Julius Vengelis • Audrius Dubietis

LASER PHYSICS

Lecture notes

VILNIUS UNIVERSITY PRESS

Faculty of Physics

Julius Vengelis • Audrius Dubietis

LASER PHYSICS

Lecture notes

VILNIUS UNIVERSITY PRESS
2023

Reviewed by:
assoc. prof. dr. Domas Paipulas (Vilnius University)

Bibliographic information is available on the Lithuanian Integral Library Information System (LIBIS) portal ibiblioteka.lt

ISBN 978-609-07-0800-2 (Digital PDF)

Foreword

Lasers and laser-related technologies are an integral part of modern science and are widely used in many areas of everyday life. The number of Nobel Prizes awarded to relevant discoveries related to lasers or using lasers as a tool in the past 60 years is extraordinary compared to any other technology or device. This proves that lasers are unique, incredibly versatile (in terms of generated pulse durations, wavelength, power and other parameters) devices pushing frontiers in many fields of science ranging from investigations of ultrafast phenomena on a femtosecond time scale to material science, nonlinear optics, frequency metrology, laser cooling, medicine, etc. Apart from shedding light on fundamental scientific problems, lasers have also found their place in our everyday life. Laser pointers, bar code readers, information processing devices (DVDs, Blu-Ray devices), laser range finders, laser surgery tools, laser material processing systems (laser engraving, cutting, welding and other devices) and many more. Therefore it is important to understand principles of their operation and key technological developments that enable such broad field of applications.

The guiding theme of this textbook is principles of laser operation with the intent for the reader to arrive at basic understanding how the laser works, what are its key components and what are the technological innovations that allow to achieve unique characteristics of laser radiation (e.g., generate and amplify the shortest pulses, convert laser wavelength by nonlinear optical methods). The readers will also be able to understand why laser radiation is unique compared to other natural or artificial light sources, e.g., the Sun, light bulbs, LEDs, etc. Chapter 1 provides a brief historic retrospect of laser development, introduces the key principles of laser operation, outlines the basic characteristics of laser radiation, their measurement tools and laser safety. The relevant properties of photons, light waves, laser beams and laser pulses, as well as coherence properties of laser radiation are discussed in Chapter 2. Chapter 3 is devoted to optical resonators: resonator algebra, stability conditions, resonator modes, sources of resonator energy losses and their importance. Chapter 4 addresses the fundamental light-matter interaction processes that govern laser operation, laser gain, amplification bandwidth, laser energy level systems and pump sources. Chapter 5 introduces the regimes of pulsed laser operation: free running, Q-switching and mode-locking along with methods of their practical implementation Chapter 6 focuses on the generation and amplification of femtosecond laser pulses, their measurement techniques. Chapter 7 gives a brief overview of solid-state, gas and semiconductor lasers, while detailed descriptions of miscellaneous lasers and their practical applications are intentionally left for students' seminar

presentation. Finally, Chapter 8 introduces the principles of laser wavelength conversion via nonlinear optical effects: second harmonic, sum frequency generation and optical parametric amplification. Simply put, these chapters are intended for undergraduate students gaining a deeper understanding of laser physics. The short english-lithuanian glossary at the very end is added to help native lithuanian students for correct usage of laser physics terms.

Contents

This page is intentionally left blank

1 Principles of lasers in a nutshell

In this chapter readers will briefly get acquainted with key principles of lasers: what is a laser, how it works, what is so special about light coming out of a laser and what quantities are used to characterize laser radiation. In addition, a brief historic overview and a short note about laser safety is also presented.

1.1 What is a Laser?

The first acquaintance that most of us had with lasers was seeing how different is light from the laser and other light sources. Let us compare light patterns coming from an incandescent light bulb, light emitting diode (LED) flashlight and a green laser pointer (Fig.1).

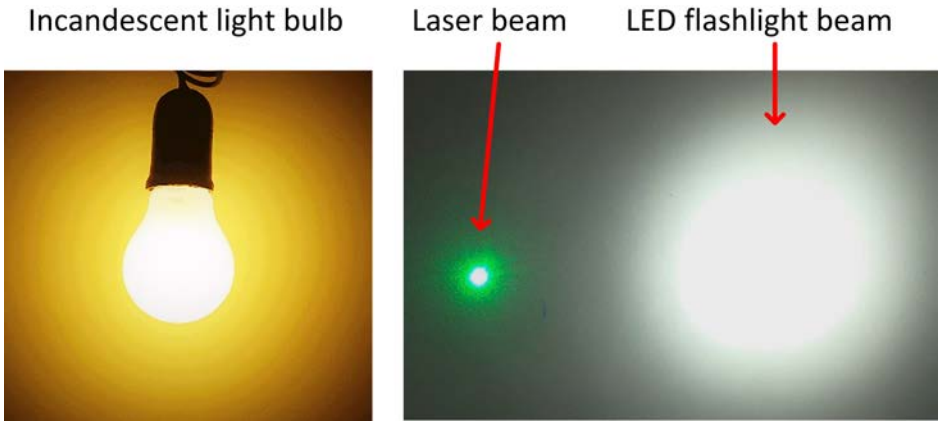


Figure 1: Comparison of light patterns coming from three different light sources. The left image is adapted from Wikimedia under GFDL 1.2 license.

Stark differences between light coming from a laser and the other two sources are immediately visible allowing to recognize distinctive features of laser radiation:

- **Laser light has low divergence (is highly directional)**: laser beam is almost aligned to the direction of propagation, with little angular spread from the point of origin. In fact, if the authors had chosen a He-Ne laser instead of a cheap laser pointer, the laser beam in the photo would have been even smaller, i.e., with even lower divergence. For He-Ne laser typical full angle beam divergence can be as low as 1 mrad.
- **Laser light is nearly monochromatic**: in general, laser emits light with very narrow spectrum. When it passes through a dispersive prism,

we see negligible dispersion since its spectrum consists of essentially a single color (see Fig.2). Note that some LEDs have certain colors which can even be defined by a specific wavelength (e.g., 700 nm " red" LED), however, such specification states only the central wavelength with spectral bandwidth being several tens of nanometers, whereas lasers are able to emit light such that all the emitted power is concentrated within extremely narrow spectral range.

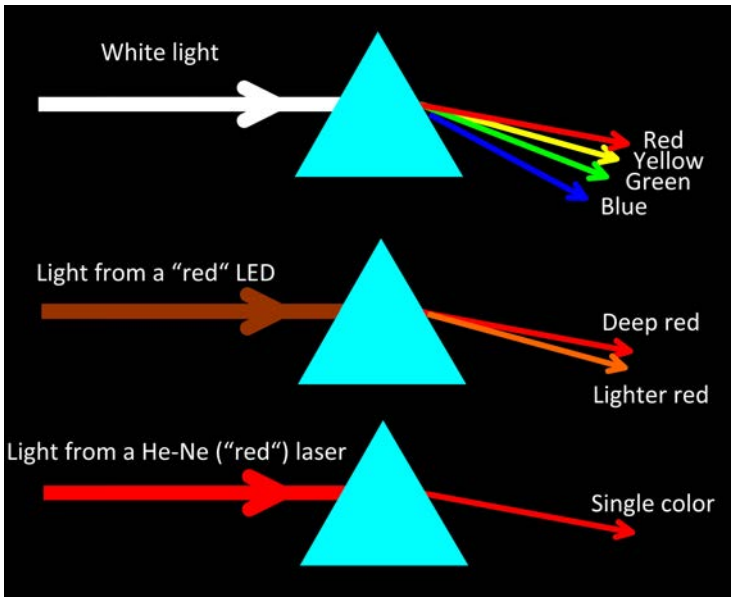


Figure 2: The concept of monochromaticity can be understood using a dispersive prism. White light consists of many different wavelengths which can be distinguished when light passes through a dispersive prism. Light coming from a LED has a relatively narrow spectrum, so dispersion effect after passing the prism is small. Laser radiation has very narrow spectrum and dispersion after passing the prism is negligible.

- **Laser light is coherent**: emitted waves are in phase in time and in space. In practice we can check this by observing interference when passing a He-Ne laser beam through an interferometer (e.g., Michelson interferometer) and viewing the interference pattern on a screen (see Fig.3). High contrast fringes are clearly visible and changing length of one interferemeter arm causes the fringes to move. Temporal coherence is defined by the monochromaticity of light, whereas spatial coherence is defined by low divergence of the light beam.

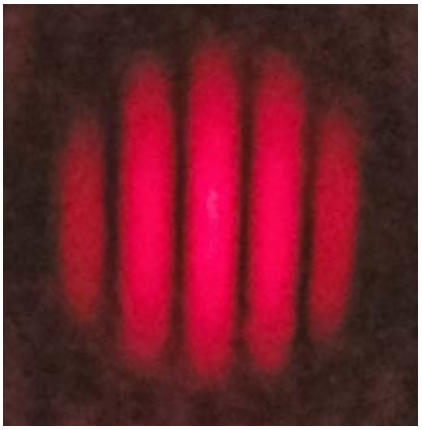


Figure 3: Interference pattern of a collimated He-Ne laser beam after passing Michelson interferometer.

