average power is required. Additionally, the laser system must be robust and stable to sustain the demands of the production floor.

The Spirit® 1030-100 laser from MKS Spectra-Physics (shown in Figure 192) is a high-power industrial fs laser used in high-precision industrial manufacturing. This laser offers impressive versatility and performance, enabling a variety of applications. High average power (> 100 W) and high pulse energy (> 100 μJ) at a wavelength of 1030 nm, combined with a high PRF (up to 10 MHz) and short pulse duration (< 400 fs), pushes fs micromachining applications to desirable throughput levels with low cost-of-ownership. The user-configurable burst mode further enables processing with increased ablation efficiency and thus increases throughput and quality for certain materials. Additionally, the integrated SHG module (see Chapter 1, Section I.A.6 for details on harmonic generation) offers an output power of > 50 W at a wavelength of 515 nm, which is useful in machining stacks of materials with different optical properties. The Spirit 1030-100-SHG laser system operating at wavelengths of 515 and 1030 nm has been used for cutting 75 μm thick ribbons of PI and PET plastics, typically used for flexible

Figure 192. Spirit® 1030-100 high power industrial fs laser.

OLED displays. Figure 193 shows microscope images of PI and PET plastics machined using the Spirit 1030-100-SHG laser. Cutting speeds exceeding 1 m/s for both plastics were achieved when operating the laser at 100 W average power at 1030 nm. It was also possible to achieve high machining quality with minimal thermal damage (HAZ < 50 μm). The demonstrated cutting speed and quality meets the requirements of OLED display manufacturers.

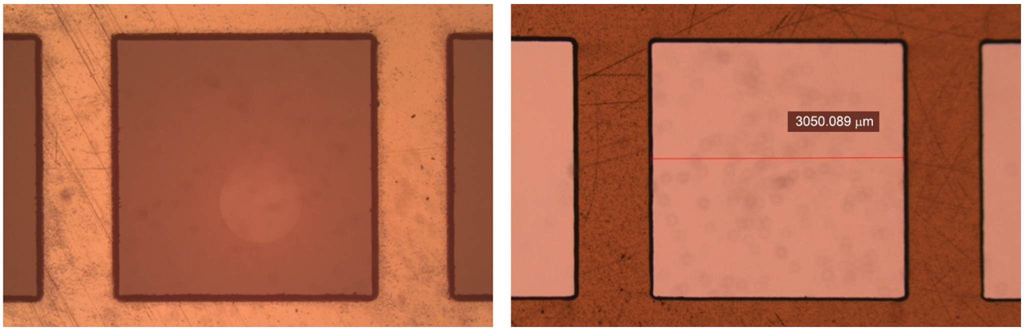


Figure 193. PET (left) and PI (right) polymer 75 um thick films cut with a Spectra-Physics Spirit® 100 W IR fs laser. Cutting speed was > 1 m/s for PET and 650 mm/s for PI. Results show minimal HAZ.

3. Solar Panels

Crystalline-Silicon (c-Si) solar cells are the predominant solar cell technology used in fabricating solar panels because of the availability of silicon, well-developed manufacturing processes, and the high conversion efficiencies that can be achieved. Passivated Emitter Rear Contact (PERC) technology has been shown to generate absolute cell efficiency gains of a percentage point or more compared to conventional cells. In a conventional solar cell, there is an aluminum metallization layer that makes contact across the full area of the back of the cell. PERC technology involves creating a dielectric passivation layer on the rear side of the cell with openings to allow electrical contact to the metallization layer. PERC cells achieve higher efficiencies for a few reasons. First, the passivation layer significantly reduces electron recombination near the back of the cell, where the electrons would otherwise experience a strong attraction to the aluminum metallization layer. This allows more electrons to reach the front-surface emitter, thereby increasing current density in the cell structure. Secondly, the passivation layer enhances the cell's ability to capture light, particularly at longer wavelengths, because it causes multiple reflections, thereby enhancing the probability of absorption. In this way, the absorbing length of the cell is effectively doubled and current density is further improved. Finally, the passivation layer reduces heating of the backside metallization layer by reflecting IR light out of the cell. IR light that is not absorbed by the silicon ($\lambda > 1180$ nm) would otherwise be absorbed by the aluminum and increase the heat in the cell. Cells are more efficient when operating at a lower temperature.

Lasers are used for creating openings in the dielectric passivation layer for ohmic contact. Aluminum paste is screen printed to this surface and a subsequent thermal annealing process alloys the aluminum with the laser-exposed silicon to form a good ohmic contact. A typical 6 inch PERC cell will have about 155 laser-scribed lines which are longer than 156 mm, 40-50 μm wide, and are separated by 1 mm. The aggregate length of the PERC scribes on a single wafer is approximately 25 m. The target processing time can be as fast as one second per wafer, which equates to a 25 m/s required scribing speed. Fast two-axis galvo scanners as well as spinning polygon scanners can achieve such speeds. MKS Spectra-Physics has demonstrated high-quality solar cell scribing for PERC processing as shown in Figure 194.

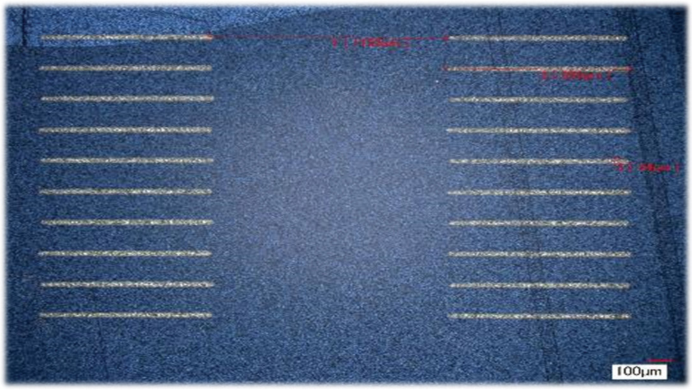


Figure 194. Multi-crystalline silicon solar cell scribed with a Spectra-Physics laser for PERC processing.

Industrial PERC processing also requires a laser that can keep up with the high scan speeds demanded. For this, MKS has tested the Spectra-Physics Quasar® 532-75 high-power hybrid-fiber laser with TimeShift™ programmable pulse technology as shown in Figure 195. The laser offers both high power and high PRF along with short pulse widths below 5 ns. For thin film removal, shorter pulse durations have been shown to be more energy efficient compared to longer pulse durations. With the Quasar laser, the shortest pulse durations are generated at the highest operating frequencies, which is ideal for PERC processing. Less heating and reduced risk of thermal damage to the underlying c-Si lattice are additional benefits of short-pulse processing. Further throughput improvement is also possible via beam-splitting. When a Quasar laser is operated at 850 kHz PRF, corresponding to one wafer per second (WPS) throughput, the available pulse energy is approximately three times more than necessary. Therefore, a three-beam split can effectively triple the throughput of a single laser thus allowing for throughputs of three WPS. Irrespective of the final system configuration, the Quasar laser platform and its unique TimeShift™ pulse-tailoring technology has the power, speed, and flexibility to meet the demands of industrial PERC solar cell manufacturing.

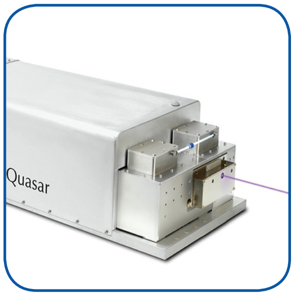


Figure 195. Spectra-Physics Quasar® high power pulsed ns UV and green hybrid fiber laser.

C. Future Directions

Laser micromachining has rapidly become a key manufacturing process for a wide range of applications. Micromachining laser technology is trending to higher laser powers for higher throughput, to UV and shorter pulse widths for improved machining quality and higher precision, and to lower cost and cost-of-ownership for achieving a reduced cost per manufactured part. Laser micromachining has the potential to continue supplanting existing mechanical processes while providing improved machining quality at lower costs, thereby enabling new applications not previously possible.

One example of a potential future application of laser micromachining is to create superhydrophobic and superhydrophilic surfaces. Superhydrophobic surfaces have water droplet contact angles exceeding 150 degrees, making them extremely difficult to wet, while superhydrophilic surfaces exhibit the opposite behavior. In recent years, it has been shown that both hydrophobic and hydrophilic surfaces can be fabricated by fs laser machining. The extreme water repellency of the lotus flower has served as inspiration for a variety of biomimetic designs where hydrophobic surfaces are needed. Figure 196 compares images of a rose petal surface (similar to the lotus flower surface) and a superhydrophobic surface machined on polypropylene using a fs laser from MKS Spectra-Physics. The imitation of the surface morphology by laser machining leads to a one-to-one reproduction of the wetting properties of rose petals. Furthermore, it is possible to create regions with varying functionality across the surface using lasers. Both superhydrophobic and superhydrophilic surfaces can be created across the surface in micropatterns. Such micropatterns offer exciting possibilities for the design of biomedical and microfluidic devices. As the throughput and cost of laser processes continue to improve, surface functionalization could potentially be applied to a wide range of applications. One such application involves maintaining clean surfaces for automotive parts, solar panels, and even architectural structures by reducing the ability of substances to adhere to these surfaces.

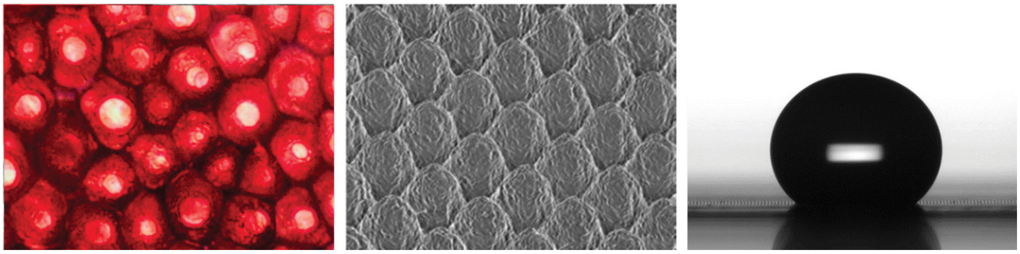


Figure 196. Optical image of a rose petal surface (left) and scanning electron microscopy image of a superhydrophobic surface on polypropylene machined using an MKS Spectra-Physics fs laser (middle). Water droplets placed onto this machined surface (right) are strongly repelled with a contact angle of over 150 degrees.

III. Macroprocessing

Laser macroprocessing is defined as the machining of metal parts with thicknesses greater than a mm to produce large features with multi-mm dimensions. The process uses high power lasers with typical average powers ranging from a few hundred W to many kW. Key applications in macroprocessing are metal cutting, metal welding, and additive manufacturing of metals.

The laser cutting industry is expected to grow significantly over the next few years. Much of the growth is driven by the fact that the application space is constantly increasing. For example, the automotive industry is producing laser-cut and laser-welded car body parts, and the construction industry is using lasers to cut construction components. The plastics industry uses lasers to cut textiles and packaging materials. The biggest advantage to laser use is that any arbitrarily-shaped cut can be achieved, including 3D shapes, without the typical limitations of mechanical tools. Unlike mechanical processes, producing well-defined cutting edges is easily achievable because the laser process is non-contact and therefore wear-free. Hence, even the hardest or most abrasive materials can be processed without the need for tool replacements. Because the laser process is non-contact, no force or mechanical stress is applied to the processed part. This is especially important for the processing of brittle or soft materials as well as in high-speed cutting applications where material movement exceeds 100 m/s, e.g. when cutting paper. Additionally, laser systems reduce production time and tool costs because there is minimal set-up time and no need to produce individual tools when punching metal sheets.

In the metal cutting arena, lasers compete with plasma- and water jet cutting. However, the edge quality and ability to control power in lasers is superior to that of plasma cutters. Water jet has the best edge quality on very thick metal sheets and takes the lead when material thicknesses exceed 20 mm. With thinner materials, the greater flexibility of the laser is exploited. In the past, the CO_2 laser was the workhorse for metal cutting. This was primarily due to its price advantage compared to solid-state lasers, e.g., the Nd:YAG laser, despite the fact that the latter delivers a better material absorbance. However, this has changed with the introduction of cost-effective disc and fiber lasers, causing CO_2 lasers to be supplanted in many applications. Nonetheless, for organic materials like plastics or wood, there is no alternative to the CO_2 laser because of the strong absorption at the FIR operational wavelength of 10.6 μm.

Laser welding is ubiquitous and has become the preferred technology in many industries over the past few years. It can be used to interconnect a wide variety of organic and inorganic materials. Production sectors dealing with mobility in the broadest sense, such as the automotive, shipbuilding, and aerospace industries are increasingly replacing bolted assemblies with welds. Laser welds provide permanent connections that save weight and reduce risks associated with nuts and bolts that can loosen or break over time. Welding is also commonly used in other applications including gas-tight welding of heart pacemakers, welding fine jewelry, and welding stainless steel in heat exchangers for white goods/appliances or in heater/cooler systems. Also, in contrast to traditional welding processes relying on electrical discharges, lasers can produce a minimal HAZ because it is possible to control the laser beam more precisely.

In laser metal welding, three major processes are currently in use. One involves conventional welding optics which have a rather short focal length, e.g., 100 – 200 mm. With the introduction of a filling wire, the joined parts are melted at a joint to allow both metals to mingle before cooling and becoming one solid part. It is important that there are no gaps in the joint between the parts to ensure the weld is effective. When using a filling wire, there are three components interacting. The two loose parts and the filling wire are melted in the focal spot and joined. The wire interacts with each of the materials and helps facilitate the connection between them. Another welding process, known as remote welding, uses a three-axis galvanometric scanning system with a long focal length, e.g., 0.8 – 1.5 m, that can be located far away from any obstructing parts. This is important for welding large parts like door panels where clamping fixtures are needed to ensure proper alignment but tend to interfere with the motion system used to position the laser. Such scanning systems allow for fast beam steering and generate many welding dots over a long distance in short time. Remote welding allows manufacturers to save valuable production time and achieve a higher throughput. Fiber lasers and disc lasers are typically used for laser welding and remote welding applications, while CO_2 laser welding is preferred for certain applications because of the special weld seam characteristics provided by these lasers.

The third welding process, the welding of plastics, has different laser requirements. Diode lasers and CO_2 lasers dominate this application process. CO_2 laser wavelengths are absorbed by any plastic material regardless of whether the plastic is transparent or not. Diode laser wavelengths are only absorbed by colored plastics. This difference in absorbance leads to a welding strategy where one transparent part and one colored part can be joined by steering the laser beam through the transparent material onto the colored plastic to melt the colored plastic and join them together. One major use of laser welding in the automotive industry is for welding headlight and taillight assemblies. Laser welded housings for electronic components are also found in many applications.

Additive manufacturing is often a confusing topic because different technologies are combined in this application. The processes can involve melting materials and delivering them through a nozzle and depositing the melt layer by layer. This process is typically referred to as 3D printing. Processes that use a laser to melt or fuse powder are typically referred to as selective laser melting or laser sintering. Laser sintering can be used with different materials such as sand, polymers, or metal powder. Sintering sand is used for mold making in mold casting processes. Powder polymers are typically used for either mold making or to create components in rapid prototyping applications. Liquid polymers are used in laser stereolithography with a UV laser curing selected areas of the plastic to build a solid part. Model building for technical design and medical device applications are the main areas of interest for laser stereolithography processes. Metal powders were originally used with additive manufacturing processes to manufacture tools for molds in injection molding applications. However, metal powders are now also used in direct manufacturing of many functional parts, including parts for vehicles, bionic designs, and medical implants.

Today, large additive metal manufacturing machines can use up to four lasers at a time to produce either one large part or to produce multiple parts in parallel. While initially used to produce prototypes, additive manufacturing is increasingly being used by companies to efficiently produce complicated 3D structures. The benefit of this approach over traditional metal manufacturing techniques that selectively remove metal to produce a structure is the significant reduction in waste material.

A. Laser and Photonics Requirements and Implications

For macroprocessing applications, high power IR lasers with kW output powers are needed to perform metal cutting, welding and additive manufacturing processes. Most of these lasers are *CW* lasers, i.e., not pulsed, while some are quasi-*CW* where the laser is pulsed in ms timescales to increase the peak output power for a given average power. One particularly important parameter in these applications is the brightness of the laser beam as represented by its BPP (beam parameter product). BPP is the product of the beam diameter in mm and beam divergence in mrad (see Chapter 1, Section II.B.1 for details). Higher brightness levels (or lower BPP) are necessary for metal cutting applications, while metal welding applications require less brightness (or higher BPP). Figure 197 maps laser power versus BPP for various macroprocessing applications.

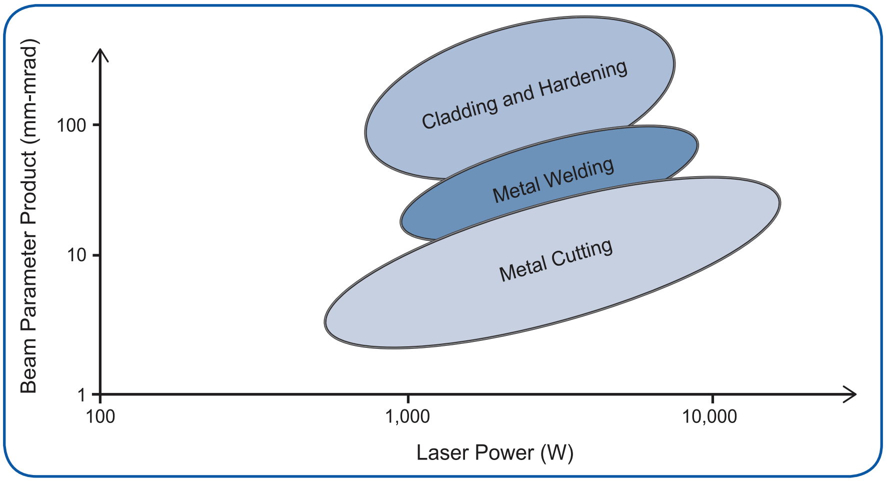


Figure 197. Macroprocessing applications mapped on beam-parameter-product vs. laser power [288].

The beam delivery section of a laser machining system transfers the laser beam from the laser cavity to the workpiece. For CO_2 laser-based 2D-machines, at least two moveable mirrors are required for guiding the laser beam to any point on the workpiece. In modern machines, and especially in 3D machines, the beam delivery section requires additional mirrors due to the more complicated design of the structure being manufactured and to enable other features incorporated in the system. The cutting head of a CO_2 laser-based 2D-machine includes a ZnSe focusing lens, which focuses the laser beam on the workpiece. For fiber-laser-based machines, the fiber core itself guides the beam to the cutting head, which includes a collimator and a focusing lens.

B. Macroprocessing Applications Using MKS Products

1. Optics for CO_2 Laser Machines

Carbon dioxide molecules are excited by a gas discharge in the cavity of a CO_2 laser. Mirrors are placed at both ends of the discharge tube such that the laser beam is reflected many times to build up the laser beam intensity. Each mirror has some transmittance: the output coupler transmits the usable laser beam while the rear mirror transmits a very small portion of the beam for power measurements and beam diagnostics. In order to gain sufficient intensity to produce several kW of output laser power, the total length of the laser cavity needs to be several meters. Covering this distance with one discharge tube is very problematic and impractical. Therefore, the laser cavity is split up into several discharge tubes working in series. To reduce the mechanical footprint as much as possible, the optical axis of the laser beam within the cavity is "folded" several times by using highly reflective mirrors.

In most CO_2 laser machines, several mirrors are used to deliver the laser beam from the cavity to the cutting head. Reflectance of these mirrors should be as high as possible, i.e., minimum absorption and scattering, to minimize laser power losses. In addition to laser power and beam mode, beam polarization also affects cutting quality. For the best cutting quality, circular polarization is required. This is achieved by converting the linear polarization emitted from the cavity to circular polarization using a series of phase-shift mirrors. Zero-phase-shift mirrors guide the beam to the cutting head while maintaining its linear polarization. The addition of one 90° phase-shift mirror at a specific orientation converts the linear polarization to circular. Examples of MKS Ophir's IR optics used in CO_2 laser machines are shown in Figure 198.

Figure 198.
Ophir IR optics used in CO_2 laser machines.

2. Optics for Fiber Laser Machines

In contrast to CO_2 lasers, fiber lasers do not require cavity optics nor beam delivery optics. These functions are performed by the fiber itself. The output coupler and the rear mirror are embedded in the fiber core, and the beam is delivered by the fiber to the cutting head. Due to the relatively small core diameter of a fiber (typically 50 – 100 µm), beam divergence is quite significant. Therefore, a collimator lens is used to collimate the beam after it emerges from the fiber. The cutting head of a fiber laser system includes a lens that focuses the beam on the workpiece. Examples of MKS Ophir's collimation optics for fiber lasers are shown in Figure 199.

Figure 199.
Ophir collimation optics for fiber lasers.

The diameter of the laser spot size on the workpiece depends on the focal lengths of the collimator and focusing lenses. It is necessary to adjust the laser spot diameter to achieve the optimal cutting conditions for a particular application. Sheet material and thickness, laser power, and other laser system parameters determine the optimized spot diameter required for cutting quality and speed. Typically, the spot size should be increased as sheet thickness increases. In CO_2 laser-based machines, the operator is tasked with replacing the focusing lens in order to adjust the beam size. In fiber-laser-based machines, replacing the focusing lens is not recommended because of the extreme sensitivity of these machines to dust particles, which may penetrate the cutting head while switching lenses. The best solution for varying the fiber laser spot size is to use a continuous zoom lens in which the focal length varies by moving internal optical elements. In this manner, the cutting head remains sealed and free of dust particles. Another advantage of using a zoom lens is the ability to adjust the spot size continuously across the full zoom range. An example of an MKS Ophir zoom lens used in fiber-laser-based machines is shown in Figure 200.

Figure 200.
Ophir zoom lens used in fiber-laser-based machines.

3. Power Measurement

Advanced laser systems usually undergo meticulous testing right up until commissioning. However, this close monitoring often stops once the systems are put into production. Real-time monitoring during normal operation affords enormous economic and ecological optimization potential. In the past few decades, production processes – particularly in manufacturing – have been steadily streamlined and automated. In many areas, laser systems have been replacing mechanical tools because they are not subject to wear and require much less maintenance. Despite these advantages, it is important to regularly verify the laser beam quality via reproducible measurement methods, even during productive operation. Unfortunately, many empirical measurement methods that enjoy widespread use do not provide accurate or standardized results. Only when performing optimally can laser systems guarantee the most cost-effective production of high-quality components. Even the smallest deviations in beam adjustment or focal position may lead to reduced part quality, massive cost increases, and pollution of the environment in a variety of ways, including increased energy consumption and use of process gases.

The distribution of the power and energy density in the beam is considered a key parameter in most laser applications; it determines the effectiveness of laser processing in the material. This parameter is calculated by dividing the emitted power or energy by the cross-sectional area of the focused beam (details regarding such laser parameters are given in Chapter 1, Section II.A.1). The higher the power or energy density in the focus, the more efficiently the laser processing operation performs. An unexpected increase in focal spot size can severely impact the beam's power density leading to deleterious effects such as:

- The travel speed of the part must be reduced to compensate for reduced power density
- The quality of the machined part in the cutting or welding process suffers
- Production times and power consumption increase, as does the need for expensive gases used in processing
- The HAZ is larger, requiring more post-process finishing treatments like straightening, deburring, or polishing
- Under certain circumstances, an undetected loss in product quality can lead to diminished strength – a defect that, once recognized, can lead to costly recalls

Furthermore, one should not underrate the economic environmental impacts of consuming more processing gases. Significant energy is expended in the production of gases such as argon, which negatively affects the overall sustainability and end price of the manufactured product. At the same time, as more processing gases are consumed, more processing by-products – both gaseous and particulate – are emitted. Such emissions can reduce the quality of the optics, resulting in further reduction in beam quality. Figure 201 depicts the impact of low beam quality on costs per part.

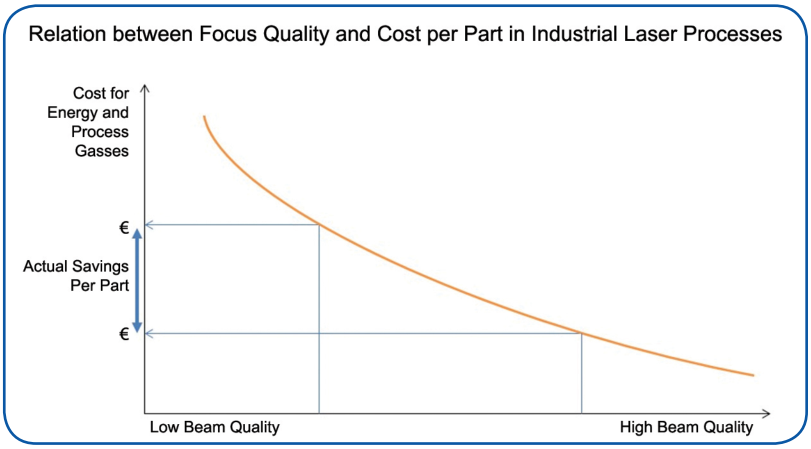


Figure 201. The relationship between beam quality and unit cost in industrial laser processes.

Typically, the first indicator that the performance of a laser system has deteriorated is a decrease of laser power in the focused beam. MKS Ophir has developed Helios®, a compact laser measurement system (shown in Figure 202), specifically for use in automated industrial cutting and welding processes. The Helios system measures high power industrial lasers – such as diode, fiber, and Nd:YAG lasers – with powers from 100 W to 12 kW and energies from 10 J to 10 kJ. The laser is set to a pulse of between 0.3 seconds and several seconds. The Helios then measures the energy and exposure time of this sample of the power to calculate the average power. The short measurement time obviates the need for water cooling and, therefore, the sensor can be kept to a compact size. Details regarding the basic operation of thermal sensors such as the Helios can be found in Chapter 1, Section II.A.2. This compact measuring device has a robust metal housing and an automatic motorized shutter that protects the gauge. This ensures that the sensor remains clean in any environment and is also protected against minor physical disturbance. It can be directly integrated into laser systems or production cells to perform automated measurements for the entire optical laser system. By connecting the Helios system via PROFINET® industrial Ethernet or RS232 serial communications, all data can be immediately transmitted, analyzed, and centrally stored. By regularly and frequently measuring the laser power, slight deviations from setpoints can be detected immediately. The operator can be directly informed by a pass/fail display. This makes it possible to take remedial measures right away, ensuring consistently high production quality.

*Figure 202.
Ophir Helios compact laser power meter for measuring high power lasers.*

4. Beam Measurement

Automated laser welding is popular within the manufacturing industry and there are new applications being developed constantly. One of the top goals for improving system performance is to ensure real-time optimization as the resulting welded joint can only be proofed by destructive material sampling or time- and cost-intensive ultrasound examination. If weak spots are detected in the welds, the bad parts need to be rejected. This is disadvantageous because all production costs to make the part have already been incurred. Ensuring a consistently high-quality weld in a proactive manner is preferred and reduces overall costs. The ability to measure the laser beam in a non-contact manner offers key advantages. A non-contact method developed by MKS Ophir is based on Rayleigh scattering. Rayleigh scattering is the scattering of light by particles that are smaller than the radiation's wavelength, e.g. by oxygen or nitrogen molecules in the air. The electric field of the laser radiation induces an oscillation in the molecule at the laser's wavelength or frequency leading to elastic scattering at that same frequency. The scattered light induced by a laser beam therefore provides a measurement of the intensity of the beam. This non-contact measurement is implemented in a compact device known as the BeamWatch® system, shown in Figure 203. MKS Ophir also offers the BeamWatch Integrated measurement system, which is specifically designed for use in automated manufacturing in the automotive industry.

Figure 203. Illustration of a laser beam passing through the Ophir BeamWatch measurement system without contact (left). BeamWatch Integrated system designed exclusively for automated production lines (right).

The BeamWatch Integrated system images the scattered laser light from the side using a telecentric lens assembly and a CCD or CMOS camera, as shown in Figure 204 . The imager detects the scattered light which provides a measurement of the intensity of the beam profile. Furthermore, since each row of the imager can detect an intensity profile, the propagation of the beam along its focusing direction can also be measured (see Figure 204). These measurements, in conjunction with accurate integrated software, allow for the calculation of the beam profile and beam-quality parameters. The latter includes ISO 13694 and ISO 11146 standards, including focus diameter, focus position, divergence, ellipticity, M^2, Beam Propagation Factor (*K*), and BPP (see Chapter 1, Section II.B.1 for details). However, due to the very weak scattered radiation, it is necessary to minimize and control for any secondary light source that could distort or add artifacts to the Rayleigh light. This is achieved through optimal placement of the individual components and light-absorbing materials in the measurement chamber. An air purge system ensures that no foreign particles (such as dust) affect the measurement.

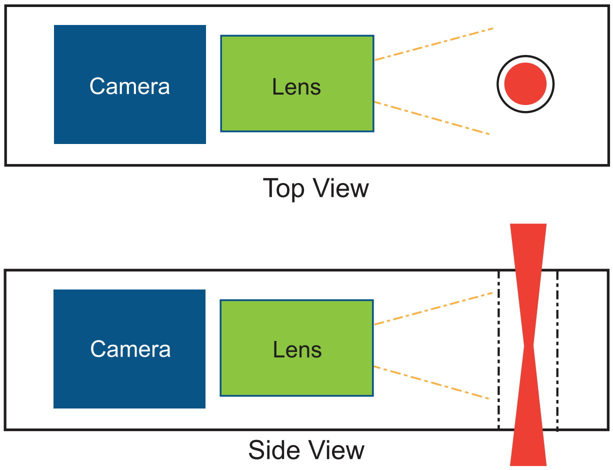


Figure 204. Ophir BeamWatch Integrated performs a non-contact measurement of the Rayleigh scattering of the beam (red circle and line) using a CCD camera with a telecentric lens assembly.

The BeamWatch and BeamWatch Integrated systems are capable of measuring beam and beam quality parameters at video rates. This permits identification of beam misalignments in real time, allowing the user to make any necessary adjustments. Furthermore, long-term drift in these critical beam parameters can be identified that would normally lead to non-optimal results. An example of this is illustrated in Figure 205 where the impact of a beam focal shift is shown. As the focal point shifts, the power density in the processing plane reduces. Since the power density depends inversely on the beam area, a doubling of the beam radius results in a four-fold reduction on the power density. Furthermore, the BeamWatch systems also possess an integrated power meter allowing for absolute measurements of the power density.

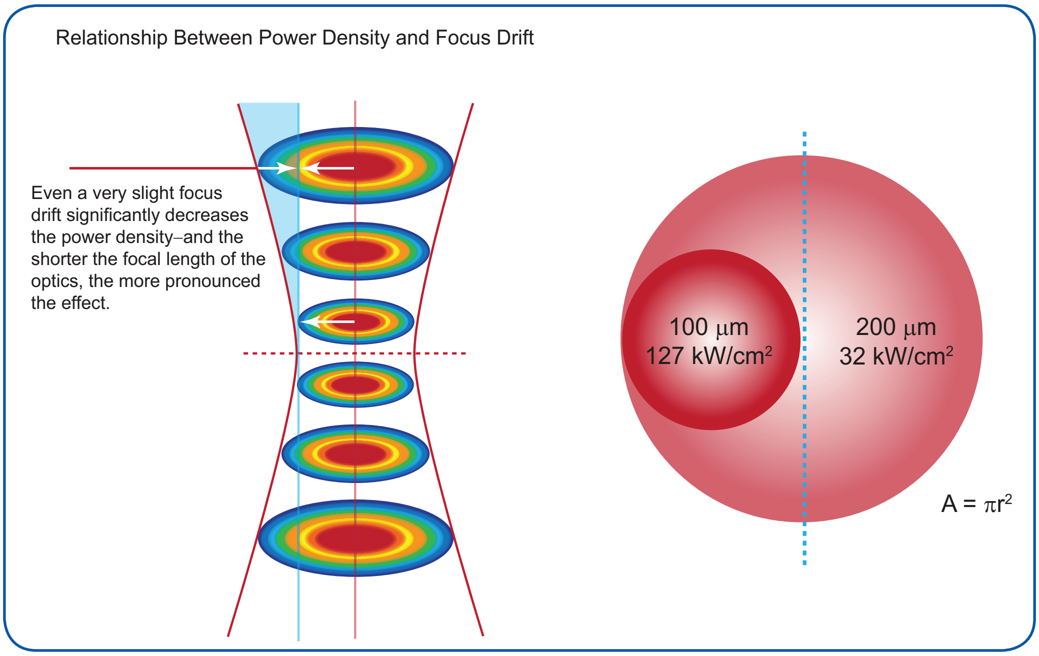


Figure 205. Diagram depicting how a focus shift influences power density (left). Illustration indicating that doubling the focus beam diameter delivers a four-fold reduction in power density (right).

The non-contact measurements provided by the BeamWatch systems can be displayed in a variety of ways (see Figure 206). Users can choose simple pass/fail displays or can pre-configure up to ten different measurement tasks, which enables flexible use with different welding stations. All the measured and calculated information can be transferred to a central platform using integrated industrial interfaces such as PROFINET®. In summary, the non-contact measurement of high-power lasers in automated production, as provided by the BeamWatch systems, has many advantages. Laser parameters can be measured rapidly and reliably, leading to a reduced number of defective parts. The documentation and storage of measurement protocols allows manufacturers to clearly identify causes of any deviation in part tolerances. Finally, maintenance costs are significantly reduced since non-contact measurements induce no wear.

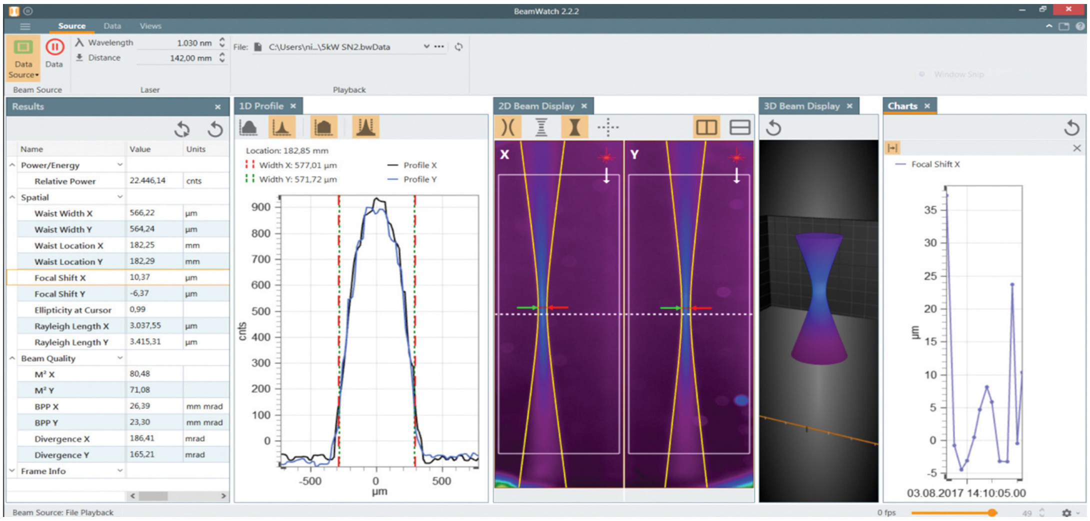


Figure 206. BeamWatch Integrated provides a full range of measurement results.

C. Future Directions

Laser macroprocessing has expanded to many facets of metal manufacturing from cutting and welding to newer processes in additive manufacturing. As laser diode cost per watt of power continues to improve (similar to Moore's Law for semiconductors), laser machining will become increasingly compelling over conventional processes in many applications. One can envision laser macroprocessing becoming the dominant process for the full range of industrial manufacturing in sectors ranging from automotive to shipbuilding.

IV. MKS Strategy in Laser Machining

MKS is uniquely positioned to take advantage of the expansion that laser machining is expected to witness in the coming years. The core pillar of MKS' Vacuum & Analysis (V&A) division strategy is Surround the Chamber. Similarly, the company's strategy in the L&M division is Surround the Workpiece℠ strategy. MKS' broad, industry-leading product portfolio not only offers customers ease of sourcing from one supplier but also the assurance of superior quality.

In a typical laser machining setup (as shown in Figure 207), laser light is generated, then carefully managed through a set of advanced optical and opto-mechanical components placed very precisely to carry the light to the workpiece. The workpiece itself needs to be moved and positioned extremely precisely so that the light shines on it at just the right spot. These motion stages are in turn placed on active vibration control tables that compensate in real-time for any external vibrations that may cause the workpiece to move, leading to imprecise processing. Since light is one of the most important elements in material processing, regular process control is required to ensure the integrity of the system. For process control, MKS provides a complete suite of products ranging from laser beam profiling and power measurement to beam quality measurement.

MKS Spectra-Physics lasers are offered at ns, ps, and fs pulse widths and wavelengths from IR to UV to accommodate a wide range of material processing applications, both in terms of materials that the lasers can process such as brittle materials, plastics, or metal, and also in terms of the process itself, such as cutting, drilling, or marking.

MKS' line of beam delivery and conditioning products includes optical components such as mirrors, waveplates, polarizing beam splitter cubes, as well as opto-mechanical components such as mounts, retaining rings, and optic positioners and holders.

The MKS motion stage product portfolio includes high precision XY stages, vertical stages, rotation stages, air bearing stages, custom-made motion systems, XPS high performance universal motion controller/driver, and LMS-Pro laser machining software. Over the decades, MKS has served customers with diverse applications including ablation, ultrafast micromachining, laser additive manufacturing, laser cutting, scribing and drilling.

Our Newport tables are an industry standard for optical tables, isolation systems, and vibration control products. We offer a broad range of broadband, tuned damped and actively damped optical tables to fit a wide variety of applications. In addition, our portfolio includes optical breadboards, vibration isolators, and custom solutions tailored to fit exact application requirements.

Our Light & Motion portfolio includes industry leading tools for measuring power or energy of an optical beam, profiling a laser, locating the position of a beam, spectral analysis, or characterizing a laser pulse. The MKS Ophir Photonics portfolio offers a diverse selection of these products. In addition, Newport brand optical meters, optical sensors, and beam characterization instrumentation are also available.

Surround the Workpiece is at the heart of MKS' laser machining strategy. Our robust portfolio of products, global reach, and focus on innovation provide the opportunity to help solve our customers' most critical issues.

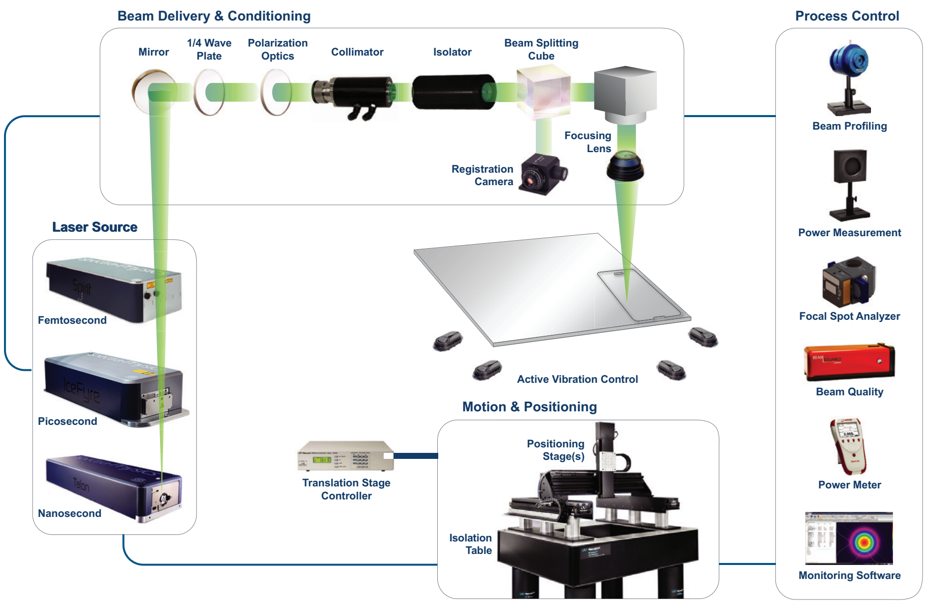


Figure 207. Surround the Workpiece solution strategy. An example laser micromachining system including beam delivery, motion and positioning, and monitoring of the laser beam delivery.

Chapter 4

Photonics in Remote Sensing

This page intentionally left blank.

I. Introduction

Remote sensing entails the use of instruments for acquiring information about an area of interest on the ground [289]. Remote sensing in photonics involves either passive or active sensing of observable signals. Passive remote sensing systems measure radiation from the earth's surface in the form of either reflected solar light or emitted thermal radiation. Active remote sensing systems emit radiation toward the target using their own energy source and detect the energy reflected from the target.

This chapter focuses on the most common passive and active remote sensing methods in photonics, that is, thermal imaging and Light Detection and Ranging (LiDAR). Spectroscopic techniques such as laser-induced fluorescence are less common and will not be covered in this book. The discussion of each method covers its basic principle of operation, the impact of photonics elements, e.g., optical components and lasers, on its operation, example applications which utilize MKS products, and future directions in the market.

II. Thermal Imaging

As depicted in Figure 208, thermal imaging techniques involve discriminating between various objects in a scene based on their respective differences in temperature [290]. The black body radiation law (see Chapter 1, Section I.C) describes the characteristics of IR radiation emitted by an object with a temperature above absolute zero. Importantly, the magnitude of the emission (termed irradiance or exitance) increases with temperature. Therefore, provided a detector is sensitive to IR light, it is possible to visualize the environment based on temperature variation. The ability to produce an image (or thermogram) based on IR radiation is called thermography. Thermograms discriminate between warmer foreground objects and their cooler backgrounds such as the visualization of warm-blooded animals or humans.

Figure 208. Image from a thermal camera illustrating the ability to discriminate between objects in the foreground and background based on their temperature differences.

Thermal imaging cameras allow for object distinction without the aid of an illumination source, thereby distinguishing it from LiDAR discussed in Section III. Thermal imaging is ideal for producing images at night or in the presence of some obscuration, such as light fog, rain, or smoke [291]. For example, forward-looking IR or FLIR cameras are used to provide night vision on military and civilian aircraft or for use in security and surveillance. FLIR cameras usually detect radiation in the mid-wavelength IR (MWIR, 3–5 μm range) and long-wavelength IR (LWIR, 7.5–14 μm range) portions of the electromagnetic spectrum. When coupled with additional cameras that can detect VIS (0.4–0.75 μm range), NIR (0.75–1.1 μm range), or short-wavelength IR (SWIR, 1.1–2.5 μm range) radiation, multispectral imaging is possible (see below). SWIR imaging is becoming increasingly popular in many industrial applications, including the inspection of solar cells, and agricultural products, and for anti-counterfeiting [292].

A. Overview of Key Markets and Applications

Thermal imaging has a long history in the areas of defense and aerospace but, commercial and industrial applications have led to a dramatic increase in its use over the past few decades. This emerging market is directed and led by various detector and camera manufacturers, such as FLIR, L3 Technologies, BAE Systems, Leonardo DRS, SCD, and Sofradir. These companies are continually making improvements and reducing costs for thermal imaging components and detectors that will result in further market growth. The main segments in the thermal imaging market are summarized in Figure 209 and are described below.

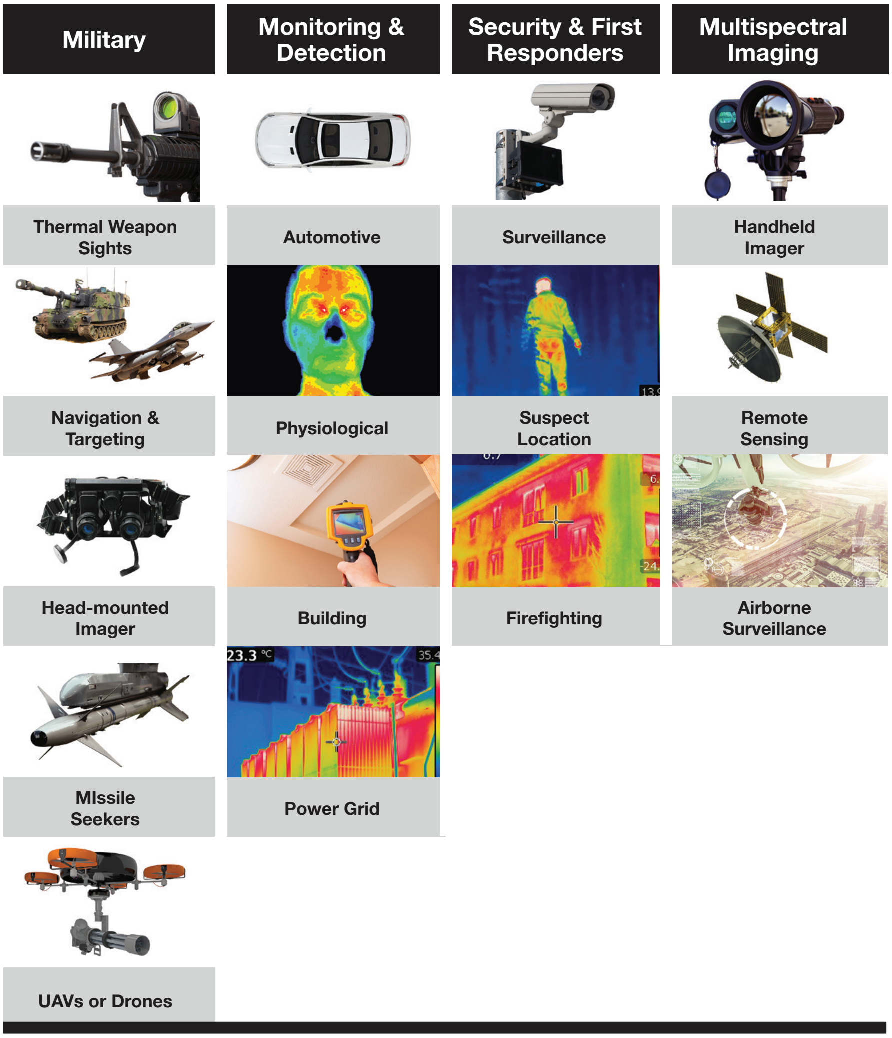


Figure 209. Thermal imaging market segments.

Military

Thermal imaging has a wide variety of military applications for operation at night, under reduced visibility, and for target acquisition. Furthermore, these applications can cover ground, air, and maritime environments. Example applications include aircraft navigation and targeting systems, thermal weapon sights (see Section II.D.3), handheld and head-mounted imagers, fire control systems for tanks, and precision missile guidance. Thermal imaging systems are frequently used in unmanned aerial vehicles (UAVs) or drones. The specific applications and requirements related to these systems are discussed in Section II.D.1.

Security and First Responders

The recent increase in terrorist activity and asymmetrical warfare has accelerated the need for IR systems in homeland security, border control, and law enforcement. These systems are utilized to identify security perimeters at various locations including airports as well as underground and above ground train stations. Law enforcement uses the technology to manage surveillance activities, locate and apprehend suspects, investigate crime scenes, and conduct search and rescue operations. Thermal imagers have helped revolutionize firefighting by enabling the detection of trapped persons as well as the location of the base of a fire.

Monitoring and Detection

Thermography is useful for monitoring physiological activities, including the detection of fever and disease. This can enable improvement in patient care as well as control at borders in the event of a contagious outbreak. Thermal imaging is ubiquitous for industrial applications, including gas leak detection, building inspection, predictive maintenance, and process control. “Smart buildings” leverage thermal activity sensors to bring intelligence to building systems, e.g., lighting, HVAC, alarms, that improves efficiency, enhances occupant comfort, and optimizes workspace management [293]. “Smart cities” use thermal imagers to enable real-time monitoring of transport systems, traffic management, and power grids with the goal of improving quality of life and optimizing city functions [294]. The automotive camera market is driven by the new trend of autonomous cars. Thermal cameras can play a key role for pedestrian recognition and obstacle warning, particularly when driving in total darkness or in tunnels (see Section II.D.2).

Multispectral Imaging

Multispectral imaging combines two to five spectral imaging bands (VIS, NIR, SWIR, MWIR, or LWIR) of relatively large bandwidth into a single optical system. This can enable increased functionality and reveal details beyond the capability of a single-band detection system. Advances in imaging components have allowed multispectral imaging to move beyond custom-built systems for laboratory and government applications to affordable, practical, commercial systems [295]. Applications range from deep-space imaging using space-based telescopes, airborne surveillance, satellite-enabled remote sensing, to handheld imagers.

B. Detectors for Thermal Imaging

The image quality produced by a thermal imaging system depends on its detector and its optical components (see Section II.C). The most important parameters for IR detectors used in thermal imaging are sensitivity, pixel pitch (pixel-to-pixel distance), and format (number of pixels). In conjunction with the imaging optics, the pixel pitch and format determine the spatial resolution and area of the target to be imaged, (see description of camera-based profilers in Chapter 1, Section II.B.2 for details). Generally, IR detectors for thermal imaging cameras are much more expensive than their VIS spectrum counterparts. Consequently, large format detectors with pixel areas of 1024×768 are typically found on higher-end imaging cameras.

Detector sensitivity typically refers to the minimum detectable signal that gives a SNR value of one. The metric used to describe detector sensitivity is the normalized (or specific) detectivity (D*). This value accounts for the impacts of bandwidth and detector area and is intuitive in that larger values represent greater sensitivity. A description of the normalized detectivity, as well as typical values for IR detectors, are given in Chapter 1, Section II.A.1. In addition to the detector sensitivity, the amount of radiation that reaches the detector ultimately determines the detection range of a thermal imaging system. The role of the optical components in efficiently delivering light to the detector is discussed in Section II.C. Other factors that contribute to the detection range include the target parameters, e.g., size, temperature, emissivity – the ratio of an object's radiance to that emitted by a blackbody radiator, and atmospheric transmission. Figure 210 shows atmospheric transmission as a function of wavelength with the relevant IR spectral regions indicated.

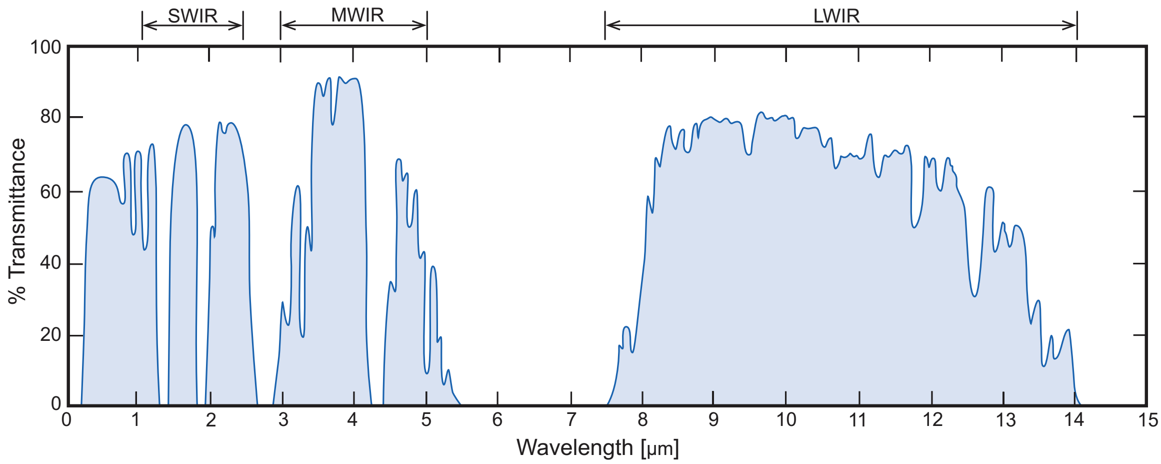


Figure 210. Spectral transmittance of the atmosphere over one nautical mile (~6000 feet) at sea level [296].