These properties make laser a unique light source, which offers a wealth of specific applications not possible with any other light source. To mention a few: laser pointers make use of low divergence of laser light; laser-induced fluorescence utilizes the fact that laser radiation is monochromatic, so we can efficiently excite fluorescence of only the desired molecules; interferometry, holography and optical sensors rely on coherence of laser light.

Summarizing the above, a following definition of a laser could be given: ***Laser*** *is a light source producing low divergence, coherent and monochromatic light in the optical region of the electromagnetic spectrum.*

Restriction to the optical range of electromagnetic spectrum, which includes ultraviolet (UV), visible (VIS) and infrared (IR) radiation is very convenient and necessary, since lasers that emit light in the microwave region due to historic reasons are called – **masers**. However, the remaining part of the definition is not very accurate since a certain degree of monochromaticity, relatively low divergence and some degree of coherence can be achieved with other light sources. Mercury vapor lamps emit a relatively narrow band of spectrum which can be described as nearly monochromatic. Specific optical systems allow sufficient collimation of light from non-laser sources effectively reducing beam divergence to a small value. Superluminescent diodes (SLD) are devices that have some properties of LEDs and some properties of laser diodes including a certain degree of coherence; such sources of light are used in low coherence interferometry. Another example is Young's interference experiment where he used one slit to make a point light source from a burning candle light and another set of two slits to make two point sources. Light coming out of the two slits was coherent enough to observe interference when

the beams were overlapped on the screen. Although none of the presented examples can match low divergence, monochromaticity and coherence of lasers, they prove that the initial definition of a laser is rather quantitative, depending on quantitative description of what we call low divergence, coherence and monochromaticity.

A truly accurate description of what is a laser comes from the **principle of its operation**, i.e. how light with such properties is generated:

> ***Laser** is a **quantum light generator** emitting light through a process of optical amplification **based on stimulated emission of radiation**.*

The definition above can be explained as follows. "Quantum light generator" means that laser is a device converting one form of energy to another (the definition of term "generator") via certain quantum process – optical amplification based on stimulated emission of radiation. In the next section of this chapter we will discuss the basic principles of this quantum process. The rigor and accuracy of this definition is that it is tied to qualitative definition of principle of operation rather than a certain quantitative measure. This is important since different types of lasers can generate light with specific properties, e.g. femtosecond laser radiation has a relatively broad spectral bandwidth (several tens of nanometers or even more), whereas a stack of high power laser diodes can emit light with considerably high divergence, nonetheless in both cases the final definition of a laser is correct.

1.2 Historical perspective

The fundamental background behind laser generation (optical amplification) theory was laid in 1916, when A. Einstein theoretically predicted the existence of a novel emission phenomenon – **stimulated (induced) emission** of radiation, which is caused by a presence of external radiation field[1,2]. This discovery had utmost importance since stimulated emission of radiation lies on the physical basis of laser generation process. The phenomenon of stimulated emission was first experimentally observed by R. Landenburg in 1928. It was called "negative absorption" to reflect the fact that light and matter interaction is stimulated by external radiation field, but occurs in the opposite way to absorbtion. At that time stimulated emission was thought to have little practical use since the energy levels of an atom are populated accord-

[1] J. Hecht, Short history of laser development, Applied Optics **49**, F100–F121 (2010)

[2] W. Zinth, A. Laubereau, W. Kaiser, The long journey to the laser and its rapid development after 1960, European Physical Journal H **36**, 153–181 (2011)

ing to Boltzman law and absorption is the dominating effect. The idea of optical amplification using stimulated emission in gases was first formulated by Russian physicist V. Fabricant, but it still was not clear how to realize this idea in practice. After World War II, W. Lamb and R. C. Retherford noticed that population inversion can be achieved during Nuclear Magnetic Resonance (NMR) and later stimulated emission of radio waves was observed. Population inversion at that time was interpreted as "negative absolute temperature" as would follow from the Boltzmann distribution. In 1951 C. H. Townes conceived an idea that electromagnetic radiation could be **amplified via stimulated emission in a resonator**, and in 1954 Townes and J. P. Gordon invented the first device operating on that principle, the **maser**, which emitted electromagnetic radiation in the microwave (millimeter wave) range. The term "maser" is an acronym for Microwave Amplifier based on Stimulated Emission of Radiation. In 1964 A. Penzias and R. Wilson using maser as a sensitive microwave detector, discovered the cosmic microwave background radiation, which is the relic radiation from the Big Bang. For this fundamental discovery they received the Nobel Prize in Physics in 1978.

In 1958 C. H. Townes and A. L. Schawlow theoretically predicted the possibility to amplify electromagnetic radiation in the optical range and put forward the key principles of laser generation[3]. They also foresaw that a Fabry-Perot etalon could be used as a simple resonator in the optical range. What was still missing, was to find suitable materials and methods to create population inversion, which became a subject of very intense research by many groups at the time. At present, it is widely recognized that **the first laser was invented in 1960** by T. Maiman, when he reported the laser action in a ruby crystal (where natural chromium impurities exist) with aluminum coated ends serving as a resonator mirrors. To achieve the population inversion, Maiman used a white light flashlamp as a pump source. The word "recognized" is used here by purpose, since Maiman described his results very vaguely and it might be that he only observed the stimulated emission below the laser generation threshold. These doubts are also based on the fact that *Physical Review Letters* rejected Maiman's manuscript, but he managed to publish a short message in *Nature*[4]. It can be claimed with a confidence that the first laser, where light amplification and directional radiation in the form of a coherent beam, was reported a few months later by R. J. Collins, D. F. Nelson, A. L. Schawlow, W. Bond, C. G. B. Garrett and W. Kaiser[5]. At

[3] A. L. Schawlow and C. H. Townes, Infrared and optical masers, Physical Review **112**, 1940 (1958).

[4] T. H. Maiman, Stimulated optical radiation in ruby, Nature **187**, 493-494 (1960).

[5] R. J. Collins, et al., Coherence, narrowing, directionality, and relaxation oscillations in the light emission from ruby, Physical Review Letters **5**, 303 (1960).

the end of 1960 A. Javan, W. R. Bennett and D. R. Herriott demonstrated the first gas laser, where they used helium and neon gas mixture which was pumped by an electric discharge[6]. Interestingly, the first He-Ne laser emitted radiation at 1.15 μm wavelength and the well-known He-Ne laser emitting 632.8 nm radiation was demonstrated several years later. Even today He-Ne lasers are widely used in many laboratories.

Soon after invention of the first laser, a variety of lasers, which used solid-state materials doped with rare-earth metal ions as active media, were demonstrated. However, most of these first solid-state lasers operated only at cryogenic temperatures. In 1961 the first Nd:glass laser was demonstrated, which later became one of the most important solid-state lasers. Subsequent observations of lasing in various active media led to a laser boom. In 1962 the first semiconductor laser was invented. The semiconductor laser operated only at cryogenic temperatures due to contemporary state of the art of semiconductor technologies. In the same year, the first chemical laser was invented, which afterwards served as the basis of laser weapons. In 1964 the first CO_2 laser was demonstrated and soon became one of the most important technological lasers. At the same year, the first fiber laser was demonstrated, where the active medium was an optical fiber with a rare-earth metal doped core[7]. In 1966 organic dye lasers were invented, serving as the main ultrashort pulse laser sources for a long time until the invention of femtosecond Ti:sapphire laser in 1991. In 1970 efficient excimer lasers were demonstrated. The first pulsed lasers were able to generate light pulses with durations ranging from a few microseconds to a few nanoseconds. The first picosecond laser pulses were generated soon after, when passive mode-locking was discovered in 1967, while the first femtosecond pulses were produced in 1974. Although the first lasers were pumped by flashlamps or electrical discharges, the idea to use other lasers as pump sources was proposed in 1963, when it was noticed that radiation from GaAs laser diode is efficiently absorbed by Nd-doped solid-state media. However, this idea became widely employed in practice only after semiconductor laser technology was sufficiently developed, giving rise to the whole generation of the so-called diode-pumped solid-state (DPSS) lasers.

The first lasers were called "optical masers" and the term "laser" became widespread later. Both masers and lasers are based on the same principle, but the acronym "laser" stands for "Light Amplification by Stimulated Emis-

[6] A. Javan, W. R. Bennett, Jr., and D. R. Herriott, Population inversion and continuous optical maser oscillation in a gas discharge containing a He-Ne mixture, Physical Review Letters. **6**, 106 (1961).