Sensors capable of detecting IR radiation include thermopiles, pyroelectrics, bolometers, and low-bandgap semiconductor photodiodes (see Chapter 1, Sections II.A.2 and II.A.3 for details). While arrays of pyroelectric sensors and microbolometers are used for thermal imaging, the use of photodiode arrays are more common (see Chapter 1, Section II.B.2). This is mainly due to the greater detectivity for these types of sensors as well as the maturity of the semiconductor manufacturing processes that enable them. Since these focal plane arrays (FPAs) are made of photodiodes, their IR spectral sensitivity depends on the bandgap of the semiconductor material being used. The most common types of FPAs are constructed using InSb, InGaAs, HgCdTe and quantum-well-based photodiodes. IR-sensitive FPAs are considerably more expensive than Si-based FPAs, resulting in higher-end models being used mainly in military applications. Furthermore, detection in the MWIR and LWIR often requires cryogenic cooling to achieve reasonable detectivity levels.

C. Optical Components for Thermal Imaging

Optical components play a critical role in thermal imaging both in terms of performance and cost. In particular, lenses are the most essential components for imaging (see Chapter 1, Section III.A.3 for more information). Some widely-used terms related to thermal imaging optics include the effective focal length (EFL), the field-of-view (FOV), the entrance pupil, and the F/# (the ratio of the EFL to the entrance pupil diameter). Details regarding the definitions of these terms and how they relate to imaging systems can be found in [170]. Generally, the main performance parameters for lenses or systems of lenses are:

- Modulation transfer function (MTF) – measures the ability to transfer (or reproduce) modulation (or contrast) in an object to the image formed by the optical component; this is a function of spatial frequency (or resolution)
- Transmission – ratio of transmitted radiation relative to incident radiation

- Relative illumination – illumination at any point on an image relative to the maximum on-axis illumination
- Distortion – image deformation caused by a change in magnification across the FOV

All these parameters are a function of the optical lens design. Consequently, lens design is crucial in terms of achieving the desired image quality. The following discusses how thermal imaging optics are designed and manufactured to meet performance requirements.

1. Lens Types and Design

There are many lens types used in thermal imaging systems. The types of lenses chosen for a particular system can depend on (among others): the spectral range of interest, e.g., NIR, SWIR, MWIR, or LWIR, the number of FOVs required (e.g., single, dual, continuous zoom), whether or not they need to be athermal (no need for refocusing following a temperature change), and their focus mechanism, e.g., fixed, manual, or motorized. Once the system requirements are established and the lens specifications are drafted, the optical designer must supply an optical design that meets specification requirements. Additionally, the designer must be cognizant of manufacturability and cost. The stages of optical design include the following: (1) establishing the basic geometry, (2) optimizing performance, fit, and cost, (3) building a tolerance set, and (4) documenting the design and performance.

The principle method of arriving at a good optical design is optimization, which is varying the numerical design parameters to achieve a desired imaging outcome such as a tight focus or a flat field. Design parameters can be extensive owing to, for example, the number of lenses involved, their shapes or curvatures, the optical materials that make up the lenses, whether there is an intermediate focus, etc. Consequently, as a starting point in the optimization, the optical designer will draw on previous experience to lay down a basic geometry. Often, a previous design that shares some common specifications to the system under consideration can be used as a starting point. For instance, if a previous design has the same FOV but a different focal length, then a simple scale up or scale down of that design may result in the final target design.

Optimization is based on a merit function. This function can include terms that directly impact imaging performance, such as optical aberrations (see Chapter 1, Section III.A.3 for details). It can also include terms for controlling design geometry, such as the thickness of elements or the spacing between them. The optical designer adjusts the merit function during the optimization process to achieve the design objectives, which include performance, size, and cost. If it becomes apparent during optimization that objectives will not be achieved based on the initial basic geometry, then the optical designer will need to modify the geometry and start the optimization process again. Such modifications could include changing a spherical surface to an aspheric one, adding an optical element, or changing the lens material.

In order to complete an optimization, a designer must have a tolerance set in addition to the optical design itself. A tolerance set can include the sensitivity limits of the manufacturing process used in lens element fabrication. These tolerances might include variations in the radius of curvature, thickness, surface irregularity, and edge thickness. There are also tolerances related to the mechanical assembly of a system. These include decenter, tilt, and distance between elements. It is the job of the optical designer to design a system that is not overly sensitive to tolerances which could result in manufacturing difficulties.

It is critical to properly document the optical design. This can include a tabulation of the numerical data that thoroughly describes a lens (a prescription) as well as a cross-sectional illustration of the lens (a layout). Documenting the performance can mean putting together several charts describing the MTF, distortion, and illumination of a system as well as minimum performance requirements at final inspection. Finally, additional data may also be required for the mechanical design such as required focus motion, back focal length, maximum FOV footprint, and thermal compensation.

2. Materials, Surfaces, and Coatings

Lenses for thermal imaging systems are made of materials with low absorption in the IR spectral region. The main materials used to make these lenses include Ge, Si, ZnS, ZnSe, CaF_2, and chalcogenide glasses. Mirrors can also be used as optical components for thermal imaging systems. These components have surfaces that are mainly made of metals, such as aluminum and copper. More information about mirrors and their surfaces can be found in Chapter 1, Section III.A.1.

There are four main types of optical surfaces used to construct optical components: plano, spherical, aspherical, and diffractive. Plano and spherical surfaces are made to accommodate paraxial wavefronts, that is, weakly focusing geometries. Aspherical surfaces are designed to correct for wavefront errors such as spherical aberrations and therefore are more appropriate for stronger focusing geometries. Diffractive surfaces are designed to correct for chromatic aberrations.

Forming the underlying shapes for these optical components is accomplished using grinding and centering. Grinding is the initial process used to fabricate the shape of the component. Centering involves cutting the edge of the component to give it its final shape and diameter and adjusting the optical axis to the mechanical one. Once the shape of the component is complete, the optical surface must be polished. Chemical mechanical polishing, used for plano and spherical surfaces, involves the use of a polishing slurry. The slurry consists of chemicals that react with the surface and soften it. The slurry also includes powders that mechanically scrape the surface when abrasive pads are applied to it. Aspherical and diffractive surfaces are typically polished by diamond turning. In this approach, the component is mounted and spins while a diamond tool carves the surface, resulting in the desired structure.

The parameters and procedures used to assess the quality and flatness of optical surfaces are described in Chapter 1, Section III.A.1. For IR optical components, surface flatness or irregularity has a standard accuracy of $\lambda/2$ at 633 nm, which can be quite demanding at certain IR wavelengths. Surface quality standards typically require scratch and dig values of 80–50 for the LWIR region and 60–40 for the SWIR region. Surface curvatures typically require accuracies of 0.1% of the radius value or better, while standard surface roughness values are in the 25 nm RMS range. In addition to high-quality optical surfaces, lens positioning is also critical in thermal imaging systems. This can include absolute positions in the optical housing and the relative positions of two lens surfaces. Measurement tools such as high accuracy calipers, micrometers, and comparators are used to meet the required specifications.

Almost all optical surfaces associated with IR optical components are coated. The coatings serve two purposes: improve spectral performance, i.e., transmittance and reflectance, and withstand environmental conditions. Details regarding general coating characteristics and application technologies are given in Chapter 1, Section III.A.2. Optical coatings specifically for IR optics fall into one of the following categories: high efficiency, high durability, and hard carbon. High efficiency coatings have superior spectral performance with low environmental durability while high durability coatings sacrifice spectral performance for increased environmental durability. Hard carbon coatings have single layer diamond-like coatings and possess very high environmental durability, while their spectral performance can be tailored based on the target application.

3. Assembly and Mechanical Design

Prior to assembly of the thermal imaging system, the specification requirements are verified for all the pre-assembled parts. During assembly, all elements, both mechanical and optical, must be kept free of potentially damaging debris. The main goal is that the optical elements are properly placed according to the optical and mechanical design requirements. There are several calibration and assembly methods designed to meet these requirements which vary in complexity depending on the system. For instance, systems that are made to function in extreme environmental conditions, e.g., temperature, shock, vibrations, rain, will undergo more rigorous assembly protocols.

More sophisticated assembly and mechanical design efforts are required when optical components are no longer in fixed positions relative to one another, such as in a zoom lens. In this case, internal lenses are moved to different positions with the goal of achieving good focusing for all designed FOVs. Positioning mechanisms may either be static or dynamic and can require a high degree of accuracy with smooth movement and zero backlash to function properly. Details regarding the mechanisms used for movement as well as the drivers and controllers used for motorized movement are discussed in Chapter 1, Section III.F.2. For motorized imaging systems, a controller uses an encoder along with a look-up table to properly position the lenses to achieve optimal focusing according to design specifications. IR imaging lenses are made of materials that can be sensitive to the surrounding temperature, yielding slight changes in shape and/or position with changes in temperature. To compensate for this, the controller can measure the ambient temperature and reposition the lenses to maintain optimal focusing.

4. Testing

Tests on the resulting opto-mechanical assembly are carried out according to a product quality assurance program. These can include mechanical and optical tests and, if necessary, electrical and control tests. The program is usually divided into three main phases and each has an expressly different purpose. The design verification test (DVT) phase typically includes an extended series of tests aimed at verifying the assumptions made during the design. DVTs also determine whether the product meets all the specification requirements. The qualification test phase involves testing the durability of the product under extreme environmental conditions, either in storage or when fully operational. These tests are often performed on a prototype. The final phase is the acceptance tests procedure (ATP). The purpose of these tests is to ensure that the unit-under-test is properly assembled and calibrated and meets the main requirements of the specifications. The ATP can be performed during the transition to production or at the production line.

One of the most important tests for a thermal imaging system is the quality of the wavefront that the lens assembly transmits. Sometimes referred to as a "wavefront error" test, it measures the irregularity of the transmitted light from the scene through the lens assembly to the detector. The measurement is performed using a commercial interferometer (see coherence portion of Chapter 1, Section I.A.3 for details). The spectral region appropriate for the particular lens assembly dictates which laser wavelength is used for the interferometric test: NIR lenses use 0.6328 µm from a HeNe laser, SWIR lenses use 1.52 µm from a HeNe laser, MWIR lenses use 3.39 µm from a HeNe laser, and LWIR lenses use 10.6 µm from a CO_2 laser. The results of the measurements are given in peak-valley RMS values (either in distance or fractions of a wavelength), and various analyses of lens quality can be obtained from this measurement. The most common analysis produces the MTF, which is considered the main quality parameter of a lens or lens assembly.

D. Thermal Imaging Applications Using MKS Products

1. Lightweight MWIR Zoom Lenses for UAVs

When equipped with high-performance thermal imaging systems, UAVs, also known as drones, lend themselves to a wide range of defense, government, and commercial applications. In the areas of defense and government, UAVs are used for military and police surveillance, border control, monitoring, and search-and-rescue operations. Commercial drones with thermal imaging capabilities are playing a prominent role in inspecting electrical power lines, oil pipelines, and other infrastructure. Thermal imaging drones are also being used to assist in firefighting operations by locating and assessing fires, even when visibility is poor.

As UAV technology is implemented for increasingly varied and sophisticated tasks, there has been a call to improve imaging performance. Accordingly, detector manufacturers are working on improving resolution while maintaining relatively small footprints to accommodate usage in smaller drones. While detector resolution should improve imaging performance, this must be accompanied by improvements in lens quality. To match the performance of state-of-the-art detectors, lenses with lower F/#'s and minimal aberrations are necessary to improve low-light performance. Furthermore, lenses must also possess long focal lengths for use when UAVs capture images from large distances.

There are also unique requirements for optical elements to be used in drones. The three most important factors to consider when developing optical components for UAV payloads are summarized by the acronym SWaP, which stands for size, weight, and power consumption. UAV payloads, especially for smaller commercial UAVs, impose strict size and weight restrictions. Furthermore, power consumption must be reduced to help lower fuel usage, thus maximizing flight time. The challenge thus falls on lens manufacturers to design and produce compact, lightweight lenses while not compromising image performance. Various solutions are required to meet these challenges, including the use of innovative optical and mechanical designs, exotic materials, free-form optics, and unique lens coatings.

Diamond turning technology can be used to produce aspheric and diffractive surfaces with exceptional levels of accuracy and quality. Aspheric lens surfaces provide improved optical performance over their spherical counterparts for IR imaging. Aspheric-diffractive lens surfaces allow for correction of both chromatic and spherical aberrations. Consequently, lenses produced by diamond turning can reduce the number of optical elements, leading to a reduction in overall size and weight. Continuous zoom lenses, which maintain focus under changing conditions, can also reduce lens size and weight compared to the use of multiple single-FOV lenses. A continuous zoom lens also enables better imaging adaptability during UAV operation. The use of durable, anti-reflective coatings also improves optical performance with minimal size/weight impact.

Figure 211. Ophir LightIR continuous zoom lens for UAV-based thermal imaging applications.

MKS Ophir is a worldwide leader in the design and manufacture of high-performance LWIR and MWIR lenses that meet the strict requirements of the UAV industry. Working in collaboration with defense and commercial customers, MKS Ophir designs IR lenses with unparalleled optical performance while also considering the SWaP constraints of the UAV payload. This is accomplished using advanced manufacturing technologies including diamond turning processes, free-form optics fabrication techniques, and advanced lens coating. The resulting continuous zoom and athermal lenses possess superior IR image quality, are lightweight with a small footprint, and can function in extreme environmental conditions. MKS Ophir's LightIR (see Figure 211) is a high-performance, IR thermal imaging lens designed specifically for use in UAV payloads. The LightIR is a motorized, continuous zoom lens and is the lightest and most compact lens of its type available on the market today. The specifications for the LightIR lens are shown in Table 14.

Property	Value	
Optical	**WFOV**	**NFOV**
Focal Length	20 mm	275 mm
F/#	5.5	
Cold Stop to FPA Distance	19.4 mm	
Cold Stop CA	3.52 mm	
Average Transmissions (3.6-4.9 μm)	80%	
Back Focal Length	23.52 mm in air	
Distortion	< 5%	
Minimum Focus Range	5 m	50 m
Nuc (by defocus)	Yes	
Mechanical		
Focus Mechanism	Motorized	
Focus Time (minimum range to ∞)	≤8 sec	
Zoom Mechanism	Motorized	
Zoom Time (NFOV to WFOV)	≤5 sec	
Through-zoom Boresight	0.5 mm on detector plane	
Weight	264 g	
Max. Dimensions	58 mm x 67.1 mm	

Table 14. Specification sheet for Ophir LightIR 20-275 mm motorized continuous zoom lens.

2. Automotive Lenses for Night Vision Systems

Automotive night vision systems use thermal imaging technology to allow drivers to detect pedestrians and provide a clear view of the road, even when vision is obstructed by environmental conditions such as darkness, smoke, or fog [297]. For maximum performance and minimal collision risk, thermal imaging accuracy, quality, and long-distance object detection are critical to provide the driver with enough response time. The key to meeting these requirements is the use of high-sensitivity and high-resolution optics.

MKS Ophir has earned its reputation as a world-leading designer and supplier in the field of thermal imaging optics for the automotive market. MKS Ophir's superior athermalized lenses increase pedestrian recognition software performance, allowing a greater ability to anticipate potential hazards. Crafted with years of experience and knowledge, these IR thermal imaging lenses feature the highest quality components and materials, designed especially to meet the needs of the industry. As the sole provider of IR thermal optics for the European automotive market, MKS Ophir's lenses are integrated in the night vision systems of top European cars.

3. Thermal Weapon Sights

A thermal weapon sight is a device that combines a compact thermal imager and an aiming reticle [298]. These devices can be mounted on a variety of small arms and are often used by hunters. The thermal sight can be quite useful in places when vision is obstructed by environmental conditions such as darkness, smoke, or fog. The sight makes it easy for a user to locate any source of heat, such as an animal or vehicle, against its lower-temperature background. MKS Ophir produces a variety of lenses used in thermal weapon sights.

III. Light Detection and Ranging (LiDAR)

LiDAR is an active remote sensing technique that is similar to RADAR but, instead of using radio waves as a radiation source, it uses laser pulses. In this technique, a laser source emits pulses that are directed towards the target of interest, such as a terrain landscape. The pulses encounter the terrain and a portion of the laser energy is reflected back to a sensor located near the source. By measuring the round-trip travel time of the emitted laser pulses, the LiDAR system can determine the distance between the sensor and the mapped terrain. LiDAR can generate a dense, three-dimensional (3D), geo-referenced point cloud, i.e., a set of data points in space, for the reflective terrain landscape when combined with a Global Positioning System (GPS) and an Inertial Measurement Unit (IMU). Compared to traditional photogrammetric approaches, LiDAR is less sensitive to the current weather, time of the year, or time of the day during which data is collected. Furthermore, this technique can generate high-resolution 3D topographic surface information more rapidly due to its ability to penetrate vegetation [299].

LiDAR has become an established method for generating dense and accurate elevation data across landscapes, shallow-water areas, and project sites. LiDAR systems are often placed in aircraft where data can be rapidly collected over large areas. LiDAR data is also collected from ground-based stationary and mobile platforms for street mapping and autonomous driving applications. These collection techniques are popular within the surveying and engineering communities since they can enable the accurate development of railroads, roadways, bridges, buildings, breakwaters, and other shoreline structures [300]. This section mainly discusses airborne applications of LiDAR systems.

The basic principles of LiDAR are shown in Figure 212. The airborne LiDAR system is comprised of three major time-synchronized components: a laser scanner unit, a GPS, and an IMU. The laser scanner is composed of a laser range finder unit, which is based on time-of-flight distance measurement techniques, and a beam deflection device that creates the desired scanning pattern. The GPS provides the absolute position of the sensor platform (plat), and the IMU records the angular attitude of the platform (including roll, pitch, and yaw/heading). This enables the system to generate the aircraft's absolute position (X, Y, Z) at any given time. The position is synchronized using the detector's recording system for each recorded reflection [301]. The required recording speeds and the amount of collected data require strong, real-time computation capabilities on board the aircraft.

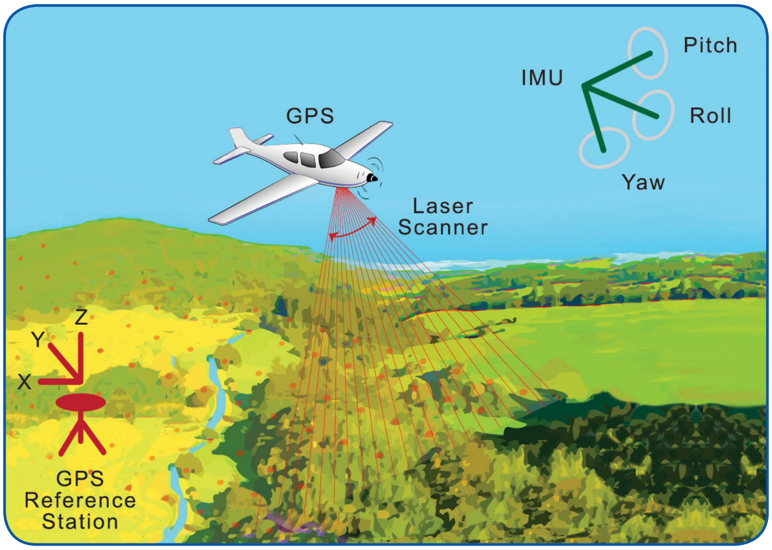


Figure 212. Basic principles of airborne LiDAR data collection [301].

Most commercial airborne sensors are based on the LiDAR principle of pulse round-trip time measurement. The airborne sensors are typically mounted on a fixed-wing aircraft (usually augmented with simultaneous digital imagery and aimed at large areas) or a helicopter (typically used for smaller areas with high-resolution mapping). The short laser pulse (typically a few ns in duration) travels from the sensor through the atmosphere and is then reflected by one or more objects on the ground that are illuminated by the laser beam. The elapsed time between emission and arrival is used to compute the distance between the sensor and the target by dividing the recorded time by two and multiplying it by the group velocity of the light pulse (approximately 3×10^{8} m/s).

LiDAR has several advantages as a remote sensing technique, including high accuracy, large point density, and extensive coverage area. Furthermore, end-users can resample regions of interest quickly and efficiently. This gives rise to a technique that can map discrete changes at very high resolution, cover large areas uniformly and accurately, and produce rapid results [302].

A. Laser Requirements for LiDAR Applications

Among the required components of a LiDAR system, the laser plays the most significant role in the overall system performance. Hence, when determining data acquisition requirements for a LiDAR system, it is usually the laser specifications that determine the system cost, performance, and the feasibility of an application. The key laser parameters that contribute to the performance of a LIDAR system are outlined below.

1. Laser Wavelength

Three different wavelength regions are used in LiDAR systems: NIR excitation at 1064 nm using either DPSS or Yb-doped fiber lasers, VIS excitation at 532 nm produced by frequency-doubling a 1064 nm laser, and SWIR excitation at 1550 nm using Er-doped fiber lasers. Details regarding these types of laser systems can be found in Chapter 1, Section I.A.4. Each wavelength has a unique set of advantages and disadvantages that depend on the target reflectance and absorbance, background radiation, atmospheric transmission, and eye-safety issues.

For airborne topographic mapping, 1064 nm is the most commonly-used wavelength. A major advantage of this wavelength is the abundance of commercially-available laser sources and light detectors. Another advantage is that the detectors can be Si-based, and therefore offer higher gain and lower cost than alternative GaAs-based photodetectors. Furthermore, this wavelength generates high reflectance from the most commonly-mapped targets, e.g., vegetation and snow. A major disadvantage is this wavelength's potential to be hazardous for the eyes (see Chapter 1, Section I.A.7 for more details on control measures for laser hazards). This limits the radiance that can be used for the laser beam requiring either laser power reduction or beam expansion to reduce the hazard. Another disadvantage is the large background noise experienced in this part of the spectrum, particularly from the spectral irradiance of the sun.

For bathymetry applications, i.e., high-resolution mapping of the sea bottom and coastal areas, a 532 nm laser source is often used because it represents the best compromise between high transmission in pure water and limited backscattering from submarine particulates. Figure 213 shows an application where two wavelengths are used in LiDAR bathymetry. In this case, an NIR pulse (typically at 1064 nm) is reflected from the water surface while 532 nm light penetrates the water surface and is reflected from the sea bottom.

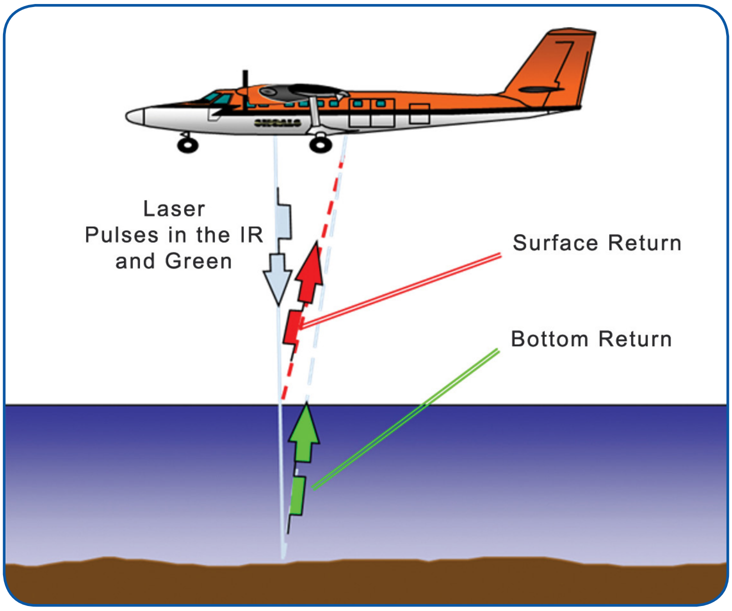


Figure 213. Basic principles of bathymetric LiDAR. Image courtesy of Teledyne Optech.

Eye-safe lasers are becoming increasingly popular in high-performance compact LiDAR systems for civil and commercial applications. SWIR lasers operating at 1550 nm are generally more eye-safe at higher power levels and are typically used when solid bodies need to be detected, e.g. in topography mapping and obstacle avoidance. Furthermore, atmospheric transmission is quite good at this wavelength. Military applications also utilize these sources, as night vision devices are relatively insensitive to this wavelength. However, detection at 1550 nm requires the use of InGaAs or Ge photodetectors which are more expensive and have lower detectivity than Si detectors (see Chapter 1, Section II.A.2). Another disadvantage of this wavelength is that it experiences strong water absorption. This significantly reduces the reflectance from certain objects, such as snow and vegetation, and limits LiDAR usage.

2. Pulse Repetition Rate

The pulse repetition rate and pulse energy determine the applicable sampling rate of a LiDAR system. These parameters determine what will be the densely-spaced network of highly accurate geo-referenced elevation points produced by the system. High pulse repetition rates enable faster acquisition of data and/or higher point cloud density. The higher the point cloud density (typically given in points per m^2), the better the achievable resolution with a LiDAR system. For instance, 4 points/m^2 correlates to 0.5 m of ground sample distance. If pulses are strong enough, faster scanning can enable an aircraft to fly at higher altitude while still effectively mapping the terrain. This yields higher swath widths, which accelerates mapping throughput, thereby reducing time and flight costs.

LiDAR system throughput has advanced considerably in recent years. Early commercial units were capable of operation at 10 kHz (10,000 points per second) and were large and bulky. Newer systems are more compact, lighter, and can process multiple laser returns in air, allowing for pulse rates to exceed 1 MHz. Multiple returns occur when a pulse strikes a target but is not completely reflected or absorbed, resulting in a portion of the pulse continuing to lower objects where it can also be reflected (see Figure 214). Multiple return systems can increase the amount of collected data significantly and enhance the ability to map 3D structures such as forest canopy, tree crowns, and other vegetation features [300].

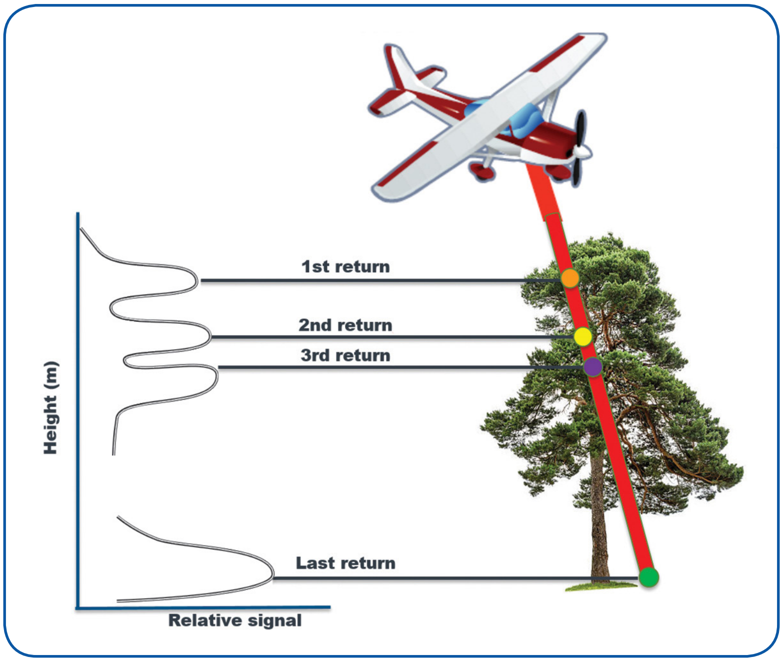


Figure 214. Multiple returns from a single laser pulse in airborne LiDAR [302].

3. Laser Pulse Width

The laser pulse width (see Chapter 1, Section I.A.5 for details on pulsed lasers) determines the range or vertical target resolution. This resolution (δ_R) can be determined by $\delta_R = c\tau/2$, where c is the speed of light in vacuum and τ is the pulse width (duration). For example, when $\tau = 2$ ns, $\delta_R = 30$ cm, which means that the distance between two objects must exceed this value to be accurately identified as separate targets. Current LiDAR sensors use laser pulse widths in the 2-5 ns range, mainly limited by the bandwidth of the receiver. The vertical range resolution can be significantly improved by full-waveform (FW) LiDAR systems, which image a scene by emitting laser pulses in a particular direction and capturing the entire temporal envelope of each echo [303, 304]. In such a method, FW systems capture more detailed physical information and characteristic properties of the 3D scenes compared to conventional LiDAR systems. However, the collected datasets are very large as there is a need to record the entire digitized backscattered laser pulses with a very high sampling rate (1-2 GHz).

4. Laser Power and Beam Divergence

The maximum distance from which data can be measured is generally important to LiDAR users. Factors affecting the maximum range are laser peak power (see Chapter 1, Section II.E for details), target surface diffuse reflectance, and the amount of ambient light coming from the target surface. The reflected laser power must be sufficient to overcome the detector's SNR and trigger the pulse detector. Detectors typically have some limiting threshold that is set to mask out noise from ambient light. Therefore, for high altitude mapping applications, there is a need for lasers that can generate high peak power pulses (in the tens of kW range) over a wide range of repetition rates (hundreds of kHz range).

The LiDAR minimum spot/footprint size at the target region is directly related to the flying height above the ground and the laser beam divergence [299]. For example, for an aircraft height of 1000 m above ground and a beam divergence of 0.3 mrad, the spot size on the ground would be 30 cm. The same spatial resolution can be achieved (while doubling the throughput) if the plane height is increased to 2000 m and the beam divergence is reduced to 0.15 mrad. These relationships typically result in trade-offs.

For instance, higher elevation requires larger pulse repetition rates (to maintain spatial resolution), but this typically results in a reduction in pulse energy which reduces system SNR. Moreover, lower divergence values require larger beam expansion optics due to the laser brightness limitation known as the etendue conservation law [305]. Modifying the LiDAR transmission optics provides control over the target spot size and the laser radiance on the target surface. The latter affects the SNR because of the amount of reflected signal compared to the ambient sunlight illumination.

5. Other Laser Parameters

Additional key laser parameters that affect the LiDAR performance include the laser spectral width, wall-plug efficiency, and system footprint and weight.

Spectral Width

As discussed above, 1064 nm is the most common wavelength for airborne LiDAR systems due to the laser/detector availability and high reflectance from common targets. One major limitation of this wavelength is the background noise created by the spectral irradiance of the sun. To improve the SNR in this wavelength regime, LiDAR receivers employ a narrow bandpass filter. It is therefore essential that the spectral width of the laser be sufficiently narrow (e.g., < 0.1 nm). Narrowband filters used in airborne LiDARs are often based on thin-film coatings which have proven to be robust enough to withstand the broad (and often harsh) set of environmental conditions.

Implementing these filters is important when working with narrow linewidth master-oscillator power-amplifier (MOPA) based pulsed lasers. In this MOPA architecture, the central wavelength is determined by a seeding laser source (often a laser diode) which determines the spectral characteristics of the emitted pulse. The linewidth can be tailored by choosing the right laser diode architecture and designing the subsequent amplification chain accordingly. The central wavelength is affected by the thermal conditions of the diode, such as temperature stabilization, drive current (see Chapter 1, Section I.B.4 for details), and the pulse repetition rates. This can make wavelength control difficult, therefore the bandwidth of the filter must be made wide enough to accommodate wavelength shifts.

This requirement suggests that the bandpass filters used by LiDAR detection systems must be designed with a broad and uniform transmission spectrum. Additionally, sharp spectral edges are desired to maintain narrowband operation and optimally reduce spectral noise. By utilizing multilayer thin film coating techniques (see Chapter 1, Section III.A.2 for details on optical coatings and filters), it is possible to achieve these characteristics to optimize system performance. In typical applications, the laser is required to deliver over 90% of its pulse energy within a specified bandwidth of < 2 nm.

Efficiency, Footprint, and Weight

LiDAR systems are mounted on aircraft, UAVs, and even drones. Therefore, it is crucial to keep the laser source as compact, lightweight, and as efficient as possible. It also needs to be robust enough and perform under constantly changing environmental conditions, e.g., temperature, humidity, vibrations, shocks, repeated takeoff and landing events. This combination of parameters limits possible laser sources and makes fiber lasers a natural candidate. Their modularity, scalability, high efficiency, and inherent robustness make fiber lasers attractive for LiDAR applications over bulk solid-state laser systems.

B. LiDAR Applications Using MKS Products

1. MKS Laser Sources for LiDAR Applications

The MKS Spectra-Physics VGEN-SP-NL lasers incorporate state-of-the-art fiber laser technology to provide a broad range of performance specifications to meet the highly-demanding LiDAR requirements with respect to wavelength, operating altitude, target spot size, scan rate, and range accuracy. The key laser parameters and corresponding LiDAR attributes are described in Table 15.

Laser Parameter	Value Offered by MKS Products	Key LIDAR Attributes
Wavelength	1064 nm 532 nm 1550 nm	• 1064 nm for various applications (eg., wide-area/urban/corridor mapping) • 532 nm for Marine/bathymetric surveying • 1550 nm for eye-safe applications
Repetition Rate	Up to 2 MHz	High spatial sampling rates
Pulse Width	< 3 ns	High resolution
Avg/Peak Power	Up to 20 W/40 kW	Maximum distance, sampling rate
Beam Divergence	< 0.3 mRad	Minimum spot size
Line Width (FWHM)	< 0.1 nm	High SNR
Wall Plug Efficiency	> 20% for 1064 nm > 15% for 532 nm > 8% for 1550 nm	Suitable for various airborne platforms
Weight	< 3 kg	

Table 15. Spectra-Physics laser specifications for LiDAR applications.

The combination of high average and peak power levels over a wide range of repetition rates (see Figure 215), coupled with short pulse widths, small beam divergences, and narrow linewidth capabilities, offers full flexibility for LiDAR systems. Such lasers enable state-of-the-art performance with respect to data acquisition rate versus flying altitude, accuracy, and outstanding point density for various airborne LiDAR applications. Since most LiDAR applications work with a fixed pulse width, higher peak power translates to large pulse energy. Power scaling of these pulsed fiber lasers increases the throughput by extending these large pulse energies to higher pulse repetition rates.

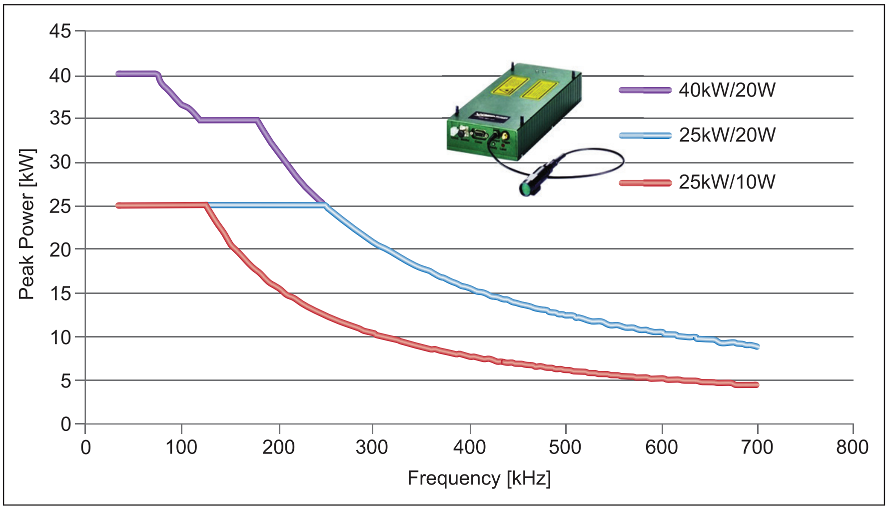


Figure 215. Typical peak power vs. repetition rate for various Spectra-Physics lasers operating at 1064 nm. The legend gives the maximum peak power followed by the average power for each laser system.

2. Applications of LiDAR

The accuracy and detail provided by LiDAR makes it an essential component of many applications that help local authorities and other land management organizations in their areas of responsibility. Below are examples of LiDAR applications that use MKS Spectra-Physics lasers.

Wide-area, Corridor and City Mapping, and Urban Planning

Accurate terrain mapping is important for highway corridor planning and design, environmental impact assessment, and infrastructure asset management. LiDAR data can be used to create highly detailed digital city models and identify areas such as new buildings and extensions that were not previously captured on base maps. This can be critical for emergency services and building inspectors that need to have access to the most current information available. Figure 216 are examples of wide-area terrain and city mapping detected from airborne LiDAR.

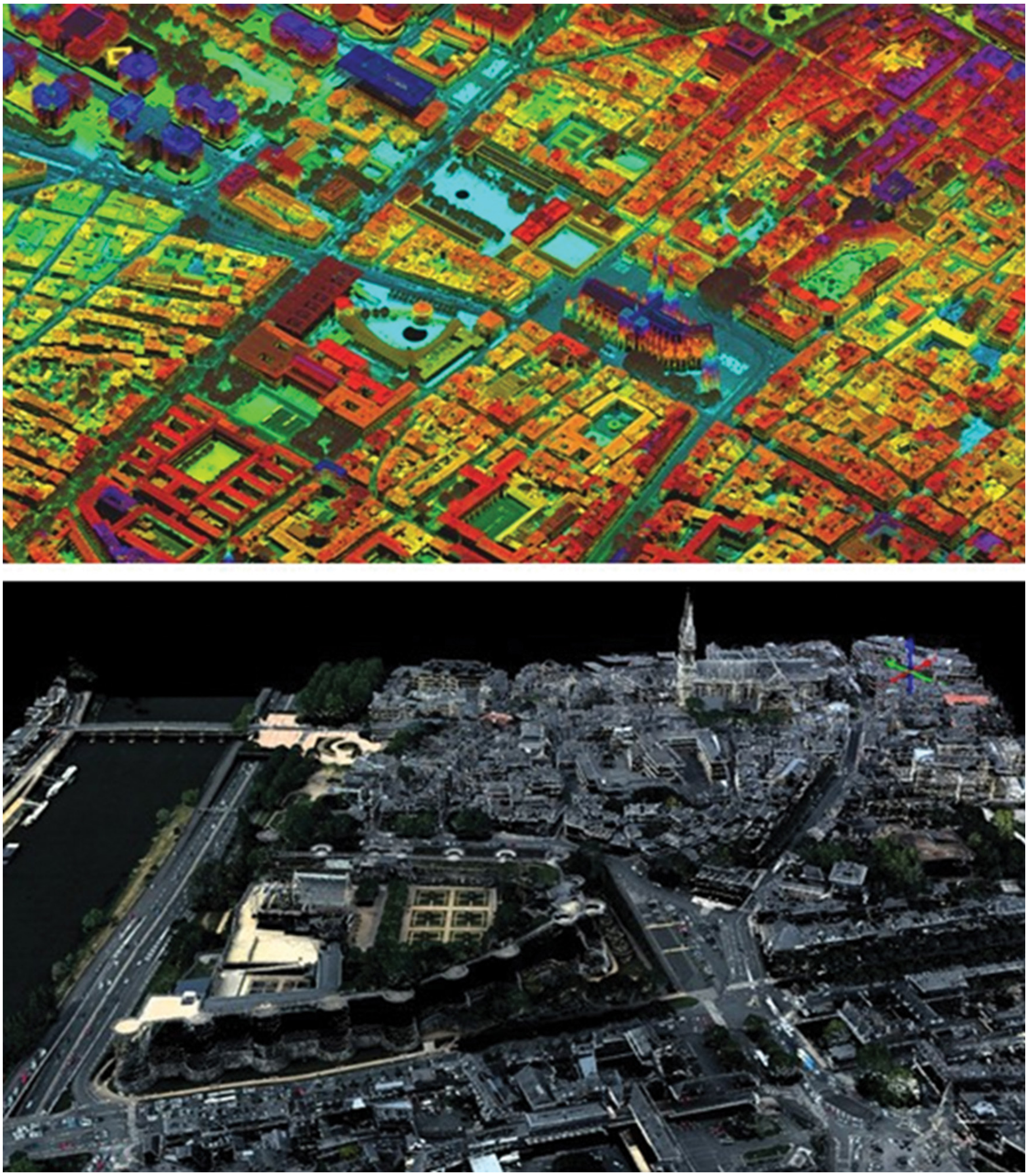


Figure 216. LiDAR-based urban mapping using Spectra-Physics lasers. Images courtesy of Leica Geosystems (top) and Teledyne Optech (bottom).

Vegetation Mapping

Accurate vegetation mapping is critical for natural resources management, forest inventory, ecological analysis, fire management planning, and conservation planning. LiDAR is unique in its ability to measure the vertical structure of forest canopies and individual tree crowns. Figure 217 shows vegetation mapping detected from airborne LiDAR. The data was acquired with a dual channel instrument employing 1064 nm and 532 nm lasers concurrently. By combining the calibrated reflectance values from the respective channels, the determination of additional attributes like tree species or natural versus man-made objects is supported.

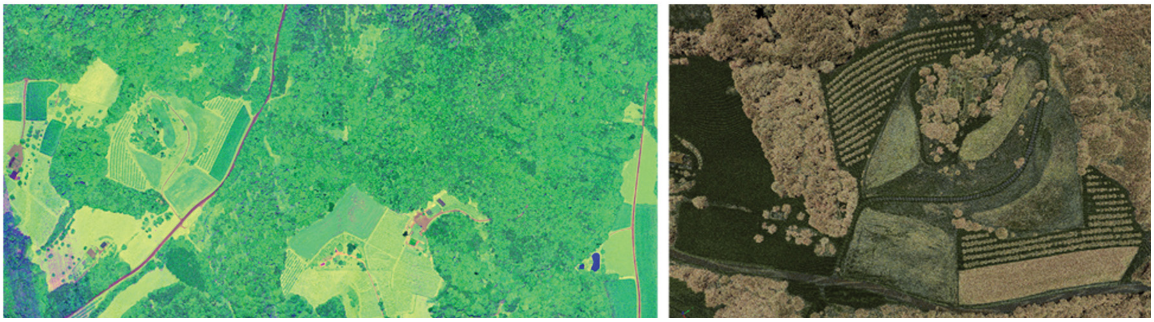


Figure 217. LiDAR-based vegetation mapping using Spectra-Physics lasers. Images courtesy of Riegl Laser Measurement Systems GmbH.

Power Lines

Airborne LiDAR is a proven technology for providing accurate elevation models for transmission lines, allowing utilities to measure the shape of the ground below the transmission line, the position of the towers and poles, the sag on the wires, and the up-growth of any vegetation or other types of incursions. Figure 218 shows an example of power lines detected from airborne LiDAR.

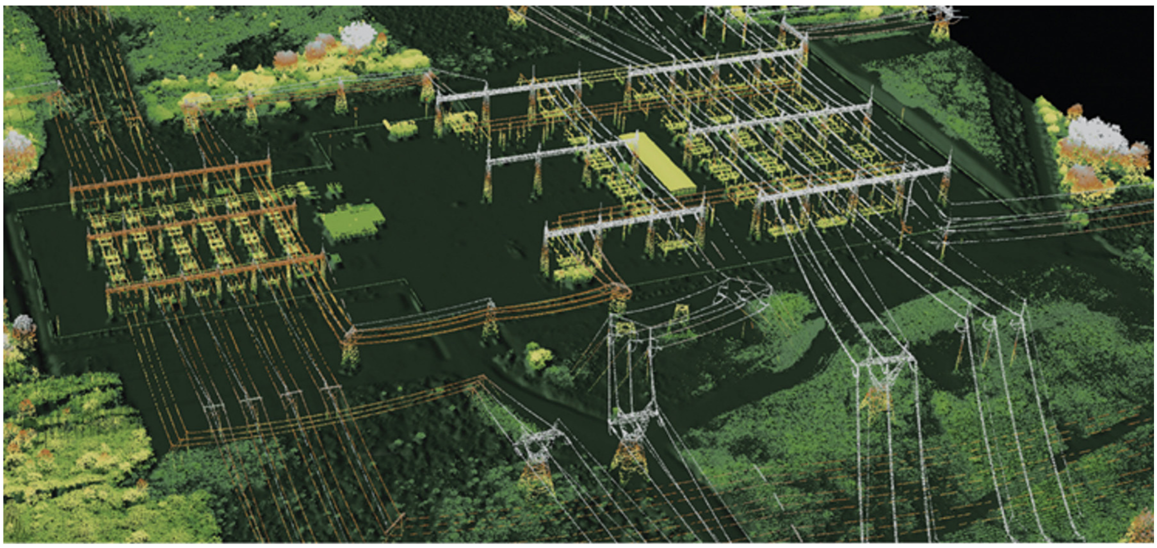


Figure 218. LiDAR-based power line mapping using Spectra-Physics lasers. Image courtesy of Teledyne Optech.

Coastal Mapping

Airborne LiDAR bathymetry has proven to be an efficient tool to monitor coastal environments. Important benefits of this technology are the ability to survey seamlessly across the land-sea boundary and to map extremely shallow areas with complex and irregular topography. An example of coastal mapping detected from airborne LiDAR is shown in Figure 219.

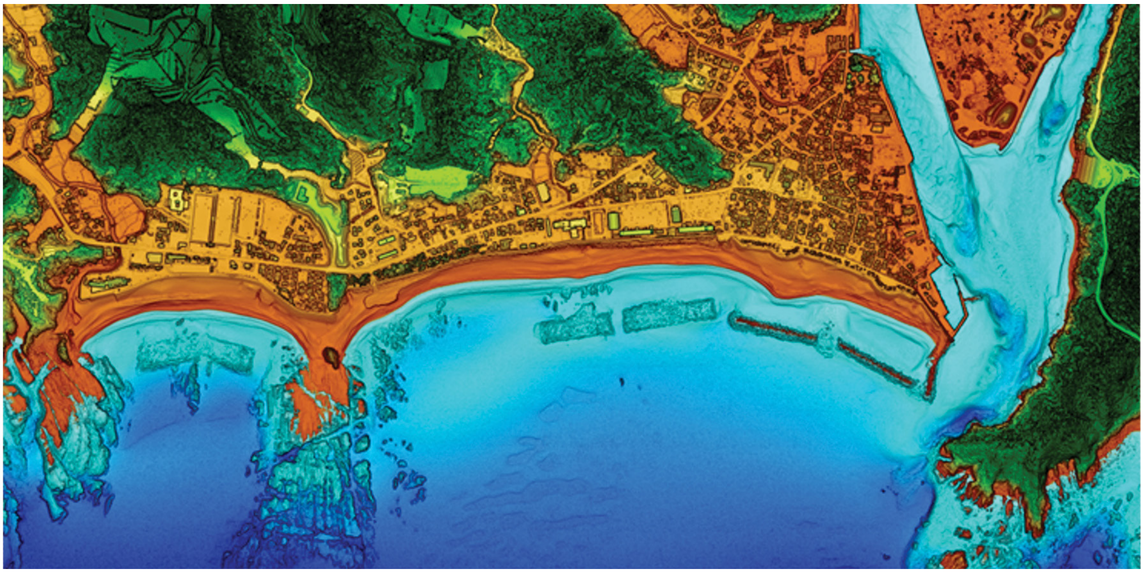


Figure 219. LiDAR-based coastal mapping using Spectra-Physics lasers. Image courtesy of Leica Geosystems.

Archaeological and Historical Surveys

Recently, several ancient civilizations' structures have been discovered using LiDAR, showing the added value of novel LiDAR mapping techniques for historical, archeological, and anthropological research. Among those discoveries are ancient Roman roads located in northern England and huge Mayan cities hidden beneath the dense jungles of northern Guatemala [306-308].

The discoveries in the jungles of Central America that were made possible through novel LiDAR 3D terrain mapping tools revealed a huge number of Mayan structures buried deep beneath the jungle of Guatemala's Petén region. In addition to hundreds of previously unknown structures, the LiDAR images show raised highways connecting urban centers and quarries. Complex irrigation and terracing systems supported intensive agriculture capable of feeding masses of workers who dramatically reshaped the landscape. This discovery suggests that this 1200-year-old civilization was more advanced than previously imagined and was comparable to sophisticated cultures such as those in ancient Greece or China.

C. Future Directions

1. General Trends

LiDAR is a rapidly growing technology that continues to advance in terms of power, accuracy, and speed. Such advancements should result in new application spaces for LiDAR technology. For example, reducing the cost-per-point through higher effective pulse rates is the best way to address large-area, high-point-density projects in the future. Most recent developments in LiDAR technology aim at increasing the pulse repetition rate either for a single scanner or by operating two scanners simultaneously. MKS Spectra-Physics lasers currently offer state-of-the-art performance with up to 2 MHz pulse repetition rates, while next generation LiDAR will require even higher repetition rate capabilities.

In addition to higher repetition rates, LiDAR applications will require higher power lasers while maintaining near-diffraction-limited beam quality to achieve higher altitudes and acquisition rates. Furthermore, shorter pulse durations (~1 ns) will be necessary to achieve higher precision. Together with improvements in pulsed fiber lasers, airborne LiDAR will continue to show remarkable advancements resulting in positive changes for various industries, such as increased efficiency in resource management, more effective infrastructure planning, and better preparation for natural disasters.

2. Pulse-on-Demand Capabilities

Designing lasers with forward-thinking features which emerge from the requirements of LiDAR mapping concepts has proven to be an excellent means of paving the way for more advanced and efficient LiDAR systems. This understanding keeps the MKS Spectra-Physics fiber laser designs for LiDAR applications aligned with future market demands. An example of this is the Pulse-On-Demand (POD) capability, which enables the pulse repetition rate to be externally controlled and modulated to match the scanning method with practically zero delay time. Emitting pulses on-demand while maintaining fixed pulse energy (denoted as the E-Pulse™ feature) enables the data acquisition system to work with a predetermined detection threshold, thereby keeping SNR optimized.

Incorporating POD capabilities enables the laser to be rapidly gated, permitting the system to turn the laser emission on and off while maintaining a fixed pulse energy. This allows the scanner full control over emission periods. Therefore, the system can be operated without laser emission over unwanted ranges, reducing the risk of regulatory eye-safety issues. Other POD capabilities include future enhancements, such as utilizing frequency sweeping and chirping for dynamically mapping moving objects. Previous work has shown that by using such frequency shifting techniques, it is possible to map moving objects [309].

3. Pulse-Shaping Capabilities

Other features which could receive further development include pulse energy modulation (while keeping the pulse repetition rate constant) or even intra-pulse waveform shaping [310]. For example, work conducted at the Naval Research Lab has shown that a modulated pulse LiDAR system can improve search rates by a factor of 10 to 100 and, for underwater layer mapping, the performance improvements include high resolution depth imaging and contrast enhancement of the same order. The topics above represent just a few examples of the close collaboration between MKS Spectra-Physics lasers and LiDAR customers. This collaboration enables MKS to supply tailor-made solutions for the rapidly-growing demands in this challenging field and to remain the leading provider of lasers for this market.

This page intentionally left blank.

Chapter 5

Photonics in Communications

This page intentionally left blank.

I. Introduction

A. Interconnectivity and Data Communications

There has been a tremendous increase in the volume of internet-based digital data communications over the past decade, with global cloud data center traffic reaching 8.2 zettabytes (8.2 x 10^{21} bytes) in 2017 alone. This is the equivalent of streaming 2.7 trillion hours of high-definition movies. Figure 220 shows the demand for internet data communications is expected to increase to approximately 20 zettabytes by 2021.

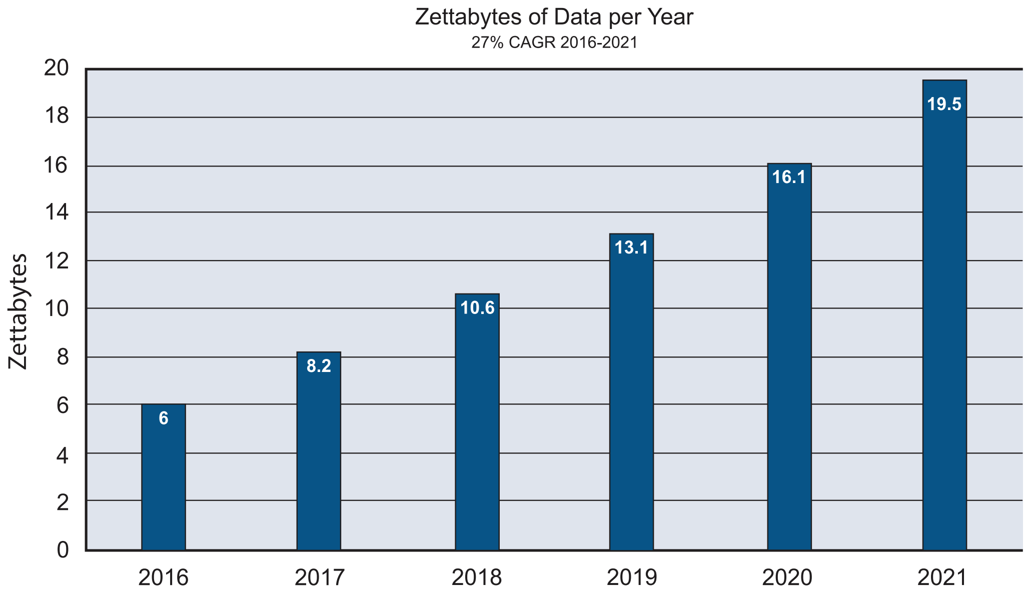


Figure 220. Global Data Center IP Traffic Growth (source: Cisco Global Cloud Index, 2016-2021 [311]).

Both consumer and non-consumer demand are driving this growth. On the consumer side, internet traffic arising from social media, e.g., Facebook, Twitter, Snapchat, LinkedIn®, etc., has exploded, as has the use of streaming media platforms such as NETFLIX®, Hulu®, YouTube®, and others. E-commerce platforms like Amazon®, Alibaba, Ebay® and others have gained significant acceptance in a very short period. Furthermore, day-to-day business activities, as they evolve to leverage the internet, are contributing to increased internet traffic via services such as Amazon Web Services, Microsoft®, Google®, and Baidu. Internet traffic originating from non-consumer sources is also growing and has the potential to surpass consumer-generated traffic. Non-consumer traffic includes sources such as Big Data and Artificial Intelligence (AI), e.g., IBM, Google, Amazon, Microsoft, and the Internet of Things (IoT). The IoT is expected to contribute a massive volume increase in internet traffic. The IoT is the network of physical devices, vehicles, home appliances and other items with embedded electronics, software, sensors, and actuators. The rapid growth of IoT is due to the fact that internet-based connectivity between devices such as smart home appliances and smart sensors enables automated control that will significantly improve the operational efficiency and safety of these devices. For instance, smart cars are projected to cause far fewer traffic accidents than human-controlled vehicles. Another example of the way that internet connectivity between devices can improve human safety is provided by the systems used for earthquake warnings. These systems currently offer about a 10-second warning prior to the arrival of an earthquake. This allows a signal to be broadcast over the internet to shut down facilities such as electricity, gas, and water in the affected area. Since most of the earthquake damage is caused by fires and flooding, the application of IoT for such safety protocols can reduce injury and damage.

B. Optical Fiber Networks

The obvious advantages provided by the interconnection of different devices has created demand for more data centers and greater bandwidth to handle high-speed internet data transfer. As this demand has increased, older technology based on copper wires has reached the limits of its speed and bandwidth. As a consequence, data transfer technology has migrated to optical fibers for digital data communications networks that employ optical transceivers (Figure 221) to generate and receive signals. Optical fiber networks offer several significant advantages:

Figure 221.
An optical transceiver.

- Bandwidth — Optical fibers have much higher bandwidths than copper wire, thus possessing a greater capacity for data transmission.
- Transmission speed — Optical fibers transmit data at nearly the speed of light, much faster than electronic transmission in copper wire.
- Lower signal loss — Signal attenuation over a given transmission distance is dramatically reduced with optical fibers. Copper cables cannot transmit information beyond two miles due to power loss, whereas optical transmission through fiber cables can reach nearly 25 miles [312].
- No electromagnetic interference (EMI) — Moving electrons produce an electromagnetic field that can interfere with network signals. Photons in a fiber do not interfere with other networks.
- Cost — Fewer repeaters are required and more bandwidth can be fit into existing channels.

The clear advantages of the optical fiber network have resulted in more and more data communications networks employing this technology, especially those networks and data centers that require the highest speed data transfers. Currently, most networks with data transfer rates of 10 Gbits or higher use optical fiber networks exclusively. The basic elements of such a communication system include: a source of light (photons) such as a laser diode; an optical fiber (or waveguide) to contain and physically guide the light; an optical amplifier; and a transceiver. The physical mechanisms governing the operation of these elements can be found in various sections of Chapter 1: Section I.B.1 describes laser diodes, Section III.D discusses fiber optics, and Section II.C details photoreceivers. Lasers having different wavelengths (and other attributes) are available and the choice of light source should be optimized for a given application. The chosen wavelength of light must be compatible with the optical fiber to enable high speed, low loss (highly efficient) signal transmission. For example, light having a wavelength of 1.3 μm or 1.55 μm is commonly employed in optical communications because optical glass has very low absorption at these wavelengths. Finally, the system must be highly reliable, stable, repeatable, and long lived.

Bandwidth is often the limiting factor in data communication networks. Consequently, users are continually exploring different options to increase it. The four most common strategies to increase bandwidth are, in order of decreasing incremental cost:

- Increase the number of optical fibers
- Increase the number of channels — In dense wavelength division multiplexers, additional channels require more wavelengths and therefore more lasers
- Increase the data rate — This complicates signal processing and requires expensive components and assembly
- Increase the bits per symbol — This is the most economical option. Complications to signal processing can be minimized by maintaining low noise levels. This is currently achieved by modulating the signal with four levels (referred to as Pulse Amplitude Modulating – level 4 or

PAM – 4) rather than two (referred to as Non-Return Zero or NRZ). Figure 222 shows that PAM-4 sends twice as much data as NRZ.

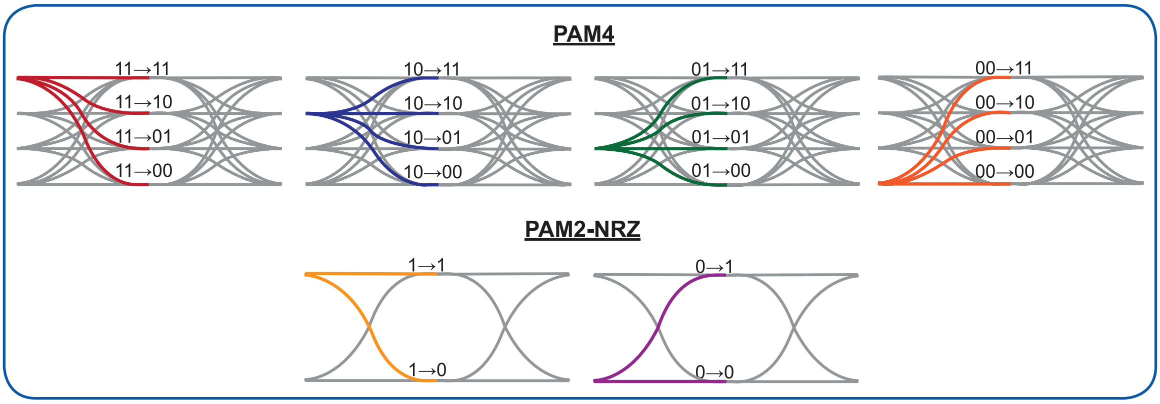


Figure 222. Eye diagram[1] comparing PAM4 signal properties with PAM2-NRZ [313].

Thus, increased bits per symbol provide the most economical route to increased bandwidth in a data center build-out or network expansion. However, all four strategies are being employed to satisfy the demand for increased bandwidth. The rapid pace of expansion and change in the industry is easily confirmed by an examination of the recent and projected sales figures for commercial optical transceiver technology. Figure 223 shows revenue tracking and forecast data for optical transceivers, which clearly reflects the rapid build-out of 100 GbE (Gigabit Ethernet, 1 × 10^9 bits per second) networks since 2015, with 200 GbE and 400 GbE build-outs expected to take off starting in 2019/2020 [314].

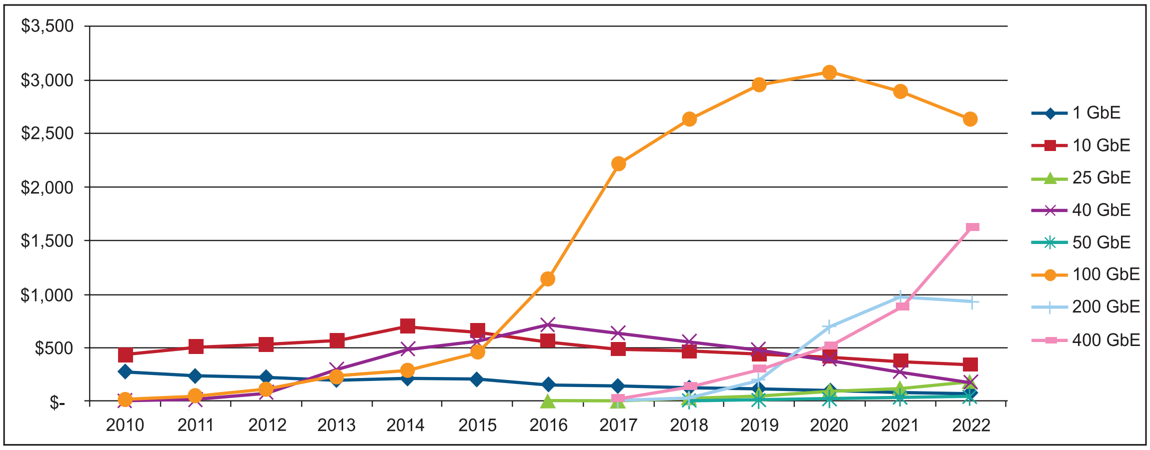
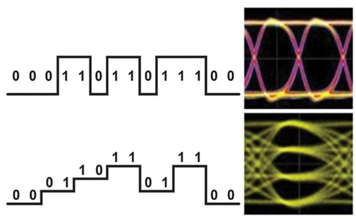


Figure 223. LightCounting forecast of optical transceiver revenue ($M) [314].

1 An eye diagram is the binary pattern that is generated while testing communication components, i.e., transceivers, on an oscilloscope at high speeds. Depending on the modulation type, the signal will have a series of 0's & 1's to be tested. For NRZ, the pattern is only one-dimensional, e.g. only 0 and 1. For more advanced modulation schemes, there will be two-dimensional binary data, e.g. for PAM-4 there are four levels of data: 00, 01, 10, 11.



Standards are also evolving to meet the needs of high-speed data communications. The Ethernet Alliance publishes a roadmap for these standards every few years and Figure 224 shows the rate at which new standards are being developed. It can be seen that the rate of development of Ethernet speed standards has accelerated threefold during the past two decades. This reflects both the rapidly increasing need for higher speeds and bandwidths and the greatly improved performance of optical technologies. The most recent standard, not yet widely adopted, is 400 GbE, published in 2017. The 800 GbE and 1.6 TbE standards are expected in the mid-2020's with adoption soon thereafter.

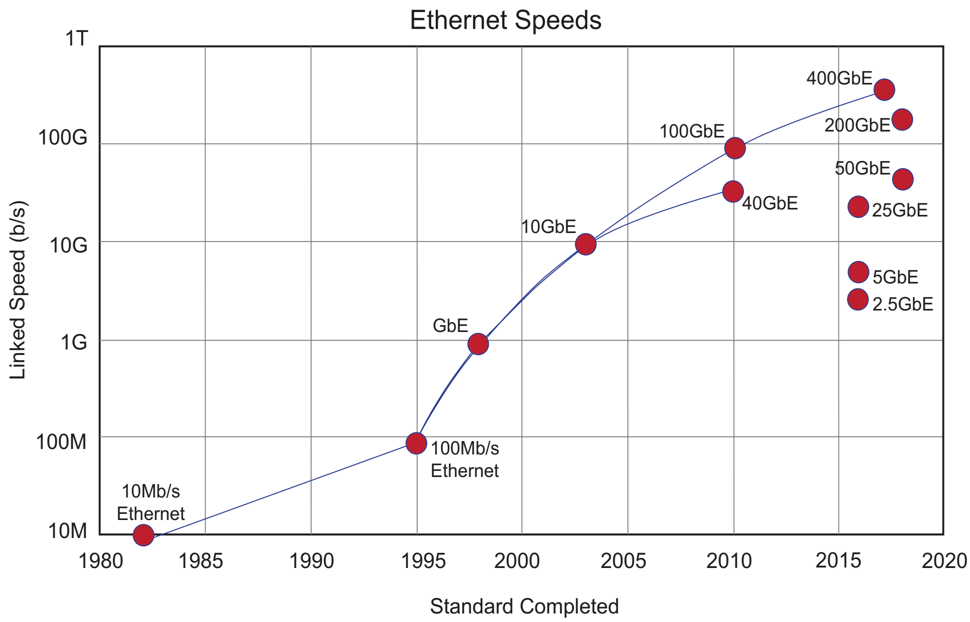


Figure 224. Development of Ethernet Speeds (image used with permission of Ethernet Alliance [315]).

The use of optical fibers has dramatically increased the transmission speed and bandwidth capabilities of digital communication systems. However, it is only possible to realize expanded capabilities when other components in the system can fully leverage them. The production of new generations of higher-speed, increased bandwidth optical fiber-based communications networks thus requires concomitant advances in the properties of other network components. These advances include more complicated form factors, advanced modulation schemes, faster optical emitters and receivers, denser wavelength division multiplexers, and significantly greater use of optical fiber.

The next generation of ultra-high bandwidth optical networks is expected to be the first to deploy silicon photonics (SiP) technology. SiP leverages semiconductor fabrication technology to integrate electronics, lasers, modulators and the receiver at the chip-level of an SiP optical transceiver. This reduces interconnect lengths which lowers SNR and improves compatibility with other system components. SiP devices have demonstrated a more than four-fold increase in optical lanes with the same form factor as older, non-photonic technology.

II. Laser Diodes and Laser Testing

Laser diodes lie at the heart of many optical communication systems. Lasers that are included in communication systems must necessarily undergo performance characterization, control, and analysis. The degree to which these devices are characterized varies, depending upon whether they are in development, undergoing validation, or have entered the mass production stage.

A. Characterization, Control, and Analysis

1. Characterization

Laser diodes are characterized by their inherent physical properties within five distinct categories, including: electrical, spatial, spectral, optical, and dynamic properties. For a detailed discussion of these physical properties of laser diodes, the reader is referred to Chapter 1, Sections I.B.2 and I.B.3. The combination of properties within these categories determines the performance and effectiveness of a given laser diode in terms of operational characteristics such as:

- Threshold current – the current at which the device begins to lase
- Output power – the average output power generated by the device
- Peak wavelength – the wavelength at which the greatest output power is observed

Prior to integration within an optical system, the device design incorporating a laser diode must be validated and the properties of the proposed laser diode have to be analyzed using appropriate instrumentation. The available tools for testing a laser diode within the categories identified above include:

- Laser diode drivers, TEC controllers, and laser diode controllers (see Chapter 1, Section I.B.4)
- Optical power meters and detector heads (see Chapter 1, Section II.A.3)
- Microscopes and vision systems
- Optical spectrum analyzers, wavelength meters or spectrometers (see Chapter 1, Section II.D.2)
- Laser beam profilers (see Chapter 1, Section II.B.2)

2. Control and Analysis

Figure 225 shows the MKS ILX Lightwave® LDC-3726 laser diode controller and ILX Lightwave® OMM-6810/OMH-6724B optical power and wavelength meter. They allow a user to control, manipulate, and measure the performance of an individual device. Using these instruments, laser diode output characteristics such as changes in output power and wavelength can be analyzed in a relatively simple setup. The LDC-3726 controller has unparalleled device protection, LIV (light, current, voltage) measurement capability, and repeatable output(s). It provides precise temperature control, which greatly improves the ability to analyze thermal effects and offers low noise current control up to 500 mA, optimal for telecommunications applications. The OMM-6810B/OMH-6742 power and wavelength meters have precise device output power measurements, are capable of measuring both output power and wavelength in a single instrument, and have remote control capability for automated test setups. Typical applications for these controllers include R&D of new products and validation of existing products under various conditions.

Figure 225. Examples of instruments used to control and analyze the properties and performance of laser diodes. An ILX LDC-3726 laser diode controller for providing current and temperature control (left); an ILX OMM-6810B power meter and OMH-6742B detector head to measure device output power and wavelength (right).

B. Reliability and Life Testing

Life test studies are used to collect laser diode lifetime data under carefully controlled operating conditions. These data are then used to develop statistical models that can predict the laser lifetime under the intended operating conditions. Within the telecommunications industry, standards for life test studies are developed and promulgated by Telcordia Technologies [316, 317]

Laser diode life test studies are type- and application-specific. They require the periodic measurement of a variety of device parameters including operating current, optical output power, threshold current and forward voltage under accelerated aging conditions. Accelerated lifetime testing can be accomplished by using higher settings for temperature, injection current or optical power. However, temperature acceleration is most common.

1. Life Test Modes

The various modes for life test studies are listed below:

- Constant current aging or ACC mode (automatic current control) — The laser current is held constant for the duration of the test.
- Constant power aging or APC mode (automatic power control) — Laser output power is held constant by continuously adjusting the current to maintain constant output power. Output power is measured either with an external photodetector or using an internal monitor photodiode if one is available within the laser package. Constant power aging is the life test mode most often used since it closely resembles the typical operational mode of a laser diode.
- Periodic sample testing — When lifetime testing laser diodes at a higher temperature, typical procedures burn the diode in at constant current and high temperature, then reduce the temperature of the laser to evaluate relative changes in output (generally, laser diodes do not lase at temperatures greater than 100°C). In this type of test, lasers are operated in constant current mode during the high temperature aging. The sampling interval may be varied over the duration of the very long-term test to reduce the amount of data collected. In these cases, sample measurements may be obtained every hour at the beginning of the test period and every few days after the test has been running for months.

2. Life Test Systems

Many modern reliability and life test systems are designed to house different devices and package types within a single system, thus facilitating faster testing, increased throughput, and captive control. Important considerations for choosing a laser diode life test system include:

- Device protection — Reliability test systems, such as the ILX Lightwave Sentinel series LRS-9434, are designed to ensure that the DUTs are protected from damaging surges/ transients, overcurrents, reverse currents, etc.
- Current control — Perhaps the most important characteristic of laser diodes is their efficiency – the degree to which they emit light when current is injected into the device. Therefore, a device under test must be subjected to a precisely known and repeatable current. This is most easily guaranteed when each device is driven by an independent current source.
- Temperature control — Laser diode properties such as threshold current are very sensitive to changes in temperature. For example, a GaAlAs laser diode operating at wavelengths around 850 nm experiences threshold increases of ~1% for every 1°C rise in temperature, while an InGaAsP laser operating at wavelengths around 1300 nm sees threshold increases of ~2% for each 1°C rise in temperature. For this reason, device reliability tests require precise and repeatable

control over temperature as well as good temperature uniformity. TECs, water cooling, air cooling and embedded temperature sensor arrays are some of the methods used to control temperature in laser diode test systems such as the MKS ILX Lightwave Sentinel series.

- Output power — A device's reliability can be evaluated by monitoring its output power along with changes in the output power over time and under different conditions. Reliability and test systems such as the ILX Lightwave Sentinel series LRS-9434 estimate reliability by monitoring device output power with either external detectors or the device's internal monitor photodiode.

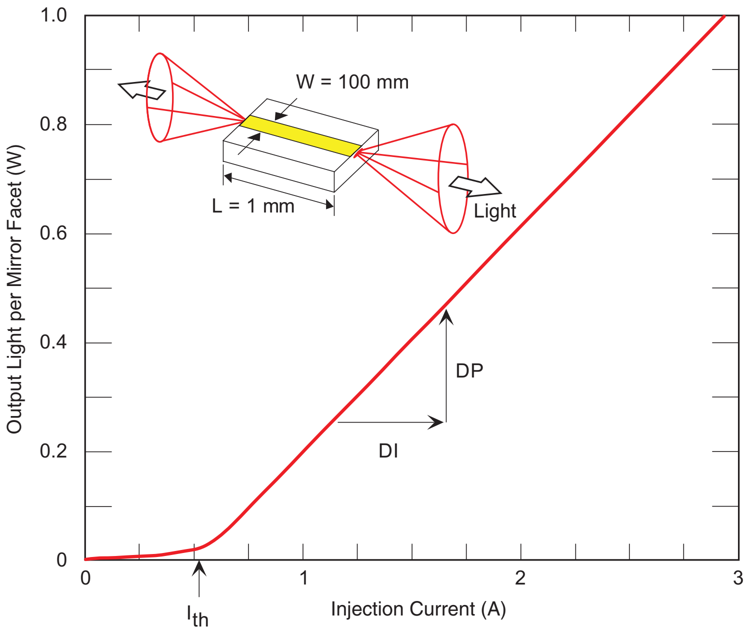


Figure 226. The threshold current is the point along the LI curve where the device starts to emit light or begins to lase. According to the Telcordia standards, it can be calculated by three different methods: second derivative, first derivative, and two-segment fit algorithms [63].

- LI curves — The LI curve (Figure 226) is one of the most useful tools for determining a device's performance and performance changes that occur over time or in relation to other devices, i.e., comparative LI curves. As such, this is one of the key graphs produced by test systems such as the ILX Lightwave® Sentinel series (Figure 227).

- Repeatability — Repeatable performance is an absolute requirement for laser diodes. Thus, life test systems must exhibit minimal variations in input currents, output power measurements, and temperature control to ensure that device changes or failures are attributable to the device rather than to inconsistencies in the test and analysis equipment.

Figure 227. ILX Lightwave LRS-9434SS Device Carrier.

- Modularity — Test systems that are modular in design are more adaptable to the many variables and modifications associated with laser diode reliability and life testing – such as differences in package type, pinout, number and type of electrical signals required, higher or lower temperature requirements, and differing output powers associated with different devices. Separating modules for different aspects of the test enables flexibility. Modular test systems facilitate troubleshooting, streamline repairs, and updating the functionality of the system. The separate modules in a modular life test system include: device carriers to house different devices, electronics to provide current and temperature controls, and software to monitor performance and modify test parameters.

- Scalability — It is advantageous to select a life test system that scales with increased production and customer growth. It is not uncommon for production and test requirements to double or even triple; the ability to expand an existing life test system to meet increased demand can save significant time and expense. Figure 228 illustrates equipment scalability in a life test system.

Figure 228. The ILX Lightwave Sentinel series single shelf LRS-9434SS (left) and Sentinel series LRS-9434 (right) illustrate scalable reliability and life test systems. Each system is capable of the same performance but with different footprints and capacities – up to 4 independent device carriers and 128 devices in the LRS-9434SS and up to 44 device carriers and 1408 devices in the LRS-9434.

C. Production Burn-In Systems

High temperature burn-in screening is used in laser diode manufacturing to screen out devices that are likely to have unacceptably short lives and to ensure that the remaining population of lasers will have a statistically acceptable level of reliability. Laser diode reliability may be broadly defined as the ability to operate the device satisfactorily in a defined environment for a specified time. Many of the issues related to laser diode reliability are revealed by the hazard rate characteristic curve for a population of lasers, also known as a “Bath Tub Curve” (Figure 229). Hazard rate is defined as the probability of failure per unit time given that the device has survived until that time [318]. The various hazard rates associated with laser diodes fall into three categories. First, infant mortality failures are caused by defects that were introduced during the manufacturing process or by intrinsic semiconductor defects. Second, external factors such as current surges and electrostatic discharge events create a constant hazard rate over the lifetime of a device. Finally, wear out failures in lasers occur that are usually caused by the growth of non-radiative, optically absorbing defects within the active region of the laser.

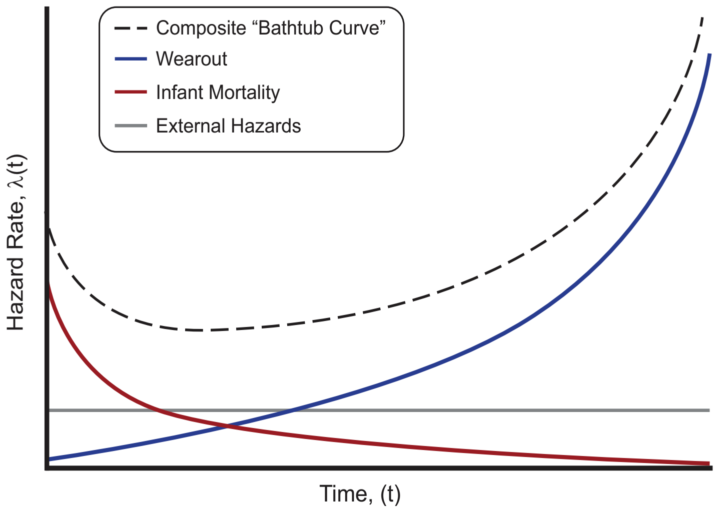


Figure 229. Hazard rate characteristic curve for unscreened laser diodes [318].

In the screening process, devices are evaluated for changes in one or more key operating parameters that are measured before and after high temperature burn-in. Commonly-measured operating parameters and screening criteria are shown in Table 16.

Operating Parameter	Symbol	Typical Screening Criteria
Threshold current	I_{th}	Change > 5 to 30%
Optical output power at specified operating current	P_{op} @ I_{op}	Change > 5 to 30%
Current required to achieve specified optical output power	I_{op}@P_{op}	Change > 5 to 30%
Slope efficiency	η	Change > 5 to 30%

Table 16. Commonly measured operating parameters and screening criteria [318].

The challenge for production burn-in testing is to achieve high throughput and failure identification at very low cost. Important considerations when choosing a production burn-in system include:

- Device protection — As with reliability and life test systems, device protection is a critical element of a production burn-in test system. System designs should ensure that devices are protected from damaging surges/transients, overcurrents, reverse currents, etc.
- Current control — To effectively test and stress devices, it is imperative that a precise and repeatable current is provided to the devices under test. This can be done by driving devices individually or in series. Systems such as the ILX Lightwave® Spartan series LTS-7410 use a series drive configuration to reduce cost, increase capacity, and improve testing throughput. Other systems utilize parallel drive to control devices individually, which simplifies failure identification to the individual device level.
- Temperature control — Temperature control, temperature uniformity, and temperature stability are all critical during device burn-in. In many instances, water cooling is needed.
- Modularity — The many variables and modifications associated with production burn-in make it advantageous to have separate modules for unique functions. Modular flexibility in the design facilitates troubleshooting and repair and ensures prolonged utility of the test system (Figure 230).
- Scalability — As with life test systems, a production burn-in test system that can scale with customer growth is desirable. As production ramps, a production burn-in test system that can easily add more capacity can save significant time and expense. Figure 231 shows a scalable production burn-in test system that can house up to 8960 devices.

*Figure 230.
Close-up of the
ILX LTS-7410.*

*Figure 231.
An example of the modularity
of the ILX Lightwave
Spartan Series LTS-7410
production burn-in
test system.*

III. Fiber Alignment

Precise fiber alignment is necessary for accurate and reliable data transmission in an optical network [319]. Most optical networks have many optical couplings and even minor (< 1%) losses at these couplings accumulate to produce significant signal loss and consequent problems in data transmission. Minimizing coupling losses is critical in these networks. Good fiber alignment produces the highest coupling efficiency and therefore the least signal loss before assembly or packaging of an optical system. Minimal signal loss results in reduced power requirements which, in turn, means fewer repeaters, lower investment costs, and reduced incidents of failure.

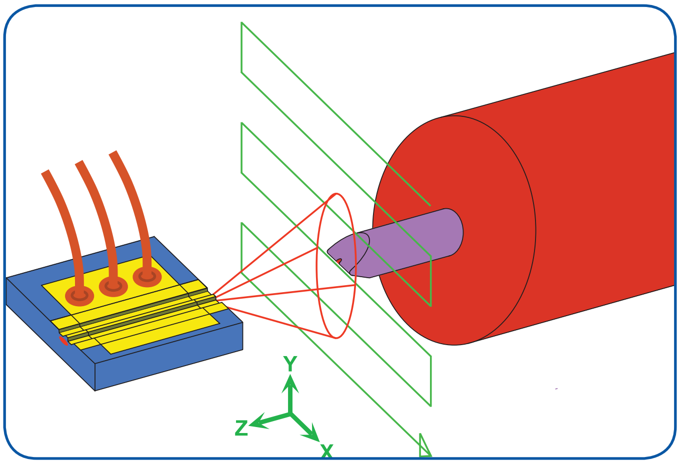


Figure 232. Coupling between laser diode and fiber [320].

The basic concepts of fiber alignment are illustrated in Figure 232. A well-characterized input beam (in this case from a laser diode) is coupled into the fiber under test and a raster scan of the fiber is used to detect first light - i.e., the output signal from the fiber indicating when the laser beam first enters the fiber. Once first light is detected, the position of the fiber is adjusted in a lateral, longitudinal, and angular coordinate system to locate the peak intensity of the output optical signal. In the simplest case, only lateral (X, Y) adjustments are necessary, while in multi-channel cases, adjustments to all six degrees of freedom (X, Y, Z, θ_x, θ_y, and θ_z) may be required (Figure 233). Effective fiber alignment requires the adjustment of several important motion parameters using a precision motion control device and an effective search algorithm suitable for a given application.

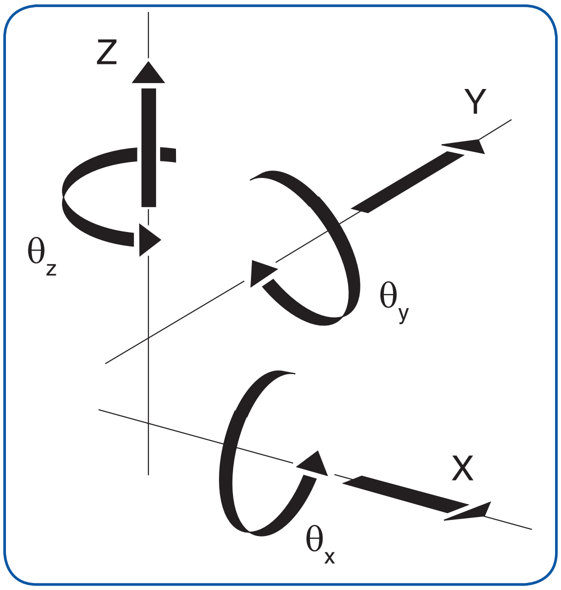


Figure 233. Motion degrees of freedom [214].

A. Motion Parameters

When using motion control systems for fiber alignment, the motion parameters considered for each axis critically affect the alignment process. The following are the primary parameters for consideration when selecting a motion controller for the location of peak power in fiber alignment procedures [321]:

- Minimum Incremental Motion (MIM) – This is the smallest increment of motion that a device can consistently and reliably deliver. It should not be confused with resolution, which is based on the smallest controller display value or smallest encoder increment. Rather, MIM is the actual physical performance of the controller that enables adjustment of the fiber position while searching for the position at which peak power is achieved. MIM of a motion controller can range from 100 nm to 1 nm. While a smaller MIM may align the fiber closest to the maximum peak power, this ability is achieved at significant costs in terms of alignment speed and power increments. XMS stages (Figure 234) are designed with optimized MIM and speed characteristics. They are capable of 1 nm MIM and 300 mm/s speed, making them ideal stages for alignment applications.

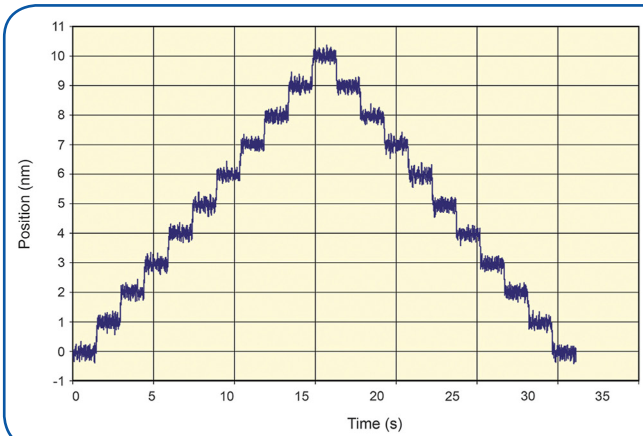


Figure 234. 1 nm MIM of an XMS linear stage (left) [214]; XMS50-S linear motor stage (right) [322].

- Repeatability – The repeatability parameter defines a motion control system's ability to repeatably position. It can be unidirectional (always approaching the target position from the same direction) or bidirectional (approaching the target position from either direction). Bi-directional repeatability typically ranges from 1 µm to a few nm in fiber alignment systems. This parameter is important for quickly finding the peak power location for similar device designs. The XMS stage shown in Figure 234 has 80 nm bi-directional repeatability.
- Position stability – Position stability is a measure of the motion system's capability to stay at a position within a defined window of time and error. Aligning fibers for assembly steps such as bonding relies on the positional stability of the fibers after the peak power has been located. Position stability requirements can range from 0.5 µm to a few microns. Figure 235 shows the step and settle performance of an MKS stage 250 ms after being moved. This stage exhibits less than 20 nm variation in position stability after settling.

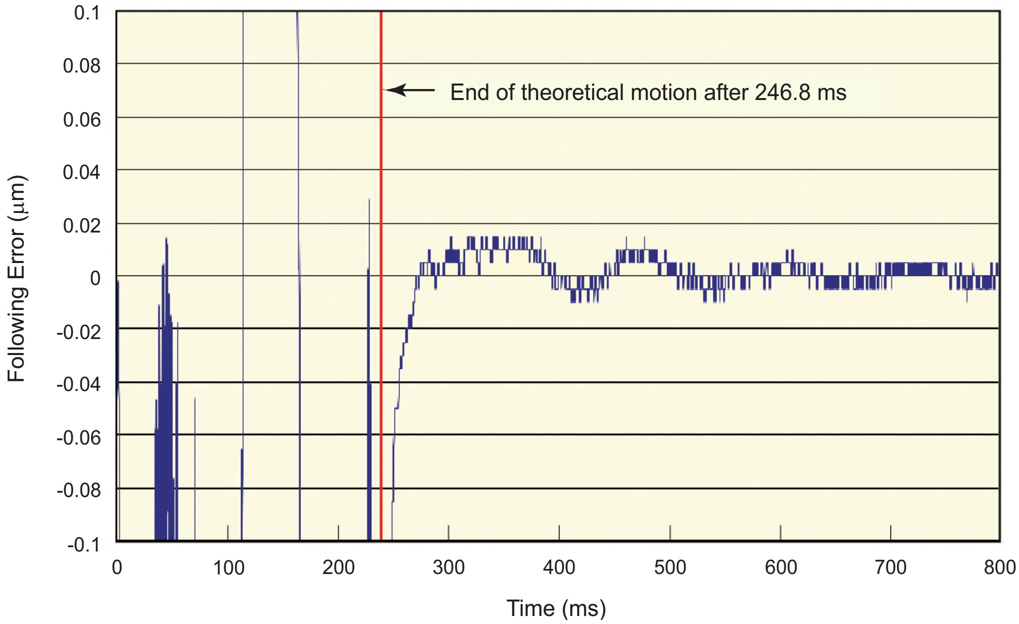


Figure 235. Step, settle, and stability at position.

- Other motion parameters — Other parameters that influence the effectiveness of a motion control system include: axis alignment, location of the gimbal point, system stiffness, pitch/yaw, thermal considerations, fixture design, Abbe error, etc.

Details regarding the fundamentals of motion control can be found in Chapter 1, Section III.F.1. MKS also has an available metrology primer that provides further information on this topic [214].

B. Search Algorithms

In addition to understanding critical motion parameters, efficient fiber alignment requires the selection of a positional search algorithm appropriate to the application and to the step in the alignment procedure. Specific search algorithms are available for finding the first light, i.e., the periphery of a light beam, after which different algorithms that are faster and more precise are used to find the peak power location. The choice of the second algorithm depends on whether the beam has a Gaussian distribution or top hat profile with multiple peaks. Some algorithms can be used to profile both types of beams and can also be used in parallel.

First light search methods include:

- Raster scan — This is the simplest search method. It scans along one axis and indexes by a certain distance along another axis, then repeats the cycle. It is one of the quickest methods for finding the first light of the beam. The concept is shown in Figure 232 where the green line shows how the raster proceeds.
- Spiral scan — Figure 236 illustrates this first light method, which tracks in the general area of the beam using a spiral motion by synchronizing the motion of the X and Y axes.

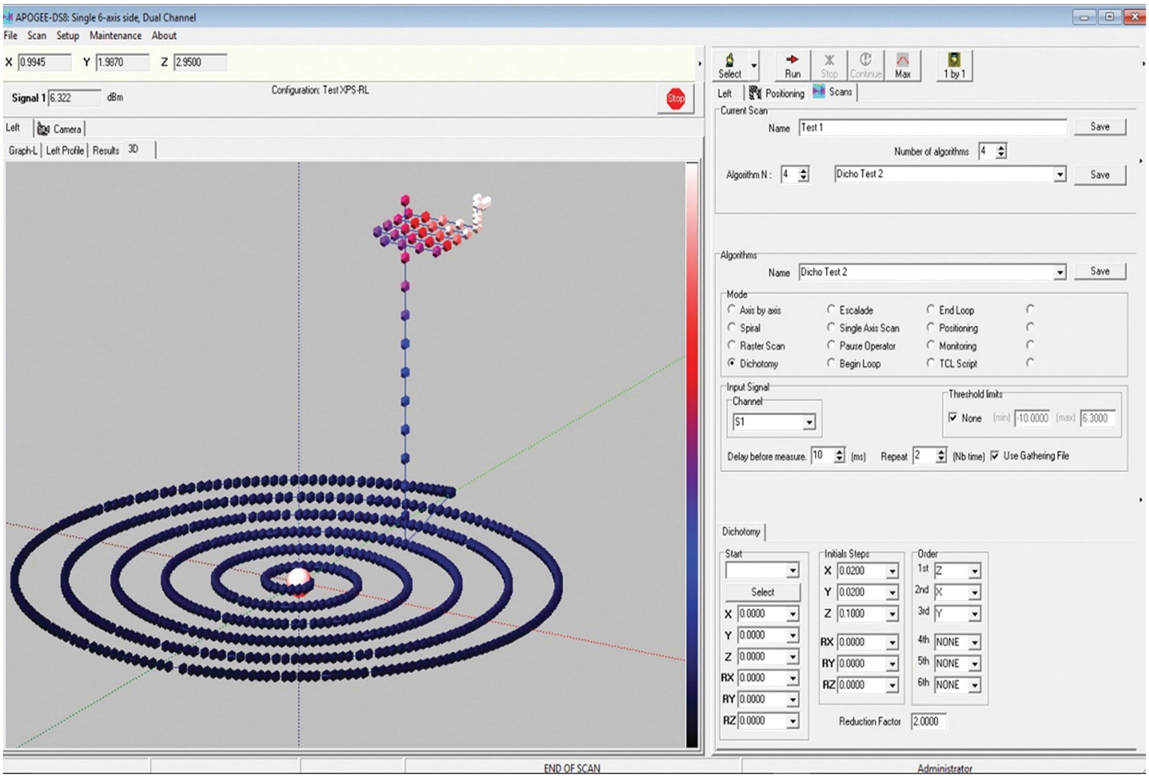


Figure 236. Spiral search for first light (image used with permission of GBC&S Consulting Services) [323].

After first light has been located, other algorithms are available that can find the peak power location:

- Hill climb — This is a 2D search method based on finding the highest power within a certain path (Figure 237). The direction of the climb favors the location of the higher power region. The hill climb method is most effective when the beam has a Gaussian profile and when power quickly increases. The hill climb method, by itself, is not effective in finding peak power with flat beam profiles.

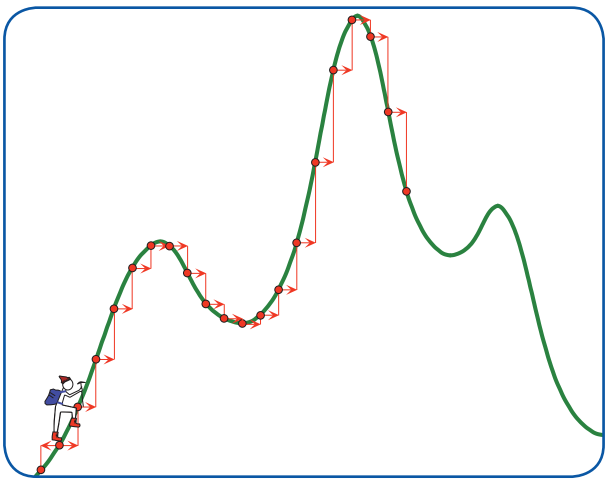


Figure 237. Hill climb algorithm [320].

- Centroid — The centroid search (Figure 238) moves along one axis and finds the peak, then from that peak, moves along the second axis to get to the final peak. Centroid is useful with top-hat or multi-peak profiles.

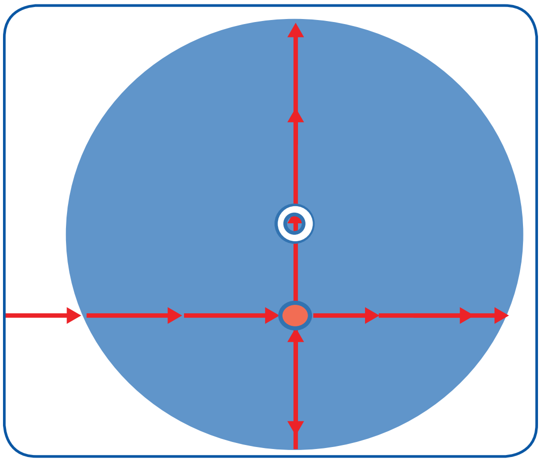


Figure 238. Centroid search algorithm.

- Dichotomy — Starting with large increments, a dichotomy search (Figure 239) initially searches one axis at a time until a peak is identified. The search then returns to the first peak. Within this first peak, another search cycle is performed using finer steps to find the maximum peak location.

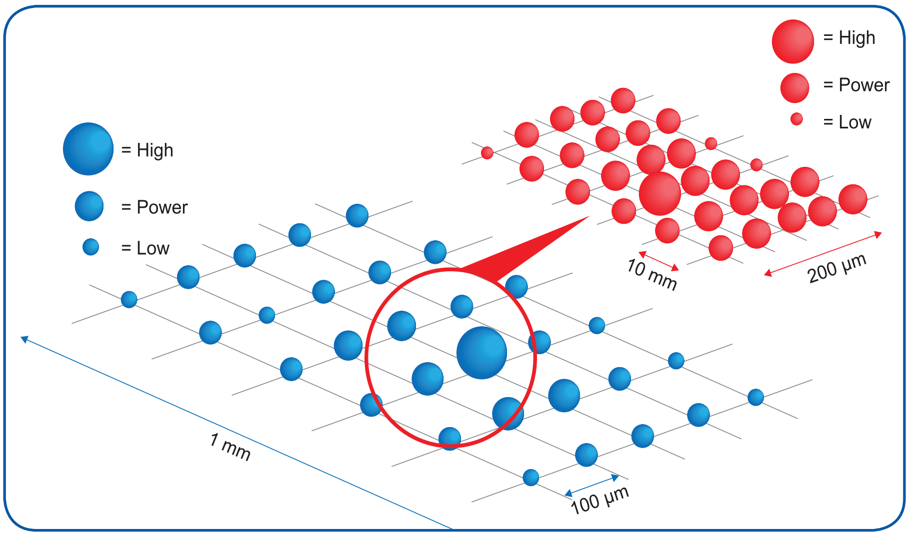


Figure 239. Dichotomy search algorithm (image used with permission of GBC&S Consulting Services) [323].

- Other algorithms — There are many other algorithms that can be used either alone or in combination. A user should experiment with different combinations to find the most effective method for a specific beam profile or set of device characteristics.

C. Motion Control Systems

Motion control systems can be customized in many ways, depending on the DUT. Figure 240 shows different configurations of fiber alignment systems, ranging from a simple, single-channel, single-side setup to a more complex double-sided, multiple-channel setup that includes machine vision, adhesive delivery/curing, and pick and place. Other setups may have horizontal and vertical beam inputs and outputs. Every configuration requires a unique set of motion products, depending on the performance required for that device. The basics of motion control systems can be found in Chapter 1, Section III.F.2.

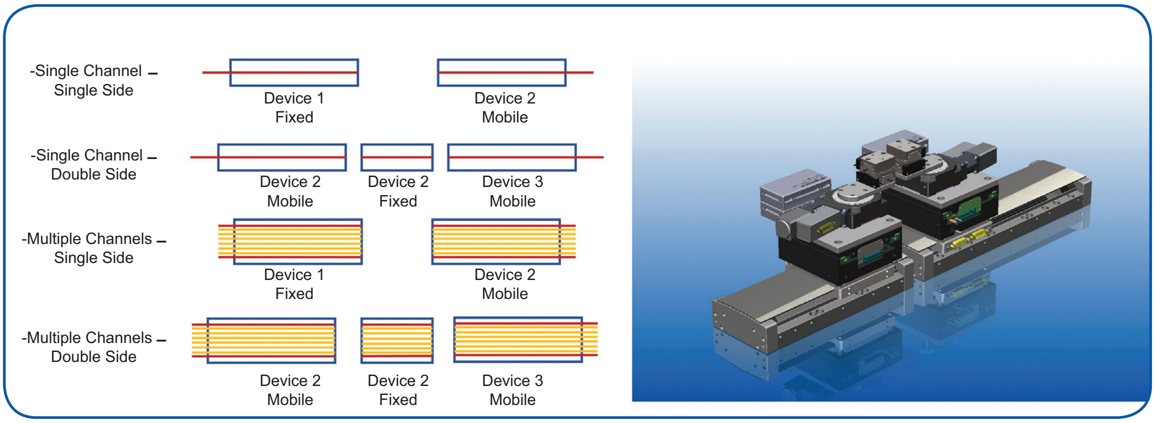


Figure 240. Configurations of fiber alignment systems (left, image used with permission of GBC&S Consulting Services [323]); example of a single-sided configuration with load/unload (right) [324].

Different configurations are available for the motion control systems used in fiber alignment, ranging from simple manual stages suitable for small scale and R&D applications to full-featured automated production systems with high precision motorized stages, pick and place automation, dispensing and curing systems, machine vision, etc. The following systems illustrate, in part, this range of configurations:

- Manual stages — Mechanically, manual stages are the simplest and the least costly motion control systems that can provide precise linear or rotational motion. They are most often used in academic or industrial R&D centers and are occasionally found in low volume production environments. Figure 241 shows the proven, stable MKS ULTRAlign™ 562 manual stage. The 562 can be motorized by adding NSA12 (Figure 241a) or TRA (Figure 241b) motorized actuators. Long travel precision motorized actuators [325] are also available for longer strokes or faster speeds. These actuators are commonly used in R&D and low volume production applications.

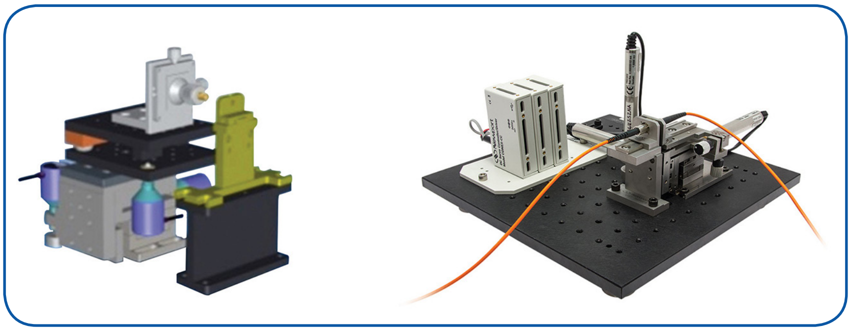


Figure 241. Single fiber, single-end configuration with 562 manual stages and NSA12 actuators (left); single fiber, single-end configuration with 562 manual stages and CONEX-TRA actuators (right).

- Piezoelectric stages — Piezoelectric stages are compact, four- to six-axis alignment systems that are driven by piezoelectric actuators. This type of actuator is capable of 30 nm linear resolution. The MKS 8071 4-axis aligner, shown in Figure 242, is ideal for remote control of alignment positioning in R&D applications. These multi-axis motorized positioners allow high-resolution (< 30 nm) adjustment for different combinations of X, Y, Z, θ_x, θ_y, and θ_z. These stages can exert a 5 lb. (22 N) force and have exceptional long-term stability and can hold their position with no power applied [326].

Figure 242.
8071 4-axis aligner driven by Picomotor™ Piezo Actuators.

Motorized stages, driven by electromagnetic motors, are commonly used in production environments to enable the automation of alignment processes. They typically feature sub-µm MIM and repeatability, which greatly improves the quality and performance of telecom devices assembled using these stages. The drive technologies that are commonly used in motorized precision fiber alignment systems include:

- Linear motor stages with direct read encoder — The XMS linear motor stages are the highest precision standard stages with 1 nm MIM capability when used with the XPS-D motion controller. The XMS can quickly and easily search within a 10 µm diameter area of the beam region exhibiting the highest power. Figure 243 shows a double-sided, multi-channel configuration with 5-axis adjustments using XMS stages.