[7] C. J. Koester and E. Snitzer, Amplification in a fiber laser, Applied Optics **3**, 1182–1186 (1964)

sion of Radiation", highlighting the fact that, unlike masers which amplify microwaves, lasers amplify the electromagnetic radiation in the optical range. Interestingly, the term "laser" was first used in 1959 in a conference paper[8] by a graduate student G. Gould, who at that time worked on a doctoral thesis about the energy levels of excited atoms of thallium. Two years before, he coined some general ideas about how lasers could be built and where they could be used in a lab notebook which he notarized. In 1959 he filed a patent application with the high-tech company he worked in. The U. S. Patent Office denied his application and granted it to Bell Labs, which resulted in a twenty eight year-long "patent war". Eventually, in 1987 G. Gould won the first patent lawsuit and was issued patents for optically pumped and gas-discharged lasers. However, the historical question about assigning the credits for laser invention is difficult since patents tell little about how the ideas arose and spread among researchers. C. H. Townes, N. Basov and A. Prochorov were awarded the Nobel Prize in Physics in 1964 for their input which led to invention of the laser. Another key researcher in the field, A. Schawlow, was awarded the Nobel Prize in Physics in 1981 for laser applications in spectroscopy. The beginning of the laser era in 1960 gave birth to a new branch of optics – nonlinear optics. Since the invention of the laser, a huge variety of new laser materials was developed, the power of lasers increased enormously and the duration of emitted pulses decreased significantly, yielding a coherent light with unprecedented intensity. The very first pulsed lasers generated light pulses with a duration of a few microseconds, while modern ultrashort pulse lasers can provide light pulses with a few femtosecond duration – a billion of times shorter! The pulse energy and power grew tremendously, and at present lasers are reliable turn-key devices, without which modern physics, biology, chemistry, medicine and materials science laboratories and future progress of these sciences is unimaginable. Interestingly, the laser was not invented for a specific purpose, the skeptics even claimed that invention of the laser was "not a solution of a specific problem, but, on the contrary, a solution that created many problems". Despite this, lasers gradually and imperceptibly became an integral part of many scientific, industrial, technological, military and many other fields, so it can be concluded that laser is the most important optical device invented in the last six decades.

[8]G. Gould, The LASER, Light Amplification by Stimulated Emission of Radiation, in *The Ann Arbor Conference on Optical Pumping* (University of Michigan, USA, June 15-18 1959)

1.3 Light Amplification by Stimulated Emission of Radiation

The term "**laser**" is an acronym of a phrase "**L**ight **A**mplification by **S**timulated **E**mission of **R**adiation". The same is true for "**maser**", only the first word in the phrase is "**M**icrowave" instead of "**L**ight". We will now discuss key principles of laser operation and stimulated emission phenomenon.

To understand key principles of laser generation, we must first get acquainted with principal components of a laser:

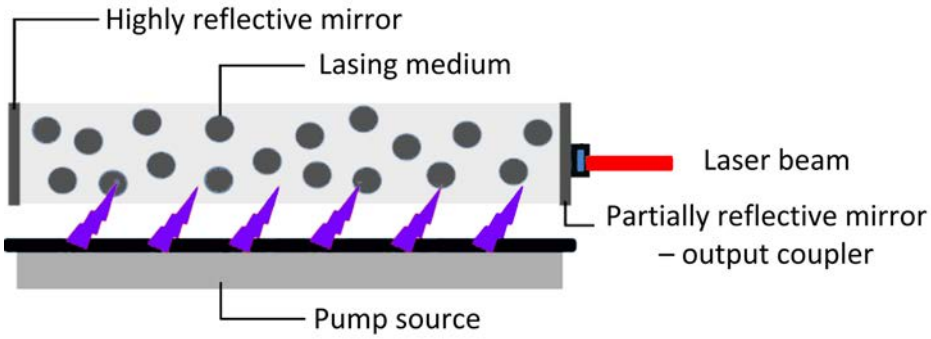


Figure 4: Principal components of a typical laser.

1. **Lasing medium (gain medium)**: a medium which is excited and emits laser radiation via stimulated emission process. It is also called **active medium**. It can be a solid-state medium (crystal, glass), semiconductor, liquid (organic solvents or dyes) or gas medium (He-Ne, CO_2). Lasers are usually named after the lasing material, e.g., Nd:YAG laser means that the lasing material is YAG (yttrium aluminum garnet) crystal doped with Nd^{3+} ions, rhodamine 6G laser means that the lasing material is a rhodamine 6G dye, He-Ne laser means that the active material is a mixture of He and Ne gas (Fig.5).

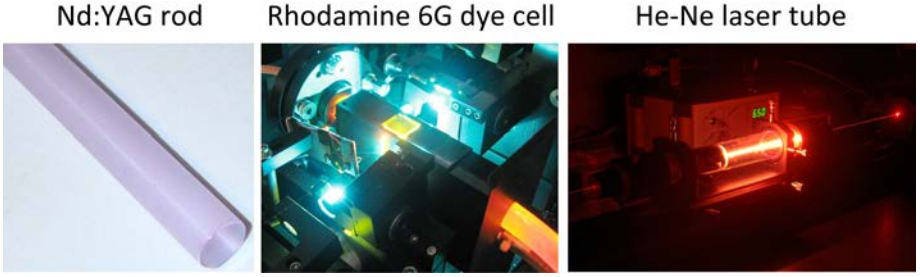


Figure 5: Examples of different lasing materials. Images adapted from Wikimedia under CC BY 1.0, CC BY 2.5 and CC BY 3.0 licences respectively.

2. **Pump source**: source of energy for excitation of the active medium and subsequent laser radiation emission. In other words, energy of the pump source is converted into laser radiation. Pump sources are classified into **optical** (e.g., flashlamp, laser diode), **electrical** (e.g., electric current or electric discharge) and **chemical** (exothermic chemical reaction).

3. **Optical resonator (optical cavity)**: a part of the laser that provides **positive feedback**: returns some part of emitted radiation back to the active medium. Optical resonator consists of at least two mirrors between which the lasing medium is placed (Fig.4). These mirrors reflect laser radiation back to the active medium, where it is amplified. One of the mirrors is made partially reflective, allowing a fraction of light to leave the cavity producing the laser output beam. This mirror is called the **output coupler**.

Another important aspect behind laser operation principles is understanding what happens in the gain medium when it is excited. The simplest example is a two-level atom energy system depicted in Fig.6.

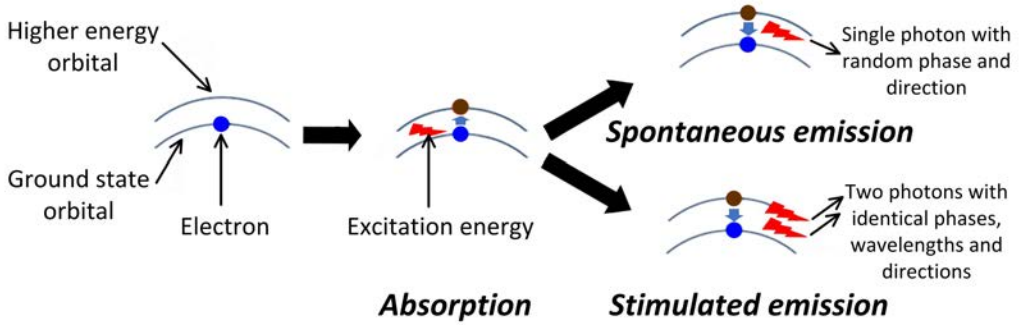


Figure 6: Possible ways of light and atom interaction.

An atom becomes excited when it absorbs a certain portion of energy. An electron from the ground state orbital jumps to a higher energy orbital absorbing amount of energy equal to the difference between the ground and higher state orbital energies. This energy is called excitation energy and the light-matter interaction process is called **absorption**. Another light matter interaction process takes place when the electron spontaneously returns from the higher energy orbital to the ground state orbital, emitting a photon with energy equal to the energy difference between higher and ground state orbitals. Such process is called **spontaneous emission**. The term "spontaneous" is used to emphasize the fact that it is not triggered by any external factor

and the phase, direction and polarization of the emitted light wave are **random**. Apart from the two aforementioned processes, yet another interaction process is also possible. This process is called **stimulated emission**. In this case when an electron is in the higher energy orbital (the atom is excited), another photon can trigger the electron to jump to the ground orbital state emitting a photon. The stimulated emission process occurs only when the triggering photon energy is equal to the energy difference between higher energy state and ground state orbitals. More importantly, the emitted and corresponding light wave has **identical** phase, direction, polarization and wavelength to the initial triggering photon.

Laser operation is based on light amplification by stimulated emission of radiation. Let us discuss this in more detail.