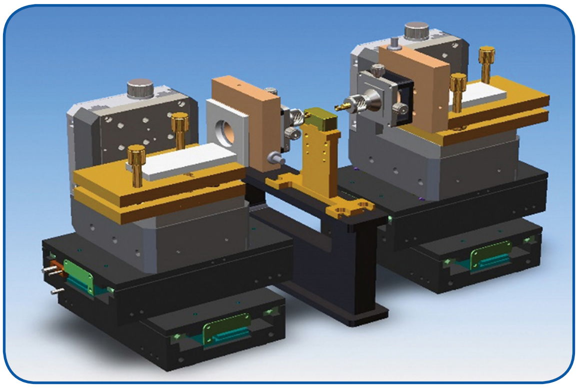


Figure 243. Double-sided configuration with VP-25 and XMS stages.

- XYZ assembly with ball screw drives — The VP-25XA family of stages was specifically designed for fiber alignment applications. These compact stages are available with either a 100 nm or 10 nm MIM and in left and right versions for single or double-ended configurations. The VP-25 has field-proven reliability in production environments. Figure 244 shows a 100 nm VP-25XA-XYZ. The basic concepts of fiber alignment using VP-25 stages are shown in [327].
- Hexapods — A hexapod is a parallel kinematic, mechanical device that uses six actuators, all moving in parallel, to provide six-axis range of motion in a Cartesian coordinate system. All axis moves are interdependent, meaning that the motion in one axis reduces the travel in the other axes. A hexapod is capable of complex combinations of linear and angular motions and is particularly useful for critical rotation adjustments. Figure 245 shows the HXP50

Figure 244.
VP-25XA-XYZL integrated specifically for fiber alignment.

hexapod which can be used to align a beam in both the horizontal and vertical axis. The operation of the HXP50 in fiber alignment is shown in [328].

Hexapods are usually more compact than a stack of stages, as the HXP50 demonstrates. The positioning requirements of fiber alignment make the HXP50 an excellent solution for most, if not all, cases. HXP hexapods feature the following innovations that are advantageous in fiber alignment applications:

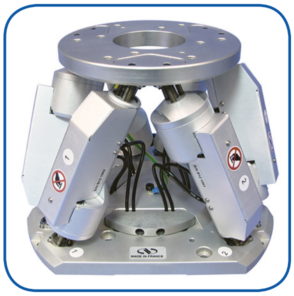


Figure 245. HXP50 Hexapod with horizontal and vertical beam paths.

- Work and Tool coordinate systems — A significant innovation in MKS Hexapods is the concept of Work and Tool coordinate systems. These programmable coordinate systems enable the independent manipulation of the Work (sample or device) or Tool (cutter or beam). The Work and Tool coordinate systems are illustrated in Figure 246.

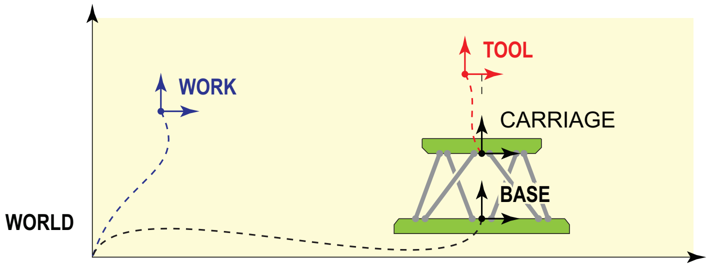


Figure 246. Work and Tool coordinate systems transformation of axes. Prior to the introduction of MKS HXP hexapods, the user had to calculate the transformation of the actuator motions into hexapod coordinates. The HXP hexapods come with transformation formulas which are embedded in the firmware, so the user simply sends commands in a Cartesian coordinate system, making the motion of the actuators quite transparent. The reader can imagine a path in a familiar coordinate system.

- RightPath™ Trajectory Control — While hexapods are ideal for point-to-point motion, they encounter some difficulties in scanning applications where a path, whether linear, rotary or arc, is to be followed. RightPath™ trajectory control [329] eliminates the need for a linear or rotation stage for short strokes. In Figure 247, the blue line illustrates the motion of a hexapod when commanded to move from one point to another in the X-axis without RightPath trajectory control. The typical deviation from a straight line could be in the order of more than a millimeter. With RightPath trajectory control, the runout is controlled to a couple of microns, similar to a linear stage. This innovation enables the Hexapod to follow linear, rotational, or arc moves with minimal trajectory run-out.

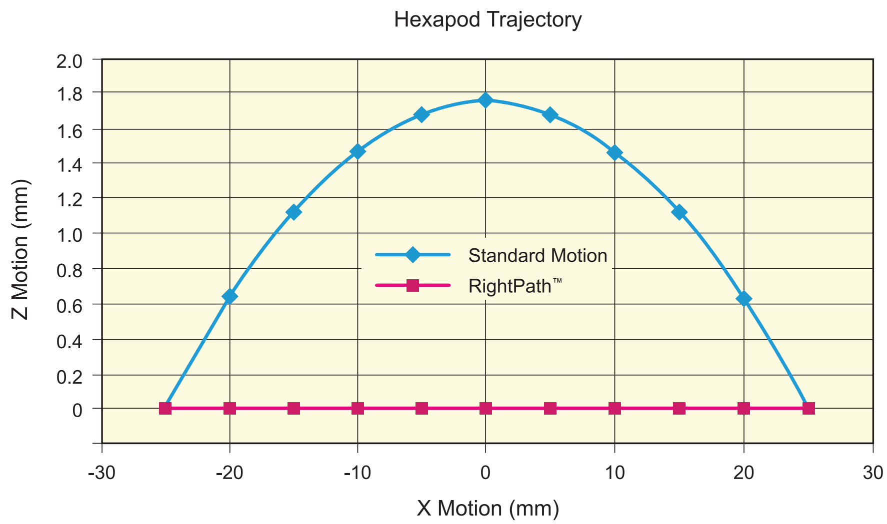


Figure 247. RightPath^{TM} trajectory showing runout [329].

- HexaViz simulation – HexaViz is simulation software that allows customers to simulate loads, motions, and potential collision for all MKS HXP hexapods. HexaViz (Figure 248) is very useful for simulating the actual devices and fixtures that will be needed in a fiber alignment application that uses an MKS HXP hexapod. Unlike other hexapod simulation software available in the market, HexaViz is free to download [330].

Figure 248. HexaViz GUI with HXP200 Model [329].

In addition to the device, device fixture or holder, light source, and motion control system, a complete fiber alignment system includes the following components:

- Detectors measure the power of a beam. Together with a power meter, they monitor the optical signal and detect the highest transmitted power. A typical detector and power meter combination from MKS is the 818-SL detector and the 1936-R power meter. An alternative detector is MKS' 3A-IS-IRG. If beam profilers are needed to characterize the shape of the beam, an SP928 is recommended (Figure 249). The Newport 1830-R and a 918D-IS-IG detectors (Figure 250) are recommended for production environments. Figure 251 shows a fiber alignment system with a power meter and detector.

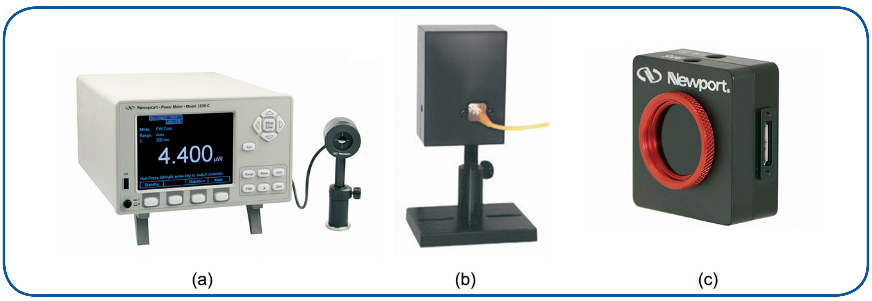


Figure 249. (a) MKS 818-SL detector and 1936-R power meter; (b) Ophir 3A-IS-IRG detector; (c) Ophir SP928 beam profiler.

- Power meters are matched with detectors for the specific wavelength, the power range measured, and a minimum data transfer rate of 2 kHz to achieve fast alignment and productivity. Figure 252 shows the Starbright and Juno handheld power meters available from MKS Ophir.

- Vision system are used to detect the proximity of devices and the rough alignment of fiber ends. Since the intensity of a beam is proportional to the inverse of the square of the distance between the source and detector, smaller gaps result in greater transmitted power. However, very small gaps present a higher risk for collision. A vision system allows a very small gap to be used. Depending on the magnification and available lighting, the fiber ends can be adjusted to be almost touching, maximizing the transmitted power.

Figure 250. Newport 918D-IS-IG detector and 1830-R power meter.

- Dispensing/bonding systems dispense an accurate volume of liquid epoxy, evenly apply it over the interface of two materials and cure it using UV light. They typically depend on the devices and materials that are being assembled. The curing process in these systems requires positional stability but at a lower degree of precision than for alignment. Therefore, a less precise motion control system can be used in dispensing and curing the glue.

- Laser welding is a bonding method that employs highly localized heating to attach two parts together. For example, laser welding is used to attach the output fiber, lenses, and the laser diode in a package. The process is automated, using pick-and-place within an enclosure to ensure both speed and safety.

- Pick-and-place automation is used for high volume, high speed production. The advantage of employing pick-and-place automation lies in the repeatability and reproducibility achieved in the devices. Components that are to be assembled are inserted into fixtures and, after assembly, the complete device is removed for packaging. During the unload process, a parallel sorting step can be performed that relies on pass/fail criteria to identify parts that can be packaged, reworked, or scrapped.

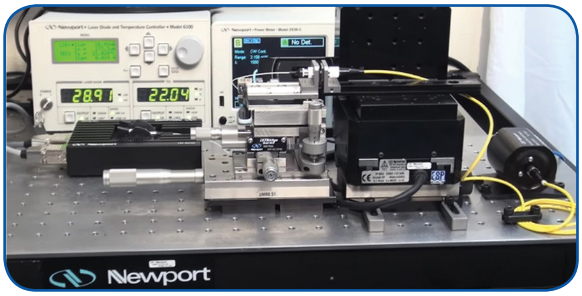


Figure 251. Single-end fiber alignment system with ULTRAlign™ 562, VP-25XYZ, power meter, and detector [331].

Figure 252. Ophir's Starbright Power Meter (left); Juno USB power meter (right).

The challenge for faster and more accurate fiber alignment continues to evolve. 400 GbE production may start in 2018 and 5G mobile communication standards are near completion, paving the way for adoption of this protocol in cellular networks. Autonomous driving technology will rely on mobile networks that require significantly higher speeds. This means that SiP technology may soon move into the mainstream. SiP will require significantly higher precision and speed in fiber alignment owing to the very small size and very high density of grating couplers on wafer surfaces.

Table 17 provides a quick reference guide for component selection in fiber alignment systems.

	Research and Development	Assembly/Production	Final Test
Laser Sources	LDC3726 and LDM Mount	LDC3900 Modular LD Controller	1784 VCSEL Fiber coupled laser source
Motion	CONEX-TRA/CONEX-LTA 562 Manual Stages Ultra Align Precision XYZ Stages	XMS/VP Linear Stages HXP50 Hexapod	Not Applicable
Laser Diode Tester	Sentry Single Shelf LD Tester Benchtop	Not Applicable	Sentinel LRS9434 with Burn-in
Power Detector	3A-IS-IRG 818-SL/DB	PD300-IRG 918D-IS-IG	PD300-IRG 918D-IS-IG
Power Meter	StarBright Power Meter 1936-R Power Meter	Juno 1830-R Power Meter	StarLite Power Meter 2936-R Power Meter
Wavemeter	OMM 6810 Power/Wavelength Meter	WM-1210 Wavemeter	Not Applicable
Beam Profiler	SP928, XC-130 Beam Profilers	Not Applicable	SP928, XC-130 Beam Profilers
Photoreceiver	Not Applicable	Not Applicable	1544 High Speed 1474A

Table 17. Quick reference guide for the selection of components for fiber alignment.

IV. Transceivers

The devices that transmit and receive data, known as transceivers, are some of the primary enablers of growth in the datacom market. MKS provides optical-to-electrical (OE) converters that are important for the design, manufacture, and testing of optical transceivers. These OEs enable the production line testing of optical transceivers by converting incoming photons into an electrical current that can be measured using conventional production test equipment such as oscilloscopes.

A. Fundamentals of OE Conversion

OE converters typically employ photodiodes to absorb incoming photons and convert them into electrical current. Photodiodes are devices in which *p*-type, *n*-type, and intrinsic semiconductors are arranged as shown in Figure 253. This structure is known as a *p-i-n* type diode, where "*i*" represents the intrinsic undoped layer between the *n* and *p* layers. The intrinsic layer absorbs photons and transfers their energy to electrons in the atoms. This transfer of photon energy frees electrons from the atoms and simultaneously generates holes. The electrons migrate to the *p*-type region where they are collected by the anode (positive terminal) of the device while the holes move towards the *n*-region and the cathode (negative terminal). In this way, the incoming photons are converted to an electrical current.

The mechanisms for photocurrent generation in *p-n* and *p-i-n* photodiodes are discussed in Chapter 1, Sections II.A.2 and II.C.2.

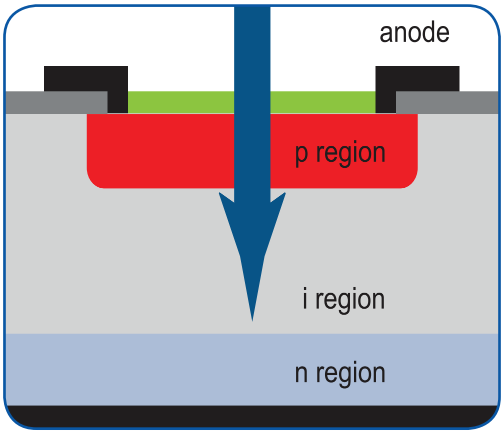


Figure 253. Schematic drawing of p-i-n photodiode (the green layer is an anti-reflection coating).

A major consideration in the selection of photodiodes used in OE applications is the semiconductor material employed. The material determines the wavelength range over which the device will produce usable current and Table 18 shows these ranges for semiconductor materials that make up common photodiodes. Within this range, the semiconductor will generate higher current at certain wavelengths and this is referred to as the device's spectral responsivity. Figure 254 shows a typical responsivity curve for an InGaAs photodiode. The choice of semiconductor material also determines, in part, the dark current of the device which dictates the achievable SNR and detector sensitivity. Details about spectral responsivity as well as detector noise and its related metrics can be found in Chapter 1, Sections II.A.1, II.A.2 and II.C.1. Since most communication applications employ SWIR laser light, the strong responsivity and low dark current for InGaAs explains why this has become the most popular material for photodiodes used in OE applications.

Semiconductor Material	Wavelength Range	Dark Current
Silicon (Si)	200-1100 nm	Medium
Germanium (Ge)	800-1600 nm	High
Gallium Arsenide (GaAs)	400-900 nm	Low
Indium Gallium Arsenide (InGaAs)	500-1800 nm	Lowest

Table 18. Wavelength range and dark current capabilities of common photodiode semiconductor materials.

The physical structure of a photodiode has a strong influence on its electrical characteristics. For example, small diameter photodiodes can be advantageous in optical communications applications that require bandwidths exceeding tens of GHz because current rise/fall time is inversely proportional to the device active area diameter (see Chapter 1, Section II.C.1 for details). However, smaller diameters also require that incoming light be directed to a much smaller surface area, e.g., a 30 GHz *p-i-n* photodiode has an active diameter of only 20 µm. Coupling the light becomes increasingly challenging as network speeds increase since further reduction in the photodiode active areas requires tighter focusing of the optical beam. Furthermore, a laser beam focused to a very small spot size greatly increases the power density at the photodiode surface. This high density can physically damage the photodiode (see Chapter 1, Section II.A.1) and can cause other undesirable non-linear effects in the device. Consequently, power levels should always be considered as part of communication system design.

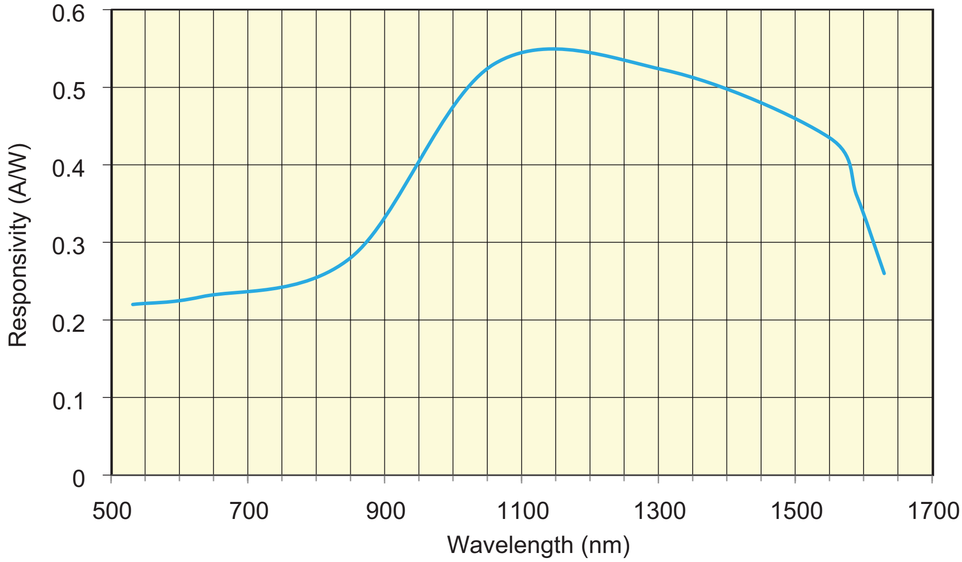


Figure 254. InGaAs Responsivity Curve.

B. Advanced Photodiode Designs

Conventional photodiode designs are configured so that light is incident normally on the surface. Unfortunately, this approach has bandwidth limitations given the restrictions imposed by high power densities and small spot sizes. For this reason, advanced designs now employ waveguide structures (Figure 255) that permit absorption of light and generation of electron-hole pairs along the beam propagation path. This allows device designs that have much larger areas for light absorption while maintaining fast current rise/fall times. Even though waveguide structures are more complicated and expensive to manufacture, the increasing demand for devices with faster speeds and greater bandwidth ensures that waveguide photodiodes will find increasing use in the future.

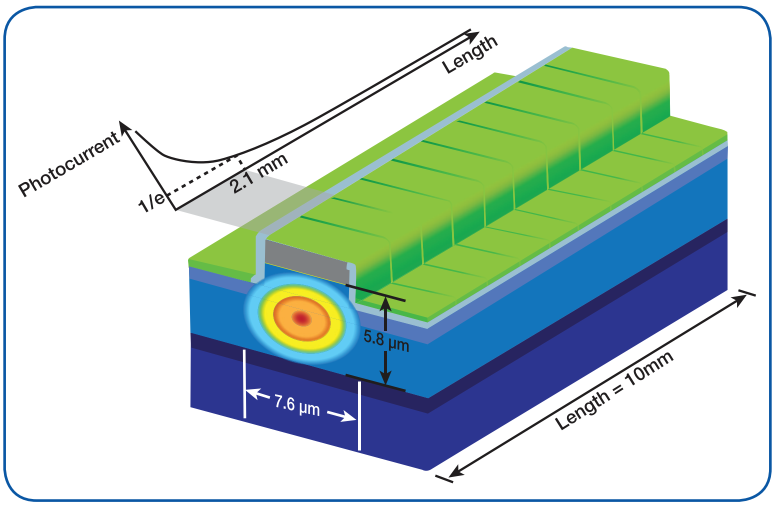


Figure 255. An optical waveguide.

APDs are also used in data communications networks. A cross-sectional view of an APD is shown in Figure 256. These devices employ an avalanche effect to internally amplify the photocurrent signal [332]. APDs are ideal for low intensity signals and for communication over long distances where attenuation diminishes the signal.

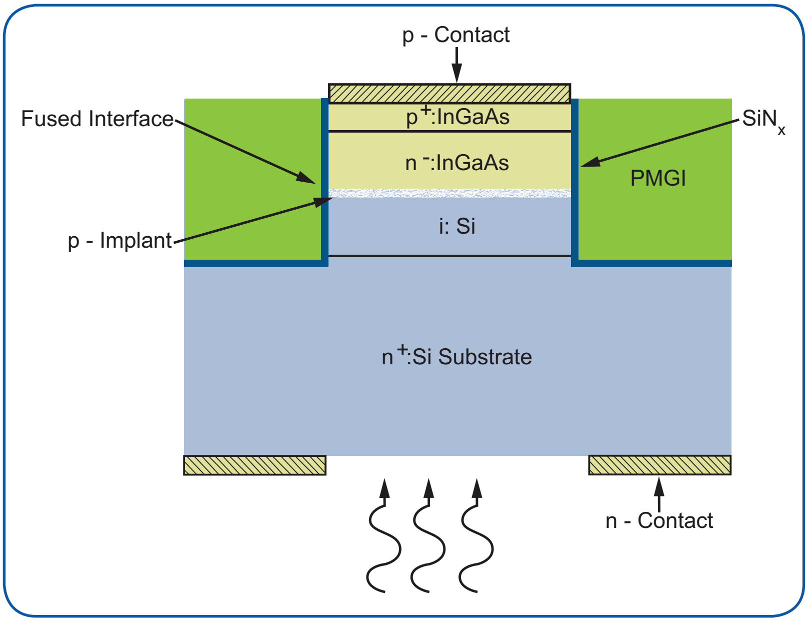


Figure 256. Schematic cross-section of wafer-bonded Si/InGaAs heterojunction APD [332].

C. OE Device Design

Following the selection of an appropriate photodiode, its known electrical characteristics can be used in the design of the OE device. Figure 257 shows a schematic of a typical OE circuit with the photodiode shown on the far left of the circuit diagram. Introduction of a resistor (R1) in series with the photodiode generates a basic photodetector circuit that can be monitored with an ammeter. The addition of a transimpedance amplifier can convert the signal from current to voltage and allows amplification of the signal (see Chapter 1, Section II.C.1). The voltage output can be easily detected with a voltmeter or, as is more typical in communications, with an oscilloscope. In some cases, multi-stage transimpedance amplifiers are used to amplify the signal further.

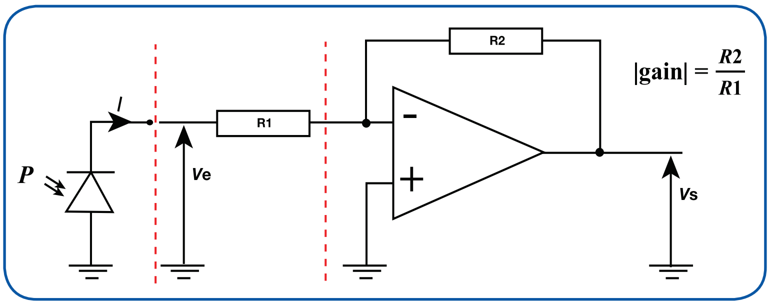


Figure 257. Electrical schematic of typical OE converter.

D. OE Optical Design

The effective coupling of photons into fast photodiodes can be extremely challenging. As noted above, a 30 GHz photodiode has an active diameter of only ~20 µm. In communications applications, a typical single-mode optical fiber propagating light at 1550 nm has a 9 µm core diameter and an NA of ~1.4 (see Chapter 1, Section III.D for more details). This means that the optimal focus distance of the fiber from the photodiode surface is around 80 µm and that there is a tolerance of ± 0.25 µm in both X and Y positioning (see Figure 258). Positioning of the photodiode and fiber for optimal coupling requires automation and the best available motion control products (see Section III.C). Positioning becomes even more challenging in applications that require multimode fibers with 50 µm core diameter. In this case, advanced optics are required to focus the beam and these additional optics make the process even more sensitive to misalignments.

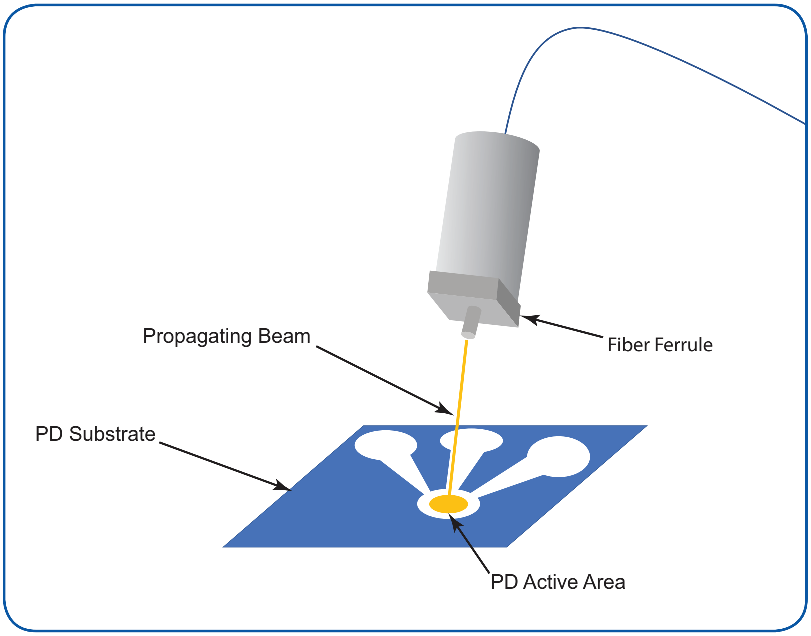


Figure 258. Depiction of the coupling of light from a single-mode fiber to a photodiode (PD) surface.

E. Application of OEs in Optical Transmitter Testing

OE converters are mainly used in the testing of optical transceivers. In a typical testing apparatus (Figure 259), they are paired with fast oscilloscopes to measure the performance of the transceiver. Eye diagrams (as shown in Figure 222) are generated as the sample data sent to the transceiver is compared against the actual input signal to determine errors. The testing apparatus has an OE module attached in front of the electrical oscilloscope and it is important that both have sufficient bandwidth to support the DUT. As optical testing has become more widespread, oscilloscope manufacturers have begun to build OEs directly into their oscilloscopes, eliminating the need for an external converter.

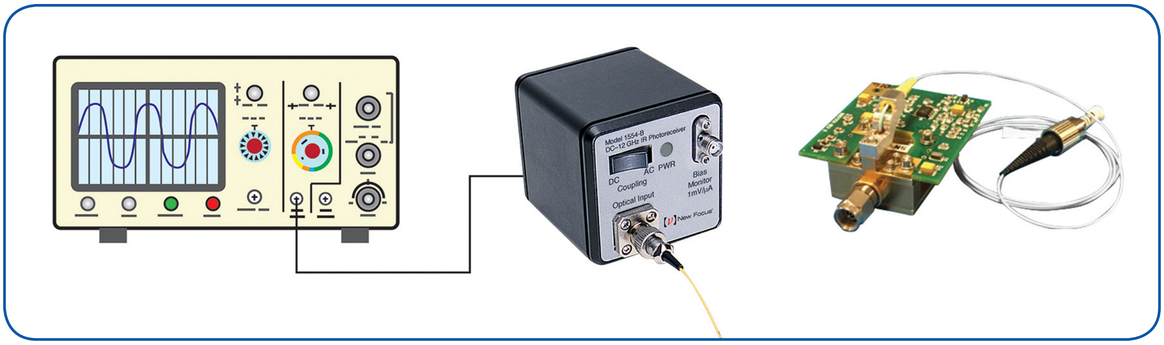


Figure 259. MKS high speed photoreceivers serve as the front-end OE converters for many electrical instruments.

F. Summary of Transceiver Requirements

A summary of transceiver and associated OE converter requirements is given in this section.

Bandwidth

Currently, the 100 GbE build out requires optical transceivers that can deliver and receive four channels of 25 Gbit/s NRZ data streams. An OE and oscilloscope combination must possess sufficient bandwidth to test each of these channels during the manufacturing process. This means that the frequency response of each component at the -3dB point (the point where there is a 50% loss in signal) must be > 25 GHz. The MKS 1484-A OE meets this specification as shown in Figure 260.

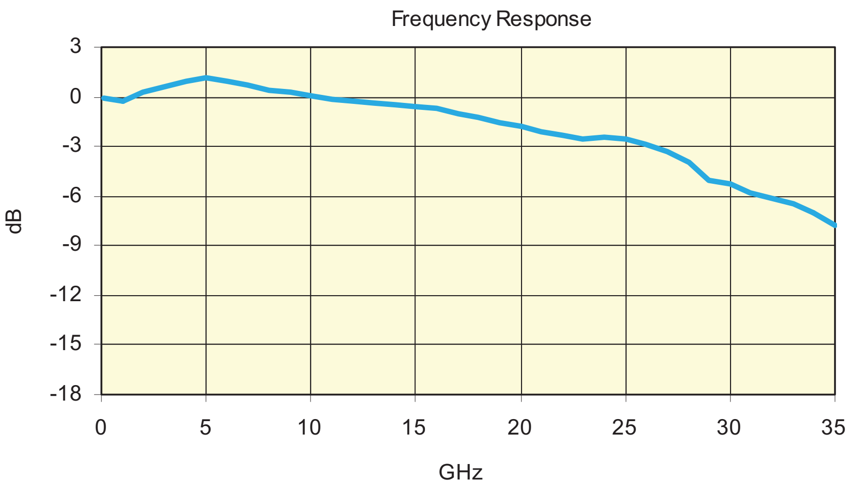


Figure 260. Frequency response curve of the model 1484-A OE converter.

Since 400 GbE became standardized with 25 Gbit/s streams still being used, the PAM-4 modulation scheme was implemented since it provides twice as much data per symbol and 8 channels. Nonetheless, there is a desire to return to 4 channels of data streams so that existing infrastructure can continue to be used. However, this will require that the data rate increases from 25 Gbit/s per lane to 50 Gbit/s per lane. The OE and oscilloscope bandwidth needed for such a configuration is therefore > 50 GHz. MKS OEs are meticulously tested and tuned for optimal frequency response performance for any given application. All frequency response and pulse response data are provided for every MKS OE.

Wavelength

Most 100 GbE data center build outs are driven by the use of a VCSEL operating at 850 nm. InGaAs photodiodes, can also be used for OE conversion at this wavelength owing to their extended responsivities (Figure 254) while still being optimal for SWIR applications. Since 400 GbE is being driven at ~1310 nm, InGaAs will continue to be the semiconductor material of choice for these applications. MKS OEs utilize both internal and external *p-i-n* photodiodes with advanced anti-reflection coatings to help improve responsivity.

Noise

While the lowest possible noise is generally preferred, the noise requirements depend on the application. Long haul communications require much lower noise specifications than short haul metro or even communications within a data center. Long haul communications typically use dual-matched receivers for coherent detection to help remove background noise from the signal. Short haul metro or data center communications do not require coherent detection, but still require that the NEP be on the order of a few tens of $pW/Hz^{1/2}$. Low noise requirements mean that it is advantageous for an OE manufacturer to minimize dark current noise in the photodiode and any other sources of electrical noise throughout the circuit. MKS OEs uses proprietary methods to minimize noise. These start with the optical signal and include meticulous calculations and testing of every electrical component.

Gain

Depending on the application, OE gain requirements may vary from 10 to 1500 V/W. Multiple amplifiers in series can be used to generate more gain. However, a tradeoff exists if the initial signal does not have a low NEP, since amplifiers have bandwidth limitations and are not easily combined at higher bandwidths, i.e., > 15 GHz. Rather, it is best to use a single amplifier that is designed for high bandwidth applications. MKS OEs utilize state-of-the-art amplifiers that have been rigorously tested for quality and reliability.

Chapter 6

Photonics in Analytical Instruments

This page intentionally left blank.

I. Introduction

Modern scientific and technological standards demand that the analytical instruments used in regulatory, industrial production, health science, and R&D environments perform to ever more demanding specifications for repeatability, precision, accuracy, and detection limits. One of the most effective ways that instrument manufacturers have been able to achieve these specifications is with photonics technology. Photonics applications in analytical instruments can range from simple optical assemblies employed in chromatography detectors to highly complex opto-mechanical subsystems used in flow cytometry instruments.

Specific examples of photonics applications are as follows: Lasers used for sample ionization in Matrix Assisted Laser Desorption/Ionization (MALDI) mass spectrometers. The Inductively Coupled Plasma-Optical Emission Spectrometer (ICP-OES), the analytical industry standard technique for elemental analysis, employs optical components (see Chapter 1, Section III.A) and vibration control technologies (see Chapter 1, Section III.G.2). In health sciences, DNA sequencers, Polymerase Chain Reaction (PCR)/ digital-PCR and flow cytometry and cell analysis instruments (techniques that are often used to make critical disease diagnosis decisions) employ ultrafast, *CW*, and pulsed UV lasers (see Chapter 1, Section I.A). Materials science and surface characterization studies employ lasers and optical components in tools for laser-based particle characterization and in imaging instrumentation such as optical and confocal/ multiphoton microscopes. Separation sciences employ optical filters with high transmission and low background noise (see Chapter 1, Section III.A.2), replicated mirrors, (e.g. retroreflectors, asphericals, plano mirrors – see Chapter 1, Section III.A.1), and diffraction gratings (e.g. plane, concave and echelle gratings – see Chapter 1, Section II.D.1). There are numerous applications of photonics technology in analytical science and these are just a small representation. While a comprehensive review of all photonics applications in analytical instrumentation is outside the scope of this document, many articles and reviews can be found in the current literature.

This chapter focuses on selected photonics applications in materials and chemical analytical tools. The application of photonics technology in analytical instrumentation in the life and health sciences is addressed in Chapter 7 – Photonics in Life and Health Sciences.

II. Matrix Assisted Laser Desorption/Ionization (MALDI)

A. Introduction

MALDI is a soft ionization technique used in mass spectrometry (MS) that produces rapid and efficient ionization of a wide variety of molecules (see Figure 261). MALDI uses a laser energy absorbing matrix to produce ions from molecules having molecular weights ranging from 100's to 1000's of Daltons with minimal fragmentation. It has found increasing application over the past 30 years, especially for the analysis of biomolecules, e.g. proteins, peptides, DNA, and polysaccharides, and other large organic molecules, e.g. polymers, dendrimers and other macromolecules. These molecules tend to be fragile, undergoing excessive fragmentation when ionized by more conventional methods.

MALDI is similar in character to ElectroSpray Ionization (ESI) in that both techniques produce low molecular fragmentation on ionization. A primary difference between MALDI and ESI is that the former typically produces ions with a net single charge and this enables simple determination of molecular mass for most compounds. However, it can also limit the ability to analyze the largest macromolecular proteins. MALDI ionization requires three steps:

- A sample is mixed with a suitable matrix material and applied to a metal plate. As the mixture dries, crystals are formed on the sample material, that are critical for efficient ionization.

- A pulsed laser beam impinges on the sample, causing desorption of the sample and matrix material. The matrix material breaks down in this process to produce gas phase ionic species.
- Finally, the analyte molecules are ionized via protonation or de-protonation in the hot plume of ablated gases. The ions are then accelerated into a mass spectrometer system for mass analysis.

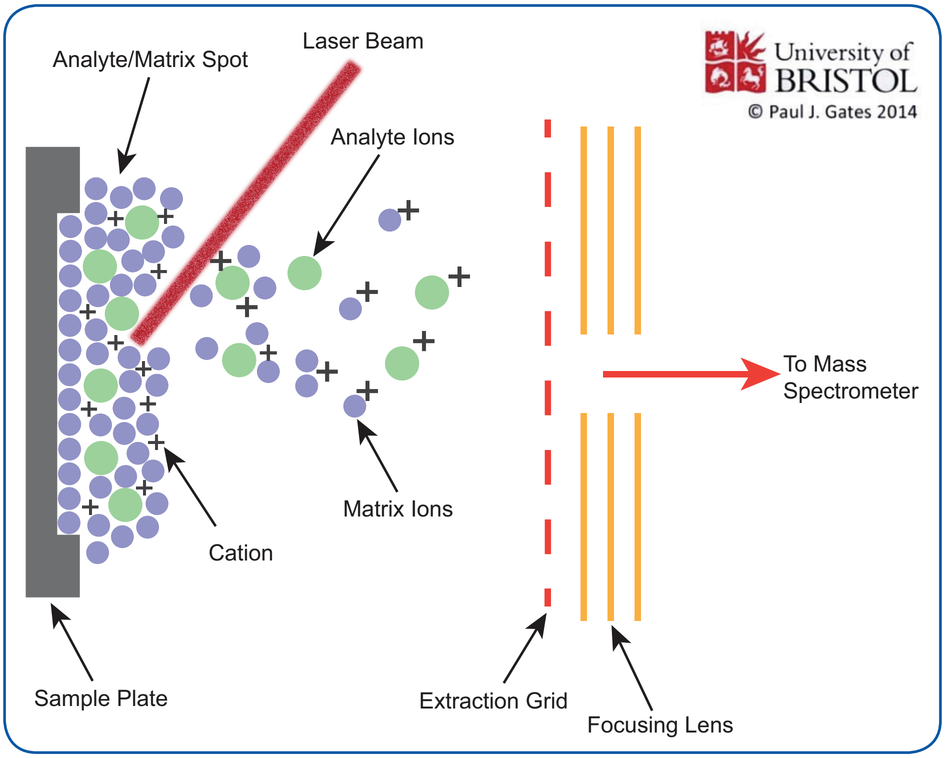


Figure 261. The MALDI process [333] (Figure used with permission of Dr. Paul Gates, University of Bristol UK).

While MALDI units are most commonly coupled with Time-of-Flight (ToF) mass spectrometers (Figure 262), in recent years they have been associated with a wider variety of instrument types that enable their use in both advanced research, medical and clinical applications. MALDI units have been used with:

- Laser technologies
  - UV gas lasers, e.g., nitrogen laser operating at 337 nm
  - DPSS Nd:YLF or Nd:YAG lasers with frequency-tripling (operating at 349 nm and 355 nm, respectively)
- ToF MS
  - Axial ToF — affordable performance, speed, ease-of-use
  - Orthogonal-acceleration ToF — higher performance, specificity, flexibility
  - Ion mobility spectrometry orthogonal-acceleration ToF — highest specificity
- Tandem quadrupole MS
  - Industry standard quantitative platform (despite limited MALDI linearity)

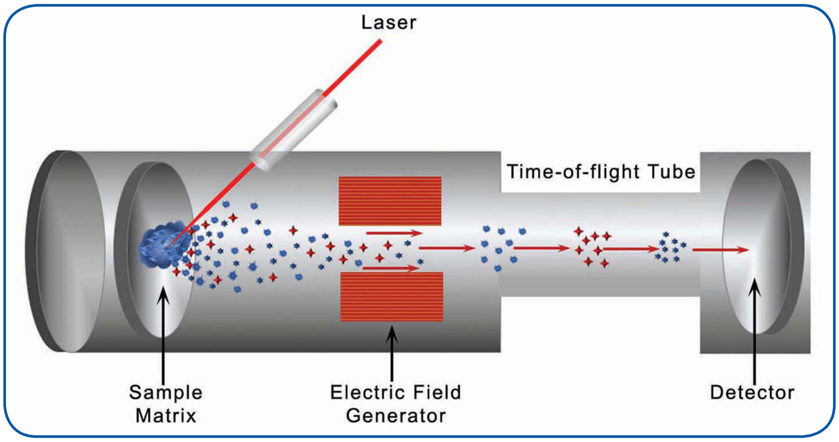


Figure 262. MALDI MS analysis, with an axial ToF detector [334].

Typical applications for MALDI MS include, but are not limited to:

- Biochemistry and molecular biology, e.g., proteomics, glycomics
- Organic chemistry
- Polymer analysis
- Microbiology
- Medical and diagnostics

B. Impact of Photonics on MALDI

The incorporation of photonics technologies into MALDI MS instruments dramatically improves the sensitivity, throughput, and ease of use of these systems. This has led to their increased deployment in biological and polymer research where MALDI systems have provided an attractive alternative to older, more complicated high-performance Liquid Chromatography (LC/MS) units. Additionally, improvements in laser technology have enabled high-speed analysis, which significantly increases the throughput of protein identification in proteomics research laboratories. More recently, as ESI-LC/MS has become the preferred method for proteomics, MALDI MS has found wider application in microbiological, pharmaceutical development and medical-related applications. The key advances in laser technologies and their consequences for MALDI methods are:

- Tunable wavelengths — more efficient ionization of biomolecules
- Increased repetition rates — speed of analysis
- Improved lifetime and cost of ownership — reduced complexity and increased reliability

MALDI MS units have been increasingly used in imaging applications where the highly localized nature of laser ablation allows the determination of the concentrations and spatial locations of compounds of interest in a wide range of sample types, especially biological tissues. The key photonics technological drivers for advanced imaging of compounds in tissues include:

- Laser repetition rate (> 1 kHz) — this enables fast processing of large tissue/sample sections
- Laser energy profile — this ensures minimal denaturation of a sample below the ionization threshold which, in turn, improves sensitivity, particularly for spot sizes below 20 μm
- High precision optics and motion control — this provides the ability to reliably measure spatial locations under 10 μm

Imaging applications are currently driving the equipment requirements in research environments within the life science and clinical communities. The critical equipment characteristics for MALDI units in these applications include:

- High resolution imaging — sub-20 μm spatial resolution
- Sample stage control — fine X/Y control, high robustness and reproducibility
- Speed of acquisition — using small spot size, GHz pulse rate, top-hat energy profile
- Flexibility — coupling of lasers by optical fiber

C. MALDI Applications Using MKS Products

1. Laser and Optics Technologies

- Spectra-Physics® Lasers
  - Innovative DPSS lasers — Explorer® One™ (Figure 263)
- Ophir® Optronics
  - Optical components — Lenses, gratings

Figure 263.
Explorer® One™ Laser.

2. Custom Vacuum Design and Manufacturing

- MKS Custom Vacuum Solutions
  - Design for manufacturing
  - Vacuum manufacturing expertise
  - Contract manufacturing service
  - Vacuum subsystems and higher-level assembly and test (Figure 264)

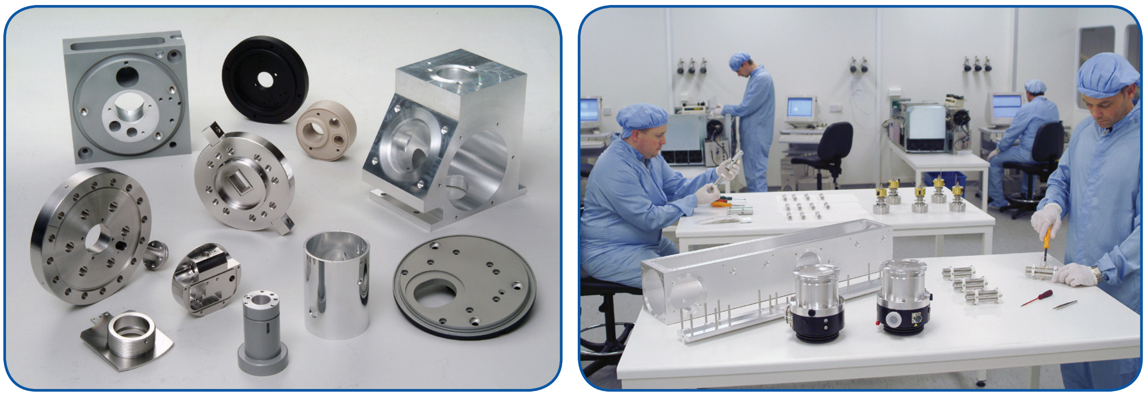


Figure 264. Custom Vacuum Solutions. High precision vacuum components (left); high level assembly and test expertise (right).

3. System Design

- MKS ISB
  - Opto-mechanical system design and manufacturing

- MKS MS Solutions
  - Electro-mechanical, Quadrupole MS
- MKS Newport opto-mechanical components and motion control
  - Mounts, assemblies, brackets, breadboards, rails, beam routing, laboratory supplies (see Chapter 1, Section III.B)
  - Linear stages, motion controllers, rotation and tilt stages, hexapods, alignment stages, actuators and adjusters (see Chapter 1, Section III.F.2)

III. IR Absorption Spectroscopy

A. Introduction

Most materials absorb electromagnetic radiation in the IR spectral region at wavelengths (between 0.8 µm and 14 µm) that are characteristic of the material's molecular structure. IR absorption spectroscopy is a common chemical analysis tool that measures absorption of an IR beam that has been passed through a sample. The position of the absorption peaks in an IR spectrum (Figure 265) are characteristic of the sample's chemical composition or purity and the intensity of an absorption peak is proportional to the concentration of the species for which that peak is a characteristic.

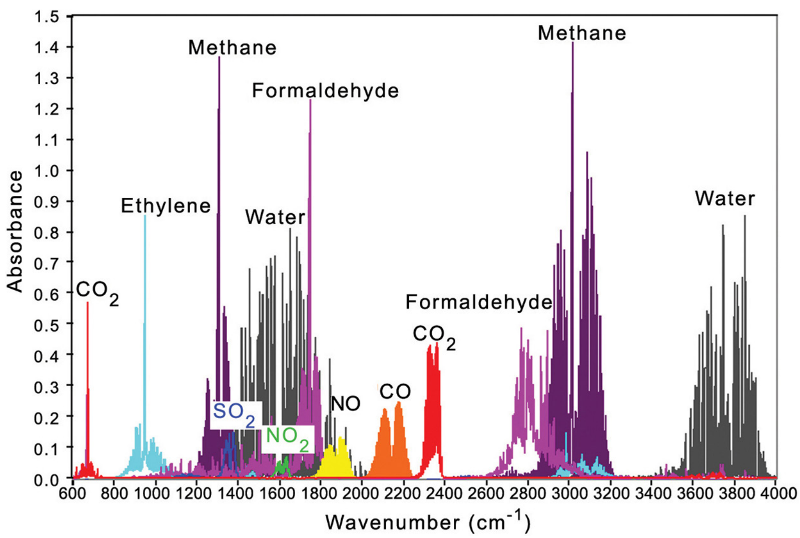


Figure 265. IR transmission spectrum of various gases and vapors.

IR spectroscopy can be used for both qualitative and quantitative non-destructive analysis of gases, liquids, pastes, powders, films, and surfaces. The absorption spectrum of a molecule provides a unique "fingerprint" of absorbances that can be used to deduce a sample's chemical composition and species concentrations. Figure 265 shows IR absorption spectra for some of the gases that are routinely measured in stack emissions monitoring applications. Each of the gas species labelled can be seen to exhibit a unique pattern of IR absorption.

IR spectrometers typically consist of a broadband IR light source, a wavelength separating device and a detector, as shown in Figure 266. Liquid or gas samples are typically contained in a sample cell. Solid samples can be analyzed using absorption or reflectance spectroscopy, in situ or in a standoff measurement system, or as a pressed disc of powder diluted by an IR transparent material or diluted in pastes commonly referred to as mulls.

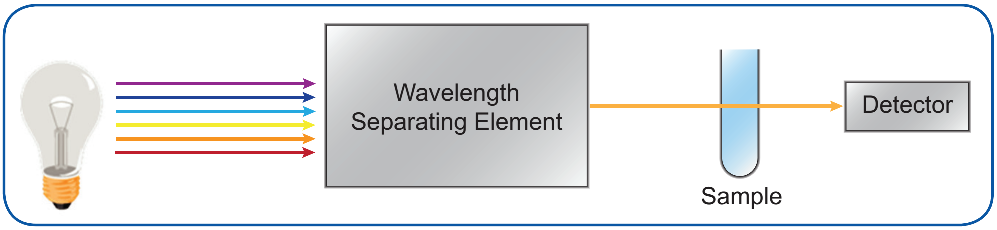


Figure 266. Major components of an IR spectroscopy instrument.

B. IR Spectrometers

Popular IR spectrometers used for routine chemical identification and quantitative measurement include: grating-based/dispersive IR spectrometers, FTIR spectrometers, and filter-based or non-dispersive IR (NDIR) instruments. The choice of spectrometer for a particular application is driven by requirements such as sensitivity, matrix complexity, packaging, and cost requirements.

1. Dispersive or Grating Spectrometer

Dispersive spectrometers (Figure 267) use a grating to disperse and separate the wavelengths of broadband light. There are two types of dispersive spectrometers: monochromators and spectrographs (see Chapter 1, Section II.D.2 for details). The former uses a single-element photodetector and a rotating grating assembly, while the latter uses a fixed grating assembly and a photodetector array. The advantage of dispersive spectrometers lies in their simplicity, which enables hardware miniaturization while retaining the ability to scan relatively wide spectral ranges. However, compared to FTIR or NDIR instruments, the throughput (or optical etendue) is limited since only a small portion of the source light ends up on the photodetector. Because of this, dispersive spectrometers are typically used for VIS and NIR spectral regions rather than for the MIR region where the radiation has lower photon energy.

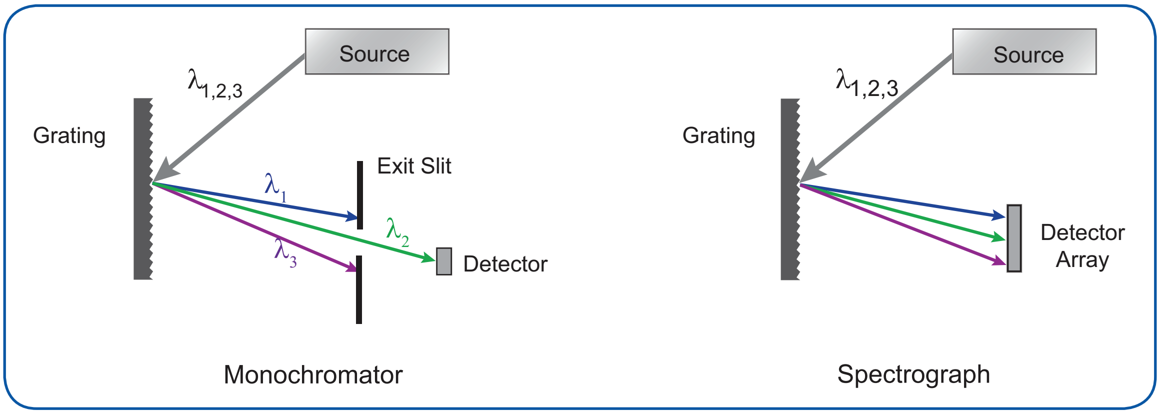


Figure 267. A simplified schematic of a monochromator (left) and a spectrograph (right).