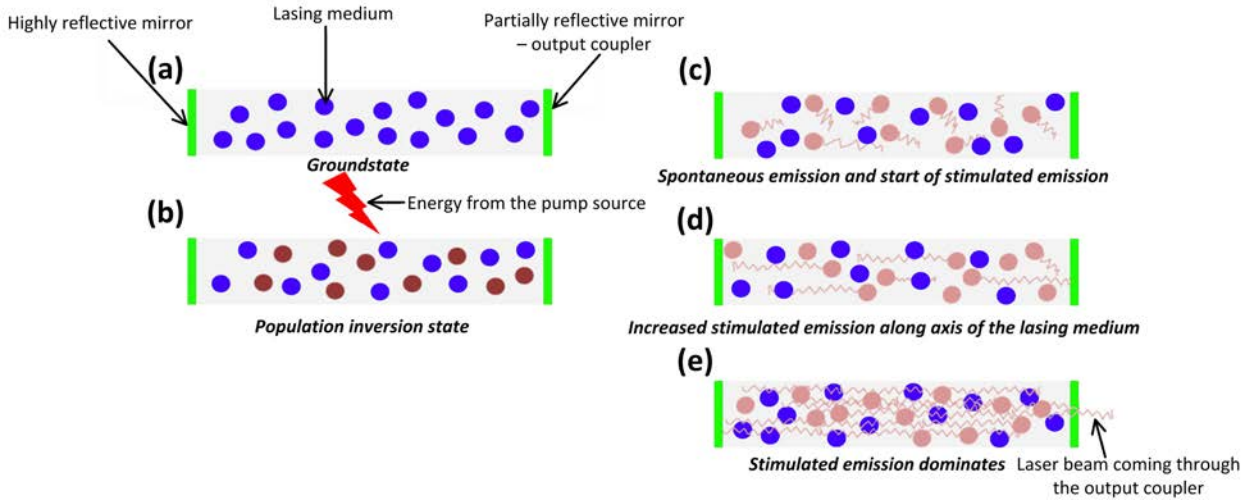


Figure 7: Principal view of stimulated emission of radiation and lasing.

Consider that the lasing medium is at ground state (Fig.7a). When a pump source (e.g., an electric flash) is turned on, some energy is transferred to the lasing medium and some atoms become excited (denoted by red circles in Fig.7b). The first key moment is spontaneous emission that starts from the excited atoms and the emitted photons can then trigger stimulated emission of other excited state atoms (Fig.7c). The second important moment is that the amount of excited electrons (population of an excited energy level) must be greater than the amount of electrons in the lower (ground state) energy level. Such state is called **population inversion**. The term "inversion" indicates that this kind of population is impossible to achieve under usual conditions: according to Boltzman distribution, the population of higher energy levels decreases exponentially. When the emitted light (both

due to spontaneous and stimulated emission) reaches resonator mirrors, it is reflected back to the lasing medium. This increases stimulated emission (Fig.7d) since after each reflection more and more light in the lasing medium has the same wavelength, phase, direction and polarization. The stimulated emission starts dominating (Fig.7e) and the number of photons with the same wavelength, phase, direction and polarization rapidly grows with each round-trip. After certain amount of round-trips, light amplification coefficient (power increase per unit length of the lasing medium) **exactly compensates losses** in the resonator. This equilibrium state is called **laser generation threshold**. When it is exceeded, laser generation starts. Since one of the resonator mirrors is partially reflective (output coupler), some light always leaves the resonator in the form of a **laser beam** and can be used for intended purposes.

The population inversion is necessary in order to achieve amplification of light (lasing) due to the following reasons. Atoms are complex systems with multiple energy levels, so photons that are emitted during spontaneous and stimulated emission can be absorbed by other atoms in the groundstate – this is called reabsorption. If population inversion condition is not achieved, this is exactly what happens. In the opposite case (population inversion), more photons are emitted than reabsorbed and the net result is amplification of light by stimulated emission or, in other words, **lasing**.

It is important to note that here we described laser operation when its dynamics (how electron population changes) are not interfered with by any external means. This is the free-running regime of a laser. Practically used lasers usually have components in the resonator which modulate lasing dynamics, e.g., enable generation of ultrashort pulses. This will be discussed in detail later.

1.4 Quantifying laser radiation

There is a great variety of lasers in terms of their emitted radiation characteristics which also determine potential applications of a particular laser. We will now discuss how to characterize laser radiation in practice. For more clarity, it helps to distinguish between the relevant parameters of lasers that emit light continuously (**continuous wave regime – CW**) and lasers that emit light in the form of short bursts that are called laser pulses (**pulsed regime**) as shown in Fig.8.

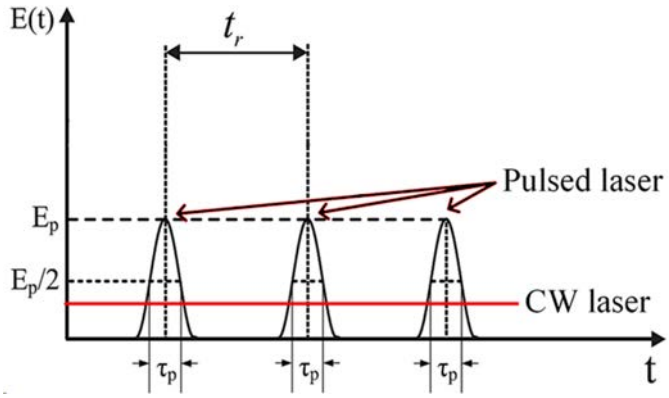


Figure 8: Time evolution of light electric field emitted from pulsed (black lines) and continous wave (CW) lasers (red line). τ_p is the pulse duration, t_r is the pulse repetition period (inversely proportional to pulse repetition rate).

For lasers that operate in CW regime we can define these main parameters:

- **Wavelength**: emission wavelengths of modern lasers covers an extremely broad spectral range from the vacuum UV to far IR (Fig.9), e.g., F_2 excimer laser emits radiation of 157 nm wavelength, while AlGaAs quantum cascade laser emits radiation of up to 250 μm wavelength[9]. The diversity in emission wavelengths in fact defines an extremely broad range of laser applications. There are two important features to note regarding laser wavelength. First, **the same gain medium can emit several different wavelengths**, e.g., He-Ne laser has these lasing wavelengths: 543.36 nm, 593.93 nm, 604.61 nm, 611.8 nm, 632.8 nm, 1.15 μm; 3.39 μm. This is related to the fact that atoms/molecules have many energy levels or their combinations that are suitable for lasing. The particular lasing wavelength is selected by choosing an appropriate resonator mirrors which reflect the desired wavelength. Second, **wavelength of several lasers, e.g. Ti:sapphire, Fe:ZnSe, Fe:ZnS, can be continuously tuned over a certain wavelength range**. However, these are rare cases related to specific features of the gain medium.

[9]C. Walther, M. Fischer, G. Scalari, R. Terazzi, N. Hoyler, and J. Faist,"Quantum cascade lasers operating from 1.2 to 1.6 THz", Applied Physics Letters **91**, 131122 (2007)

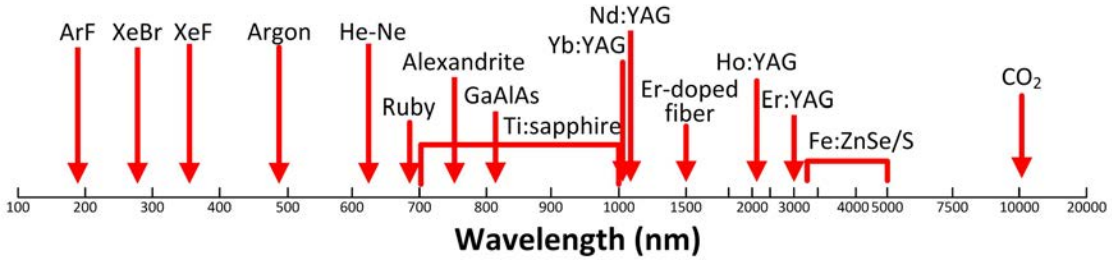


Figure 9: Emitted radiation wavelengths of various lasers. Lasers with continuously tunable radiation wavelength are denoted by square bracket marking.

Wavelength of light (including laser radiation) is measured using a device called **optical spectrometer** (usually simply called "spectrometer"). This instrument separates different wavelengths by refraction in a prism or diffraction by a diffraction grating and measures their intensity. The measured light intensity as a function of wavelength (or frequency) is called the **optical spectrum**. Image of a spectrometer and example of He-Ne laser radiation spectrum are depicted in Fig.10.

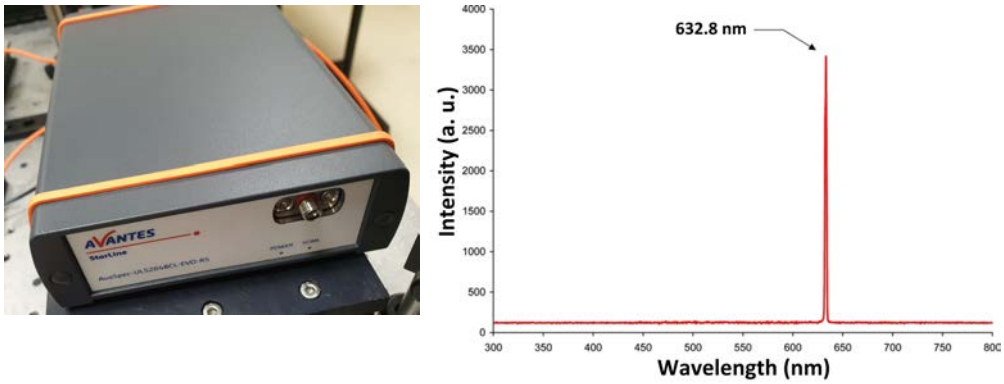


Figure 10: Optical spectrometer (left) and example of measured He-Ne laser spectrum (right).

- **Power**, that is an amount of energy per unit time. It is important to note that when characterizing pulsed laser radiation the term "power" is usually used to quantify the **average power** of laser radiation. Various lasers have average powers ranging from milliwatts to megawatts, e.g., solid-state microlasers produce output power of several tens of milliwatts while chemical lasers typically produce output power of hundreds of kilowatts and the most powerful versions can exceed megawatt power level. The term, **peak power**, is also used to characterize pulsed laser radiation and will be discussed below with other parameters of

pulsed laser radiation.

The power of laser radiation is measured using a **power meter**. The main types of these devices are: photodiode-based, thermal and pyroelectric. In **photodiode-based power meters** the electric current signal proportional to the power of laser radiation is generated in a photodiode. Such detectors are fast, sensitive and have linear response function over a broad range of powers: from fraction of nanowatts to several milliwatts. On the other hand, they can be easily damaged if too much power is applied and their response function differs significantly for different wavelengths of light. **Thermal power meters** are based on the heating effect of laser radiation: a temperature gradient between laser radiation absorbing area and a peripheral reference area (where heat is dissipated) is measured using thermopiles. Thermal power meters can measure tens of watts of power, with water-cooled versions being able to measure even greater powers. They are also durable, accurate and their response is almost constant for different wavelengths of light. However, such power meters are quite slow (response time is several seconds) and require thermal equilibrium with the environment. An example of a thermal power meter consisting of a sensor and measuring device is shown in Fig.11.

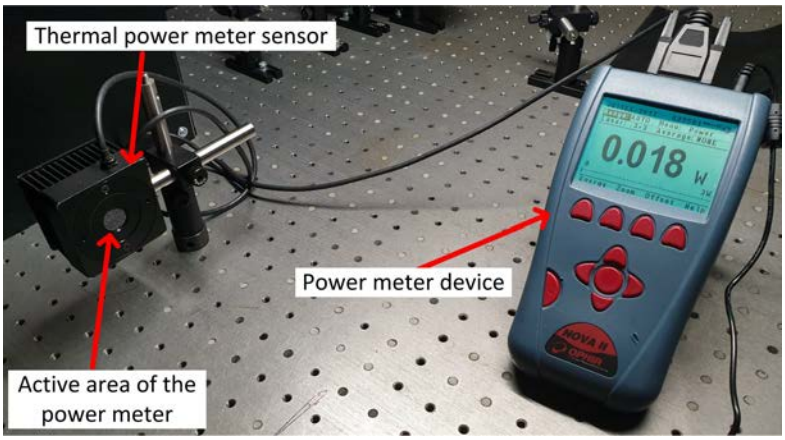


Figure 11: Thermal power meter. It consists of two connected parts: power sensor and power meter device which displays the measured power and is used to control the power meter.

Pyroelectric power meters, as the name suggests, are based on the pyroelectric effect: temperature gradient induced by the incident laser radiation in the pyroelectric material generates current signal which is converted into a voltage signal that is proportional to the power of

laser radiation. Such power meters are very fast, broadband (response is independent of wavelength in a broad spectral range, usually from 200 nm to 20000 nm) and do not require thermal equilibrium with the environment. However, pyroelectric power meters are fragile, less accurate, cannot measure CW laser radiation power and are sensitive to vibrations. Therefore, they are usually used to measure low repetition rate laser power for each pulse.

- **Beam size** is usually measured as a diameter of the laser beam at a certain power/intensity level (Fig.12).

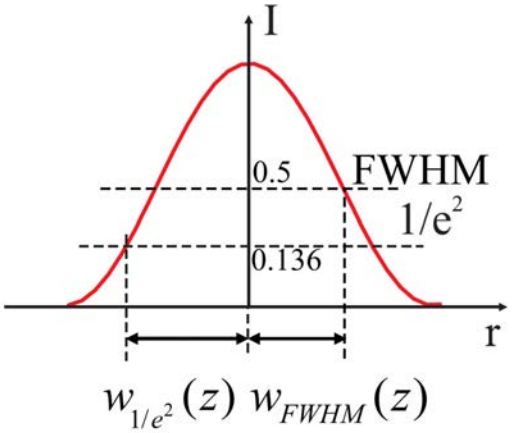


Figure 12: Laser beam size measured at different intensity levels: w_{1/e^2} – radius at 13.6% of maximum intensity level; w_{FWHM} – radius measured as half of beam diameter at full width at half maximum intensity level.

The beam diameter can differ in the horizontal (x) and vertical (y) directions; such beams are called elliptical since the beam spot on a screen looks like an ellipse instead of a circle. Commonly used intensity level values: FWHM – full width at half maximum, $1/e^2$ – $1/e^2$ (13.6%) of maximum intensity level, $1/e$ – $1/e$ (36.8%) of maximum electric field amplitude level. There is also 4σ notation defined as 4 times the standard deviation of the horizontal or vertical marginal distributions respectively. It is the ISO international standard definition for beam width. Despite the fact that 4σ notation is the international ISO standard for beam width definition, other beam width definitions are also commonly used. Beam size can be measured in several different ways but the simplest one is to use a charge-coupled device (CCD) camera which displays laser beam intensity distribution and estimates its size (Fig.13). Professional beam width measurement software usually provides beam width values at various intensity levels. Although

CCD camera is a simple and straightforward way to measure beam profile and size, great care has to be taken not to damage it. Therefore neutral filters that attenuate laser radiation are usually placed in front of the CCD sensor.

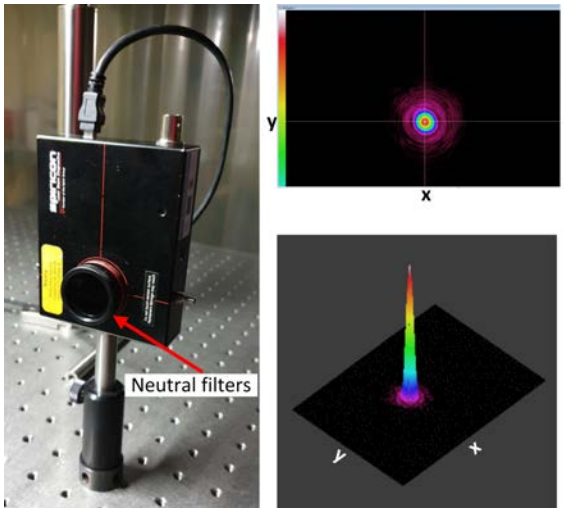


Figure 13: CCD camera (left) and example of measured laser beam intensity profile in 2D (top right) and 3D (bottom right). Color coding indicates the intensity.

Laser beams usually have intensity distribution described by a Gaussian function and are therefore called **Gaussian beams**. Properties of Gaussian beams will be discussed in more detail in Chapter 2. Another important aspect to note is that laser beam size tends to change with propagation. If no additional optics is used, beams coming out of a laser either diverge (diameter increases) or converge (diameter decreases). If, using specific optical systems, size of the propagating beam is maintained constant in the near field, such laser beam is called **collimated**.

- **Intensity** is defined as power per unit area:

$$\boxed{I = \frac{Power}{Area}} \tag{1}$$

As with the term "power", when characterizing pulsed laser radiation it is important to clearly distinguish whether by using the term "intensity" we mean **average intensity** or **peak intensity** of laser radiation. Unlike "power", the term "intensity" in the context of pulsed lasers usually means peak intensity unless the text states otherwise.