2. FTIR Spectrometer

An FTIR spectrometer (Figure 268) generates a spectrum by modulating the IR radiation in the time domain using interference to produce an interferogram that is then subjected to a Fourier transform. In a Michelson interferometer, the most common interferometer used in FTIR, the incoming beam of light is split into two identical beams using a beamsplitter (a partially reflecting mirror). Each of these beams travels a different route and they are recombined before arriving at a detector. The path difference, the difference in the distance travelled by each beam, creates a phase difference between them. This recombined beam is the interferogram, a modulated signal as a function of the path difference. Performing a Fourier transform on the interferogram produces the spectrum of the incoming beam. A more thorough explanation of the fundamentals and operation of FTIR may be found in [335].

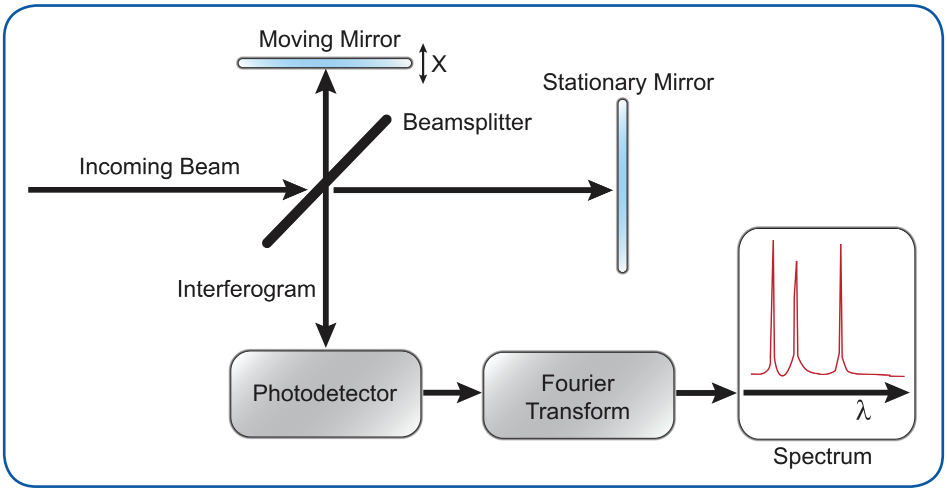


Figure 268. A simplified schematic of FTIR with a basic Michelson interferometer.

The FTIR spectrometer has several advantages over a traditional dispersive spectrometer. First, it provides fast measurements over a wide wavelength range. A modern FTIR instrument such as the MKS MultiGas™ FTIR analyzer can perform a scan in as little as 200 ms. With each scan, it covers the entire MIR wavelength region between approximately 2 μm and 16 μm, depending on the material of the optics and the type of photodetector employed in the instrument. Second, and perhaps most important, is the fact FTIR spectrometers have high optical throughput or etendue. An FTIR does not use a slit to control the wavelength resolution of the instrument. As a result, the spectra produced by FTIR spectrometers are generally much “sharper” than those produced by dispersive spectrometers under the same conditions. This is important in quantitative analysis where SNR generally determines the sensitivity of the measurement. Another advantage of the FTIR spectrometer is its wavelength precision and stability. An FTIR spectrometer normally uses a laser to control the position and velocity of the moving mirror and to trigger the collection of data points throughout the scan. A well-designed FTIR instrument provides a very high unit-to-unit repeatability, eliminating the need for cumbersome and expensive individual unit calibration.

3. Non-Dispersive IR (NDIR) Analyzer

NDIR analyzers are filter-based instruments designed for specific measurement applications. For example, NDIR analyzers are the industry standard method for measuring the concentrations of CO and CO_2 in stack emissions monitoring. Instead of scanning the wavelength and generating spectra, NDIR instruments generally capture the absorption of discrete wavelengths relevant to the chemical species being measured. They do so by employing optical filters that transmit light at selected narrow-band regions. Figure 269 shows the basic concept of an NDIR instrument.

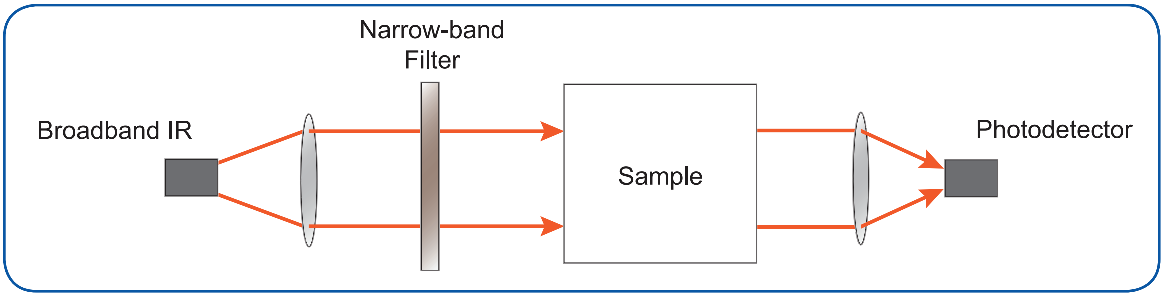


Figure 269. Basic diagram of NDIR showing its major components.

The filter used in NDIR instruments is usually an interference type, called an “etalon,” which is essentially a Fabry-Perot interferometer. It is typically made of a thin-film spacer separating two thin-film reflectors. Wave interference occurs in the etalon, and the waves that are in phase constructively interfere and are transmitted through the filter. The rest of the waves interfere destructively and are therefore ‘blocked.’ The broadband IR source is generally blackbody radiation produced by a heated filament such as tungsten or Kanthal.

The main advantage of an NDIR instrument is the simplicity of its hardware. This makes NDIR instruments both low-cost and rugged, making them ideal for industrial applications. MKS Process Sense™ is an example of an NDIR analyzer used in semiconductor process applications.

4. Tunable Filter Spectrometer (TFS™)

TFS™ spectrometer refers to an MKS NDIR instrument that provides wavelength scanning capability. The wavelength scanning is generally achieved by adjusting the gap distance between the two thin-film reflectors in the Fabry-Perot element. MKS' TFS™ spectrometer adjusts this gap distance by adjusting the incident angle of the light by rotating the filter element. As depicted in Figure 270, the filter transmission wavelength is altered as the incident angle is adjusted. As the incident angle increases, the transmission scans to lower wavelengths.

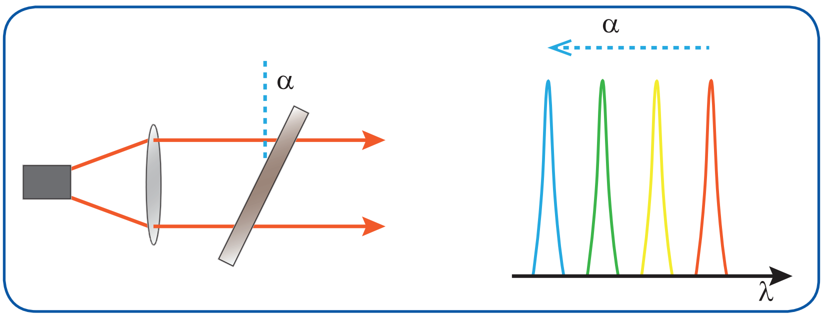


Figure 270. Concept of a TFS spectrometer. Change in the incident angle (left) produces a varying wavelength transmission (right).

C. IR Spectroscopy Applications Using MKS Products

MKS Instruments offers several analytical tools for emissions and process monitoring applications based on IR spectroscopy.

1. Emissions Monitoring

Gas emissions monitoring is normally required at emission sources such as power plant stacks and chemical manufacturing facilities. Accurate monitoring data is a critical part of regulatory compliance and environmental protection programs designed to control and regulate airborne pollutants such as carbon monoxide, nitrogen oxide, sulphur dioxide, and ozone, as well as greenhouse gases such as methane and carbon dioxide. MKS' MultiGas™ 2030 system (Figure 271) is a high-speed, high-resolution FTIR-based gas analyzer designed to simultaneously monitor more than 30 emission gases in under a second.

2. Chemical Agent and Toxic Industrial Chemical Detection

There is a growing demand in the safety and security industries for rapid and reliable detection of chemical agents and toxic industrial chemicals. MKS Instruments' FTIR-based AIRGARD® system (Figure 272) can detect parts per billion (ppb) levels of most chemical warfare agents (CWA) and toxic industrial compounds within 20 seconds. It is the industry standard, a fixed monitoring of CWAs in critical infrastructure buildings.

3. Semiconductor Process Monitoring

Semiconductor chemical-vapor deposition process chambers must be periodically cleaned to remove deposited build-up on the chamber walls and internal components. Optimal cleaning times for different processes depend on a complex relationship between variables such as the thickness of the build-up, the interior temperature of the chamber components, deposition/sputter ratios, and the chemical composition of the materials to be removed. MKS' Process Sense™ endpoint sensor (Figure 273) monitors the effluent from the chamber cleaning process in real-time using an NDIR method. For example, Process Sense™ monitors are used with fluorine-based chamber cleaning processes to monitor the by-product silicon tetrafluoride (SiF_4) from chambers employed for silicon-based deposition processes (including poly silicon, silicon dioxide, and silicon nitride). The sensor reports the SiF_4 level remaining in the effluent from the chamber in real-time, allowing the user to rapidly detect the endpoint of the cleaning process and avoid over-etching of the chamber components that might lead to damage or other maintenance issues.

4. Heating Value Measurement of Hydrocarbon Gases

Knowledge of the heating value (BTU content) of a fuel is an important parameter for optimal control of processes such as power generation, petrochemical manufacturing, and flare control. Traditionally, gas chromatography (GC) has been the analytical technique most commonly used to determine the heating value of fuels. However, GC analyses are slow (updates require minutes) and GC instrumentation typically requires significant maintenance. MKS' Precisive® Gas Analyzer (Figure 274) is the industry's first all-optical analyzer that performs quantitative speciation of hydrocarbon gases. Using patented TFS™

Figure 271.
MKS MultiGas™
2030 FTIR Analyzer.

Figure 272.
MKS AIRGARD®
Analyzer.

Figure 273.
MKS Process Sense™
Endpoint Detector.

Figure 274.
MKS Precisive®
Gas Analyzer.

analyzer technology, the Precisive analyzer is a simple and rugged NDIR instrument that provides high accuracy BTU measurements in often dirty and hazardous industrial environments.

5. Modular FTIR for Analytical Measurements

MKS Newport offers a flexible, high performance family of FTIR models (Figure 275) that employ a scanner designed for both routine analytical and non-traditional applications. The 8035 family of FTIRs uses a modular design that facilitates application-specific configurations using a variety of component options. The Michelson interferometer in the 8035 uses potassium bromide beamsplitters and windows. Furthermore, the retroreflector and beamsplitter are mounted together to ensure accurate alignment while simultaneously desensitizing the system to vibrations and temperature fluctuations. This "unibody" configuration of the retroreflector/beamsplitter requires only minimal realignment when interchanging system components. The 8035 has selectable resolution, starting with 0.5 cm^{-1}, and a very broad spectral range, depending on choices of optics, IR sources, detectors and beamsplitters.

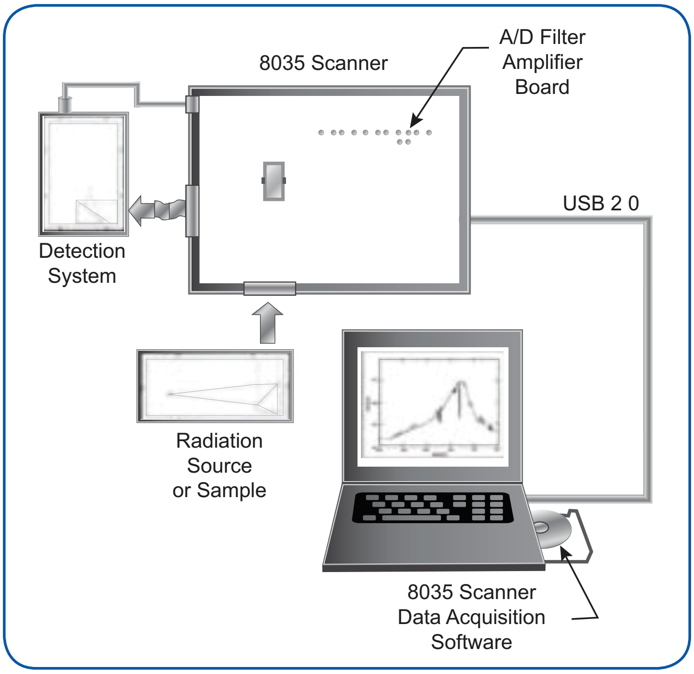


Figure 275. MKS Newport Modular 8035™ FTIR.

IV. Flow Cytometry

A. Introduction

Flow cytometry is an analytical technique that can rapidly measure the properties of individual cells or particles as they pass through a beam of light, typically a laser. A flow cytometer takes a sample of cells, transitions them into a single stream and uses lasers and/or light sources to excite biomarkers or labels on the cells to count the number of relevant constituents. The properties measured include relative particle size, relative granularity or internal complexity, and relative fluorescence intensity.

The method requires several photonics components, starting with lasers and other light sources with small, tightly focused beams that can illuminate a single size cell or particle (~30 μm diameter). Laser light with specific wavelengths is needed to excite labels or fluorophores on a cell. Dichroic mirrors/filters with precision coatings are needed for narrow reflection/transmission bands and high optical density for out-of-band rejection. All photonics components must be precisely aligned and controlled with μm spatial resolution.

A flow cytometer (Figure 276) consists of the three major component systems. The first is the fluidics system that transports sample particles to the laser beam in a narrow, single particle wide stream. The second is the optical system composed of lasers or light sources that illuminate the particles in the sample stream as well as optical filters and beamsplitters that direct the post-sample light signals to optical detectors for counting and processing. The third system is the electronics and signal processing equipment that converts the detected optical signals to electronic signals for processing and analysis.

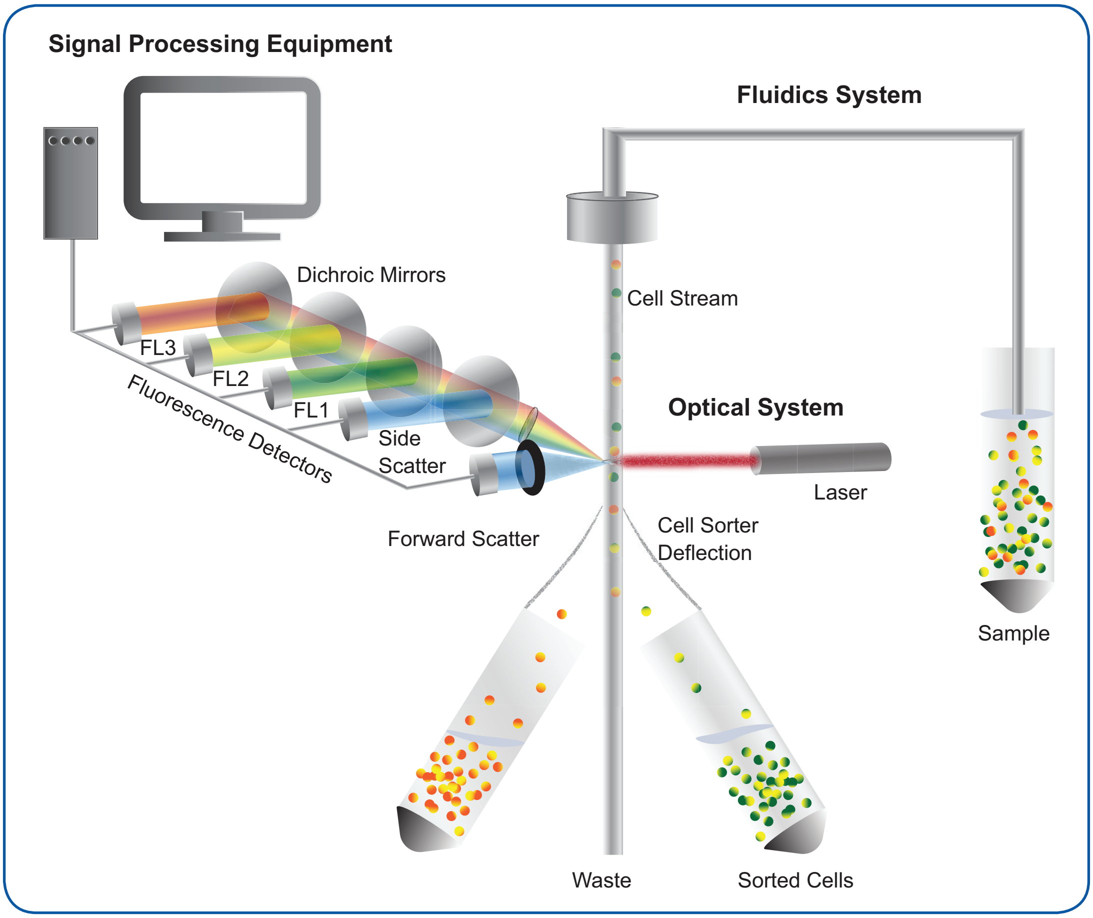


Figure 276. Schematic of a flow cytometry system.

B. Flow Cytometry Systems

1. Fluidics System

The fluidics system creates a stream of single particles that can be interrogated individually by the instrument's detection system. The sample is focused by the Bernoulli effect (Figure 277), creating a stream of particles in single file using a method called hydrodynamic focusing. Under optimal conditions (laminar flow), there is no mixing of the central fluid stream and the sheath fluid.

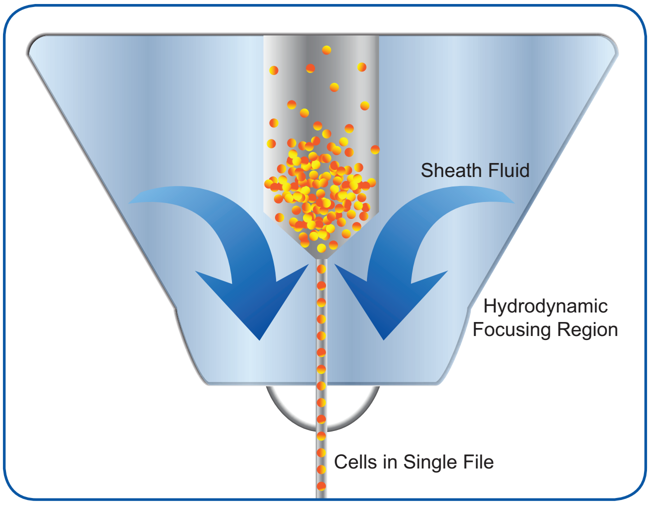


Figure 277. Hydrodynamic focusing produces a single stream of particles.

2. Optics and Detection

Following hydrodynamic focusing, the particle stream passes through one or more focused laser beams and light scattering or fluorescence emission occurs. The optical output is collected as forward or side scattered light (FSC or SSC) by a PMT or photodiode and the optical data is used to characterize the cell properties. FSC data provides an estimate of a particle's size while SSC provides information on the relative internal complexity of a cell. By combining the FSC and SSC information with fluorescence labeling, it is possible to differentiate cell types in a heterogeneous population such as blood.

The detectors in most flow cytometers are usually PMTs. The specificity of detection is controlled by optical filters. There are three major filter types: long pass filters that transmit light above a cutoff wavelength; short pass filters that transmit light below a cutoff wavelength; and band pass filters that transmit light within a narrow band of wavelengths (see Chapter 1, Section III.A.2 for details). These are dichroic filters that block light by phased reflection, allowing only certain wavelengths of light to pass through while interfering with other wavelengths (Figure 278).

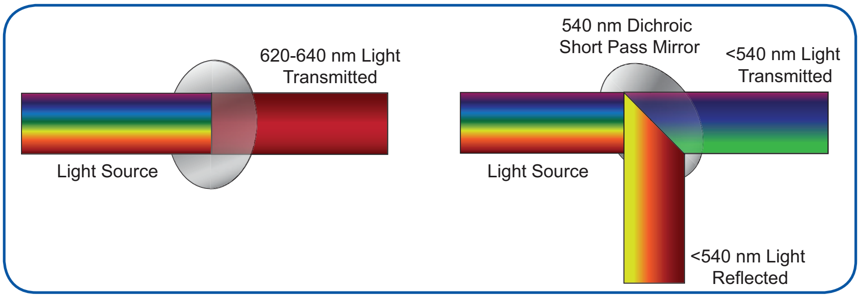


Figure 278. Dichroic optical filters.

When placed at an angle to the oncoming light, a dichroic filter acts as a mirror, allowing it to perform two functions: transmitting specific wavelengths in the forward direction and reflecting the remaining light at a 90° angle. This allows the light path to be passed through a series of filters. The precise choice and order of the filters can be arranged so that multiple signals can be detected simultaneously (Figure 279).

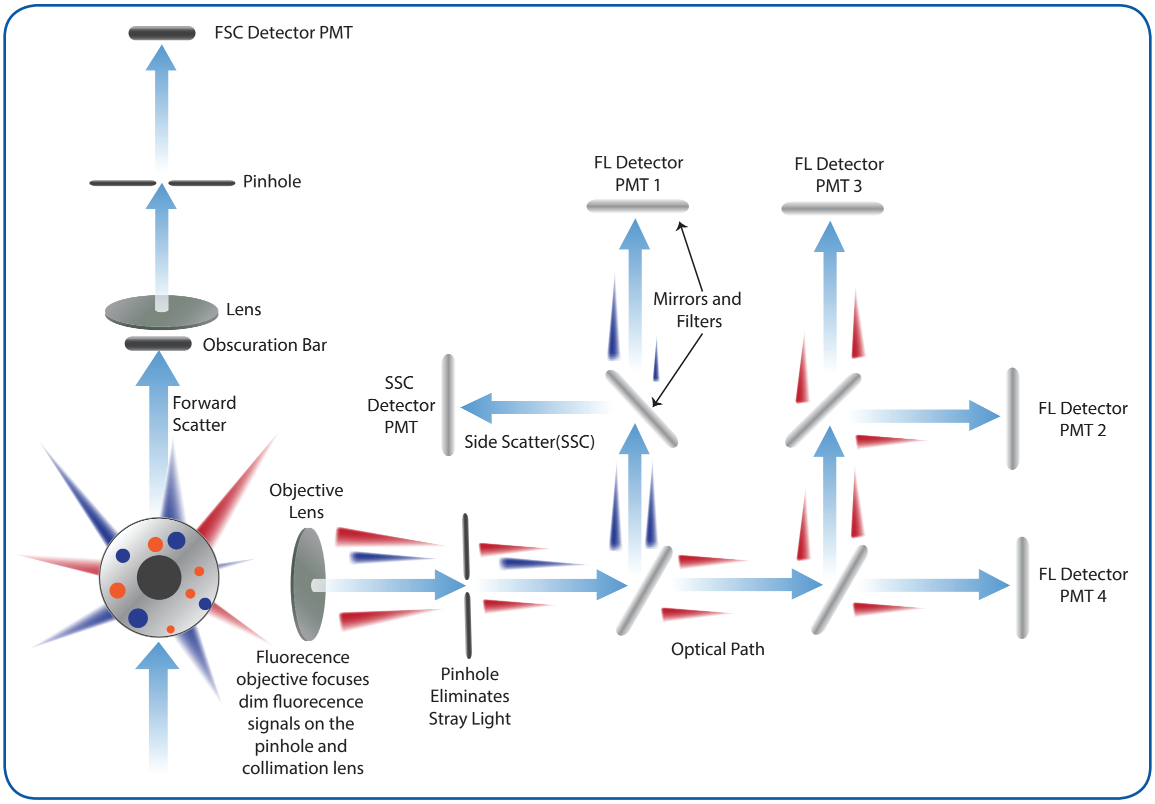


Figure 279. Schematic overview of the optical configuration of a typical flow cytometer setup.

3. Signal and Pulse Processing

Every time a particle passes through the interrogation point, a signal pulse is generated in every detector in Figure 279 with the current from each PMT proportional to the intensity of the scatter or fluorescence signal generated by the cell. These pulses can be mapped by plotting signal as a function of time.

Not all signals that are generated correspond to a particle of interest. PMTs are extremely sensitive and detect signals from irrelevant sources such as stray light, dust, very small particles and debris. The number of these irrelevant pulses can be orders-of-magnitude higher than the number of pulses that are generated by particles of interest. It is therefore desirable and necessary to set a threshold below which non-essential data are ignored. This is done by designating a trigger channel, usually a forward scatter detector, and setting a threshold signal intensity in that channel for recording scattering events. Any pulse that fails to exceed the threshold level is ignored in all detectors; any pulse that surpasses the threshold level is fully processed by the electronics.

The electronics process fluorescence signals for display, analysis, and interpretation. The analog current from the PMT is typically digitized and the pulse height, area, and width determined. The height and area, or maximum and integral, respectively, are used to measure signal intensity since their magnitudes are proportional to the number of photons that reach the PMT. The width of the pulse is proportional to the time that the particles spend in the laser and this can be used to distinguish between single particles or closely interacting particles and doublets.

C. Applications of Flow Cytometry

The ability to simultaneously measure multiple parameters on a cell-by-cell basis is the most powerful attribute of analytical flow cytometry, making it suitable for a wide range of applications. Most commonly, it is used to determine the presence of antigens either on the surface or within cells. In addition, flow cytometry may be used for the analysis of DNA or RNA content, and for functional studies on cells. The following broad range of technological activities employ flow cytometry analytical tools:

- Medicine
  - Hematology
  - Oncology
  - Immunology
- Genetic testing
- Biochemistry and molecular biology, e.g., proteomics, glycomics
- Marine science
- Biosynthesis
- Cell health and biology (including stem cells)
- Screening
- Cell cycle analysis
- BioProcess

D. Flow Cytometry Applications Using MKS Products

MKS Instruments provides equipment for flow cytometry applications at the component, sub-assembly and fully integrated systems level:

- Opto-mechanical components
  - High quality opto-mechanical components
  - Motion stages, controllers, and software
  - Laser power sensors and power meters, beam profilers
- Laser technologies
  - MKS Spectra-Physics offers several lasers for flow cytometry applications. Three examples are shown in Figure 280 with the Vanguard™ and Millennia® being high-power lasers used in high-throughput applications. The Excelsior® is a family of low-power lasers with 14 different wavelengths. More information about these lasers for flow cytometry applications is available in [336].

Figure 280. Spectra-Physics lasers for flow cytometry: Vanguard™ (left), Millennia® (middle), and Excelsior® (right).

- Opto-mechanical and electromechanical design and manufacturing
  - Figure 281 shows a complete flow cytometry system that was designed and manufactured by MKS ISB.

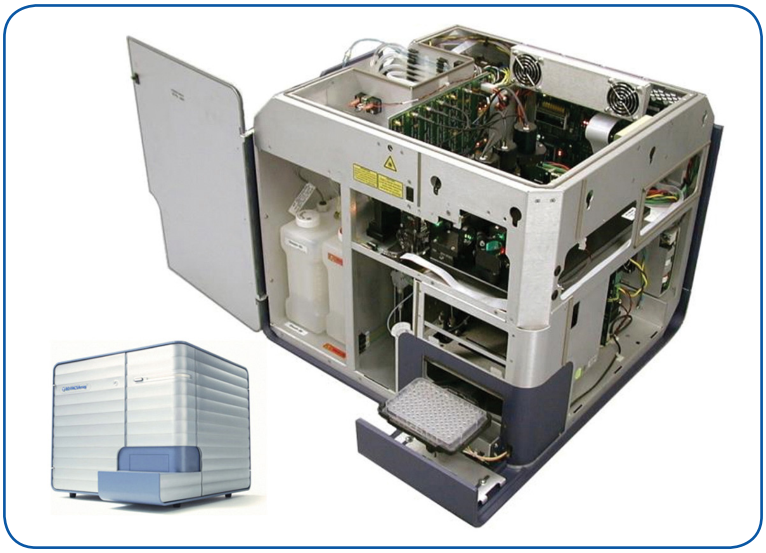


Figure 281. Flow cytometer system.

V. Tunable Light Source (TLS)

A. Introduction

A conventional lamp-based TLS produces monochromatic light with wavelengths that range from the UV to the NIR. TLS finds use in applications requiring coordinated operation of a light source and a detection instrument with data acquisition and processing. The TLS is versatile, being both a broadband and high-resolution monochromatic light source which makes it suitable for applications such as the study of wavelength-dependent chemical or biological properties and wavelength-induced physical changes in materials. These light sources can also be used in color analysis and to measure the reflectivity of materials for quality purposes. The key components of a lamp-based TLS are a highly stable light source and a monochromator. Arc lamps and QTH lamps are commonly used in a TLS.

This section describes different TLS systems including their performance specifications and targeted applications.

B. Bright Emissions vs. Monochromatic Throughput

Arc lamps (Figure 282) are excellent sources of *CW*, broadband light (see Chapter 1, Section I.C for details). They consist of two electrodes (an anode and a cathode) separated by a gas such as neon, argon, mercury, or xenon. Light is generated by ionizing the gas between the electrodes. The bright emission from the short arc between the anode and cathode means that these lamps can be considered as high intensity point sources of light that can be collimated using the proper lens configuration. Arc lamps offer the advantages of long lifetime, superior monochromatic throughput (particularly in the UV range), and a small divergence angle. They are particularly well-suited for fiber coupling applications. Xe arc lamps, in particular, have a relatively smooth emission curve in the UV to VIS spectral region, with characteristic wavelengths emitted from 750-1000 nm. Their sun-like emission spectrum and ~5800 K color temperature make them a popular choice for solar simulation applications.

Figure 282. Lamp housing with arc lamp.

QTH lamps produce light by heating a filament wire with an electric current (see Chapter 1, Section I.C for details). The hot filament wire is surrounded by a vacuum or inert gas to prevent oxidation. QTH lamps produce continuous blackbody spectrum and exhibit very accurate color reproduction. They are a popular alternative to arc lamps owing to their stable, high output intensity and to the smooth spectral emission lines in their output curve (Figure 283). QTH lamps are preferable for radiometric and photometric applications and as excitation sources of VIS to NIR light. QTH lamps are also easier to handle and install. Selecting the most appropriate lamp type is a matter of deciding which performance criteria are most important when assembling a TLS.

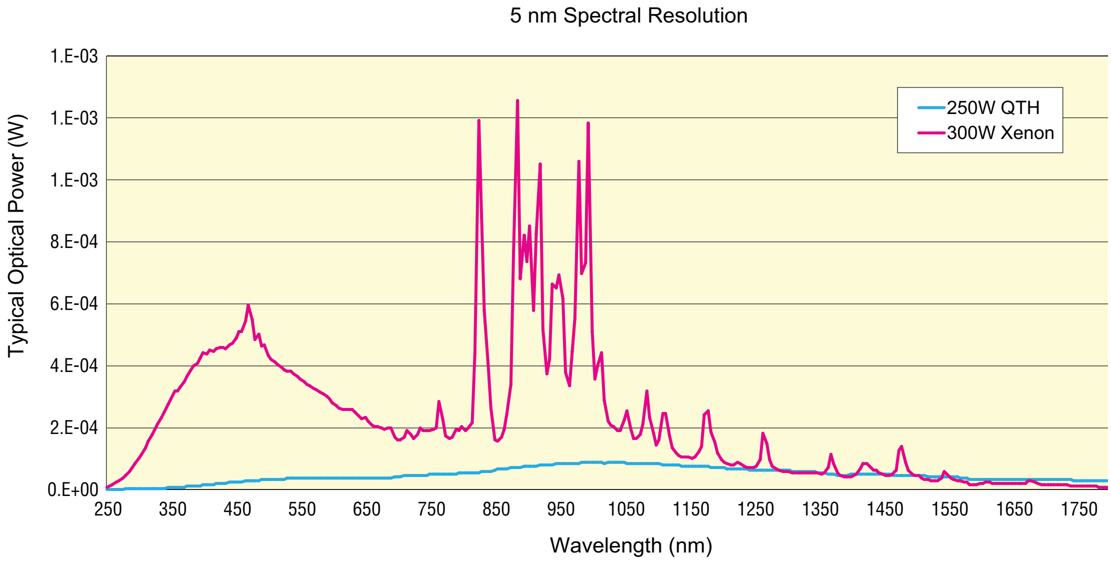


Figure 283. Spectral irradiance of QTH and Xe lamps.

C. Monochromators

Monochromators (Figure 284) use diffraction gratings to spatially isolate and select a narrow band of wavelengths from broadband radiation (see Chapter 1, Section II.D.2). They can be used to create quasi-monochromatic light and to obtain highly precise spectral measurements. A high precision stepper motor is typically employed to select the desired wavelength and to quickly switch between diffraction gratings without reducing an instrument's performance.

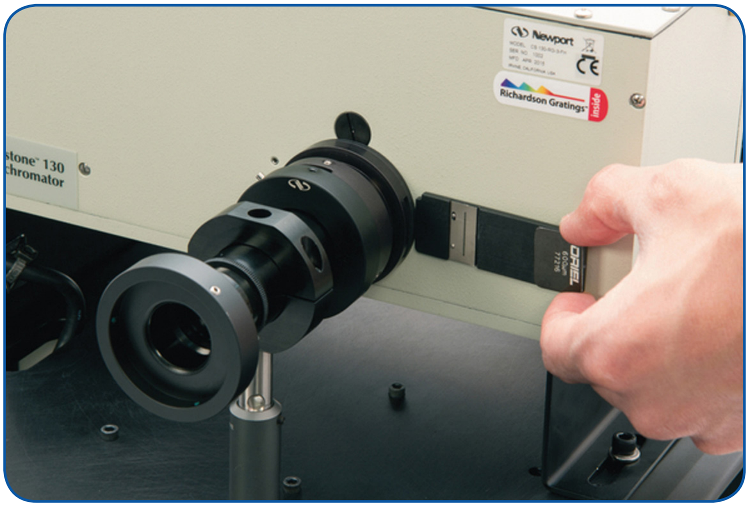


Figure 284. Monochromator with slit at exit port.

The slit width in a monochromator is usually selected to achieve the bandpass or spectral resolution required by a given application. Determining the best slit width requires a trade-off between light throughput and the resolution demanded by the application. A larger slit width allows greater light throughput; however, more light throughput results in poorer resolution. In operation, focused light enters the monochromator through the entrance slit and is redirected by the collimating mirror toward the grating. The grating directs the light toward the focusing mirror, which then redirects the chosen wavelength toward the exit slit. Note that both the input and output ports of a monochromator must use the same slit width. At the exit slits, quasi-monochromatic light is emitted.

D. Application of TLS – Measuring Quantum Efficiencies

A TLS provides an ideal light source for the measurement of device quantum efficiency (QE) over different wavelengths. QE measurements determine the ratio of charge carriers (electrons or holes) generated in a material to the number of photons impinging on its surface. In practice, this is done by subjecting a solar cell to a known flux of light from a calibrated TLS while simultaneously measuring its output electrical current. External quantum efficiency (EQE) is the ratio of the number of photons incident on a solar cell to the number of charge carriers generated. Internal quantum efficiency (IQE) is more complex in that it considers the internal efficiency — that is, the losses associated with the photons absorbed by non-active layers of the cell. EQE provides a direct measure of how much output current will be generated per incident photon at a given wavelength, while IQE includes the losses from non-active layers of the material to calculate a net QE — a much truer efficiency measurement. (Quantitatively, IQE = EQE/(1 – R), where R is the reflectivity of the solar cell.) The IQE is an indication of the capacity of the active layers of the solar cell to make good use of the absorbed photons. It is always higher than the EQE, but should never exceed 100%, except in the case of multiple-exciton carrier generation. Accurate QE measurement requires exact, quantitative knowledge of the light flux incident on the DUT and of the current generated by that light flux. The TLS output must always be calibrated using a detector having

NIST-traceable calibration to ensure accurate knowledge of the incident light flux in QE measurements; the current should be similarly measured with equipment calibrated to known standards. QE measurements provide critical understanding of the conversion efficiency as a function of wavelength in materials research and solar cell design. With this data, solar cell composition and topography can be modified to optimize conversion over the broadest possible range of wavelengths.

Figure 285 illustrates how a TLS is used to illuminate the solar cell in an IQE measurement. All components of the measurement system are software-controlled, including the monochromator and data acquisition components. The measurement of QE in 10 nm wavelength steps is typical and this requires a monochromator slit width of hundreds of microns. The slit width must be reduced if small wavelength increments are desired. For example, the slit width for 5 nm wavelength increments is approximately half that for 10 nm increments. However, the output optical power of the monochromator is reduced by more than 50% if the slit width is halved and this negatively impacts QE measurement since the solar cell under test will produce lower output currents that require extremely sensitive measurement equipment that often exhibit poor SNR. This will make errors more likely in the QE measurement.

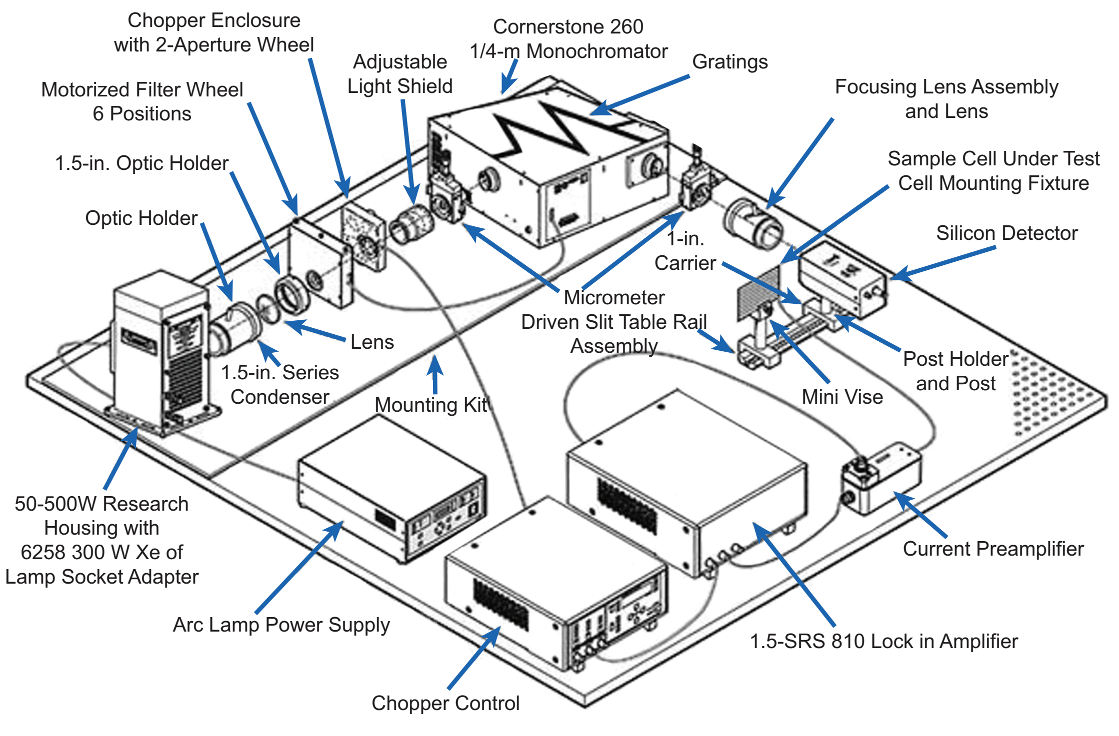


Figure 285. QE measurement system layout.

Arc lamp sources are generally the better choice for QE measurements made with 5 nm or lower wavelength increments since the arc size produces better monochromator throughput. Conversely, a QTH lamp is the better choice if greater than 0.1% light stability is required. However, a QTH lamp cannot achieve wavelength increments as precise as an arc lamp. Balance between the optical power and the resolution in a QE measurement is important since it strongly impacts the quality of the measurement. This makes the selection of lamp type and monochromator specifications an extremely important consideration for TLS design. To be considered a suitable component for most spectroscopic applications, high-output power and stability, long lifetime of the lamp, and broadband spectral emission with high-resolution capability are required in a TLS.

VI. Solar Simulators

A. Basics of Solar Radiation

The surface temperature of the sun is approximately 5800 K; this means that the electromagnetic spectrum of radiation from the sun is similar to that of a 5800 K blackbody (see Chapter 1, Section I.C for details), with the exception that it includes fine structure due to absorptions by cool gases in the solar periphery (Fraunhofer lines). The solar irradiance on the earth's outer atmosphere when the sun and the Earth are 1 Astronomical Unit (the mean earth-sun distance of 149,597,890 km) apart is 1360 W m^{-2}. This value, called the Solar Constant, is the total integrated irradiance over the entire electromagnetic spectrum. Figure 286 shows the electromagnetic spectrum of the solar radiation outside of the Earth's atmosphere. The range shown, 200 to 2500 nm, includes 96.3% of the total irradiance reaching the Earth with most of the remaining 3.7% at longer wavelengths.

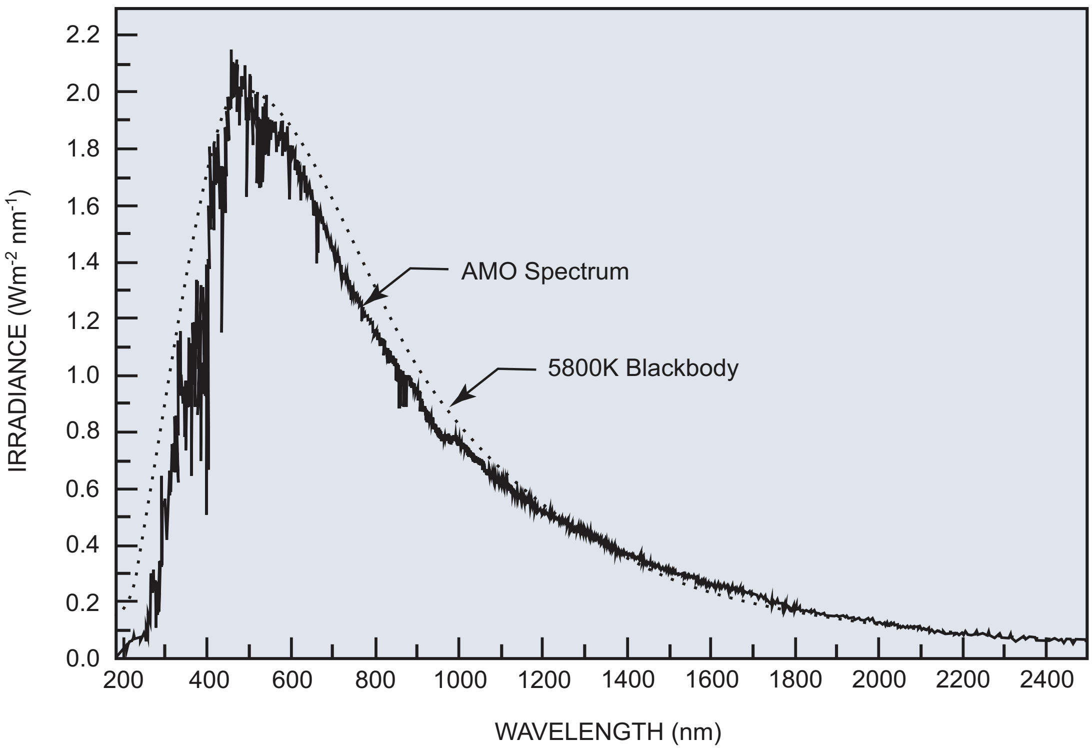


Figure 286. Spectrum of the solar radiation outside the earth's atmosphere compared to the spectrum of a 5800 K blackbody.

The electromagnetic spectrum of the solar radiation at the Earth's surface is influenced by several contributing factors, including direct and diffuse radiation. Direct radiation is energy received at the Earth's surface in a direct path from the sun, while diffuse radiation is energy received from light scattered from the sky and reflected by surroundings (Figure 287). The total radiation measured at the surface is called the Global Radiation where the direction of the target surface must be defined for global irradiance. For Direct Radiation, the target surface is orthogonal to the incoming radiation. The sun is a spherical source, about 1.39 million km in diameter. The average distance to Earth is 1 astronomical unit. The direct portion of the solar radiation is collimated with an angle of approximately 0.53 (full angle), while the "diffuse" portion is incident from the hemispheric sky and from ground reflections and scatter. The "global" irradiation, the sum of direct and diffuse components, is essentially uniform.

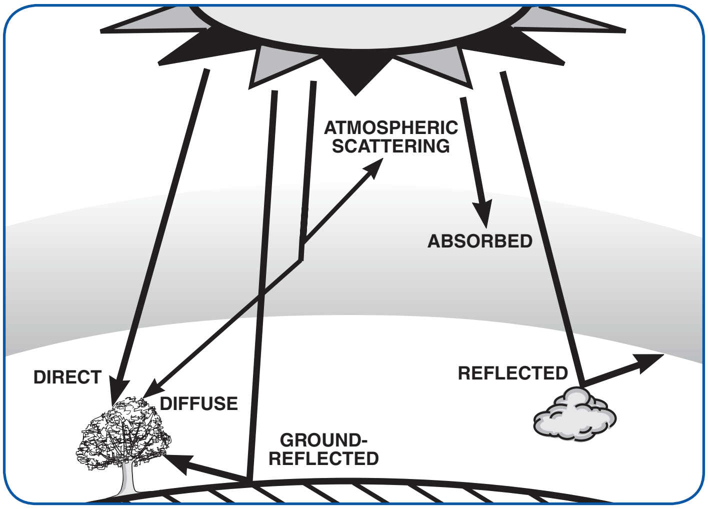


Figure 287. The total global radiation on the ground has direct, scattered, and reflected components.

The measured solar radiation incident upon the earth's surface is influenced by the amount of atmosphere through which it must pass. At any location, the length of this path to reach ground level changes as the day progresses, and there are obvious changes in ground solar radiation levels during the day as the angle of the sun changes. In addition, the shape of the spectrum can change throughout each day due to changing absorption and scattering length. Because it passes through no air mass, the extraterrestrial spectrum is called the Air Mass 0 (AM 0) spectrum. With the sun directly overhead at noon at the Earth's equator, solar radiation passes straight through the atmosphere. The solar spectrum at sea level under these conditions is called the Air Mass 1 spectrum. The global radiation with the sun directly overhead is similarly called Air Mass 1 Global (AM1G) radiation. The atmospheric path for any zenith angle is simply described relative to the overhead air mass (Figure 288). The actual path length can correspond to air masses of less than 1 (high altitude sites), to very high air masses just before sunset. MKS Solar Simulators use filters to simulate spectra corresponding to air masses of 0 and 1.5, the values on which most comparative test work is based.

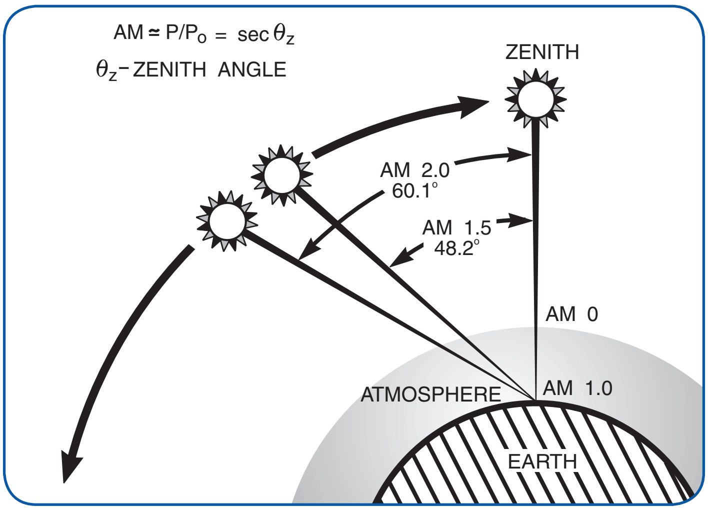


Figure 288. The path length, in units of Air Mass, changes with the zenith angle.

B. Xenon-based Solar Simulators

Solar simulators provide the closest spectral match to the solar spectrum available from any source. While not exact, the match is better than needed for most applications. Figure 289 shows the optics of a typical solar simulator. The Xenon arc lamp at the heart of the device has a small, high radiance arc that allows efficient beam collimation. It emits a spectrum similar to a 5800 K blackbody with occasional line structure. The system design features include low F/# collection geometry, optical beam homogenization and filtering and finally, collimation. The simulator produces a continuous output with a solar-like spectrum in a uniform, collimated beam. Beam collimation simulates the direct terrestrial beam and allows for characterization of radiation induced phenomena.

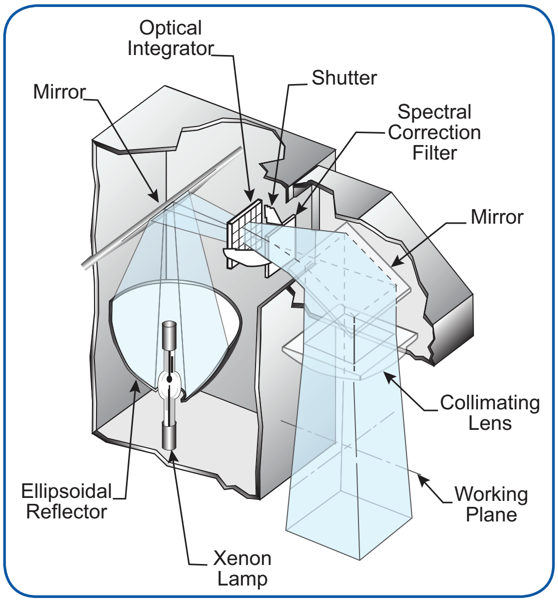


Figure 289. Light path of a solar simulator.

The Xenon lamp output differs from the solar spectrum in the 800 to 1100 nm wavelength range due to the intense atomic emission lines of Xe gas. MKS solar simulators use filters to reduce the influence of these atomic emission lines. The impact of the residual IR mismatch depends on the application. The AM Direct and AM 1.5 filters modify the VIS and UV portion of the spectrum for a better match to the standard terrestrial solar spectra.

C. UV Solar Simulators

UV-enhanced solar simulators are useful for work requiring an intense UV source without the complications of VIS and IR heating. This is especially important for biological work on live subjects. UV-enhanced simulators are like full spectrum solar simulators; however, their spectral distribution is modified to reduce the radiation in the VIS and IR regions of the spectrum. The UVA and UVB regions are of particular interest because natural sunlight includes UVA and UVB radiation.