In addition to the above parameters, characterization of pulsed laser operation involves a set of supplementary parameters:

- **Pulse duration** is usually defined as width of a pulse temporal shape at FWHM power/intensity level (see Fig.14 and Fig.19). A huge variety of pulsed lasers cover a broad range of possible pulse durations: lasers operating in free-running regime generate microsecond duration pulses, but lasers operating at advanced regimes generate nanosecond (1 ns=10^{-9} s), picosecond (1 ps=10^{-12} s) and even femtosecond (1 fs=10^{-15} s) pulses. Moreover, using nonlinear optical techniques and femtosecond laser systems, attosecond (1 as=10^{-18} s) duration pulses can be generated with current world record standing at 43 as![10] There is no universal method for laser pulse duration measurement: it essentially depends on the pulse duration itself. Long pulse (several nanoseconds or longer) duration can be measured simply by using a photodiode and observing the generated pulse temporal profile on the oscilloscope screen (Fig.15 left).

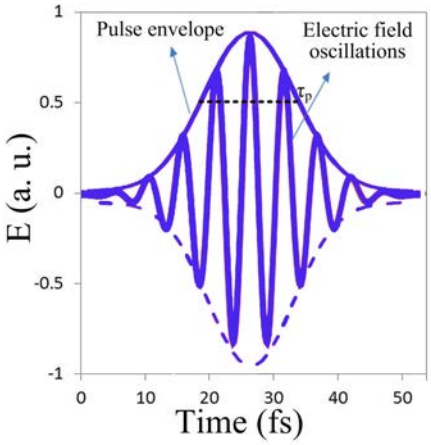


Figure 14: Electric field oscillations in an ultrashort light pulse and its duration estimation at FWHM level (τ_p). Thick blue line above the oscillations represents what is called the pulse envelope.

[10]T. Gaumnitz et al, Streaking of 43-attosecond soft-X-ray pulses generated by a passively CEP-stable mid-infrared driver, Optics Express **25**, 27506-27518 (2017).

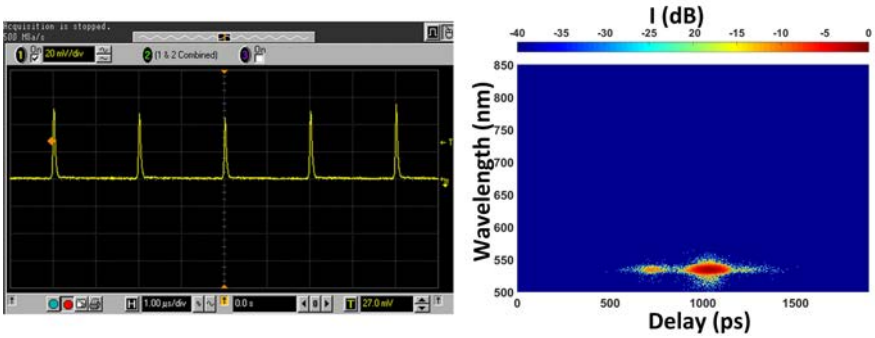


Figure 15: Signal in an oscilloscope screen corresponding to temporal profiles of laser pulses detected with a photodiode (left); example of spectrogram of a complex laser pulse measured with a streak camera (right).

For shorter duration laser pulses more sophisticated devices/techniques have to be used, **ultrafast photodiodes** are used for pulse durations down to 50 ps, they are expensive and require ultrafast oscilloscopes and special cables. **Streak camera** is used for pulse durations from nanoseconds to 500 fs, it is an extremely expensive device, though it is also able to measure pulse spectrum (Fig.15 right). **Correlation techniques** are used to measure pulse durations ranging from hundreds of picoseconds to 30 fs, these are relatively cheap and straightforward methods but cannot measure complex pulses. **Time-frequency techniques** comprise a broad range of very complex methods that enable characterization of pulse duration (to the shortest possible durations), shape, phase and more. In the context of laser pulse duration, there are several widely used definitions to classify the temporal width of laser pulse: **long pulses** are pulses whose duration is several tens of ns and longer; **short pulses** are- pulses with duration from tens of ns to 100 ps; **ultrashort pulses** are pulses which duration is less than 100 ps.

- **Pulse repetition rate** is the number of pulses emitted per second. It is the inverse of pulse repetition period t_r (Fig.8). Different lasers typically have pulse repetition rates from tens of Hz to 100 MHz, though in specific cases it can be from below 1 Hz to hundreds of GHz. Pulse repetition rate is usually measured using a photodiode and estimating the temporal interval between pulses (looking at the oscilloscope screen – Fig.15 left).

- **Pulse energy** is the ratio of average power (P) and pulse repetition rate (f_{rep}):

$$\boxed{E_p = \frac{P}{f_{rep}}} \tag{2}$$

- **Peak power** is the maximum power of a single pulse. It is defined as the ratio of pulse energy (E_p) and pulse duration (τ_p):

$$\boxed{P_p = \frac{E_p}{\tau_p}} \tag{3}$$

  Peak power of ultrashort laser pulses can vary in the range from MW to PW, however, we should have in mind that such extreme power is achieved only for a very short moment of time (duration of the pulse).

- **Central (carrier) frequency** is defined as a frequency of maximum spectral intensity (example is shown Fig.16). The shorter is the pulse, the broader is its spectrum. Ultrashort laser pulses have a very broad spectrum (unlike CW lasers whose spectrum is extremely narrow), thus the term central frequency is used. In the wavelength domain we can also define the **central (carrier) wavelength**.

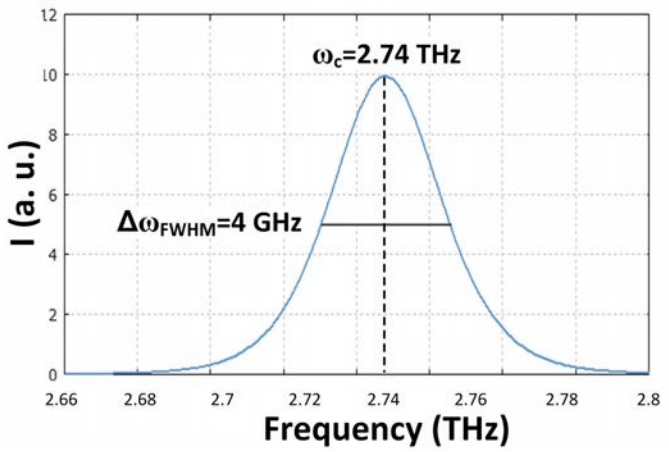


Figure 16: Spectrum of the ultrashort laser pulse and its main characteristics: central frequency (ω_c) and spectral width at FWHM level ($\Delta\omega_{FWHM}$).

Laser spectra usually have intensity distribution described by a Gaussian function and are therefore sometimes called **Gaussian spectra**. In this case the central frequency matches the spectrum peak. For

complex non-Gaussian spectra the central frequency is defined as:

$$\nu_{centroid} = \frac{c}{\lambda_{centroid}}, \tag{4}$$

where $\nu_{centroid}$ is center of mass of the spectrum in the frequency domain, called **spectral centroid**:

$$\nu_{centroid} = \frac{\sum_{n=0}^{N-1} \nu_n I_n}{\sum_{n=0}^{N-1} I_n}, \tag{5}$$

here ν_n is central frequency of a discrete bin, in this case – frequency step of the used spectrometer and I_n is the weighted coefficient in this case – measured intensity value of the frequency component.

- **Spectral width (bandwidth)** is the width of the pulse spectrum. As already mentioned, ultrashort pulses have a relatively broad spectrum, so their spectrum width is usually defined at FWHM level (Fig.16). The estimation in the case of Gaussian spectrum is simple as Gaussian spectrum is symmetrical (in the frequency domain) and smooth.

  Spectral bandwidth is inversely proportional to pulse duration; the shorter is the pulse, the broader is its spectrum and vice versa. Ultrashort pulse spectrum width, when its central wavelength is in the visible range, in some cases can exceed hundreds of nanometers. The spectrum of such pulse covers the entire visible range and the laser appears to be generating white light. Such lasers are called **white light lasers**.

- **Peak intensity** is the peak power per unit area:

$$\boxed{I = \frac{P_p}{Area}} \tag{6}$$

  Due to unique properties of laser radiation which enable to achieve extremely high peak power and focus the beam to a spot size that is close to wavelength, modern laser systems can achieve significantly greater peak intensities than any other light source.

1.5 Laser safety classes

Laser radiation is especially harmful to the eyes and skin. It can cause temporary or permanent retina and skin damage, which could severely disrupt eyesight and cause skin burns. In the case of repetitive exposure to laser radiation eye or skin illnesses can develop. Laser radiation effect to the eye or skin cells depends on laser source parameters:

- **Wavelength of radiation**. The effect of ultraviolet, visible and infrared radiation to biological systems is very different. Due to large photon energy, the UV radiation produces photo-ionization, while the IR radiation has thermal effect. Visible light has the largest penetration depth in living tissues, since their main component is water, which is fully transparent in the visible range, therefore visible radiation can affect deeper tissues.