UV-enhanced solar simulators typically are available in three configurations:

- UVA, UVB, and UVC — this source produces UV from 210 to 400 nm
- UVA and UVB — this simulator produces UV from 280 to 400 nm and has a UVC blocking filter
- UVA — this system produces UV from 320 to 400 nm and has UVB and UVC blocking filters

D. UV Solar Simulator Applications Using MKS Products

Photobiology studies the effect of UVA and/or UVB, VIS, and IR radiation upon living systems. In biological systems, light must first be absorbed by either the skin or an agent on the skin for a biological photochemical event to occur (the Grotthaus-Draper Law). The chromophores in drug products and the DNA in dermal tissue are targets for such photochemical reactions. In the case of drug products, photochemical reactions can produce photoirritation and/or photoallergy responses (see below) when a photoactive chemical enters the skin by dermal penetration or systemic circulation and becomes excited by appropriate UV or VIS light photons. Photosafety testing is suggested for any chemicals that absorb light in the range from 290-700 nm and for any substance that may be applied topically or can reach the skin and/or eyes by systemic exposure (oral or intravenous). There are four basic results or endpoints that photosafety testing addresses:

- Phototoxicity — sometimes referred to as photoirritation, an acute light-induced skin response to a photo-reactive chemical
- Photoallergy — an immune reaction to a chemical initiated by the formation of photo-products; this can be a byproduct of exposure to an antigen
- Photogenotoxicity — a genotoxic response after exposure to a chemical which is photo-activated by UV or VIS light
- Photocarcinogenicity — the potential for a chemical to promote skin tumor formation in combination with exposure to UV light

MKS Instruments' Sol-UV Solar Simulators are used in a variety of commercial and research dermatological applications. In commercial applications, MKS UV solar simulators have been used by skin product manufacturers to verify their products' compliance with regulations for Sun Protection Factor (SPF) testing under the United States Food and Drug Administration (FDA), the European Cosmetic Toiletry Perfumery Association (COLIPA), and the Japan Cosmetic Industry Association (JCIA).

Large area solar simulators are used to project UVA and UVB (280 - 400 nm) light onto a sample or a volunteer. By using an exposure controller for timed exposures or dose-controlled exposures for in-vitro testing of samples or in-vivo testing of volunteers, skin care product manufacturers can quickly determine the effectiveness of their products (Figure 290).

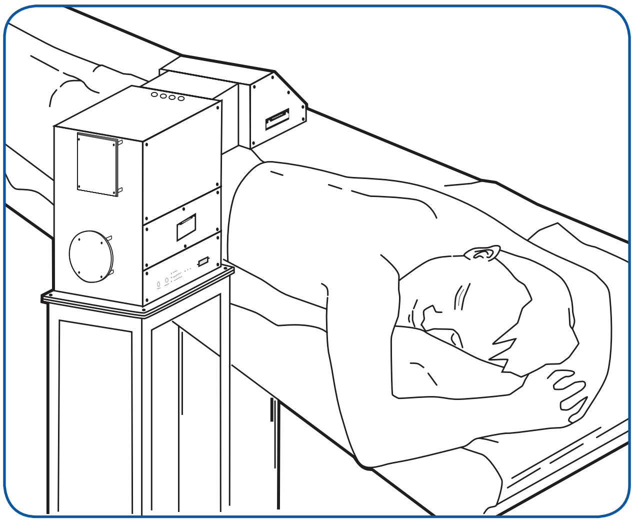


Figure 290. MKS Sol-UV Solar Simulators, used to verify skin product compliance with FDA, COLIPA and JCIA regulations.

Many diverse classes of drugs (including antimicrobials, nonsteroidal anti-inflammatory drugs, antidepressants, anticonvulsants, diuretics, and antihypertensives) have been reported to cause photoirritation in humans. For pharmaceuticals or cosmetic chemical components, basic photosafety testing is performed using in-vivo animal or in-vitro cell culture studies where samples exposed to UV radiation are compared to samples without UV exposure and to controls without the chemical exposed to the same dose of UV light. Most test methods call for an irradiation spectrum that approximates the solar spectrum by using appropriate filters to remove the UVC component but pass the UVA and UVB. Since specific guidance in intensities is not usually supplied, the intensity of UVA and UVB in tests can differ significantly, depending on the light source used. MKS Instruments' Sol-UV Solar Simulators provide a defined spectral output certified to be compliant to COLIPA and FDA CFR Part 201.327, ISO 24444:2010(e) (Figure 291). The use of a standardized light source facilitates the comparison of experimental data by eliminating the variables associated with using unstandardized and undocumented light sources having different spectral profiles in the UV.

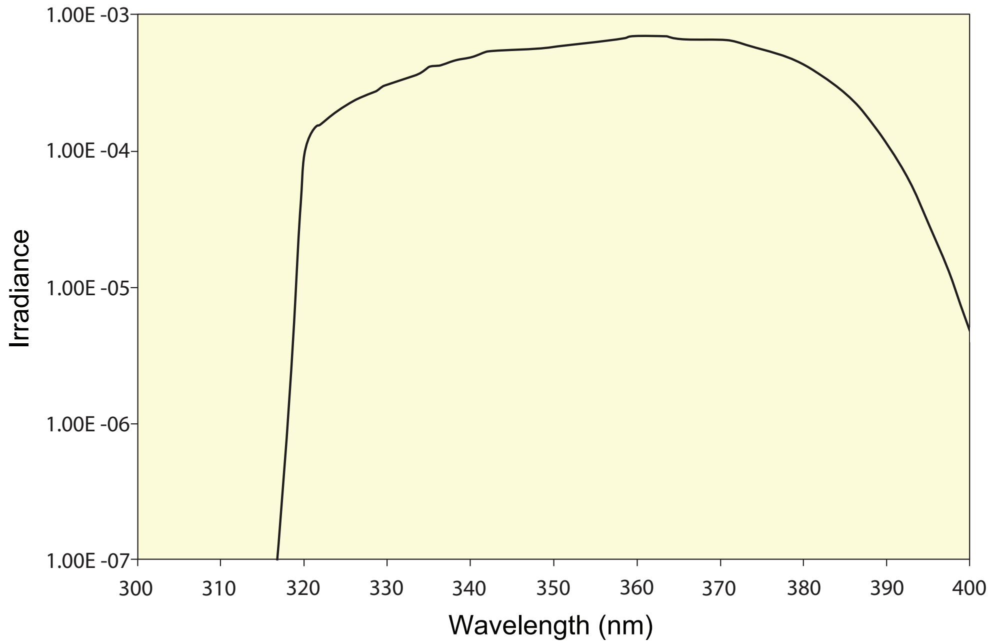


Figure 291. COLIPA and FDA compliant spectral irradiance curve.

Along with spectral consistency, uniformity of irradiance over the work area must be maintained, as so-called hot spots can lead to errors in delivered dosage. The Sol-UV Simulator delivers UV radiation with less than 5% non-uniformity across the entire work area. Temporal stability is also important for minimizing dosage errors. MKS' experience in producing ultra-stable power supplies and feedback controllers guarantees the best short- and long-term stability and assures that the output light is stable over time, minimizing errors in delivering the desired dosage.

This page intentionally left blank.

Chapter 7

Photonics in Life and Health Sciences

This page intentionally left blank.

I. Introduction

Life sciences involve the scientific study of organisms, such as microorganisms, plants, animals, and human beings. Health sciences are pivotal to improving the quality and standard of life. Together, life and health sciences cover a wide range of applications in health care, agriculture, medicine, and the pharmaceutical and food science industries [337].

Optical techniques provide valuable tools for life and health sciences in the areas of visualization, measurement, analysis, and manipulation [338]. Lasers, LEDs, and photonics components play key roles in these optical techniques and Chapter 1 – Introduction to Photonics provides a review of these components. Lasers have been applied to imaging techniques such as optical coherence tomography, a high-resolution imaging technique based on low coherence interferometry, as well as photoacoustic imaging, which is described in Section II.A. Confocal laser scanning fluorescence microscopy has provided highly detailed 3D images of biological structures. Two-photon microscopy techniques have not only enhanced the capabilities of fluorescence microscopy but have also opened new possibilities for performing highly localized photochemistry within cells. The limit of optical resolution in a conventional microscope, the so-called diffraction limit, is on the order of half the wavelength of the light used to image. Recent progress in super-resolution microscopy has enhanced resolution to beyond the diffraction limit. These microscopy techniques are widely used to identify neurodegenerative diseases and cancer. Section II of this chapter explores these laser-based microscopy techniques in detail.

The application of photonics to health care is changing the practice of medicine and surgery and offering new approaches to both therapy and diagnostics. Because of their unique capability to remove controlled amounts of material with extremely low collateral thermal damage, ultrafast lasers have been used for precise manipulation of tissue and individual cells in a wide area of biomedical research applications, including the processing of teeth and other hard tissues [339], and nanosurgery in developmental biology and neuroscience [340]. Furthermore, lasers have enabled many new therapies, ranging from the treatment of kidney stones to the removal of skin lesions. Lasers and LEDs have been applied to phototherapy for dermatological conditions and photodynamic cancer therapies. Lasers have also been applied to general surgery such as endovascular surgery and gastrointestinal surgery. Some fields, such as ophthalmology, have completely integrated lasers into clinical practice. Section III discusses several applications of lasers in ophthalmology.

Flow cytometry, an optical diagnostic technique, is a critical tool for monitoring viral loads in AIDS patients and for guiding their therapy. In biotechnology, lasers have become essential parts of all the systems used for DNA sequencing, ranging from commercially-available systems to more research-based capillary electrophoresis systems. Chapter 6 – Photonics in Analytical Instruments reviews the instrumentation used in these applications.

Laser tweezers represent a creative application of photonics to biology [338, 341] that was recognized with the 2018 Nobel Prize in Physics [342]. Normal optical forces are minuscule on the scale of larger organisms but they can be significant on the scale of macromolecules, organelles, and even whole cells. Laser tweezers thus afford an unprecedented means for manipulation at microscopic scales. These techniques have led to trapping and manipulation of single living cells, organelles within cells, single biological molecules, and the measurement of mechanical forces and elastic properties of cells and molecules [343]. Laser tweezers can tow a bacterium through water faster than it can swim, halt a swimming sperm cell in its track, or arrest the transport of an intracellular vesicle.

II. Advanced Bioimaging

Optical microscopes have been instrumental in the study of the life sciences for centuries. Since the invention of the laser, many advanced bioimaging techniques using microscopy have been created with progress accelerating greatly in the last two decades. Limitations in optical penetration depth, due to the scattering of light, originally limited studies to thin samples and, by necessity, processes taking place outside of the organism (ex vivo). In recent years, in vivo techniques have been developed that can visualize within the living organism or cell and have led to greatly increased understanding of cell function. Any bioimaging technique requires generating a signal from each portion of the cell or organism through some method of contrast. Some of these techniques require an exogenous contrast agent (originating outside of the organism or cell) to be added to the sample under investigation, such as a fluorescent dye. These agents can be specifically engineered and targeted to particular molecules. In other cases, the signal can be generated with endogenous contrast agents (those that originate naturally from within the organism or cell). This route has the advantage of not perturbing the original tissue micro-environment and limits possible cell toxicity.

Traditional microscopy techniques provide excellent spatial resolution, typically below 1 μm, in the lateral directions (within the plane of the sample). However, the axial resolution, into the depth of the sample, is typically much poorer. Multiple advanced bioimaging techniques have been developed to improve the axial resolution and allow a true 3D representation of the sample to be reconstructed. Chief among these techniques is confocal microscopy, where the axial resolution is improved by the addition of a confocal aperture that discriminates against light emanating from different depths in the sample. Confocal microscopy is described in Section II.C.

A wide range of laser-based nonlinear microscopies have been developed in the last two decades that make use of the high peak powers available from ps and fs laser sources. In each case, the excitation is due to an intensity-dependent effect which is most efficient at the focus of the laser beam. This can improve the axial resolution to about 1 μm without the use of a confocal aperture. Sections II.B, II.C, and II.D describe nonlinear microscopy techniques and a summary of the various mechanisms that enable these techniques is presented in Table 19. Chapter 1, Section I.A.6 describes various types of wavelength conversion and generation processes that are pertinent to these techniques. All of the techniques require a high peak power ultrafast laser and some require two synchronized ultrafast lasers at different wavelengths. Some techniques are applicable to endogenous contrast agents while others require the addition of a fluorescent dye. Figure 292 shows the energy level diagram for each mechanism which reveals how the laser light interacts with the system to produce a measurable signal. These diagrams will be referred to in the following sections.

Technology	Fluorescent Dye	Pulse Duration	Number of Synchronized Wavelengths
Two-Photon Fluorescence (2PF)	X	fs	1
Second Harmonic Generation (SHG)		fs	1
Three-Photon Fluorescence (3PF)	X	fs	1
Third Harmonic Generation (THG)		fs	1
Coherent anti-Stokes Raman Scattering (CARS)		ps or fs	2
Stimulated Raman Scattering (SRS)		ps or fs	2
Sum-Frequency Generation (SFG)		fs	2

Table 19. Comparison of mechanisms used for nonlinear microscopy techniques.

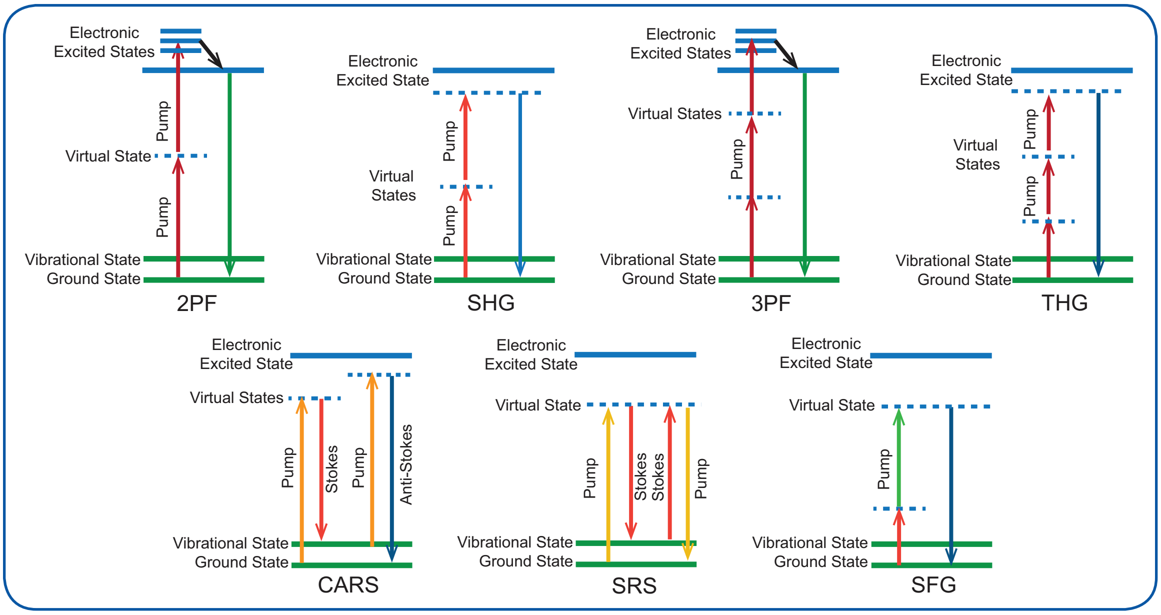


Figure 292. Energy level diagrams for the mechanisms described in Table 19.

A. Photoacoustic Microscopy

Photoacoustic microscopy (PAM) is a hybrid in vivo imaging technique that detects optical contrast via the photoacoustic effect [344, 345]. Unlike the pure optical microscopy techniques discussed in the following sections, PAM takes advantage of the weak acoustic scattering in tissue and thus breaks through the optical depth penetration limit (~ 1 mm) in soft tissues. Since ultrasonic scattering in tissue is three orders of magnitude weaker than the optical scattering, PAM can effectively image at depths up to a few millimeters.

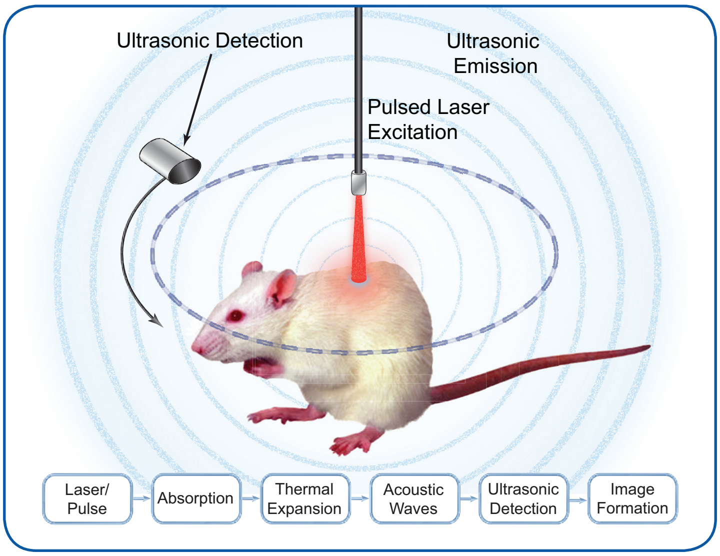


Figure 293. Photoacoustic microscopy using a pulsed laser excitation source [344].

In PAM (see Figure 293), a laser beam is focused using an objective lens and the light passes through a water tank that is in contact with the sample to be imaged. This light is absorbed at the focus, heats the sample, and induces an initial pressure rise which propagates through the tissues as a wide-band acoustic wave. An ultrasonic transducer then detects the acoustic wave. The lateral resolution is determined by the focus of the laser beam, and is typically a few microns, while the axial resolution is about 15 μm. Both endogenous and exogenous contrast agents have been used successfully in PAM imaging. The laser sources typically have ns pulse durations with wavelengths dictated by the absorption of the contrast agents. A representative laser source for PAM is a low repetition rate (10-30 Hz) ns-pulsed Nd:YAG laser. Such a system typically operates at 1064 nm, but this wavelength can be converted to 355 nm via THG and produce pulse energies of several hundreds of millijoules. This source is then used to pump an OPO, which generates ns pulses throughout the VIS portion of the spectrum (see Chapter 1, Section I.A.6 for details regarding wavelength conversion and generation).

Figure 294. Spectra-Physics primoScan OPO, when pumped by Spectra-Physics Quanta Ray® Laser, produces widely tunable ns pulses suitable for PAM imaging.

MKS provides ideal laser sources for PAM. The Spectra-Physics® primoScan series (see Figure 294) features a compact OPO design that provides tunability from 190 nm to 2750 nm. This fully-automated OPO delivers high wavelength conversion efficiencies and good beam quality throughout the entire tuning range with repetition rates from 10 Hz to 50 Hz. The primoScan OPO can be pumped by a range of Spectra-Physics Quanta Ray® INDI, Lab, and Pro Nd:YAG lasers with peak energies reaching 750 mJ, the highest output energy level available from a commercial OPO. The primoScan includes fast wavelength switching, allowing access to ten pre-selectable wavelengths each second and the ability to easily access all tunable beams from a single output port. In addition to PAM imaging, the primoScan is ideal for applications involving materials analysis, narrowband and broadband spectroscopy, laser-induced fluorescence, remote sensing, LIDAR, multiphoton interactions, and combustion studies.

B. Super-Resolution Microscopy

The axial resolution in standard confocal microscopy is limited by the microscope objective and the wavelength of light. Even by using short-wavelength light and an oil immersion objective, the resolution of a standard fluorescence system is limited to around 200 nm [346]. Recently, super-resolution microscopy techniques have been developed which break this limit and allow imaging with an axial resolution down to a few tens of nm. This work was recognized with the 2014 Nobel Prize in Chemistry [347, 348].

One super-resolution approach is called stimulated emission depletion (STED) microscopy. This technique uses a confocal microscope and a diffraction limited excitation beam that is scanned across the sample. A second beam (the depletion beam) is spatially shaped like a donut and, when overlapped with the excitation beam, forces the illuminated portion of the sample to emit light at one wavelength. The non-depleted region then emits light at a different wavelength, which is filtered and measured. Because the central hole in the donut can be made much smaller than the diffraction limit, the non-depleted area used for imaging will be smaller as well.

A second approach to obtain super-resolution is classified as localization microscopy. Two of the most popular techniques are called photoactivatable localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM). A fluorophore is a fluorescent chemical compound that can emit light upon excitation. When measuring the spatial distribution of a single fluorophore through the microscope, the center of the distribution can be determined with greater accuracy than its width. This is the basis for improved resolution using localization microscopy. This approach requires that a small number of fluorophores are emitting in a single captured frame and that there is no overlap with neighboring emitters. In order to make up for the necessarily weak signal strength, the super-resolution image is built up from several hundred thousand images of the localized positions of the individual fluorophores. This can take several minutes and therefore may not be fast enough to capture processes in live cells. The PALM and

STORM techniques use slightly different methods to control the behavior of the fluorophores in order to keep the emission well-separated in time.

MKS offers a large range of *CW* and quasi-*CW* lasers of low to moderate output powers that can be used for super-resolution microscopy and other bioimaging applications. The Spectra-Physics Excelsior® family (see Figure 295) of low-power *CW* solid-state and semiconductor lasers offers customers a complete range of wavelengths and power levels. DPSS lasers are available at a variety of wavelengths: 473, 505, 515, 532, 542, 561, 594, and 1064 nm. For applications requiring a 488 nm laser, Spectra-Physics offers an Excelsior laser based on highly reliable, externally frequency-doubled diode laser technology. The Excelsior laser is the world's smallest 488 nm laser head with a 40% narrower width than other commercially available lasers. In addition to these wavelengths, laser diodes are also available at 375, 405, 440, 642, and 785 nm. Excelsior lasers provide state-of-the-art performance with the smallest footprint in their class. The Spectra-Physics Excelsior product line provides a consistent platform with the same mechanical footprint over a wide range of power levels and wavelengths. This allows for full component interchangeability among laser heads and controllers within a given architecture, requiring no additional adjustment or optimization.

Figure 295. Spectra-Physics® Excelsior® One™ laser delivers CW output at a wide variety of wavelengths in the VIS and NIR.

C. Two-Photon Fluorescence Microscopy

Conventional wide-field microscopy provides micron-level resolution laterally across a sample. In contrast, resolution into the depth of the sample, sometimes notated as the Z direction, can be considerably poorer due to signals that originate from various depths within a sample leading to obscuration of its origin. In a confocal laser scanning microscope, the fluorescence generated from a focused laser is imaged onto a pinhole, or confocal aperture, which blocks the fluorescence that is generated above or below the focal plane before reaching the detector. The intensity of the fluorescence is then measured as the beam is scanned horizontally across the sample to build up a 2D image. Finally, the sample is sequentially moved in the Z direction, allowing for 3D image construction with a spatial resolution close to the diffraction limit. This technique of confocal laser scanning microscopy dates back to 1969 [349].

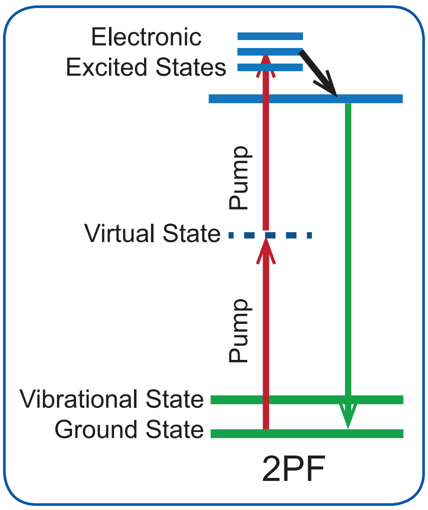


Figure 296. Energy level diagram showing 2PF of a fluorescent dye.

Confocal microscopes are essential tools in today's biomedical research laboratories, but they do have limitations. Due to losses at the pinhole and scattering of the excitation light, confocal microscopy techniques are best applied to thin specimens (< 20 µm). While the out-of-focus excitation does not contribute to the 3D image, it does increase the likelihood of photodamage in living specimens. In addition to the potential photodamage caused by typically-used VIS and UV wavelengths, they also do not penetrate deeply into tissue due to scattering. In 1990, workers at Cornell University [350] demonstrated the use of two-photon excitation with laser scanning microscopy [351-354], denoted here as 2PF microscopy. In 2PF, two photons are simultaneously absorbed to cause a higher energy electronic transition in a fluorescent molecule as shown in Figure 296.

Since the nonlinear process of 2PF has a low cross-section for excitation, a high peak power laser source is required as well as a high NA microscope objective. This combination produces a small focal volume where the intensity is high enough to drive the 2PF process. The NIR wavelengths required for 2PF are about twice as long as those used in conventional confocal microscopy and have the important benefit of reduced scattering and thus deeper penetration. To achieve the high peak power required for 2PF without the high average power that would damage the sample, mode-locked fs lasers are required (see Chapter 1, Section I.A.5 for details). A typical laser for 2PF microscopy produces 120 fs pulses every 12 ns (or a repetition rate of 80 MHz). Thus, the peak power is 100 kW while the average power is 1 W since the laser is only emitting radiation 0.001% of the time. Since 2PF occurs only in the focal plane, the sample is not excited in regions above or below it. Therefore, the confocal aperture is no longer required to get high resolution in the Z direction, which reduces losses in collection of the fluorescence signal. As a further benefit, damage to the sample is decreased due to the longer excitation wavelength and the localized excitation. In summary, the benefits of 2PF include deeper penetration and reduced photodamage, which are conducive for the imaging of live cells.

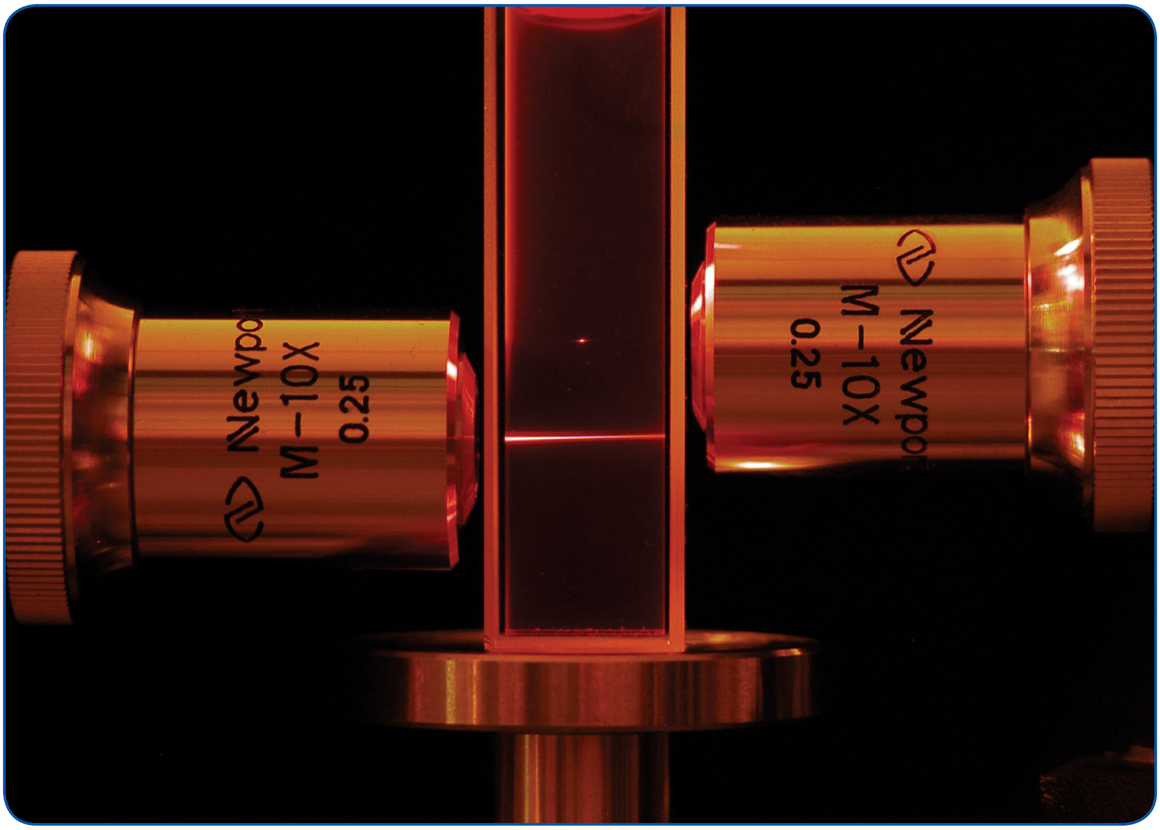


Figure 297. Comparison of one-photon fluorescence (lower beam) and 2PF (upper beam) in a fluorescent dye cell [351].

An experimental demonstration of the localized excitation possible for 2PF is shown in Figure 297. The dye cell contains a solution of an organic chromophore known to be an efficient emitter of 2PF. The objective on the left focuses 405 nm light from a *CW* laser diode while the objective on the right focuses 800 nm light from a mode-locked fs laser. When the solution is excited by a one-photon absorption process at 405 nm, fluorescence is generated everywhere along the propagation axis. For two-photon absorption at 800 nm, fluorescence is observed exclusively at the focal point of the objective [351].

A schematic of a microscope system for 2PF is shown in Figure 298. A laser is focused to a tight spot in the specimen plane and scanned in a raster over the sample. When the laser focus overlaps with fluorescent molecules in the sample, fluorescence is generated selectively in the tiny focal volume and detected by photodetectors. The signal is spatially-mapped to generate individual pixels of an image by a data acquisition computer. The principal differences between confocal and 2PF microscopes are the laser and the fluorescence detection path. In 2PF microscopy, all fluorescent photons collected by the objective constitute useful signal as the detector pinhole is not required.

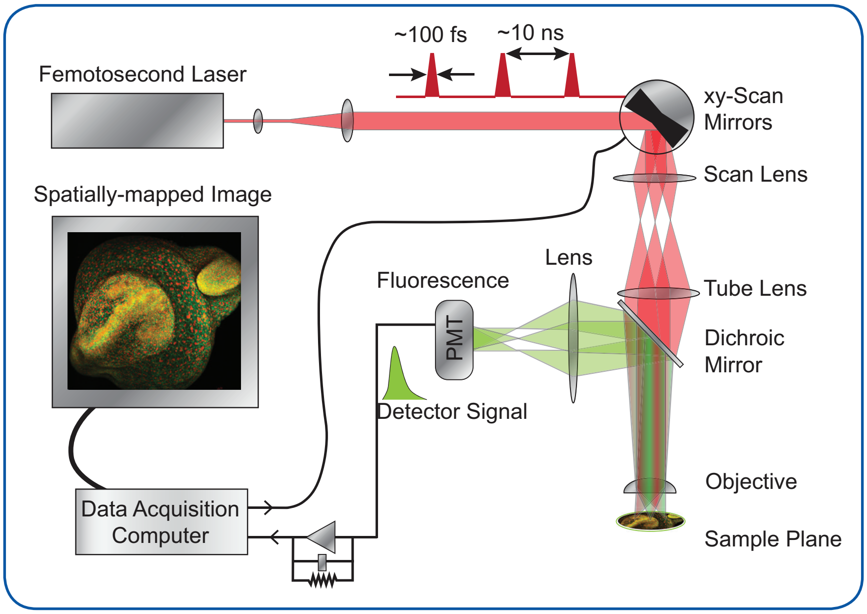


Figure 298. Schematic of a 2PF microscope [353].

Starting in the mid-1990's, the laser system of choice for 2PF was the mode-locked Ti:Sapphire laser. These lasers typically produce 100 fs pulses at an 80 MHz repetition rate with 1 W of average power and a wavelength tunable from 680 to 1050 nm. This tunability means they are ideally suited for generating 2PF for a large number of fluorescent dyes [355]. As 2PF microscopy grew in popularity over the next 15 years, several improvements were implemented. Multiple manufacturers offered microscopes optimized for 2PF, which included improved optics for longer wavelengths and sophisticated scanners, detectors, and software. The laser sources became more robust, compact, and easy to use. For instance, the pump laser was integrated together with the Ti:Sapphire laser in a single-box system and wavelength tuning was performed through an automated computer interface. A typical layout of a commercial 2PF microscope system is shown in Figure 299.

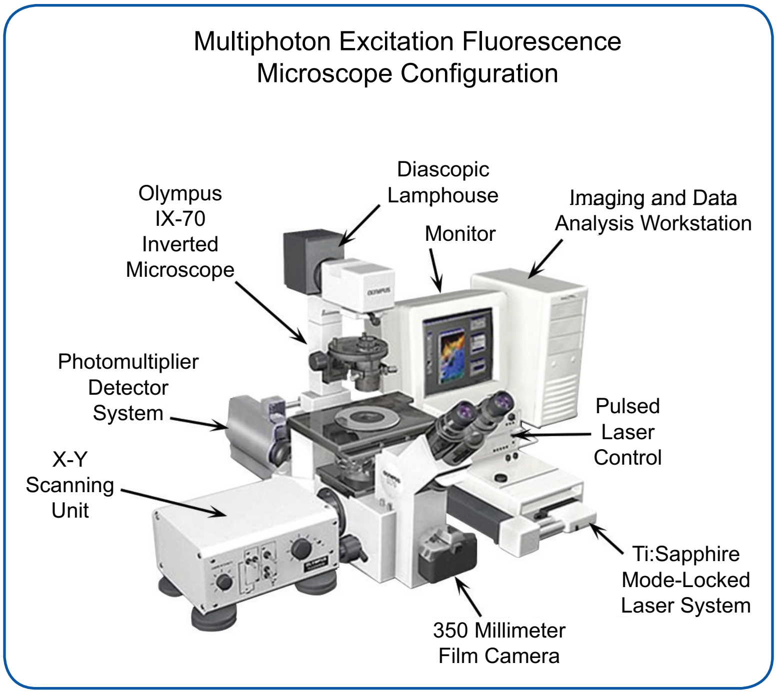


Figure 299. A commercial 2PF microscope system [356]. Image courtesy of Olympus Corporation.

One challenge in the development of 2PF microscopy was to deliver 100 fs pulses to the focal plane of the sample. The pulses produced by the laser pass through multiple optics, including sophisticated multi-element microscope objectives and acousto-optic modulators where the dispersion of the materials broadens the pulses. This effect produces 200-300 fs pulses at the focus and, since two-photon excitation scales inversely with the pulse duration, this greatly reduces the intensity of the fluorescence. Fortunately, the dispersion of the materials used in the beam path can be pre-compensated with a sequence of four prisms. Automated dispersion compensation was later added to single-box laser systems.

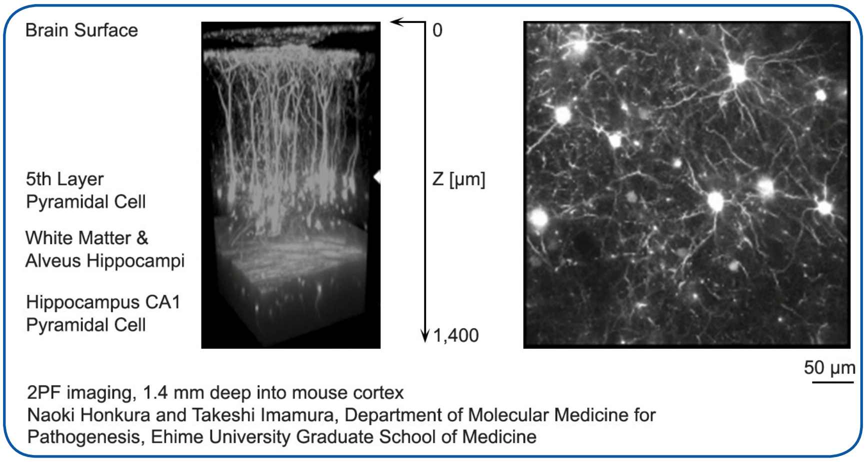


Figure 300. A 3D reconstructed image of a mouse cortex up to a depth of 1.4 mm using 2PF microscopy.

The penetration depth in 2PF microscopy is improved relative to one-photon excitation but is still limited to less than 1 mm in many biological samples due to scattering. Scattering decreases greatly at longer wavelengths and so there was a desire to move to longer wavelengths to improve the penetration depth [357]. A new generation of sources that could tune from 680 nm to 1300 nm was introduced in 2011. These sources were based on NIR pump lasers and tunable oscillators. They produced similar pulse durations and average powers as the Ti:Sapphire lasers and included automated tuning and dispersion compensation in a single box. Most importantly, they nearly doubled the tuning range of Ti:Sapphire lasers, including the longer wavelengths desired for deeper penetration [358]. As shown in Figure 300, these longer wavelengths allow imaging of a mouse hippocampus at a depth of 1.4 mm. In Figure 301, fs pulses at 980 nm were used to excite the dye enhanced green fluorescent protein (eGFP) used to label the cell membrane while pulses at 1041 nm excited the dye mCherry which labeled the cell nuclei. This demonstrates both the use of multiple wavelengths to excite different dyes and the monitoring of live samples over a period of eight hours without damage to the sample.

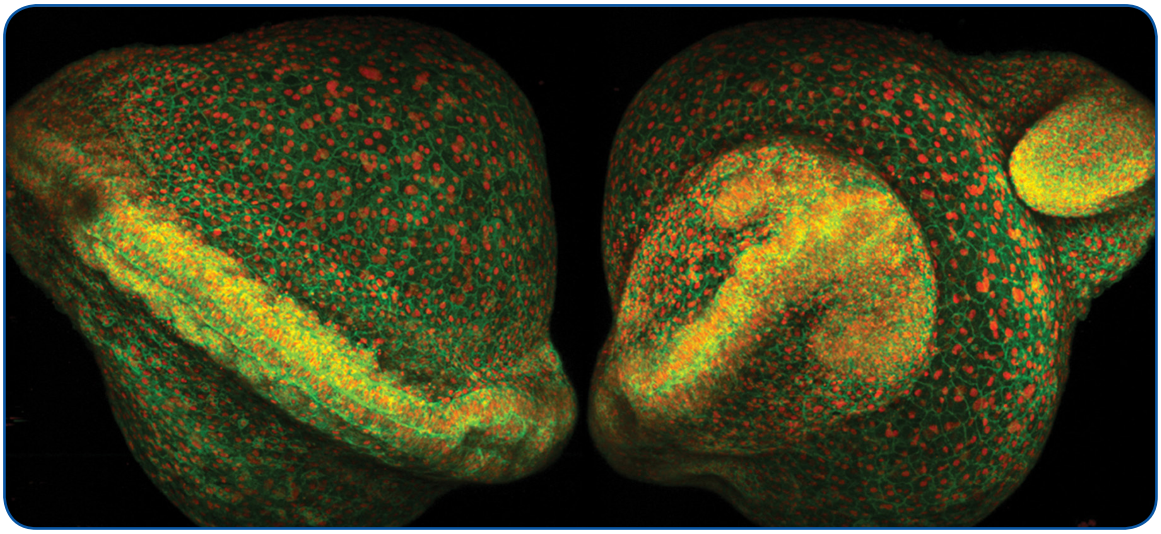


Figure 301. Embryo development monitored for eight hours in a live zebra fish embryo using 2PF microscopy.

In SHG microscopy, the fluorescent dye is not required as the second harmonic signal is generated from the sample itself [359]. When a material is identical to its mirror image, it is said to possess inversion symmetry. SHG arises only from materials that lack inversion symmetry. For example, thin membranes can be probed with SHG microscopy. Several key endogenous protein structures, notably collagen, also give rise to intense SHG. SHG microscopy on a laser-scanning system has proven to be a powerful and unique tool for high-resolution, high-contrast, 3D studies of live cell and tissue architecture. Unlike 2PF, SHG suffers no inherent photobleaching or toxicity and does not require exogenous labels. SHG microscopy is just one of several nonlinear microscopies that do not require labeling together with others such as THG and SFG microscopy.

Figure 302. Spectra-Physics MaiTai® DeepSee™ laser produces broadly tunable fs pulses in the NIR.

For nearly 20 years, the MKS Spectra-Physics MaiTai® laser (see Figure 302) has been the workhorse of bio-imaging labs around the world. It is based on traditional Ti:Sapphire technology and offers a 690 nm - 1040 nm tuning range with up to 2 W of output power and a 100 fs pulse duration. Following its initial release in 1999, it was the first laser to provide automatic wavelength tuning. Then, in 2007, it was the first laser to integrate automatic dispersion control through its DeepSee™ laser technology. In addition to its application in multiphoton microscopy, the MaiTai laser is also used in the areas of time-resolved photoluminescence, nonlinear spectroscopy, optical computed tomography, surface SHG, terahertz imaging, semiconductor metrology, materials processing, and laser amplifier seeding.

The InSight® X3™ (see Figure 303) is the third generation of MKS Spectra-Physics' industry-leading laser platform that is specifically designed for advanced multiphoton microscopy applications. The InSight X3 laser features a broad 680 nm to 1300 nm continuous, gap-free tuning range from a single source, nearly doubling the tuning range of legacy Ti:Sapphire ultrafast lasers. The system delivers high average and peak power levels across the tuning range, including the critical NIR wavelengths above 900 nm, which allow for the deepest penetration in vivo imaging. With MKS Spectra-Physics' integrated patented DeepSee™ laser technology, the industry standard dispersion pre-compensator, fs pulses are optimally delivered through a microscope to the sample for maximum fluorescence and penetration depth.

Figure 303. Spectra-Physics InSight® X3™ Laser produces fs pulses tunable from 680 nm to 1300 nm.

D. Three-Photon Fluorescence Microscopy

2PF microscopy has enabled impressive progress in the fields of neuroscience, embryology, and oncology. Researchers can visualize tissue morphology at a cellular level within scattering tissue even in vivo. The imaging depth of 2PF is limited by scattering to about 1 mm and there is a desire to image deeper into samples, particularly in neuroscience such as in studies of subcortical structures in the mouse brain.

Three-photon fluorescence (3PF) microscopy can provide a significant advantage over 2PF microscopy in strongly scattering samples such as the mouse brain. The fluorescence from three-photon excitation falls off as $1/z^4$, where z is the distance away from the focal plane, while the fluorescence from two-photon excitation falls off as $1/z^2$ [360]. Thus, three-photon excitation reduces the background from regions away from the focal plane and improves the signal-to-background ratio by orders of magnitude. This reduction in background is illustrated in Figure 304 where 2PF and 3PF are compared.

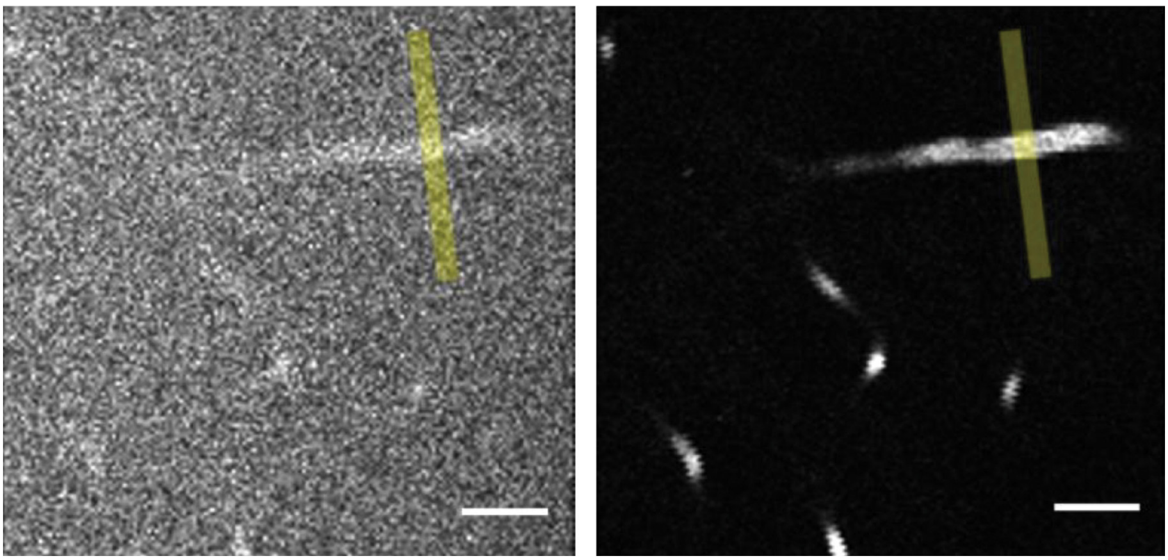


Figure 304. A comparison of images from 2PF (left, at 920 nm) and 3PF (right, at 1300 nm) microscopy of fluorescein-labeled blood vessels 650 μm deep in a mouse cerebellum. The two images have comparable signal strength and were displayed with the same contrast setting. Scale bar, 50 μm. Images reprinted with permission from [361], SPIE Publications.

Moving to longer wavelength excitation is important to reduce scatter. Figure 305 shows how the calculated penetration depth due to Mie scattering in brain tissue increases with wavelength. Mie scattering dominates when particles are larger than the incident wavelength. The same figure shows the penetration depth due to water absorption with strong absorption bands at 1.4 µm and 1.9 µm. The combination of the scattering and absorption gives an effective penetration depth. The optimal wavelengths to minimize both the scatter and the absorption are thus at 1.3 µm and 1.7 µm.

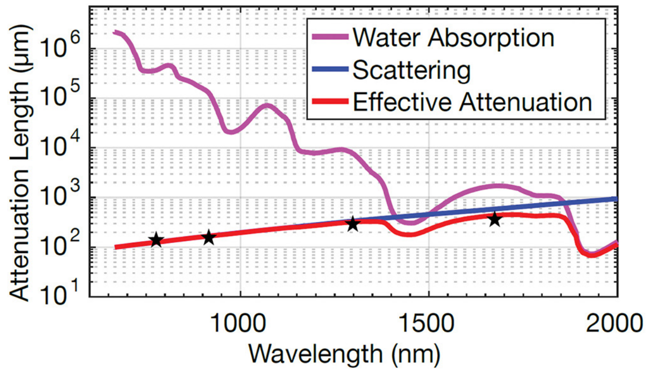


Figure 305. Effective penetration depth due to scattering and absorption in imaging where the optimal excitation wavelengths of 1.3 µm and 1.7 µm are shown (reprinted with permission from [362], The Optical Society).

Currently, there are not many fluorescent dyes where the two-photon excitation wavelength overlaps with these optimal wavelengths. However, there are dyes with a spectral overlap of their three-photon absorption in these low loss windows. Some of these dyes are the same as those used for 2PF microscopy at 900 nm, e.g., green fluorescent protein (GFP) based probes, and 1100 nm, e.g. red fluorescent protein (RFP) based probes. Over the last decade, commercial manufacturers have improved the transmission of their microscopes for longer wavelength excitation.

One remaining challenge is that the cross-section (or the probability of excitation) for 3PF is several orders of magnitude smaller than for 2PF, requiring lasers with much higher peak powers for efficient excitation. However, to prevent damage at the surface of samples, the average power must be limited. Thus, to achieve the required peak powers, laser systems that produce 100 fs pulses with higher energy but with lower repetition rates (to maintain modest average power) are required. A typical laser for 3PF microscopy produces 100 fs pulses every 1 µs (a repetition rate of 1 MHz). Therefore, 1 W of average power gives a peak power of 10 MW with the laser on only 0.00001% of the time. Laser systems that produce ultrashort pulses at this high energy are based on ultrafast amplifier systems that operate at a wavelength of 1 µm. To shift the wavelength to the target wavelengths of 1.3 µm and 1.7 µm, additional nonlinear conversion schemes are required including SHG and OPG.

In THG microscopy, the fluorescent dye is omitted and the third harmonic signal is generated from the sample itself. In a homogenous sample, the THG signal from above the focus cancels the signal generated below the focus due to phase matching. Thus, a third harmonic signal is only generated when the focus is close to a gradient in the refractive index. As a result, THG microscopy is particularly well-suited for 3D label-free imaging of transparent specimens where the membrane or other interface is of interest [363].

When a combination of techniques is utilized to interrogate a single sample, it is referred to as multi-modal imaging. Using multiple techniques allows investigators to look at different features in the same sample as illustrated in Figure 306, which shows a 3D reconstruction of a mouse brain cerebellum. Fs pulses at 1.3 µm are used to excite the tissue, which generates 3PF in fluorescein-labeled blood vessels while the same wavelength also produces a THG signal in different portions of the cerebellum, i.e., myelinate fibers.

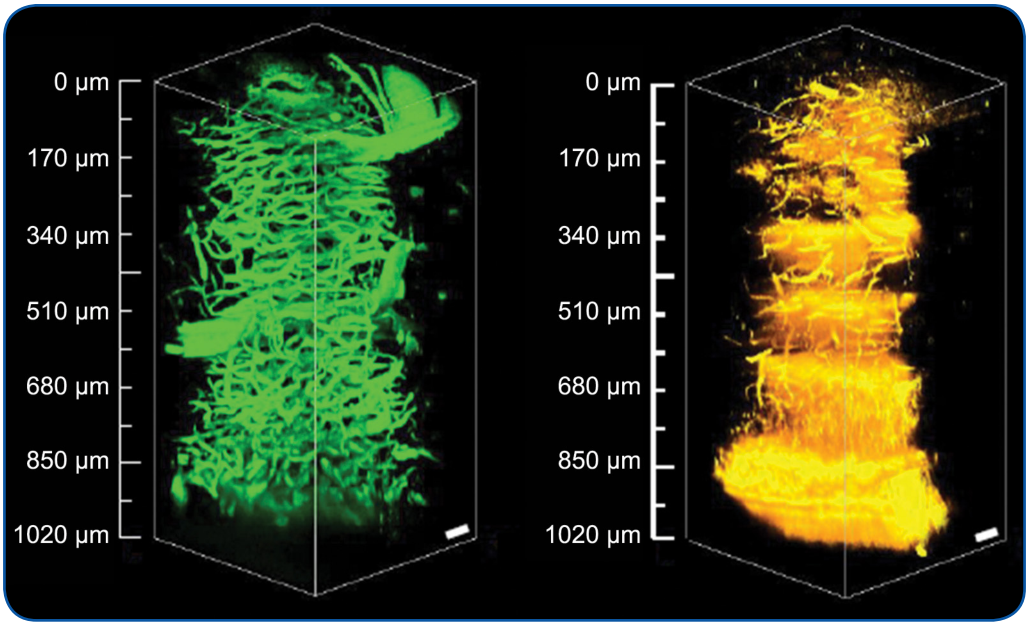


Figure 306. 3D images from a mouse brain cerebellum extending 1 mm deep into the tissue acquired via 3PF (left) and THG (right) microscopy at 1.3 µm. Scale bar, 50 µm. Images reprinted with permission from [361], SPIE Publications.

The MKS Spectra-Physics Spirit® fs laser (see Figure 307) is an all-in-one amplifier which delivers high average powers (up to 100 W), at a fixed wavelength of 1040 nm, and with a pulse duration of 400 fs. This pulse duration is longer than the Spectra-Physics InSight laser, but the reduced pulse repetition rate (0.1 to 1 MHz) allows for higher pulse energies (up to 40 µJ). These specifications open the door for new applications such as a direct source for certain biological applications or as a pump for other lasers, like the Spectra-Physics Spirit-NOPA-IR. Other applications for the Spirit laser include materials processing, medical device fabrication, microsurgery, time-resolved fs spectroscopy, strengthened glass and sapphire cutting.

The MKS Spectra-Physics Spirit-NOPA-VISIR (see Figure 308) is a family of automated non-collinear optical parametric amplifiers (NOPA) specifically built and optimized for the Spectra-Physics Spirit® and Spirit® One™ pump lasers. This combination creates a powerful, user-friendly, tunable, high repetition rate source of ultrashort pulses for in vivo, deep-tissue imaging in neuroscience, spectroscopy, and other ultrafast scientific applications. The Spirit-NOPA-VISIR laser provides an ideal platform for 3PF microscopy as it produces microjoule pulses at both 1.3 µm and 1.7 µm.