- **Laser operation regime** (continuous wave or pulsed), which determines the exposure time. In the case of pulsed regime, the strength of effects is determined by the pulse repetition rate and duration. The peak intensity (the amount of energy per unit area per unit of time) of nanosecond, picosecond and femtosecond pulses is very different, therefore, physical mechanisms of their interaction with biological tissues are also significantly different.

- **The power or pulse energy** of a laser source, or in the most general case, the fluence (power or energy per unit area), which depends on the parameters of a laser beam and character of the beam propagation (focused, collimated or scattered beam).

Cell or tissue damage due to laser radiation occurs because of photochemical reactions, thermal effects and photo-ionization. When the laser fluence increases, these effects manifest themselves in the above order and the result is different for each of these effects. Low fluence (up to 1 W/cm^2) usually causes reversible photochemical reactions which have short-term effects and do not cause any irreversible changes. Laser radiation with 10-10^5 W/cm^2 fluence produces a strong thermal effect (especially during longer exposure times) due to which the biological tissue is heated, proteins in the cells break up, and the cell eventually dies. The fluence greater than 10^{10} W/cm^2 is achieved only with short pulse lasers. Short laser pulses produce strong ionizing effect since the electric field strength can exceed electron bond energy. In the wake of intense laser radiation microplasma is formed, which can cause ablation (evaporation) of the tissue's outer layer, dissociation of intermolecular bonds (optical damage), while plasma expansion generates acoustic waves, which affect and disrupt cells and tissues that even were not directly in the radiation exposure area.

When evaluating laser radiation danger, the laser sources are divided into 4 classes. Laser safety class is marked on the laser housing or on the device part where the laser is located.

- **Class 1**: The power density of these lasers is so small that their radiation is not dangerous to skin or to the eyes in the case of any use

of the laser. Sometimes two subclasses are distinguished: class 1M laser system is a class 1 laser using magnifying optics which is incapable of causing injury during normal operation unless collecting optics are used; class 1C laser (established in July 2015) is a class 1 laser that covers laser systems that are designed for direct contact with the "objective". This can be laser systems for hair removal, reduction of wrinkles, treatment of akne, tattoo removal, etc.

- **Class 2**: Only continuous wave (CW) lasers emitting visible (400 nm–700 nm) light whose power does not exceed 1 mW fall into this class. Direct exposure to such laser radiation is considered not dangerous if the exposure time does not exceed 0.25 s (the duration of eye blinking reflex). Sometimes a subclass is distinguished: class 2M laser system is a class 2 laser using magnifying optics. It includes visible wavelength lasers incapable of causing injury in 0.25 seconds unless collecting optics are used.

- **Class 3**: This class is also divided into 3R and 3B subclasses. 3R class includes lasers whose power in the visible range (400 nm–700 nm) does not exceed 5 mW. Direct exposure to such laser radiation is considered not dangerous, but focused radiation can cause temporary vision disorder. 3B class includes CW lasers emitting in the entire optical range (ultraviolet, visible and infrared), whose power does not exceed 0.5 W and pulsed lasers emitting in the visible range, whose pulse energy does not exceed 30 mJ. Direct exposure to such laser radiation is dangerous to skin and eyes, therefore, when working with such lasers it is necessary to wear laser safety goggles and avoid direct contact of laser radiation with skin.

- **Class 4**: It is the most dangerous class. All the lasers whose power or pulse energy is greater than 3B subclass lasers fall into this class. All ultrashort pulse (picosecond and femtosecond) lasers also fall into this class. Both direct and scattered (reflected from matted or diffuse surfaces e.g., sheet of paper) radiation as well as reflections from optical or mechanical element surfaces (lenses, filters, mounts, etc.) is dangerous to skin and eyes. The most dangerous in that regard are lasers emitting ultraviolet and infrared radiation since it is not directly visible to human eye. Maximum safety precautions are necessary when working with such lasers.

Summary of Chapter 1

- Laser operation is based on stimulated emission of radiation. The term 'laser' stands for Light Amplification by Stimulated Emission of Radiation.
- Laser radiation is unique compared to light coming from other sources: it has low divergence, is monochromatic and coherent.
- Laser radiation is characterized by a set of parameters: central wavelength, spectral width, pulse duration, beam width, intensity, peak and average power.
- Any laser consists of three principal components: active medium, optical resonator and pump source.
- Laser invention has a long and rich history and is a result of many competing scientific groups rather than a single person.
- Laser radiation is harmful to the eyes and skin. According to the level of danger, the laser sources are divided into 4 safety classes.

2 Light beams and pulses

Light can be described in two ways: either as electromagnetic waves or as particles (photons) (Fig.17). From the viewpoint of laser physics, these concepts are tightly related although they describe different aspects of light. The wave description enables accurate analysis of light beam and pulse propagation phenomena: beam diffraction, pulse dispersion, and wave interference. Equations describing wave optics are derived from Maxwell equations, which describe propagation of electromagnetic radiation in any medium.

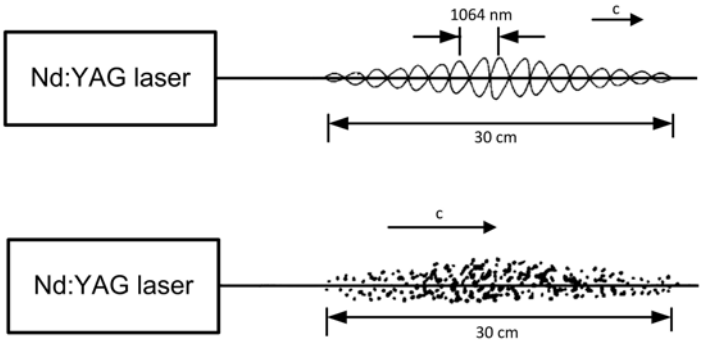


Figure 17: 1 ns duration 1064 nm wavelength Nd:YAG laser pulse pictured as waves (upper image) or particles (lower image).

In some cases wave optics can be simplified to geometrical optics, when light propagation is analyzed using only reflection and refraction laws. This assumption is very convenient when analyzing light propagation in optical resonators and determining their stability. Geometrical optics can be applied when diffraction and interference phenomena can be ignored e.g., when the size of resonator optical elements is considerably greater than wavelength of light.

However, light absorption and emission processes cannot be described by Maxwell equations. Quantum optics is based on the postulate that electromagnetic field and matter interaction is quantized: atoms and molecules absorb and emit only quantized energy. Therefore, the operation of laser medium (laser generation and amplification), laser-matter interaction, detection of radiation are described by means of quantum optics.

2.1 Photons and their properties

The smallest quantum of light is a **photon**. The energy of a photon is defined as:

$$\boxed{E = h\nu}, \tag{7}$$

where $h = 6.62 \times 10^{-34}$ J/s is the Planck's constant and ν is the frequency of the corresponding electromagnetic wave. The wavelength of light is defined as $\lambda = c/\nu$, where c is the speed of light in a vacuum. The photon propagates in a certain direction and the quantity $\vec{k} = 2\pi\vec{e}/\lambda$ is called the wavevector. Here $\vec{e}$ is the unity vector which characterizes the propagation direction. The momentum of a photon is defined as $\vec{p} = h\vec{k}/2\pi$ with and its modulus equals to $p = h/\lambda$.

Photon as any other quantum particle is a subject to the **Heisenberg uncertainty principle**, which states that relevant parameters, i.e. the position and momentum, can only be known (or determined) with a certain accuracy. In fact, the uncertainty principle poses certain limitations that are important in laser physics. Namely, light *beams* can be "diffraction-limited" due to the smallest possible photon position-momentum uncertainty, whereas light *pulses* can be "bandwidth-limited" due smallest possible time-energy uncertainty. We will now discuss this in detail.

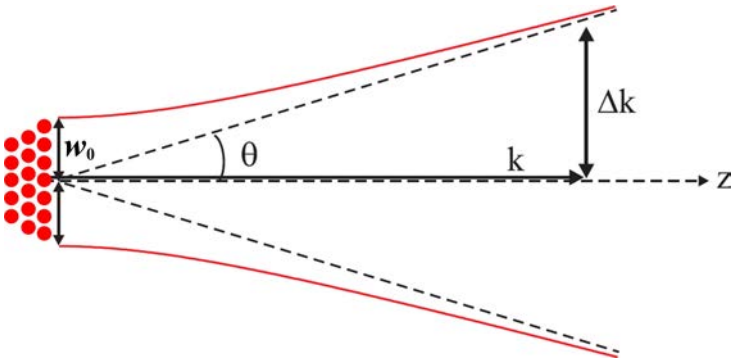


Figure 18: Diffraction of the light beam and relevant parameters. Red circles mark an idealized view of photons in the beam.