Figure 307. Spectra-Physics Spirit Laser produces high energy (µJ) fs pulses at a wavelength of 1030 or 1040 nm.

Figure 308. Spectra-Physics Spirit-NOPA-VISIR Noncollinear OPA Laser produces high energy (µJ) pulses over a broad range of wavelengths in the NIR including 1.3 µm and 1.7 µm.

E. CARS and SRS

Investigation of the molecular properties of complex systems such as living cells is a topic of wide interest. Researchers need spatial resolution to focus on the object of interest and chemical selectivity to probe specific species. The combination of optical spectroscopy and microscopy provides a direct and non-invasive approach to visualizing nucleic acids, proteins, and other molecules at work in living cells [364].

Raman microscopy is a technique that enables label-free chemical imaging. It is based on the Raman scattering effect of molecules that was discovered by C.V. Raman in the early 1930s [365]. When light with a particular wavelength or photon energy is incident on a molecule, it can be inelastically scattered. This means that part of the energy is absorbed by the molecule and the scattered photon has a lower energy than the incident photon (see Figure 309). The fraction of the energy absorbed depends on the vibrational frequencies of the molecule. All molecules have specific Raman frequencies, typically given in wavenumbers, which span from 100 cm^{-1} to 3500 cm^{-1}. A molecule's Raman spectrum is highly dependent on its chemical structure but mostly unaffected by the local environment. Therefore, Raman spectroscopy is not only specific but also robust to environmental variability [366].

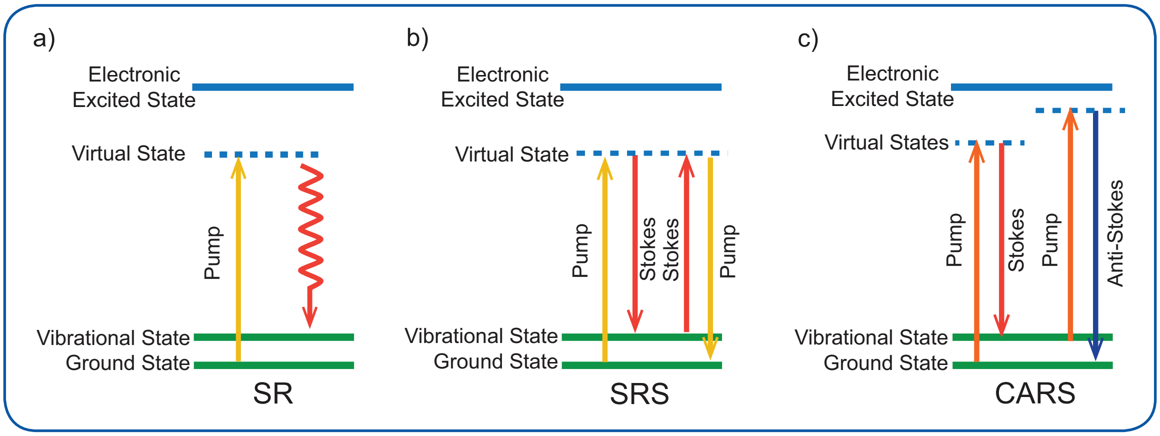


Figure 309. Energy diagrams of Raman interactions: Spontaneous Raman (SR), SRS, and CARS.

Three types of Raman processes are shown in Figure 309. In SR scattering, the pump field is inelastically scattered off molecular vibrations of the sample, generating a new red-shifted field component, called the Stokes wave. SR scattering is typically very weak and, as a result, imaging based on the SR effect is extremely slow due to long acquisition times necessary to provide images with reasonable SNR. In SRS, both the pump and Stokes frequencies are incident on the sample. If the frequency difference matches a molecular vibration of the sample, stimulated excitation of the vibrational transition occurs. SRS-based microscopy is a relatively recent development having been demonstrated in 2008 [367] and gaining commercial viability only in recent years. CARS is a four-wave mixing process that generates a new blue-shifted field, called the anti-Stokes wave. When the energy difference matches a molecular vibration of the sample, the scattering process is coherently amplified. The CARS process itself was first studied by researchers at the Ford Motor company in 1965 [368]. Multiphoton vibrational microscopy based on CARS has been investigated since the early 2000's.

In CARS microscopy, the pump and Stokes beams are tightly focused into the sample and a CARS image is generated by scanning either the sample or the laser beams. Besides possessing the same 3D sectioning capability afforded to 2PF and 3PF microscopy, CARS microscopy offers several advantages. CARS microscopy permits non-destructive molecular imaging without any labeling. This advantage is important for imaging small molecules such as lipids where labeling may significantly affect their molecular properties. The coherent amplification from the molecular oscillators results in a highly directional output, which greatly facilitates signal collection. A comparison between images of a polymeric structure taken by CARS and SR microscopy is shown in Figure 310. The CARS image possesses higher fidelity despite an acquisition time (under comparable laser power) two orders-of-magnitude shorter than for the SR image. Finally, CARS signals appear at a wavelength that is shorter than the excitation wavelengths; therefore, spectrally-separating it from the one-photon fluorescence background is straightforward.

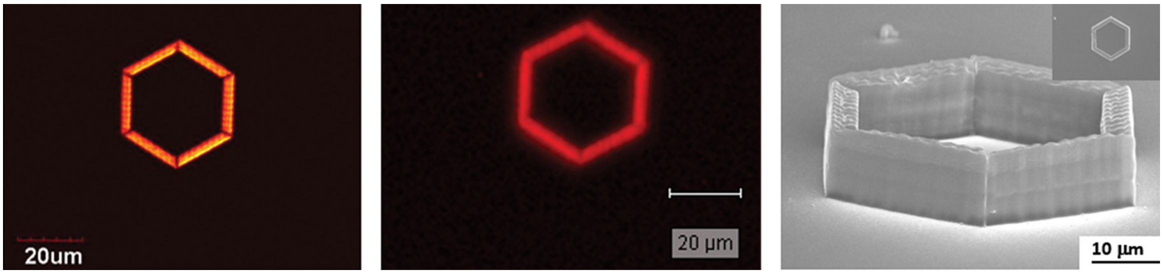


Figure 310. Imaging of a hexagonal polymeric structure by CARS microscopy (left) and SR microscopy (middle) with similar average laser powers. Acquisition of the CARS image was performed in one second while the SR image took five hours. Scanning electron micrograph of the hexagonal polymeric structure (right).

CARS microscopy requires a high peak power source similar to those required by 2PF microscopy. However, a CARS measurement, unlike 2PF, is not background-free. The non-resonant background limits the image contrast and spectral selectivity and much work has concentrated on methods to overcome this background [364]. SRS imaging is largely free from the non-resonant background inherent in CARS. Nonetheless, there are still nonlinear processes such as two-photon absorption, excited state absorption, and thermal lensing that can impact the SRS signal [369]. Not only is the SNR improved in SRS, the Raman spectral fidelity is also preserved. Furthermore, because the SRS signal has a linear dependence on concentration, it has the potential to be a powerful method for label-free quantitative determination of chemical concentration of individual species in a multi-component system [370]. These qualities make SRS microscopy a straightforward method for image interpretation and quantification.

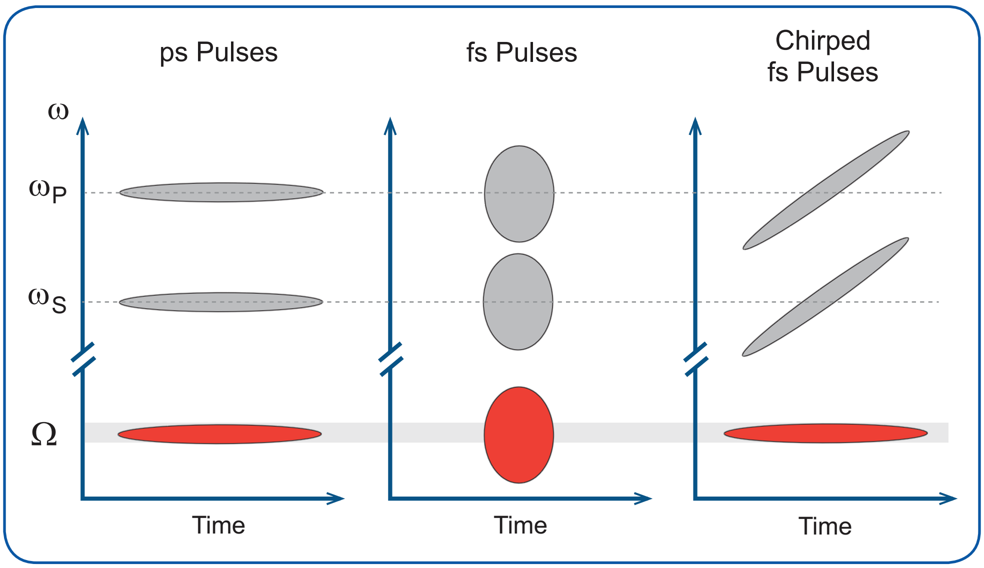


Figure 311. The impact of using different laser pulses to excite Raman transitions for CARS and SRS microscopy. The difference in laser frequencies ($\Omega = \omega_P - \omega_S$) shows efficient overlap with the width of the Raman transition (grey region) for ps pulses but not for fs pulses. Alternatively, the use of chirped fs pulses allows for good spectral overlap using ultrafast pulses [371].

While 100 fs pulses provide the high peak power needed for a strong signal from CARS or SRS, the corresponding bandwidth of these pulses is about 10 nm, which is much broader than the width of most Raman lines. This is illustrated in Figure 311. Accordingly, a large portion of the pulse energy only contributes to the generation of the non-resonant background. The line width of the Raman transition is better matched by ps pulses (see Figure 311) and much work has been done with ps lasers or OPOs [364].

A clever solution called spectral focusing allows the full bandwidth of a fs pulse to be used efficiently for these nonlinear microscopy techniques [369]. In this method, both the pump (ω_P) and Stokes (ω_S) pulses are linearly chirped as shown in Figure 311. Optical pulses where the frequency increases or decreases within the temporal pulse are referred to as chirped, in analogy with sound pulses. Therefore, in the plots shown in Figure 311, a horizontal line represents a pulse that is not chirped and the vertical width of the line represents the bandwidth of the pulse. Conversely, a tilted line indicates a linearly chirped pulse. When both the pump and the Stokes pulses are linearly chirped by the same amount, the difference between the two frequencies ($\Omega = \omega_P - \omega_S$) can be quite narrow. Thus, two broadband fs pulses can still be used to efficiently excite a narrowband Raman line. Tuning of the frequency difference between the pulses is also possible using this technique [371]. This is accomplished not by tuning the central wavelength of the laser but rather by changing the relative timing between the two chirped pulses. Spectral focusing allows fs lasers to be used for all nonlinear microscopy techniques (see Table 19) including CARS and SRS.

In SFG microscopy, the fluorescent dye is omitted and the sum frequency signal is generated from the sample itself, similar to both SHG and THG microscopy. However, in contrast to these techniques where only a single wavelength is required, two different wavelengths are required for SFG. Consequently, this technique has similar requirements to CARS in that the pulses from the two lasers must arrive at the same location in the sample at the same time. Therefore, synchronized lasers are needed for SFG microscopy.

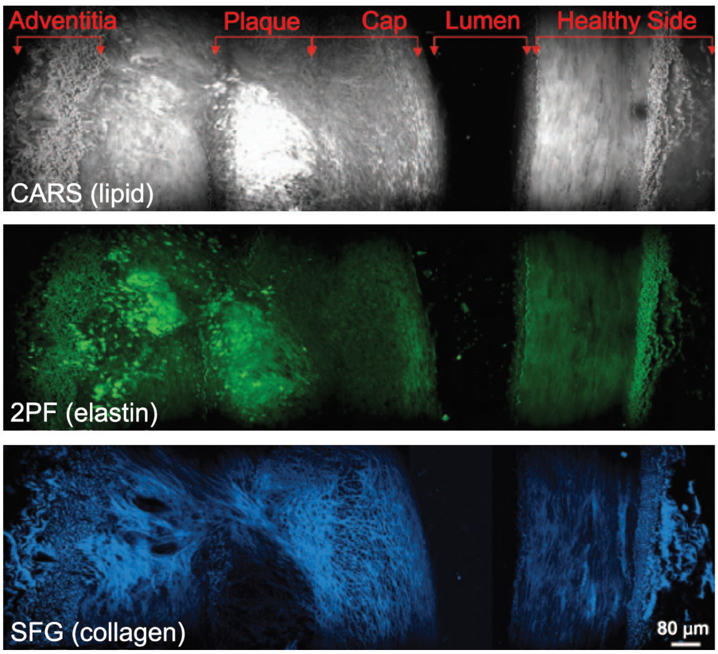


Figure 312. Cross-sectional view of an atherosclerotic plaque demonstrating multi-modal imaging [372].

An interesting application of multi-modal imaging is shown in Figure 312 where atherosclerotic plaque in an iliac artery was studied in [372] using three different modalities: CARS, 2PF, and SFG. All three techniques work without labels and the sample was studied in vivo. The plaque shows up in the CARS image due to the strong signal from lipid bodies that are enriched in C-H bonds. The 2PF image (endogenous fluorophores generate two-photon excitation fluorescence or TPEF) also shows the lipid core while the SFG image reveals a lack of collagen fibrils in the same region.

Figure 313. Spectra-Physics InSight® X3™ Laser in combination with the TRU enables spectral focusing and hyperspectral imaging.

The MKS Spectra-Physics InSight® X3™ (Figure 303) laser, when equipped with the fixed 1045 nm dual beam option, can fully support the diverse needs of multimodal imaging. The two synchronized output beams, one tunable and one at the fixed wavelength, enable simultaneous imaging of various fluorescent proteins (for example GFP and mCherry) and genetically encoded calcium indicators (GCaMP6 and jRGECO1a), SHG/THG imaging, and advanced imaging techniques such as CARS and SRS. A time-recombination unit (TRU) takes the two outputs from the InSight laser (the 1045 nm beam plus the tunable 680-1300 nm beam) and recombines them in a single beam, insuring their spatial and temporal overlap (see Figure 313). This system offers several desirable features, including easy switching between beams, fs or ps modes, and integrated attenuators for fine energy tuning. The SF-TRU is a time-recombination unit that also enables spectral focusing and, most importantly, it enables hyperspectral imaging or spatial mapping of an object with good spectral resolution. Spectral focusing allows high spectral resolution, down to 5 cm^{-1}, which is comparable with the linewidth of Raman active molecules in the fingerprint region. Hyperspectral imaging is conducted by simply changing the delay between the pump and the Stokes pulses [373].

F. Optogenetics

Optogenetics combines genetic and optical techniques to selectively and specifically modulate signaling pathways in intact living organisms using light [374]. This rapidly expanding field has been driven by developments in ultrafast laser and microscope technologies. Using a microscope, laser light can be focused precisely on specific neurons containing light sensitive ion channels, or opsins, and causes those neurons to activate and signal each other. Fluorescent proteins that are sensitive to calcium or voltage can be used to visualize this activation non-invasively. As neuroscientists develop new genetic light sensitive tools, high peak power fs lasers for two-photon activation have developed in parallel. These lasers have the high peak power necessary for activating more neurons simultaneously and with high precision.

A single human brain contains more than 100 billion neurons. These neurons form a complex network with more than 100 trillion connections, or synapses, that send electrical impulses to control muscles and organs, and perform cognitive functions. An overarching objective in neuroscience is to develop an understanding of not only the physical structure, but more importantly, the functionality of the brain, especially the means of communication among neurons. Understanding these connections can reveal the fundamental science behind neural networks, which could lead to cures for diseases such as Parkinson's and Alzheimer's.

2PF microscopy allows for visualizing neural networks in live tissue, in three dimensions, and with resolution down to a single cell. Optogenetics can be combined with 2PF microscopy (denoted as two-photon optogenetics) to yield tools that enable the activation and visualization of brain activity with single cell precision. However, the fluorescent proteins (such as GFP and many different colored variants) that have driven a renaissance in light microscopy are not amenable for two-photon optogenetics. Since GFP is a protein, it can be introduced into the genetic material of an organism and its location identifies specific cells or cell components in a living organism. Unfortunately, fluorescent proteins are not sensitive to dynamic changes in the cell and cannot directly influence the cell's or organism's behavior. Two key biological discoveries (discussed below) were required to enable the manipulation and visualization central to optogenetics.

The discovery of Channelrhodopsin-2 (ChR2) spurred the first experiments of optogenetics and opened the possibilities for controlling neuronal activity using light [374]. ChR2 is an opsin, a class of proteins comprised of different ion channels that can be opened and closed using light, i.e., photoactivation or photostimulation, and cause different signaling effects in the neurons that contain it. When an organism has been genetically engineered so its neurons contain opsins, neuronal activity can be triggered by illuminating the cell using bright light. Through photoactivation, the ChR2 channels pump cations into the cell, causing an increase of the voltage across the cell membrane (an action potential) and firing of the cell. This causes the release of neurotransmitters and signaling of neighboring neurons. This process is repeatable and can be done many times without causing damage to the cells. In order to activate a single cell, an ultrafast laser and a 2PF microscope are needed. By focusing the beam to a small spot around the neuron's membrane, sufficient power can be delivered to reach the threshold where the neuron will activate.

In addition to the use of opsins for activating cells, there are optogenetic tools that can be used to report the signal of neural activity after activation. One class of fluorescent indicators that can be used to read out this activation is a genetically encoded calcium indicator. These indicators become brighter when there are higher concentrations of free calcium ions in the cytoplasm of a neuron, which happens following an action potential in the neuron. When a neuron fires, the intracellular calcium concentration increases for around 100 – 500 ms, making it easy to detect a single firing of a neuron.

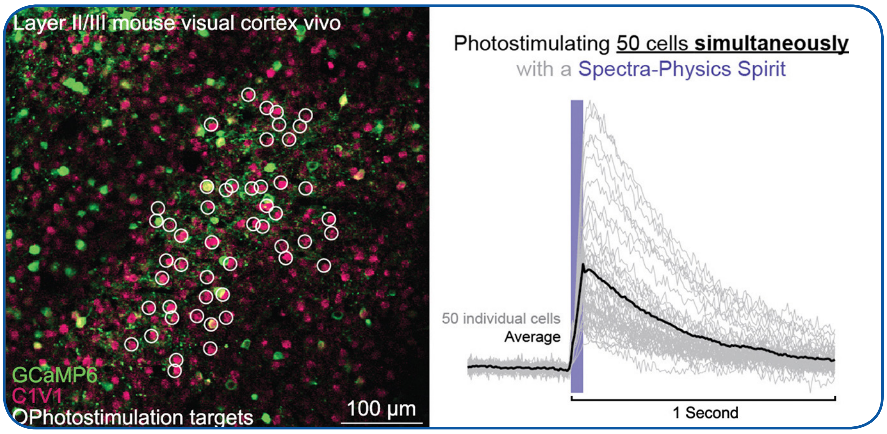


Figure 314. Simultaneous photostimulation of up to 50 neurons in vivo courtesy of Lloyd Russel, Adam Packer & Michael Hausser, UCL [374].

These discoveries enable both the activation and measurement of neurons in vivo. This requires having a single organism containing both an opsin and calcium indicator genetically encoded in its neurons. To perform these functions in concert, the photoactivation and readout need to use distinct wavelengths of light that do not spectrally overlap. For instance, one probe can be activated using light between 900 and 1000 nm, while the other probe records the response with light between 1000 and 1100 nm. Most commonly, a red-shifted opsin such as C1V1 is used along with a green calcium indicator such as GCaMP6 for visualization. High peak power lasers with lower repetition rates are ideal for use in optogenetics when using a spatial light modulator (SLM). SLM's can focus laser light onto many different spots, allowing simultaneous activation of multiple neurons. Figure 314 shows the simultaneous photostimulation and imaging of 50 neurons in vivo in a mouse visual cortex using these dyes.

Figure 315. Spectra-Physics femtoTrain™ Laser produces high energy (100s of nJ) fs pulses for use in optogenetic applications.

The MKS Spectra-Physics femtoTrain™ laser (see Figure 315) is a fixed wavelength (1040 nm) laser system delivering up to 5 W in a 10 MHz pulse train. The femtoTrain 1040-5 laser offers pulse widths below 200 fs, at an average power of 5 W to deliver 2.2 MW of peak power. It is well-suited for use in optogenetics as well as other applications such as multiphoton imaging, tissue dissection, and micro-surgery.

III. Ophthalmic Surgery

Ophthalmology is the branch of medicine that deals with the anatomy, physiology and diseases of the eye. The laser is particularly suited to ophthalmic surgery since it can provide a non-contact method of interacting with the cornea or even the interior of the eye, including the lens and the retina. Laser photocoagulation surgery is used to treat several eye diseases and has become widely used in recent decades. During this procedure, a laser is used to finely cauterize ocular blood vessels to attempt to bring about various therapeutic benefits [375]. Treatments for diabetic retinopathy and macular degeneration have also been accomplished with several kinds of lasers including ion lasers, dye lasers, and laser diodes. Recently, two forms of ophthalmic surgery have attracted great attention: laser-assisted in situ keratomileusis (LASIK), which is used to modify the cornea and correct vision, and cataract surgery, which can facilitate replacement of the lens when it becomes cloudy. Both procedures are accomplished with the help of ultrafast pulsed laser systems, as described below.

A. LASIK

LASIK is the most commonly-performed interventional procedure to correct myopia (short-sightedness), hyperopia (far-sightedness), and astigmatism [376]. It relies on optimizing the refractive properties of the cornea by removing corneal tissue through photoablation using UV ns pulses from an excimer laser. LASIK involves the creation of a thin corneal flap to give the excimer laser pulses access to deeper corneal layers. The flap is closed back to its original position after photoablation, as the preservation of the surface epithelial layer accelerates recovery. Traditionally the flap has been created mechanically using specialized mechanical cutting tools called microkeratomes. However, in recent years, these microkeratomes have been widely replaced by clinically-approved “blade-less” devices, which utilize NIR fs laser pulses for the creation of the flap.

The cornea is transparent at NIR wavelengths. However, in the focus of the beam, the peak power density can be high enough to cause laser-induced optical breakdown and photodisruption of the tissue. By using fs lasers together with a high NA objective, the damage can be spatially confined to the focal volume. The highly localized deposition of energy generates a microplasma, which can be followed by deleterious side effects such as the emission of a shock wave and the generation of cavitation bubbles [377]. However, in contrast to ns or even ps pulses, fs laser photodisruption requires much lower pulse energies, which minimizes the range of the photodisruptive shock wave and the size of the cavitation bubbles [378]. As a result, the collateral thermal damage zone is extremely limited since the thermal diffusion is in the sub-micron range [379].

The corneal flap for LASIK is created by scanning the focus of a fs laser beam at the desired depth to form a dissection plane while leaving a hinge for the flap to remain connected with the cornea. Compared to a microkeratome, the fs laser allows for flexibility with respect to the thickness of the flap. Furthermore, it enables much greater precision, which results in faster recovery times, improved visual results, and greater safety [378, 379]. The procedure was introduced first by Intralase, which has since been acquired by Abbott Medical Optics. Today, there are several suppliers for fs laser-based instruments, that have been clinically approved for cutting corneal flaps [380]. The lasers deployed in these devices are DPSS or fiber lasers delivering pulse widths of a few hundred fs at repetition rates in the high kHz to low MHz range. Pulse energies can reach a few microjoules, which typically requires pulse amplification [381]. This method for amplifying fs pulses called chirped pulse amplification was recognized with the 2018 Nobel Prize in Physics [342].

Flap cutting for LASIK was the first fs laser-based, clinically-approved vision correction procedure, but others have followed [382-384]. Currently, there is significant focus on bringing devices to market that utilize fs pulses for high-precision interventional processes involved in the transplant of intraocular lenses (IOLs) during cataract surgery, as will be described below.

B. Cataract Surgery

A cataract is a clouding of the lens inside the eye causing vision loss that cannot be corrected with glasses, contact lenses, or corneal refractive surgery like LASIK [385]. Most cataracts are associated with the aging process. Metabolic changes of the crystalline lens fibers over time lead to the development of an opacification in the lens and loss of transparency, causing impairment of vision [386]. According to the National Eye Institute (NEI), 68.3% of Americans 80 and older had cataracts in 2010. The prevalence of cataracts in the United States is expected to grow significantly in the years ahead due, in part, to the aging of the population. In 2010, roughly 24.4 million Americans had cataracts and that number is projected to grow to 50.2 million by the year 2050, according to NEI.

Modern cataract surgery is one of the safest and most effective surgical procedures performed today. More than 3 million cataract surgeries are performed in the United States every year, with the majority of these procedures producing excellent visual outcomes. In cataract surgery, the lens inside the eye that has become cloudy is removed and replaced with an artificial lens (or IOL) to restore clear vision. Today, the procedure is typically performed on an outpatient basis.

There are two main types of surgical procedures in use today. The first procedure is phacoemulsification (or phaco) and the second is extracapsular cataract extraction (ECCE). Phaco is the most commonly-performed cataract procedure in the developed world while ECCE, which is less expensive, is more frequently performed in developing countries. Phaco involves the use of a high-frequency (40 kHz) ultrasound device that breaks up the cloudy lens into small pieces, which are then gently removed from the eye with suction. This can be performed with smaller incisions than previous surgical techniques for cataract removal, promoting faster healing and reducing the risk of cataract surgery complications. After all remnants of the cloudy lens have been removed from the eye, the cataract surgeon inserts a clear IOL, positioning it securely behind the iris and pupil, in the same location the natural lens occupied. The surgeon then completes the cataract removal and IOL implantation procedure by closing the incision in the eye.

Recently, a number of fs lasers, similar to the lasers used to create the corneal flap in LASIK, has been approved by the FDA for use in cataract surgery performed in the United States. The main advantages are standardized corneal incisions, perfectly centered, round openings of the lens capsule (capsulorhexis), and lens nucleus fragmentation even in eyes with hard cataracts. Laser cataract surgery is relatively new and significantly increases cataract surgery cost, in part, due to the cost of the laser. In the United States, private insurance and Medicare cover most if not all of the associated costs. There are currently at least five companies with fs laser-based cataract surgery platforms, including Alcon LenSX, LENSAR, Abbott Medical Optics (OptiMedica), Bausch & Lomb Technolas, and Ziemer. Fs lasers have gained approval for multiple steps in cataract surgery, which reduces the need for surgical blades and other hand-held tools. These surgical steps include creating corneal incisions to allow the surgeon access to the lens, performing so-called limbal relaxing incisions to reduce or eliminate pre-existing astigmatism, removing the anterior capsule of the lens, and fragmenting the cataract so lower phaco energy is required to break it up.

The MKS Spectra-Physics Spirit® fs laser, as shown in Figure 307, is a one-box amplifier which delivers 400 fs pulses with 4 W of average power at 1040 nm. With up to 40 µJ of energy, this laser has been extensively deployed for LASIK flap cutting and has the capability of being implemented for approved cataract surgical steps.

C. Future Directions

In LASIK, NIR fs lasers are used to cut the flap in the cornea while an excimer laser is used to ablate the exposed portion. It would be desirable if both processes could be achieved with a single laser. This has recently been achieved with the introduction of the Refractive Lenticule Extraction Small Incision Lenticule Extraction (ReLEx SMILE) procedure developed by Carl Zeiss Meditec. The procedure uses a single fs laser system to cut a small lens-shaped piece (known as lenticule) within the corneal stroma and to create a small incision along the periphery of this lenticule to remove it through the incision. In a similar vein to LASIK, cataract surgery requires both an ultrafast laser to make the incision in the cornea and an ultrasound source to fragment the lens material that will be removed. Initial results have been announced that use a fs laser for both the incision and the photofragmentation of the lens [387].

Presbyopia is another age-related condition of the eye. It is the normal loss of near-focusing ability that occurs with age with the majority noticing the effects of presbyopia sometime after the age of 40. This typically manifests in difficulty in reading small print such as text messages on a smartphone. According to a global report on presbyopia issued in 2012 by Market Scope, worldwide, an estimated 1.3 billion people had presbyopia in 2011. This number is expected to increase to 2.1 billion by 2020. Currently, the most common treatments [388] are the use of reading glasses or contact lenses. A special type of contact lens correction for presbyopia is monovision, in which one eye wears a distance prescription, and the other wears a prescription for near vision. The brain learns to favor one eye or the other for different tasks. There also are surgical treatments such as the implantation of a corneal inlay. A corneal inlay is typically implanted in the cornea of the person's non-dominant eye. Like monovision, this approach helps increase the depth of focus of the treated eye and reduces the need for reading glasses without significantly affecting the quality of distance vision. The sole corneal inlay that is approved by the FDA is the KAMRA inlay [389]. Since virtually every person will get presbyopia given a long enough life span, several medical companies are working on developing a laser-based treatment for this ever-expanding market.

There is an intriguing application for ultrafast lasers currently under investigation with the potential to impact both cataract surgery and vision correction. In both LASIK and cataract surgery, ultrafast lasers cut material. This process of micro-machining the cornea requires amplified fs laser systems with pulse energies of several microjoules. Lower energy fs lasers, instead of cutting, only modify the index of refraction in transparent materials. Through the process of multiphoton absorption, the refractive index modification is localized in three dimensions due to the nonlinear nature of the process. Initial work has focused on using this approach to modify the refractive index in either contact lenses or the IOLs that are implanted during cataract surgery [390]. Researchers have been able to modify the index and write Fresnel lens structures with an optical power of several diopters into the materials used for contacts. Clerio Vision is pursuing this process with unamplified laser oscillators that produce fs pulses with nJ energies at 800 nm, while Perfect Lens is using amplified fs pulses with µJ energies at 515 nm.

More exciting, and more challenging, is to write the index change directly into the cornea of the patient. It has been reported that only 2% of those needing refractive correction of some sort have had LASIK. Thus, a less invasive procedure might find wide adoption. Fs pulses at a wavelength of 405 nm with nanojoules of energy are required for this approach. The wavelength is chosen to optimize the index change in the cornea while protecting the retina. Currently, these pulses are produced by generating the second harmonic of fs oscillators operating at 800 nm. Animal studies have been underway for several years with encouraging results. Successful implementation will depend on many factors, including longevity of the correction, successful trials and FDA approval, as well as the development of lower-cost, dedicated fs laser systems.

Glossary of Acronyms

2D	Two-Dimensional
2PF	Two-Photon Fluorescence
3D	Three-Dimensional
3PF	Three-Photon Fluorescence
ACC	Automatic Current Control
ACF	Autocorrelation Function
AI	Artificial Intelligence
AM1G	Air Mass 1 Global
ANSI	American National Standards Institute
APC	Automatic Power Control
APDs	Avalanche Photodiodes
ATP	Acceptance Tests Procedure
BBO	Beta-Barium Borate (BBO)
BJT	Bipolar Junction Transistor
BPP	Beam Parameter Product
CARS	Coherent anti-Stokes Raman Scattering
CCD	Charge-Coupled Device
CD	Critical Dimension
CEs	Conversion Efficiencies
ChR2	Channelrhodopsin-2
CIE	International Commission on Illumination
CMOS	Complementary Metal Oxide Semiconductor
CNC	Computer Numerical Control
COLIPA	European Cosmetic Toiletry Perfumery Association
c-Si	Crystalline-Silicon
CW	Continuous Wave
CWA	Chemical Warfare Agents
DARC	Dielectric Anti-Reflective Coating
DBR	Distributed Bragg Reflectors
DFB	Distributed Feedback
DFG	Difference Frequency Generation
DH	Double Heterostructure
DOF	Depth of Focus
DPSS	Diode Pumped Solid-state
DRAM	Dynamic Random Access Memory
DUT	Device Under Test
DUV	Deep Ultraviolet
DVT	Design Verification Test
EB	Electron Beam
ECCE	Extracapsular Cataract Extraction
ECDL	External Cavity Diode Lasers
EFL	Effective Focal Length
eGFP	Enhanced Green Fluorescent Protein
EMI	Electromagnetic Interference
EQE	External Quantum Efficiency
ESI	ElectroSpray Ionization
EUV	Extreme UV
F/#	f-number
FABS	Fabrication Plants
FBGs	Fiber Bragg Gratings
FDA	United States Food and Drug Administration
FIR	Far-Infrared
FLIR	Forward-Looking InfraRed
FOV	Field-of-View
FPAs	Focal Plane Arrays
FROG	Frequency-Resolved Optical Gating
fs	Femtoseconds
FSC	Forward Scattered Light
FSMs	Fast Steering Mirrors
FTIR	Fourier Transform Infrared
FW	Full-Waveform
FWHM	Full-Width at Half-Maximum
GaAs	Gallium Arsenide
GbE	Gigabit Ethernet
GC	Gas Chromatography
GFP	Green Fluorescent Protein
GPS	Global Positioning System
HAZ	Heat-Affected Zone
HW1/e^2	Half-Width 1/e^2

IC	Integrated Circuit
ICP-OES	Inductively Coupled Plasma-Optical Emission Spectrometer
IEC	International Electrotechnical Commission
IMU	Inertial Measurement Unit
IOLs	Intraocular Lenses
IoT	Internet of Things
IQE	Internal Quantum Efficiency
IR	Infrared
IRB	Ideal Rigid Body
ISB	Integrated Solutions Business
ISO	International Organization for Standards
JCIA	Japan Cosmetic Industry Association
K	Beam Propagation Factor
KD*P	Potassium Dideuterium Phosphate
KDP	Potassium Dihydrogen Phosphate
KTP	Potassium Titanyl Phosphate
L&M	Light & Motion
LASER	Light Amplification by Stimulated Emission of Radiation
LASIK	Laser-Assisted in Situ Keratomileusis
LBO	Lithium Triborate
LC	Liquid Chromatography
LCDs	Liquid Crystal Displays
LEDs	Light-Emitting Diodes
$LiNbO_3$	Lithium Niobate
LER	Line Edge Roughness
L-I	Light-Current
LiDAR	Light Detection and Ranging
LIDT	Laser-Induced Damage Threshold
LIV	Light, Current, Voltage
LWIR	Long-Wavelength Infrared
M^2	M-Squared
MALDI	Matrix Assisted Laser Desorption/Ionization
MIM	Minimum Incremental Motion
MIR	Mid-Infrared
MOPA	Master-Oscillator Power-Amplifier
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor
MPE	Maximum Permissible Exposure
MRF	Magneto Rheological Figuring
MS	Mass Spectrometry
MTF	Modulation Transfer Function
MWIR	Mid-Wavelength Infrared
NA	Numerical Aperture
NDIR	Non-Dispersive Infrared
NEI	National Eye Institute
NEP	Noise Equivalent Power
NILS	Normalized Image Log-Slope
NIR	Near-Infrared
NIST	National Institute of Standards and Technology
NLO	Nonlinear Optical
NOPA	Non-Collinear Optical Parametric Amplifiers
ns	Nanoseconds
OE	Optical-to-Electrical
OLED	Organic Light Emitting Diode
OPA	Optical Parametric Amplification
OPC	Optical Proximity Correction
OPO	Optical Parametric Oscillator
PALM	Photoactivatable Localization Microscopy
PAM	Photoacoustic Microscopy
PCBs	Printed Circuit Boards
PCFs	Photonic Crystal Fibers
PCR	Polymerase Chain Reaction
PD	Photodiode
PDA	Photodiode Array
PERC	Passivated Emitter Rear Contact
PET	Polyethylene Terephthalate
Phaco	Phacoemulsification
PI	Polyimide
PL	Pitch lapping
Plat	Sensor Platform
PMT	Photomultiplier Tube
POD	Pulse-On-Demand
PPB	Parts Per Billion
PPE	Personal Protection Equipment

PRF	Pulse Repetition Frequency
ps	Picoseconds
PSD	Power Spectral Density
PTB	Physikalisch-Technische Bundesanstalt
Q	Quality Factor
QE	Quantum Efficiency
QTH	Quartz-Tungsten Halogen
QW	Quantum-Well
ReLEx SMILE	Refractive Lenticule Extraction Small Incision Lenticule Extraction
RFP	Red Fluorescent Protein
SD FROG	Self-Diffraction Frequency-Resolved Optical Gating
SFG	Sum-Frequency Generation
SHG	Second Harmonic Generation
SiP	Silicon Photonics
SLM	Spatial Light Modulator
SNR	Signal-to-Noise Ratio
SPDT	Single-Point Diamond Turning
SPF	Sun Protection Factor
SR	Spontaneous Raman
SRS	Stimulated Raman Scattering
SSC	Side Scattered Light
STED	Stimulated Emission Depletion
STORM	Stochastic Optical Reconstruction Microscopy
SWaP	Size, Weight, and Power Consumption
SWIR	Short-Wavelength Infrared
TE	Thermoelectric
TEM	Transverse Electromagnetic
TFS	Tunable Filter Spectrometer
TIR	Total Internal Reflection
TLS	Tunable Light Source
ToF	Time-of-Flight
TPEF	Two-Photon Excitation Fluorescence
TRU	Time-Recombination Unit
UAVs	Unmanned Aerial Vehicles
UV	Ultraviolet
V&A	Vacuum & Analysis
VCSEL	Vertical Cavity Surface Emitting Laser
VIS	Visible
WPS	Wafer Per Second
ZTT	Z Tip Tilt

This page intentionally left blank.

Bibliography

Online content listed below is provided as additional reference material for the reader. While all these links were active and current as of December 2018, external links, i.e. those which MKS does not own, may have subsequently changed or may not longer be active.

[1] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Second Edition, Wiley, 2007.

[2] R. Menzel, *Photonics - Linear and Nonlinear Interactions of Laser Light and Matter*, Second Edition, Springer, 2007.

[3] W. T. Silfvast, "Lasers", *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[4] L. J. Vandergriff, "Nature and Properties of Light", *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[5] *MKS Instruments Handbook: Semiconductor Devices and Process Technology*, MKS Instruments, https://www.mksinst.com/about/MKSeBook2017.aspx, 2017.

[6] Khan Academy. "Introduction to Light," [Online]. Available: https://www.khanacademy.org/science/chemistry/electronic-structure-of-atoms/bohr-model-hydrogen/v/introduction-to-light.

[7] Online. "How Lasers Work," [Online]. Available: https://science.howstuffworks.com/laser.htm.

[8] MIT OpenCourseWare. "Laser Fundamentals I | MIT Understanding Lasers and Fiberoptics," [Online]. Available: https://www.youtube.com/watch?v=rgivGZqFcfY.

[9] R. Paschotta. "Lasers - Principles of Operation," [Online]. Available: https://www.rp-photonics.com/lasers.html?s=ak.

[10] W. Koechner, *Solid-State Laser Engineering*, Sixth Edition, Springer, 2006.

[11] O. Svelto, *Principles of Lasers*, Fifth Edition, Springer, 2009.

[12] J. T. Verdeyen, *Laser Electronics*, Third Edition, Prentice Hall, 1995.

[13] A. E. Siegman, *Lasers*, First Edition, University Science, 1986.

[14] W. T. Silfvast, *Laser Fundamentals*, Second Edition, Cambridge University Press, 2004.

[15] W. T. Silvfast, "Lasers", *Handbook of Optics - Volume 1*, M. Bass, ed., McGraw-Hill Inc., 1995.

[16] Online. "Laser," [Online]. Available: https://en.wikipedia.org/wiki/Laser.

[17] Online. "List of Laser Types," [Online]. Available: https://en.wikipedia.org/wiki/List_of_laser_types.

[18] J. Mertz, *Introduction to Optical Microscopy*, Roberts and Company Publishers, 2009.

[19] R. Paschotta. “Polarization of Laser Emission,” [Online]. Available: https://www.rp-photonics.com/polarization_of_laser_emission.html.

[20] R. Nave. “Lasers and Their Properties,” [Online]. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/optmod/qualig.html#c4.

[21] Online. “List of Laser Applications,” [Online]. Available: https://en.wikipedia.org/wiki/List_of_laser_applications.

[22] R. Paschotta. “Laser Applications,” [Online]. Available: https://www.rp-photonics.com/laser_applications.html.

[23] Lumibird. “Nd:YAG Laser Rods. ,” [Online]. Available: https://www.photonics.com/a25042/NdYAG_Lasers_Standing_the_Test_of_Time.

[24] R. Paschotta. “Fiber Lasers,” [Online]. Available: https://www.rp-photonics.com/fiber_lasers.html.

[25] J. D. Kafka, Laser Diode Pumped Fiber Lasers with Pump Cavity, United State of America 4,829,529, to Spectra-Physics, Inc., 1989.

[26] Online. “Diode-pumped Solid-state Laser,” [Online]. Available: https://en.wikipedia.org/wiki/Diode-pumped_solid-state_laser.

[27] R. Paschotta. “Solid-state Lasers,” [Online]. Available: https://www.rp-photonics.com/solid_state_lasers.html.

[28] R. Paschotta. “Gas Lasers,” [Online]. Available: https://www.rp-photonics.com/gas_lasers.html.

[29] B. Shiner. “Fiber Lasers: New Types and Features Expand Applications,” [Online]. Available: https://www.photonics.com/a25158/Fiber_Lasers_New_Types_and_Features_Expand.

[30] R. Paschotta. “Pulse Generation,” [Online]. Available: https://www.rp-photonics.com/pulse_generation.html.

[31] Online. “Pulsed Laser,” [Online]. Available: https://en.wikipedia.org/wiki/Pulsed_laser.

[32] Optcomm. “Principles Of Mode-Locking - Passively Mode-Locked Lasers,” [Online]. Available: https://www.youtube.com/watch?v=efxFduO2YI8.

[33] Debdeep Jena. “Lasers & Optoelectronics Lecture 23: Mode Locked Lasers (Cornell ECE4300 Fall 2016),” [Online]. Available: https://www.youtube.com/watch?v=Jf-hQzVsLJ8&t=1s.

[34] Debdeep Jena. “Lasers & Optoelectronics Lecture 22: Q-Switching in Lasers (Cornell ECE4300 Fall 2016),” [Online]. Available: https://www.youtube.com/watch?v=mSEqylb-Pkg.

[35] Lehrstuhl für Lasertechnik LLT. “Pulsed laser radiation I,” [Online]. Available: https://www.youtube.com/watch?v=or4T01PaMXE.

[36] R. W. Boyd, *Nonlinear Optics*, Third Edition, Elsevier Science, 2008.

[37] G. I. Stegeman and Robert A. Stegeman, *Nonlinear Optics - Phenomena, Materials, and Devices*, First Edition, Wiley, 2012.

[38] R. Paschotta. "Nonlinear Crystal Materials," [Online]. Available: https://www.rp-photonics.com/nonlinear_crystal_materials.html.

[39] R. Paschotta. "Wavelength Tuning," [Online]. Available: https://www.rp-photonics.com/wavelength_tuning.html.

[40] R. Paschotta. "Nonlinear Frequency Conversion," [Online]. Available: https://www.rp-photonics.com/nonlinear_frequency_conversion.html.

[41] Online. "Nonlinear Optics," [Online]. Available: https://en.wikipedia.org/wiki/Nonlinear_optics.

[42] R. Paschotta. "Laser Safety," [Online]. Available: https://www.rp-photonics.com/laser_safety.html?s=ak.

[43] F. Seeber, "Light Sources and Laser Safety", *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[44] Rockwell Laser Industries Inc. "Laser Standards and Classifications," [Online]. Available: http://www.rli.com/resources/articles/classification.aspx.

[45] R. Paschotta. "Laser-Induced Damage," [Online]. Available: https://www.rp-photonics.com/laser_induced_damage.html.

[46] Online. "Laser Safety," [Online]. Available: https://en.wikipedia.org/wiki/Laser_safety#Revised_system.

[47] Newport Corporation. "Laser Damage Threshold," [Online]. Available: https://www.newport.com/n/laser-damage-threshold.

[48] Laser Institute of America. "Laser Safety Resources," [Online]. Available: https://www.lia.org/resources.

[49] P. L. Derry, Luis Figueroa, and Chi-Shain Hong, "*Semiconductor Lasers*", *Handbook of Optics - Volume 1*, M. Bass, ed., McGraw-Hill Inc., 1995.

[50] R. Nave. "Light-Emitting Diode Structure," [Online]. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/Electronic/led.html.

[51] Online. "Laser Diode," [Online]. Available: https://en.wikipedia.org/wiki/Laser_diode.

[52] Online. "Direct and Indirect Band Gaps," [Online]. Available: https://en.wikipedia.org/wiki/Direct_and_indirect_band_gaps.

[53] Online. "Light-Emitting Diode," [Online]. Available: https://en.wikipedia.org/wiki/Light-emitting_diode.

[54] R. Paschotta. "Semiconductor Lasers," [Online]. Available: https://www.rp-photonics.com/semiconductor_lasers.html.

[55] R. Paschotta. "Light-Emitting Diode," [Online]. Available: https://www.rp-photonics.com/light_emitting_diodes.html.

[56] MIT OpenCourseWare. “PN Junction, Diode and Photovoltaic Cells,” [Online]. Available: https://www.youtube.com/watch?v=ZYMWsmertQk.

[57] R. H. Haitz, M. George Craford, and Robert H. Weissman, “Light-Emitting Diodes”, *Handbook of Optics - Volume 1*, M. Bass, ed., McGraw-Hill Inc., 1995.

[58] R. Paschotta. “Laser Diodes,” [Online]. Available: https://www.rp-photonics.com/laser_diodes.html.

[59] R. Paschotta. “Vertical Cavity Surface-emitting Lasers,” [Online]. Available: https://www.rp-photonics.com/vertical_cavity_surface_emitting_lasers.html.

[60] B. Kashyap. “Photonic Devices (LED, Photo diode, LASER, PIN diode) (Electron devices),” [Online]. Available: https://www.youtube.com/watch?v=hEJq8_iWsh4.

[61] B. Kashyap. “Photonic Devices (LED, Photo diode, LASER, PIN diode) Continued... (Electron devices),” [Online]. Available: https://www.youtube.com/watch?v=cnZ8-A_CH7c.

[62] Afrotechmods. “LED Basics,” [Online]. Available: https://www.youtube.com/watch?v=Yo6JI_bzUzo.

[63] Newport Corporation. “Laser Diode Technology,” [Online]. Available: https://www.newport.com/t/laser-diode-technology.

[64] Newport Corporation. “Laws of Radiation,” [Online]. Available: https://www.newport.com/t/laws-of-radiation.

[65] Newport Corporation. “An Overview of Laser Diode Characteristics,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/he9/hd7/8797049520158/AN05-Laser-Diode-Characteristics-Overview.pdf.

[66] Newport Corporation. “Laser Diode Controllers,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h8c/hcc/8797188816926/Laser-Diode-Controllers-Tighter-Control-More-Integration-and-Higher-Current.pdf.

[67] Newport Corporation. “Laser-Diode Power-Supply Basics: Performance Depends on Specification “ [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h4b/h70/8797188849694/Laser-diode-power-supply-basics.pdf.

[68] Newport Corporation. “Test and Characterization of Laser Diodes,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h8a/hde/8797287514142/Test-and-Characterization-of-Laser-Diodes-App-Note-1.pdf.

[69] Newport Corporation. “Modulating Laser Diodes,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h1b/heb/8797049913374/AN22-Modulating-Laser-Diodes.pdf.

[70] Newport Corporation. “Pulsing a Laser Diode,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/ha6/h59/8797049618462/AN11-Pulsing-a-Laser-Diode.pdf.

[71] Newport Corporation. “Optimizing TE Drive Current,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/hc2/hd2/8797049716766/AN14-Optimizing-TE-Drive-Current.pdf.

[72] Newport Corporation. “Controlling Temperatures of Diode Lasers Thermoelectrically,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/hcc/hc7/8797049389086/AN01-Controlling-Temperatures-of-Laser-Diodes-Thermoelectrically.pdf.

[73] U. Malzahn. “Driving Diode Lasers: A Straightforward Procedure,” [Online]. Available: https://www.photonics.com/a32187/Driving_Diode_Lasers_A_Straightforward_Procedure.

[74] Newport Corporation. “6100 LIV Characterization Demonstration,” [Online]. Available: https://www.youtube.com/watch?v=zO7gXkNIN0Q.

[75] R. Paschotta. “Laser Diode Drivers,” [Online]. Available: https://www.rp-photonics.com/laser_diode_drivers.html?s=ak.

[76] Online. “Thermal Radiation,” [Online]. Available: https://en.wikipedia.org/wiki/Thermal_radiation.

[77] Online. “Xenon Arc Lamp,” [Online]. Available: https://en.wikipedia.org/wiki/Xenon_arc_lamp.

[78] Newport Corporation. “Choosing a Light Source,” [Online]. Available: https://www.newport.com/g/choosing-a-light-source.

[79] Newport Corporation. “Oriel Calibration Sources,” [Online]. Available: https://www.newport.com/n/oriel-calibration-sources.

[80] Online. “Deuterium Arc Lamp,” [Online]. Available: https://en.wikipedia.org/wiki/Deuterium_arc_lamp.

[81] Online. “Halogen Lamp,” [Online]. Available: https://en.wikipedia.org/wiki/Halogen_lamp.

[82] Newport Corporation. “Tunable QTH Light Sources,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h0b/h9a/9131612110878/DS-051504-Qth-Light-Source-4pg.pdf.

[83] Newport Corporation. “Infrared Elements,” [Online]. Available: https://www.newport.com/f/infrared-elements.

[84] Newport Corporation. “Information on Spectral Irradiance Data,” [Online]. Available: https://www.newport.com/n/information-on-spectral-irradiance-data.

[85] National Institute of Standards and Technology. “Spectroradiometric Source Measurements Calibrations, “ [Online]. Available: https://www.nist.gov/calibrations/spectroradiometric-source-measurements-calibrations.

[86] Hamamatsu Photonics. “Light Source,” [Online]. Available: https://www.hamamatsu.com/resources/pdf/etd/LIGHT_SOURCE_TLSZ0001E.pdf.

[87] M. W. Davidson. “Fundamentals of Xenon Arc Lamps,” [Online]. Available: http://zeiss-campus.magnet.fsu.edu/articles/lightsources/xenonarc.html.

[88] E. F. Zalewski, “Optical Measurements,” *Handbook of Optics - Volume 2*, M. Bass, ed., McGraw-Hill Inc., 1995.

[89] Physikalisch-Technische Bundesanstalt. “Metrological Services,” [Online]. Available: https://www.ptb.de/cms/en/metrological-services.html.

[90] P. R. Norton, “Photodetectors,” *Handbook of Optics - Volume 1*, M. Bass, ed., McGraw-Hill Inc., 1995.

[91] Ophir Optronics Solution Ltd. “Laser Power Through Aperture Calculator,” [Online]. Available: http://www.ophiropt.com/laser--measurement/laser-power-energy-meters/services/power-through-aperture-calculator?r=peakpower.

[92] Ophir Optronics Solution Ltd. “Laser Power Density Calculator,” [Online]. Available: http://www.ophiropt.com/laser--measurement/power-density-calculator?r=drp.

[93] Newport Corporation. “Optical Detection Systems,” [Online]. Available: https://www.newport.com/n/optical-detection-systems.

[94] Newport Corporation. “Optical Power Measurement,” [Online]. Available: https://www.newport.com/t/optical-power-measurement.

[95] Newport Corporation. “Optical Radiation Terminology and Units,” [Online]. Available: https://www.newport.com/n/optical-radiation-terminology-and-units.

[96] Online. “Photodetector,” [Online]. Available: https://en.wikipedia.org/wiki/Photodetector.

[97] R. Paschotta. “Photodetectors,” [Online]. Available: https://www.rp-photonics.com/photodetectors.html.

[98] Newport Corporation. “Thermopile Laser Power Sensor Technology,” [Online]. Available: https://www.newport.com/t/thermopile-laser-power-sensor-technology.

[99] M. R. Shenoy. “General Characteristics of Photodetectors,” [Online]. Available: https://www.youtube.com/watch?v=N2IAs3b8H4Q.

[100] Ophir Optronics Solution Ltd. “Sensor Finder and Laser Fluence Calculator,” [Online]. Available: http://www.ophiropt.com/laser--measurement/laser-power-energy-meters/services/sensor-finder.

[101] Ophir Optronics Solution Ltd. “Photodiode Sensors for Measuring Very Low Powers,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/12063.

[102] Ophir Optronics Solution Ltd. “Laser Parameters for Choosing a Sensor: The pitfall of worst case combinations,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/11857.

[103] Ophir Optronics Solution Ltd. “Measuring Laser Power and Energy with Meters and Sensors,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/7719.

[104] Ophir Optronics Solution Ltd. “How to choose the right sensor using the Ophir Sensor Finder,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/8543.

[105] Ophir Optronics Solution Ltd. “Sensors for Measuring Laser Energy,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/2246.

[106] Ophir Optronics Solution Ltd. “Measuring Power of LEDs: UV, Visible and NIR,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/3768.

[107] Ophir Optronics Solution Ltd. “How to Choose the Right Sensor Using the LED Sensor Finder,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/10128.

[108] Ophir Optronics Solution Ltd. “Introducing the FluxGage LED Luminaire Test System,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/10143.

[109] Ophir Optronics Solution Ltd. “FluxGage LED Luminaire Test System: Setup and Use,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/11178.

[110] Ophir Optronics Solution Ltd. “Focal Spot Size Calculator for Gaussian Laser Beams,” [Online]. Available: http://www.ophiropt.com/laser--measurement/laser-power-energy-meters/services/focal-spot-size-calculator-for-gaussian-beams.

[111] Ophir Optronics Solution Ltd. “Laser Focusability (Gaussian Beam),” [Online]. Available: http://www.ophiropt.com/laser--measurement/laser-power-energy-meters/services/laser-focusability-calculator.

[112] Ophir Optronics Solution Ltd. “Laser Beam Measurement Vocabulary,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/article/9547.

[113] Newport Corporation. “Gaussian Beam Optics,” [Online]. Available: https://www.newport.com/n/gaussian-beam-optics.

[114] R. Paschotta. “M2 Factor,” [Online]. Available: https://www.rp-photonics.com/m2_factor.html.

[115] R. Paschotta. “Gaussian Beams,” [Online]. Available: https://www.rp-photonics.com/gaussian_beams.html.

[116] R. Paschotta. “Flat-top Beams,” [Online]. Available: https://www.rp-photonics.com/flat_top_beams.html?s=ak.

[117] Ophir Optronics Solution Ltd. “Beam Profiling: A Primer,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/article/8113.

[118] Ophir Optronics Solution Ltd. “A shopping list for your beam profiler,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/article/8112.

[119] R. Paschotta. “Beam Profilers,” [Online]. Available: https://www.rp-photonics.com/beam_profilers.html?s=ak.

[120] Ophir Optronics Solution Ltd. “Benefits of Beam Profiling,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/article/8110.

[121] Ophir Optronics Solution Ltd. “State-of-the-Art Report: Laser Beam Measuring Instruments,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/article/8141.

[122] Ophir Optronics Solution Ltd. “M2 Laser Measurement,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/7745.

[123] Ophir Optronics Solution Ltd. “Laser Beam Profiler Wizard,” [Online]. Available: http://www.ophiropt.com/laser--measurement/beam-profiler-wizard.

[124] Ophir Optronics Solution Ltd. “NanoScan Beam Profiler Demo,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/9424.

[125] Ophir Optronics Solution Ltd. “BeamGage Laser Beam Profiling System,” [Online]. Available: http://www.ophiropt.com/laser--measurement/knowledge-center/video/11041.

[126] Ophir Optronics Solution Ltd. “Beam Profiler Products,” [Online]. Available: http://www.ophiropt.com/laser--measurement/beam-profilers?r=drp.

[127] Newport Corporation. “Optical Detector Definitions of Characteristics,” [Online]. Available: https://www.newport.com/n/optical-detector-definitions-of-characteristics.

[128] Newport Corporation. “High Speed Detector Terminology,” [Online]. Available: https://www.newport.com/t/high-speed-detectors.

[129] M. R. Shenoy. “Semiconductor Photo-Diodes,” [Online]. Available: https://www.youtube.com/watch?v=S9jMMALKKKs.

[130] M. R. Shenoy. “Responsivity and Impulse Response,” [Online]. Available: https://www.youtube.com/watch?v=a1YhuBrbRI8.

[131] J. E. Bowers and Y. G. Wey, “High-Speed Photodetectors,” *Handbook of Optics - Volume 1*, M. Bass, ed., McGraw-Hill Inc., 1995.

[132] J. R. Willison, “Signal Detection and Analysis,” *Handbook of Optics - Volume 1*, M. Bass, ed., McGraw-Hill Inc., 1995.

[133] Newport Corporation. “Photomultiplier Tubes,” [Online]. Available: https://www.newport.com/f/photomultiplier-tubes.

[134] Newport Corporation. “Using High-Speed Photodetectors for Pulsed-Laser Measurements,” [Online]. Available: https://www.newport.com/n/using-high-speed-photodetectors-for-pulsed-laser-measurements.

[135] Newport Corporation. “Insights into High-Speed Detectors and High-Frequency Techniques,” [Online]. Available: https://www.newport.com/n/insights-into-high-speed-detectors-and-high-frequency-techniques.

[136] Newport Corporation. “A Survey of Methods Using Balanced Photodetection,” [Online]. Available: https://www.newport.com/n/a-survey-of-methods-using-balanced-photodetection.

[137] C. Palmer, *Diffraction Grating Handbook*, Newport Corporation, http://optics.sgu.ru/~ulianov/Students/Books/Applied_Optics/E.%20Loewen%20Diffraction%20Grating%20Handbook%20%282005%29.pdf, 2005.

[138] E. Hecht, *Optics*, Fourth Edition, Addison Wesley, 2002.

[139] Newport Corporation. “Diffraction Grating Physics,” [Online]. Available: https://www.newport.com/t/grating-physics.

[140] Newport Corporation. "Diffraction Grating Selection Guide," [Online]. Available: https://www.newport.com/g/diffraction-grating-selection-guide.

[141] R. Paschotta. "Diffraction Gratings," [Online]. Available: https://www.rp-photonics.com/diffraction_gratings.html.

[142] Khan Academy. "Diffraction Grating," [Online]. Available: https://www.youtube.com/watch?v=F6dZjuw1KUo.

[143] A Level Physics Online. "A Level Physics - Diffraction Gratings," [Online]. Available: https://www.youtube.com/watch?v=flPSbngdLwl.

[144] Newport Corporation. "Bandpass of a Spectrograph," [Online]. Available: https://www.newport.com/t/bandpass-of-a-spectrograph.

[145] Newport Corporation. "Resolution of Monochromators and Spectrographs," [Online]. Available: https://www.newport.com/t/resolution-of-monochromators-and-spectrographs.

[146] Newport Corporation. "Stray Light in Monochromators and Spectrographs," [Online]. Available: https://www.newport.com/t/stray-light-in-monochromators-and-spectrographs.

[147] R. Paschotta. "Spectrometers," [Online]. Available: https://www.rp-photonics.com/spectrometers.html.

[148] Newport Corporation. "Calculating Monochromator Throughput," [Online]. Available: https://www.newport.com/t/calculating-monochromator-throughput.

[149] J-C. Diels and Wolfgang Rudolph, *Ultrashort Laser Pulse Phenomena*, Second Edition, Academic Press, 2006.

[150] Ophir Optronics Solution Ltd. "Laser Peak Power Calculator," [Online]. Available: https://www.ophiropt.com/laser--measurement/laser-power-energy-meters/services/peak-power-calculator?r=Density.

[151] R. Trebino, *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*, Kluwer Academic Publishers, 2002.

[152] Newport Corporation. "Long Scan Autocorrelator," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h69/ha2/8797197205534/Long-Scan-Autocorrelator-App-Note-27.pdf.

[153] Newport Corporation. "Amplitude and Phase Characterization of Ultrashort Laser Pulses," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/hab/h1c/8797049225246/Amplitude-and-Phase-Characterization-of-Ultrashort-Laser-Pulses-App-Note-33.pdf.

[154] Newport Corporation. "The Effect of Dispersion on Ultrashort Pulses," [Online]. Available: https://www.newport.com/n/the-effect-of-dispersion-on-ultrashort-pulses.

[155] Newport Corporation. "Multimodal Ultrafast Spectroscopy System Based on a 35 Femtosecond Ti:Sapphire Chirped Pulse Amplification Laser," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h5b/h54/8990151507998/Application-Note-41-Multimodal-Ultrafast-Spectroscopy-System-Based-on-a-35-fs-Ti-sapphire-CPA-Laser.pdf.

[156] Newport Corporation. “Pulsed Radiation,” [Online]. Available: https://www.newport.com/t/pulsed-radiation.

[157] L. S. Pedrotti, “Basic Geometrical Optics,” *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[158] Newport Corporation. “Optical Materials,” [Online]. Available: https://www.newport.com/n/optical-materials.

[159] Newport Corporation. “Optical Surfaces,” [Online]. Available: https://www.newport.com/n/optical-surfaces.

[160] Newport Corporation. “Optical Mirror Selection Guide,” [Online]. Available: https://www.newport.com/g/optical-mirror-selection-guide.

[161] The Physics Classroom. “Reflection and the Ray Model of Light,” [Online]. Available: http://www.physicsclassroom.com/class/refln.

[162] EDU.photonics.com. “Mirrors: Coating Choice Makes a Difference,” [Online]. Available: https://www.photonics.com/Article.aspx?AID=25501.

[163] R. Paschotta. “Mirrors,” [Online]. Available: https://www.rp-photonics.com/mirrors.html.

[164] D. B. Murphy, Kenneth R. Spring, Matthew J. Parry-Hill, Ian D. Johnson, and Michael W. Davidson. “Interference Filters,” [Online]. Available: https://micro.magnet.fsu.edu/primer/java/filters/interference/.

[165] Newport Corporation. “Specifying Bandpass Filters,” [Online]. Available: https://www.newport.com/n/specifying-bandpass-filters.

[166] Newport Corporation. “Optical Filter Construction,” [Online]. Available: https://www.newport.com/n/optical-filter-construction.

[167] Newport Corporation. “Dichroic Filter Guide,” [Online]. Available: https://www.newport.com/g/dichroic-filters-selection-guide.

[168] Newport Corporation. “Bandpass Filters,” [Online]. Available: https://www.newport.com/c/bandpass-filters.

[169] Newport Corporation. “Light Collection and Systems Throughput,” [Online]. Available: https://www.newport.com/t/light-collection-and-systems-throughput.

[170] W. J. Smith, *Modern Optical Engineering*, Fourth Edition, McGraw-Hill Education, 2007.

[171] The Physics Classroom. “Refraction and the Ray Model of Light,” [Online]. Available: http://www.physicsclassroom.com/class/refrn.

[172] Newport Corporation. “Optical Material Properties,” [Online]. Available: https://www.newport.com/n/optical-material-properties.

[173] Newport Corporation. “Spherical Lenses,” [Online]. Available: https://www.newport.com/g/spherical-lens-selection-guide.

[174] Newport Corporation. "Optics Fundamentals," [Online]. Available: https://www.newport.com/n/optics-fundamentals.

[175] Newport Corporation. "Focusing and Collimating," [Online]. Available: https://www.newport.com/n/focusing-and-collimating.

[176] L. S. Pedrotti, "Basic Physical Optics", *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[177] Newport Corporation. "Polarization," [Online]. Available: https://www.newport.com/n/polarization.

[178] The Physics Classroom. "Polarization," [Online]. Available: http://www.physicsclassroom.com/class/light/Lesson-1/Polarization.

[179] Newport Corporation. "Polarization Optics Selection Guide," [Online]. Available: https://www.newport.com/g/polarization-optics-selection-guide.

[180] Newport Corporation. "Polarizers," [Online]. Available: https://www.rp-photonics.com/polarizers.html.

[181] Newport Corporation. "Glan-Laser Calcite Polarizers," [Online]. Available: https://www.newport.com/f/glan-laser-calcite-polarizers.

[182] Newport Corporation. "Optical Beamsplitter Selection Guide," [Online]. Available: https://www.newport.com/g/optical-beamsplitter-selection-guide.

[183] Newport Corporation. "Broadband Polarizing Cube Beamsplitters," [Online]. Available: https://www.newport.com/f/broadband-polarizing-cube-beamsplitters.

[184] Newport Corporation. "Optical Windows Selection Guide," [Online]. Available: https://www.newport.com/g/optical-windows-selection-guide.

[185] Newport Corporation. "Broadband Beam Samplers," [Online]. Available: https://www.newport.com/f/broadband-beam-samplers.

[186] Newport Corporation. "Optical Mounts: Ignore Them at Your Peril," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/hbd/h75/8797225975838/Optical-Mounts-Ignore-Them-at-Your-Peril.pdf.

[187] Newport Corporation. "Optical Mirror Mount Technology Guide," [Online]. Available: https://www.newport.com/n/optical-mirror-mount-technology-guide.

[188] Newport Corporation. "Optical Mirror Mount Guide," [Online]. Available: https://www.newport.com/g/optical-mirror-mount-guide.

[189] Newport Corporation. "Lens Mount Selection Guide," [Online]. Available: https://www.newport.com/g/lens-mount-selection-guide.

[190] Newport Corporation. "Stability of Optical Posts vs. Pedestals," [Online]. Available: https://www.newport.com/n/stability-of-optical-posts-vs-pedestals.

[191] Newport Corporation. "How to Build a Beam Expander " [Online]. Available: https://www.newport.com/medias/sys_master/images/images/he7/h44/8797226237982/Optics-How-to-Build-a-Beam-Expander.pdf.

[192] Newport Corporation. "Spatial Filters," [Online]. Available: https://www.newport.com/n/spatial-filters.

[193] Newport Corporation. "Variable Attenuator for Lasers," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h24/h0e/8797301637150/Variable-Attenuator-for-Lasers-App-Note-26.pdf.

[194] Newport Corporation. "Computer Controlled Variable Attenuator for Tunable Lasers," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h90/h3c/8797076553758/Computer-Controlled-Variable-Attenuator-for-Tunable-Lasers-App-Note-30.pdf.

[195] Newport Corporation. "Laser Beam Expander," [Online]. Available: https://www.youtube.com/watch?v=SK66JA5m4M8.

[196] N. Massa, "Fiber Optic Telecommunication," *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[197] A. Ghatak and K. Thyagarajan, "Optical Waveguides and Fibers," *Fundamentals of Photonics*, C. Roychoudhuri, ed., SPIE Press, 2008.

[198] Newport Corporation. "Fiber Optics: How to Choose an Optic for Free Space Fiber Coupling," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h04/h2a/9131726962718/Tech-Note-14-Fiber-Coupling.pdf.

[199] Newport Corporation. "Fiber Optic Coupling," [Online]. Available: https://www.newport.com/n/fiber-optic-coupling.

[200] Newport Corporation. "Photonic-crystal fibers have many uses," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h7a/hbe/8797239771166/Photonic-Crystal-Fibers-have-many-uses.pdf.

[201] Newport Corporation. "Fiber Optic Basics," [Online]. Available: https://www.newport.com/t/fiber-optic-basics.

[202] Newport Corporation. "Polarization in Fiber Optics," [Online]. Available: https://www.newport.com/t/polarization-in-fiber-optics.

[203] Newport Corporation. "Highly Germanium Doped Photosensitive Fibers," [Online]. Available: https://www.newport.com/f/high-germanium-doped-photosensitive-fibers.

[204] R. Paschotta. "Fibers," [Online]. Available: https://www.rp-photonics.com/fibers.html.

[205] R. Paschotta. "Photonic Crystal Fibers," [Online]. Available: https://www.rp-photonics.com/photonic_crystal_fibers.html?s=ak.

[206] R. Paschotta. "Fiber Bragg Gratings," [Online]. Available: https://www.rp-photonics.com/fiber_bragg_gratings.html?s=ak.

[207] A. Yariv and P. Yeh, *Optical Waves In Crystals*, John Wiley & Sons, 2002.

[208] Newport Corporation. "Practical Uses and Applications of Electro-Optic Modulators," [Online]. Available: https://www.newport.com/n/practical-uses-and-applications-of-electro-optic-modulators.

[209] Newport Corporation. "Phase and Amplitude Modulator Definitions of Characteristics," [Online]. Available: https://www.newport.com/n/phase-and-amplitude-modulator-definitions-of-characteristics.

[210] Newport Corporation. "Electro-Optic Modulator Selection Guide," [Online]. Available: https://www.newport.com/g/electro-optic-modulator-selection-guide.

[211] Newport Corporation. "Phase and Amplitude Modulator Applications," [Online]. Available: https://www.newport.com/n/phase-and-amplitude-modulator-applications.

[212] R. Paschotta. "Electro-optic Modulators," [Online]. Available: https://www.rp-photonics.com/electro_optic_modulators.html.

[213] Newport Corporation. "Manual Positioning Basics," [Online]. Available: https://www.newport.com/t/manual-positioning-basics.

[214] Newport Corporation. "Motion Basics Terminology & Standards," [Online]. Available: https://www.newport.com/n/motion-basics-terminology-and-standards.

[215] Newport Corporation. "On the Origin of Motion Control Specifications," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/he1/h94/8797225091102/On-the-Origin-of-Motion-Control-Specifications-App-Note-1.pdf.

[216] Newport Corporation. "Specifications don't always tell the whole story," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h6f/h8c/8797212868638/Motion-specifications-don-t-always-tell-the-whole-story.pdf.

[217] Newport Corporation. "Linear Translation Stage Technology Guide," [Online]. Available: https://www.newport.com/n/linear-translation-stage-technology-guide.

[218] Newport Corporation. "Stage Components Considerations," [Online]. Available: https://www.newport.com/n/stage-components-considerations.

[219] Newport Corporation. "Linear Translation Stage Terminology," [Online]. Available: https://www.newport.com/n/linear-translation-stage-terminology.

[220] Newport Corporation. "Motion Electronics Design," [Online]. Available: https://www.newport.com/n/motion-electronics-design.

[221] Newport Corporation. "Control Theory Terminology," [Online]. Available: https://www.newport.com/n/control-theory-terminology.

[222] Newport Corporation. "The Top Five Motion Application Mistakes ... and How to Avoid Making Them," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h5c/h11/8797287710750/The-Top-Five-Motion-Application-Mistakes-and-How-to-Avoid-Them.pdf.

[223] Newport Corporation. "Fundamentals of Vibration," [Online]. Available: https://www.newport.com/t/fundamentals-of-vibration.

[224] Newport Corporation. "Understanding the Compliance Curve," [Online]. Available: https://www.newport.com/t/understanding-the-compliance-curve.

[225] Newport Corporation. "About Optical Table Performance Specifications," [Online]. Available: https://www.newport.com/t/about-optical-table-performance-specifications.

[226] Newport Corporation. "Approximating Real-World Beam Deflection," [Online]. Available: https://www.newport.com/t/approximating-real-world-beam-deflection.

[227] Newport Corporation. "How to Approximate an Optical Tables "Ideal Rigid Body" Line," [Online]. Available: https://www.newport.com/t/how-to-approximate-an-optical-tables-ideal-rigid-body-line.

[228] Newport Corporation. "Vibration Control Definitions of Characteristics," [Online]. Available: https://www.newport.com/t/vibration-control-definitions-of-characteristics.

[229] Newport Corporation. "How to determine your equipment needs," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/hb4/hfe/8797087891486/Determine-Equipment-Needs.pdf.

[230] Newport Corporation. "Effective vibration reduction stabilizes laser beams," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h15/hdc/8797089300510/Effective-vibration-reduction-stabilizes-laser-beams.pdf.

[231] Newport Corporation. "Optics Based Research - The Need for Vibration Isolated Rigid Support Structures," [Online]. Available: https://www.newport.com/n/optics-based-research-the-need-for-vibration-isolated-rigid-support-structures.

[232] Newport Corporation. "Narrowband Damping with Tuned Vibration Absorbers," [Online]. Available: https://www.newport.com/n/narrowband-damping-with-tuned-vibration-absorbers.

[233] Newport Corporation. "Active Vibration Damping," [Online]. Available: https://www.newport.com/n/active-vibration-damping.

[234] Newport Corporation. "Optical Table Guide," [Online]. Available: https://www.newport.com/g/optical-tables.

[235] Newport Corporation. "Optical-table basics: from breadboards to active vibration-control systems " [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h0e/h4a/8797226041374/Optical-Table-Basics.pdf.

[236] Newport Corporation. "Proper Load Distribution Ensures Stable Operation," [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h56/h28/8797242982430/ProperLoadDistribution.pdf.

[237] Newport Corporation. "Vibration Control Technology Enables Critical Improvement in Microscopy Applications," [Online]. Available: https://www.newport.com/t/vibration-control-technology-enables-critical-improvement-in-microscopy-applications.

[238] A. Kar-Roy, D. Howard, M. Racanelli, M. Scott, P. Hurwitz, R. Zwingman, S. Chaudhry, and S. Jordan, “Mixed-signal 0.18μm CMOS and SiGe BiCMOS foundry technologies for ROIC applications,” *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 7834, no. 78340R, 2010.

[239] “Moore’s law,” [Online]. Available: https://www.britannica.com/technology/Moores-law.

[240] M. Lapedus. “Can We Measure Next-Gen FinFETs?,” [Online]. Available: https://semiengineering.com/can-we-measure-next-gen-finfets/.

[241] B. Tsai. “Inspection and Metrology to Support the Quest for Perfection,” [Online]. Available: https://www.youtube.com/watch?v=JHTh_S29yWo.

[242] Silicon Mentor. “Semiconductor Manufacturing Process,” [Online]. Available: http://www.siliconmentor.com/semiconductor-manufacturing-process/.

[243] N. Raqccah, D. Rogers, E. Tzuri, and A. Shavit, “Extending inspection limits using a DUV-laser-based bright-field inspection tool,” MICRO Magazine, July, 2006.

[244] C. Van Dommelen. “Choose the Right Lighting for Inspection,” [Online]. Available: http://www.embedded.com/print/4382211.

[245] KLA-Tencor. “Surfscan SP1 Operations and Overview Course,” [Online]. Available: http://slideplayer.com/slide/10915236/.

[246] “Fizeau Interferometer,” [Online]. Available: https://en.wikipedia.org/wiki/Fizeau_interferometer.

[247] Hitachi. “Wafer Defect Inspection Systems,” [Online]. Available: http://www.hitachi-hightech.com/global/products/device/semiconductor/inspection.html.

[248] T. K. Tran. “Defect Inspection for Advanced Technology Nodes,” [Online]. Available: https://www.nist.gov/sites/default/files/documents/pml/div683/conference/tran.pdf.

[249] E. Megged, M. Wylie, and C. Perry-Sullivan, “EUV Lithography and Overlay Control,” *YMS Magazine*, pp. 1-5, December, 2017.

[250] Ophir Optronics Solution Ltd. “Maximizing Laser Accuracy with Laser Beam Diagnostics,” [Online]. Available: http://www.ophiropt.com/laser-measurement-instruments/beam-profilers/knowledge-center/white-paper/maximizing-beam-diagnostics.

[251] R. Cappel. “Inspection Strategies for 10nm Design Node Challenges,” [Online]. Available: http://www.semi.org/en/inspection-strategies-10nm-design-node-challenges.

[252] “Lithography,” [Online]. Available: https://www.merriam-webster.com/dictionary/lithography.

[253] R.-H. M. Schmidt, “Ultraprecision engineering in lithographic exposure equipment for the semiconductor industry,” *Philosophical Transactions of the Royal Society A*, 2012.

[254] G. Stix, “Water immersion Arrangement,” *Scientific American*, vol. 293, pp. 64-67, 2005.

[255] C. Mack, *Fundamental Principles of Optical Lithography*, John Wiley & Son, 2007.

[256] Y. Wei and R. L. Brainard, *Advanced Processes for 193-nm Photolithography*, SPIE, 2009.

[257] M. Abramowitz and M. W. Davidson. "Optical Microscopy Primer," [Online]. Available: http://micro.magnet.fsu.edu/primer/anatomy/numaperture.html.

[258] C. A. Mack, "Using the Normalized Image Log-Slope," *The Lithography Expert*, vol. Winter 2001, pp. 1-6, 2001.

[259] C. A. Mack, "Using the Normalized Image Log-Slope, part 2," *The Lithography Expert*, vol. Spring 2001, pp. 5, 2001.

[260] Compugraphics. "Optical Lithography," [Online]. Available: http://www.compugraphics-photomasks.com/our-capabilities/opcphase-shift/.

[261] M. K. Kwak and L. J. Guo, "Phase Shift Lithography," *Encyclopedia of Microfluidics and Nanofluidics*, Springer-Verlag, 2014.

[262] "Phase-Shift Masks," [Online]. Available: http://spie.org/publications/fg06_p78-80_phase-shift_masks.

[263] Y. Wei and D. Back. "193 nm immersion lithography: Status and challenges," [Online]. Available: http://spie.org/newsroom/immersionlitho-intro.

[264] Monolith IC3D. "The Quad Patterning Era Begins," [Online]. Available: http://www.monolithic3d.com/blog/the-quad-patterning-era-begins.

[265] R. H. French, "Immersion Lithography: Photomask and Wafer-Level Materials," *Annual Review of Materials Research*, vol. 39, no. 1, pp. 93-126, 2009.

[266] Editor V. Bakshi, *EUV Lithography*, SPIE Press, 2009.

[267] Cymer. "EUV Sources," [Online]. Available: https://www.cymer.com/euv-light-sources/.

[268] P. French, G. Krijnen, and F. Roozeboom, "Precision in Harsh Environments," *Microsystems & Nanoengineering*, vol. 2, no. October, 2016.

[269] S. Hill. "UV Lithography: Taking extreme measures," [Online]. Available: https://phys.org/news/2011-12-uv-lithography-extreme.html.

[270] D. D. Simone, M. Mao, and G. Vandenberghe, "Novel metal-oxide photoresist materials for extreme UV lithography," *SPIE Newsroom Micro/Nano Lithography*, 2016.

[271] MKS Instruments. "Lithography in Advanced Packaging," [Online]. Available: https://www.mksinst.com/docs/ur/PackagingLithography-MBrief.pdf.

[272] S. K. Moore. "EUV Lithography Finally Ready for Chip Manufacturing," [Online]. Available: https://spectrum.ieee.org/semiconductors/nanotechnology/euv-lithography-finally-ready-for-chip-manufacturing.

[273] S. Hofman. "The era of 'Things' is coming and our industry is enabling it," [Online]. Available: https://medium.com/@ASMLcompany/the-era-of-things-is-coming-and-our-industry-is-enabling-it-faae85b0aa.

[274] M. Lapedus. "Searching For EUV Mask Defects," [Online]. Available: https://semiengineering.com/searching-for-euv-mask-defects/.

[275] T. A. Brunner, "Line-Edge Roughness performance targets for EUV Lithography," *SPIE*, vol. 10143, no. 101430E-1, pp. 5, 2017.

[276] M. Lapedus. "Next EUV Issue: Mask 3D Effects," [Online]. Available: https://semiengineering.com/next-euv-issue-mask-3d-effects/.

[277] E. Korczynski. "High-NA EUV Lithography Investment," [Online]. Available: http://semimd.com/blog/2016/11/28/high-na-euv-lithography-investment/.

[278] T. Sameshima, S. Usui, and M. Sekiya, "XeCl Excimer Laser Annealing Used In The Fabrication Of Poly-Si TFTs," *IEEE Electron Device Letters*, vol. 7, no. 5, pp. 276-278, May, 1986.

[279] G. Giordano, "Exponential growth of medical laser applications in the U.S.," *LIA Today*, January / February 2013.

[280] B. Michaels, "Shrinking stents strut their stuff," *Medical Product Manufacturing News*, vol. 27, 2011.

[281] R. R. Gattass and E. Mazur, "Femtosecond laser micromachining in transparent materials," *Nature Photonics*, vol. 2, no. 4, pp. 219-225, Apr, 2008.

[282] X. Liu, D. Du, and G. Mourou, "Laser ablation and micromachining with ultrashort laser pulses," *IEEE Journal of Quantum Electronics*, vol. 33, no. 10, pp. 1706-1716, Oct, 1997.

[283] G. Shannon and S. Hypsh. "Femtosecond Laser Processing in the Medical Device Industry," [Online]. Available: https://www.productionmachining.com/articles/femtosecond-laser-processing-in-the-medical-device-industry.

[284] R. Knappe, H. Haloui, A. Seifert, A. Weis, and A. Nebel, "Scaling ablation rates for picosecond lasers using burst micromachining," *Proceedings of SPIE, Laser-Based Micro- and Nanopackaging and Assembly IV*, vol. 7585, 2010.

[285] J. Bovatsek and R. Patel, "Ultrashort pulse laser technology for processing of advanced electronic materials," *Proceedings of LPM2017*, 2017.

[286] R. Patel, J. Bovatsek, and H. R. Goy, "Doubling Silicon Ablation Process Efficiency and Improving Quality Using High Power High Repetition Rate Green Laser with TimeShift (TM) Capability," *Journal of Laser Micro Nanoengineering*, vol. 10, no. 2, pp. 134-139, May, 2015.

[287] J. Bovatsek, "Advanced UV lasers for fast, high-precision PCB manufacturing," pp. 36-44, November, 2016.

[288] U. Brauch and F. Bammer, "Diode Lasers," *Advanced Laser Materials Processing*, D. Schuocker, ed., OMICS International, 2014.

[289] S. Prasad, L.M. Bruce, and J. Chanussot, Optical Remote Sensing: *Advances in Signal Processing and Exploitation Techniques*, Springer, 2011.

[290] "Thermography," [Online]. Available: https://en.wikipedia.org/wiki/Thermography.

[291] "Night Vision," [Online]. Available: https://en.wikipedia.org/wiki/Night_vision.

[292] Edmund Optics. "SWIR Imaging," [Online]. Available: https://www.edmundoptics.com.sg/resources/industry-expertise/swir-imaging/.

[293] ULIS. "Smart Buildings," [Online]. Available: https://www.ulis-ir.com/applications/smart-buildings.html.

[294] Device A-Lab. "Smart City," [Online]. Available: http://www.device-alab.com/smart-city/.

[295] V. C. Coffey, "Multispectral Imaging Moves into the Mainstream," *Optics & Photonic*s, April, 2012.

[296] "Infrared Window," [Online]. Available: https://en.wikipedia.org/wiki/Infrared_window.

[297] AZO Sensors. "Infrared Technology for the Automotive Industry," [Online]. Available: https://www.azosensors.com/article.aspx?ArticleID=639.

[298] The Gun Zone. "Best Thermal Imaging Rifle Scopes," [Online]. Available: https://thegunzone.com/best-thermal-imaging-rifle-scopes/.

[299] E. P. Baltsavias, "A comparison between photogrammetry and laser scanning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2-3, pp. 83-94, Jul, 1999.

[300] NEON Science. "How Does LiDAR Remote Sensing Work? Light Detection and Ranging," [Online]. Available: www.youtube.com/watch?v=EYbhNSUnIdU.

[301] B. Hofle and M. Rutzinger, "Topographic airborne LiDAR in geomorphology: A technological perspective," *Zeitschrift für Geomorphologie*, vol. 55, pp. 1-29, Mar, 2011.

[302] National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center, *Lidar 101: An Introduction to Lidar Technology, Data, and Applications*, Charleston, SC: NOAA Coastal Services Center, 2012.

[303] J. Reitberger, P. Krzystek, and U. Stilla, "Analysis of Fullwaveform LiDAR Data for Tree Species Classification," in ISPRS Symposium of Commission III "Photogrammetric Computer Vision and Image Analysis", Bonn, Germany, 2006, pp. 228-233.

[304] W. Wagner, M. Hollaus, C. Briese, and V. Ducic, "3D vegetation mapping using small-footprint full-waveform airborne laser scanners," *Int J Remote Sens*, vol. 29, no. 5, pp. 1433-1452, Mar 10, 2008.

[305] K. J. Kasunic, *Optical Systems Engineering*, McGraw-Hill Professional, 2011.

[306] "Lasers Reveal 'Lost' Roman Roads," [Online]. Available: https://www.gov.uk/government/news/lasers-reveal-lost-roman-roads.

[307] J. Fortin. "Lasers Reveal a Maya Civilization So Dense It Blew Experts' Minds," [Online]. Available: https://www.nytimes.com/2018/02/03/world/americas/mayan-city-discovery-laser.html.

[308] S. Zhang. "That Long-Lost Mayan City a Teen Found Isn't Lost … or a City," [Online]. Available: https://www.wired.com/2016/05/long-lost-mayan-city-teen-found-isnt-lost-city/.

[309] R. Diaz, S. C. Chan, and J. M. Liu, “Dual-frequency multifunction LiDAR,” *Free-Space Laser Communication Technologies XIX and Atmospheric Propagation of Electromagnetic Waves*, vol. 6457, 2007.

[310] M. Contarino. “Modulated Pulse LiDAR,” [Online]. Available: https://pdfs.semanticscholar.org/1b54/7f006c5c4c80d94af969a282fe1f6ba529cf.pdf.

[311] Cisco. “Cisco Global Cloud Index: Forecast and Methodology, 2016-2021 White Paper,” [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html.

[312] CABLExpress. “5 Reasons Why IT Professionals Choose Fiber Optic Cables Instead of Copper,” [Online]. Available: http://www.cablexpress.com/blog/5-reasons-why-it-professionals-choose-fiber-optic-cables-instead-of-copper/.

[313] H. Tritschler. “Verifying PAM4 Performance in High-Speed Designs,” [Online]. Available: http://eecatalog.com/intel/2017/09/25/verifying-pam4-performance-in-high-speed-designs/?printView=on.

[314] S. Hardy. “ICP spending on Ethernet optics to double in 2017: LightCounting,” [Online]. Available: https://www.lightwaveonline.com/articles/2017/04/icp-spending-on-ethernet-optics-to-double-in-2017-lightcounting.html.

[315] G. Halfacree. “Ethernet Alliance unveils 1.6TbE roadmap,” [Online]. Available: https://www.bit-tech.net/news/ethernet-alliances-unveils-16tbe-roadmap/1/.

[316] GRE-468-CORE. “Generic Reliability Assurance Requirements for Optoelectronic Devices Used in Telecommunications Equipment,” [Online]. Available: www.ieee802.org/3/efm/public/email/doc00005.doc.

[317] GRE-3013-CORE. “Generic Reliability Assurance Requirements for Optoelectronic Devices Used in Short-Life, Information-Handling Products and Equipment,” [Online]. Available: https://telecom-info.telcordia.com/site-cgi/ido/docs.cgi?ID=SEARCH&DOCUMENT=GR-3013&.

[318] F. R. Nash, *Estimating Device Reliability: Assessment of Credibility*, Kluwer Academic Publishers, 1993.

[319] A. Baig. “Fiber alignment and joint loss,” [Online]. Available: https://www.slideshare.net/mdabaig/fiber-alignment-and-joint-loss.

[320] Newport Corp. “Application Note 6: Fiber to Waveguide Algorithms,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h57/hc5/9135932669982/DS-03002-App-Note-6.pdf.

[321] Newport Corp. “Specifying fiber alignment system performance,” [Online]. Available: http://assets.newport.com/webDocuments-EN/images/MCTB-Fiber-Alignment.PDF.

[322] Newport Corp. “Ultra-Precision Linear Motor Stage, 50 mm Travel, 100 N Load, no cable, for XPX-D and XPS-RL,” [Online]. Available: https://www.newport.com/p/XMS50-S.

[323] G. Bouvree, *APOGEE Fiber Alignment software User’s Manual*, GBC&S Consulting Services.

[324] Newport Corp. “BR-031701 Telecom Data Brochure.pdf,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h32/h82/9041929207838/BR-031701-Telecom-Data-Brochure.pdf.

[325] Newport Corp. “LTA Precision Motorized Actuators,” [Online]. Available: https://www.newport.com/f/precision-motorized-actuators-lta-series.

[326] Newport Corp. “Nanopositioners,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h86/h72/8974577860638/Nano-Catalog-FINAL.pdf.

[327] Newport Corp. “Newport Hardware for Fiber Alignment Applications,” [Online]. Available: https://www.youtube.com/watch?v=gyc_WJ-vCXk.

[328] Newport Corp. “HXP50 Hexapod for Fiber Alignment,” [Online]. Available: https://www.youtube.com/watch?v=hP_DDio1-KI.

[329] MKS Instruments. “MKS HXP Hexapods,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/h71/h59/9134741356574/HXP-Brochure.pdf.

[330] Newport Corp. “Introducing HexaViz,” [Online]. Available: https://www.newport.com/g/hexaviz-hexapod-simulation-software.

[331] Newport Corp. “Application Note: Newport Motorized Alignment System,” [Online]. Available: https://www.newport.com/medias/sys_master/images/images/hcc/h73/8797211754526/Motion-Application-Note-Fiber-Alignment-System.pdf.

[332] J. C. Campbell, “Recent Advances in Telecommunications Avalanche Photodiodes,” *Journal of Lightwave Technology*, vol. 25, no. 1, pp. 109-121, 2007.

[333] P. Gates. “Matrix-assisted Laser Desorption/Ionization (MALDI),” [Online]. Available: http://www.chm.bris.ac.uk/ms/maldi-ionisation.xhtml.

[334] Sigma-Aldrich. “Oligonucleotide Quality Control by Mass Spectrometry,” [Online]. Available: https://www.sigmaaldrich.com/technical-documents/articles/biology/custom-dna-oligos-qc-analysis-by-mass-spectrometry.html.

[335] Newport Corp. “Introduction to FTIR Spectroscopy,” [Online]. Available: https://www.newport.com/n/introduction-to-ftir-spectroscopy.

[336] Newport Corp. “Flow Cytometry,” [Online]. Available: https://www.spectra-physics.com/search-results?addsearch=flow+cytometry.

[337] “List of Life Sciences,” [Online]. Available: https://en.wikipedia.org/wiki/List_of_life_sciences.

[338] The National Academies Press. “Optics in Health Care and the Life Sciences,” [Online]. Available: https://www.nap.edu/read/5954/chapter/5.

[339] M. Strassl, V. Wieger, D. Brodoceanu, F. Beer, A. Moritz, and E. Wintner, “Ultra-Short Pulse Laser Ablation of Biological Hard Tissue and Biocompatibles,” *Journal of Laser Micro Nanoengineering*, vol. 3, no. 1, pp. 30-40, Jan, 2008.

[340] N. Nishimura, C. B. Schaffer, and D. Kleinfeld, “In vivo manipulation of biological systems with femtosecond laser pulses,” *Proc. SPIE: High-Power Laser Ablation VI, Pts 1 and 2*, vol. 6261, 2006.

[341] I. A. Favre-Bulle, S. Zhang, A. V. Kashchuk, I. C. D. Lenton, L. J. Gibson, A. B. Stilgoe, T. A. Nieminen, and H. Rubinsztein-Dunlop, “Optical Tweezers Bring Micromachines to Biology,” *Optics and Photonics News*, vol. 29, pp. 40-47, 2018.

[342] “Press release: The 2018 Nobel Prize in Physics “ [Online]. Available: https://www.nobelprize.org/prizes/physics/2018/press-release/.

[343] A. Ashkin, “History of optical trapping and manipulation of small-neutral particle, atoms, and molecules,” *IEEE J. Sel. Top. Quant. Elec.*, vol. 6, no. 6, pp. 841-856, Nov-Dec, 2000.

[344] “Photoacoustic Microscopy,” [Online]. Available: https://en.wikipedia.org/wiki/Photoacoustic_microscopy.

[345] J. J. Yao and L. H. V. Wang, “Photoacoustic microscopy,” *Laser Photonics Rev*, vol. 7, no. 5, pp. 758-778, Sep, 2013.

[346] S. Cox, “Super-resolution imaging in live cells,” Dev Biol, vol. 401, no. 1, pp. 175-181, May 1, 2015.

[347] J. Biteen and K. A. Willets, “Introduction: Super-Resolution and Single-Molecule Imaging,” *Chem Rev*, vol. 117, no. 11, pp. 7241-7243, Jun 14, 2017.

[348] A. M. Sydor, K. J. Czymmek, E. M. Puchner, and V. Mennella, “Super-Resolution Microscopy: From Single Molecules to Supramolecular Assemblies,” *Trends Cell Biol*, vol. 25, no. 12, pp. 730-748, Dec, 2015.

[349] “Confocal Microscopy,” [Online]. Available: https://en.wikipedia.org/wiki/Confocal_microscopy.

[350] W. Denk, J. H. Strickler, and W. W. Webb, “2-Photon Laser Scanning Fluorescence Microscopy,” Science, vol. 248, no. 4951, pp. 73-76, Apr 6, 1990.

[351] Newport Corp. “Application Note: Two-Photon Excited Fluorescence Microscopy,” [Online]. Available: http://www.spectra-physics.com/assets/client_files/files/documents/AppsNote32.pdf.

[352] A. Krueger. “Multiphoton microscopy: Turnkey femtosecond lasers fuel growth of multiphoton imaging,” [Online]. Available: http://www.spectra-physics.com/assets/client_files/files/documents/LFW-Multiphoton_Micro.pdf.

[353] K. Svoboda and R. Yasuda, “Principles of two-photon excitation microscopy and its applications to neuroscience,” *Neuron*, vol. 50, no. 6, pp. 823-839, Jun 15, 2006.

[354] “Two-photon Excitation Microscopy,” [Online]. Available: https://en.wikipedia.org/wiki/Two-photon_excitation_microscopy.

[355] C. Xu and W. W. Webb, “Measurement of two-photon excitation cross sections of molecular fluorophores with data from 690 to 1050 nm,” J Opt Soc Am B, vol. 13, no. 3, pp. 481-491, Mar, 1996.

[356] Olympus. “Introduction to Multiphoton Fluorescence Microscopy,” [Online]. Available: https://www.olympus-lifescience.com/es/microscope-resource/primer/techniques/fluorescence/multiphoton/multiphotonintro/.

[357] D. Kobat, M. E. Durst, N. Nishimura, A. W. Wong, C. B. Schaffer, and C. Xu, “Deep tissue multiphoton microscopy using longer wavelength excitation,” *Opt Express*, vol. 17, no. 16, pp. 13354-13364, Aug 3, 2009.

[358] P. G. Smith and J. Klein. "Taking Multiphoton Imaging to New Depths," [Online]. Available: http://www.spectra-physics.com/assets/client_files/files/documents/Taking%20Multiphoton%20Imaging%20to%20New%20Depths.pdf.

[359] P. J. Campagnola, A. C. Millard, M. Terasaki, P. E. Hoppe, C. J. Malone, and W. A. Mohler, "Three-dimensional high-resolution second-harmonic generation imaging of endogenous structural proteins in biological tissues," *Biophys J*, vol. 82, no. 1, pp. 493-508, Jan, 2002.

[360] N. G. Horton, K. Wang, D. Kobat, C. G. Clark, F. W. Wise, C. B. Schaffer, and C. Xu, "In vivo three-photon microscopy of subcortical structures within an intact mouse brain," *Nature Photonics*, vol. 7, no. 3, pp. 205-209, Mar, 2013.

[361] M. Wang, T. Wang, C. Wu, B. Li, D. G. Ouzounov, D. Sinefeld, A. Guru, H-S. Nam, M. R. Capecchi, M. R. Warden, and C. Xu, "In vivo three-photon imaging of deep cerebellum," Proc. SPIE: *Multiphoton Microscopy in the Biomedical Sciences XVIII*, vol. 10498, 2018.

[362] M. R. Wang, C. Y. Wu, D. Sinefeld, B. Li, F. Xia, and C. Xu, "Comparing the effective attenuation lengths for long wavelength in vivo imaging of the mouse brain," *Biomed Opt Express*, vol. 9, no. 8, Aug 1, 2018.

[363] D. G. Ouzounov, T. Y. Wang, M. R. Wang, D. D. Feng, N. G. Horton, J. C. Cruz-Hernandez, Y. T. Cheng, J. Reimer, A. S. Tolias, N. Nishimura, and C. Xu, "In vivo three-photon imaging of activity of GCaMP6-labeled neurons deep in intact mouse brain," *Nat Methods*, vol. 14, no. 4, pp. 388-+, Apr, 2017.

[364] J. X. Cheng and X. S. Xie, "Coherent anti-Stokes Raman scattering microscopy: Instrumentation, theory, and applications," *J Phys Chem B*, vol. 108, no. 3, pp. 827-840, Jan 22, 2004.

[365] American Chemical Society. "C.V. Raman and the Raman Effect," [Online]. Available: https://www.acs.org/content/acs/en/education/whatischemistry/landmarks/ramaneffect.html.

[366] S. Xie. "Coherent Raman Imaging," [Online]. Available: https://bernstein.harvard.edu/research/cars-why.htm.

[367] C. W. Freudiger, W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. W. He, J. C. Tsai, J. X. Kang, and X. S. Xie, "Label-Free Biomedical Imaging with High Sensitivity by Stimulated Raman Scattering Microscopy," *Science*, vol. 322, no. 5909, pp. 1857-1861, Dec 19, 2008.

[368] P. D. Maker and R. W. Terhune, "Study of Optical Effects Due to an Induced Polarization Third Order in Electric Field Strength," *Phys Rev*, vol. 137, no. 3A, pp. A801-+, 1965.

[369] A. Zeytunyan, T. Baldacchini, and R. Zadoyan, "Module for multiphoton high-resolution hyperspectral imaging and spectroscopy," *Proc. SPIE: Multiphoton Microscopy in the Biomedical Sciences XVIII* vol. 10498, 2018.

[370] D. Fu, F. K. Lu, X. Zhang, C. Freudiger, D. R. Pernik, G. Holtom, and X. S. Xie, "Quantitative Chemical Imaging with Multiplex Stimulated Raman Scattering Microscopy," *J Am Chem Soc*, vol. 134, no. 8, pp. 3623-3626, Feb 29, 2012.

[371] L. Bruckner, T. Buckup, and M. Motzkus, "Exploring the potential of tailored spectral focusing," *J Opt Soc Am B*, vol. 33, no. 7, Jul 1, 2016.

[372] J. X. Cheng, "Coherent anti-Stokes Raman Scattering microscopy," *Appl Spectrosc*, vol. 61, no. 9, pp. 197a-208a, Sep, 2007.

[373] Newport Corp. "Spectral Focusing Timing and Recombination Unit (SF-TRU)," [Online]. Available: https://www.newport.com/medias/sys_master/root/h91/hdc/9128302247966/DS-121704-SF-TRU-Datasheet.pdf.

[374] S. Peck, "Femtosecond Lasers Advance Two-Photon Optogenetics," *Optik & Photonik*, vol. 12, no. 4, pp. 28-31, 2017.

[375] "Laser Coagulation," [Online]. Available: https://en.wikipedia.org/wiki/Laser_coagulation.

[376] A. Krueger. "Ultrafast Lasers in Biophotonics," [Online]. Available: https://www.degruyter.com/viewbooktoc/product/458320?rskey=8T62Or&result=1.

[377] T. Juhasz, H. Frieder, R. M. Kurtz, C. Horvath, J. F. Bille, and G. Mourou, "Corneal refractive surgery with femtosecond lasers," *IEEE J Sel Top Quant*, vol. 5, no. 4, pp. 902-910, Jul-Aug, 1999.

[378] T. Juhasz, G. Spooner, Z. Sacks, C. Suarez, F. Raksi, R. Zadoyan, M. Sarayba, and R. M. Kurtz, "Ophthalmic applications of ultrashort pulsed lasers," Proc. SPIE: *Commercial and Biomedical Applications of Ultrafast Lasers IV*, vol. 5340, pp. 66-75, 2004.

[379] M. Freebody. "Femtosecond Lasers Enable Bladeless Eye Surgery," [Online]. Available: https://www.photonics.com/a46013/Femtosecond_Lasers_Enable_Bladeless_Eye_Surgery.

[380] M. Lachman, "Cataract and Refractive Surgery: Hoping for a Rebound," *Medtech Insight*, vol. 11, no. 5, 2009.

[381] D. Strickland and G. Mourou, "Compression of Amplified Chirped Optical Pulses," *Opt Commun*, vol. 56, no. 3, pp. 219-221, Dec 1, 1985.

[382] M. Lachman, "ESCRS 2009: The Next Wave in Cataract and Presbyopia Devices," *Medtech Insight*, vol. 11, no. 9, 2009.

[383] M. Lachman, "Laser Cataract Surgery Comes into Focus at ASCRS," *Medtech Insight*, vol. 12, no. 5, 2010.

[384] M. Lachman, "Ophthalmology Community Focuses on Laser Cataract Surgery," Medtech Insight, vol. 12, no. 10, 2010.

[385] V. Thompson. "Cataract Surgery," [Online]. Available: http://www.allaboutvision.com/conditions/cataract-surgery.htm.

[386] "Cataract surgery," [Online]. Available: https://en.wikipedia.org/wiki/Cataract_surgery.

[387] "Wavefront shaping promises all-in-one laser cataract procedure," [Online]. Available: http://optics.org/news/9/4/41.

[388] K. Boyd. "Presbyopia Treatment," [Online]. Available: https://www.aao.org/eye-health/diseases/presbyopia-treatment.

[389] “What is the KAMRA Inlay?,” [Online]. Available: https://kamra.com/.

[390] W. H. Knox. “Nonlinear Optics: Femtosecond lasers and nonlinear optics: New approaches solve old problems in ophthalmology,” [Online]. Available: https://www.laserfocusworld.com/articles/print/volume-54/issue-03/features/nonlinear-optics-femtosecond-lasers-and-nonlinear-optics-new-approaches-solve-old-problems-in-ophthalmology.html.