Consider a light beam propagating in the z direction in the coordinate system, as schematically illustrated in Fig. 18. It can be analyzed as a collection of many photons each with a certain lateral position and propagation direction which is related to momentum of the photon. The uncertainty principle states that the photon position (or its localization) in the plane perpendicular to the propagation direction and the propagation direction

in that plane have finite precision defined by the uncertainty relations for photon position and momentum

$$\Delta x \Delta p_x \geq \frac{h}{\pi}; \quad \Delta y \Delta p_y \geq \frac{h}{\pi}. \tag{8}$$

These expressions show that the photon propagation direction in one plane will be undefined as much as is defined its localization in the x-y plane (the plane perpendicular to propagation direction). For a collection of propagating photons (the light beam) with a radius $w_0 = \Delta x$ (Fig.18) the propagation direction uncertainty for photons in the beam is

$$\Delta \vec{p} \geq \frac{h \Delta \vec{k}}{2\pi}. \tag{9}$$

For wavevector uncertainty of $\Delta \vec{k} = \vec{k}\theta$, where θ is the beam divergence angle, the uncertainty relation for the beam radius and divergence angle reads as

$$\boxed{w_0 \theta \geq \frac{\lambda}{\pi}}. \tag{10}$$

This expression can be interpreted as follows: the smaller is the beam radius (w_0), the greater is its divergence angle (θ) and vice versa. For example, taking the laser wavelength of 500 nm and the beam radius of 100 μm then, according to expression in Eq.(10), the divergence angle of such a beam cannot be smaller than $\frac{\lambda}{\pi w_0} \approx 1.6$ mrad. This is a fundamental limitation following from the aforementioned quantum mechanical uncertainty principle.

Two boundary cases can be pointed out. Firstly, if we can accurately define photon propagation direction, then we cannot define where in the x-y plane the photon is localized. In this case the light is described as an infinite plane wave which does not diffract. Secondly, if we can accurately define localization of the photon, then its propagation direction is completely undefined. In this case the light is described as a spherical wave emitted from a point source. The real light beams represent an intermediate case between these two boundary cases. Eq.(10) describes the minimum beam divergence angle for a given beam radius, such a beam is called **diffraction-limited**. Beam with an ideal Gaussian intensity distribution (Fig. 22) are diffraction-limited beams.

The time-energy uncertainty is expressed as

$$\Delta E \Delta \tau \geq \frac{h}{2\pi} \tag{11}$$

and relates energy uncertainty ΔE with characteristic time $\Delta\tau$ at which changes occur in the given system, e.g. the photon is absorbed or emitted. Having in mind that $\Delta E = h\Delta\nu$, this expression is equivalent to photon time-frequency uncertainty relation

$$\boxed{\Delta\nu\Delta\tau \geq \frac{1}{2\pi}}. \tag{12}$$

In the case of emission, $\Delta\tau$ is directly related to the lifetime of an excited energy level of an atom or molecule: the shorter is the lifetime of the excited level, the greater is spread in the energy (frequency uncertainty $\Delta\nu$) of the emitted photons or, in other words, the greater is spectral broadening of the energy level. For example, energy levels that have long lifetime (the so-called metastable levels) represent an opposite case and emit radiation with a narrow spectrum. The above considerations may be applied to light pulses as well; then parameters in expression $\Delta\nu$ and $\Delta\tau$ (Eq.(12)) have meanings of the spectral bandwidth and pulse duration, respectively. Then the expression in Eq.(12) could be rewritten in a slightly different form:

$$\boxed{\Delta\nu\Delta\tau \geq C}, \tag{13}$$

where C is a certain constant related to the shape of the pulse envelope (Fig. 14). For example, in the case of a Gaussian-shaped pulse $C = 0.441$, for sech^2 shaped pulses $C = 0.315$, for rectangular pulses $C = 0.886$, etc (Fig. 19).

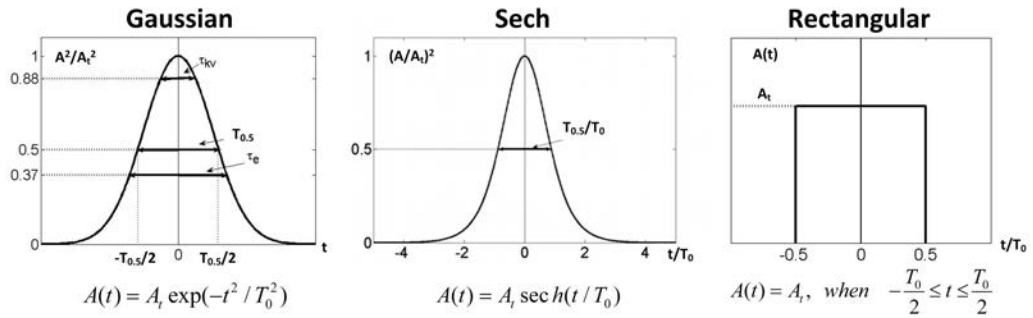


Figure 19: Examples of different pulse shapes. $\tau_{0.5}$ marks pulse duration at FWHM level. Note that some pulse shape models (such as rectangular) are purely theoretical models and in practice only approximations of such shapes are possible to achieve.

In the case of equality in Eq.(13), pulse duration exactly corresponds to its spectral bandwidth. Such pulses are called **bandwidth-limited** or

transform-limited. This also means that the shorter is the pulse, the broader is its frequency spectrum and vice versa. For example, continuous wave (infinite pulse duration) lasers are essentially monochromatic (their emission spectrum corresponds to their intrinsic spectral bandwidth), while femtosecond lasers are polychromatic, they generate a broadband frequency spectrum. For example, in the near-IR range (λ=1 μm) the spectral bandwidth of a $\Delta\tau$=1 ps pulse (FWHM level) Gaussian pulse is:

$$\Delta\upsilon = \frac{0.44}{\Delta\tau} = 4.4 \cdot 10^{11}(\text{Hz}) \tag{14}$$

or

$$\Delta\lambda = \frac{\Delta\upsilon\lambda^2}{c} = 1.47(\text{nm}), \tag{15}$$

whereas the spectral bandwidth of the same central wavelength $\Delta\tau$=10 fs Gaussian pulse is $\Delta\lambda$=147 nm at FWHM. In the visible range the spectrum of a latter pulse would essentially be white light. It is important to mention that artificial narrowing (extension) of the bandwidth-limited pulse spectrum results in an increase (decrease) of its duration. If the left hand side of expression in Eq.(13) is significantly greater than the right hand side, the light pulse is long, but its frequency spectrum is broad. Such pulses are called **phase modulated (or chirped)**. Manipulation of the pulse phase and frequency modulation and its control is the basis of temporal compression and stretching, which is widely exploited in modern laser amplifiers. Compared with other light sources, lasers are unique in their ability to generate light pulses and beams with the smallest possible uncertainty: diffraction-limited beams and bandwidth-limited pulses.

2.2 Light waves

The wave nature of light means that it can be described as electromagnetic waves: alternating electric and magnetic fields propagating in space and carrying energy. Electromagnetic wave propagation is fully described by a set of coupled equations, called **Maxwell equations**:

$$\nabla \times \vec{H} = \vec{j} + \frac{\partial \vec{D}}{\partial t}, \tag{16}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \tag{17}$$

$$\nabla \cdot \vec{D} = \rho, \tag{18}$$

$$\nabla \cdot \vec{B} = 0, \tag{19}$$

where $\vec{H}$ is the magnetic field strength, $\vec{j}$ is the electric current density, $\vec{D}$ is the electric field flux density (electric induction), $\vec{B}$ is the magnetic field flux density (magnetic induction), ρ is the electric charge density. $\times$ is the curl operator indicating vector product, while $\cdot$ indicates scalar product and $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ is the nabla operator (expressed in Cartesian coordinate system).

These equations, named after the physicist and mathematician James Clerk Maxwell, who published them in early forms in the 1860s, provide a mathematical model unifying the phenomena of electricity and magnetism and put forward the foundation of classical electrodynamics. Eq.(16) essentially states the **Ampere's circuital law**: electric current and time-varying electric fields induce magnetic fields curling around them. Eq.(17) represents the **Faraday's law of electromagnetic induction**: time-varying magnetic fields induce electric fields curling around them. The third equation (Eq.(18)) is essentially the **Gauss' law**: electric charges create electric fields. The last equation (Eq.(19)) can be interpreted as the **Gauss's law for magnetism**: magnetic fields do not diverge, they curl around. In other words, Eq.(19) indicates that magnetic charges do not exist.

For propagation of electromagnetic waves in free space (vacuum) a more practical equation can be derived from Maxwell equations. This equation describes light propagation in free space and is called **Helmholtz equation**. In one-dimensional case it is expressed as:

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = 0 \tag{20}$$

where E is the wave amplitude. Solutions of this equation are certain functions $E(z,t) = f(z \pm ct)$, which are called fundamental waves. Helmholtz equation is also valid for a harmonic wave:

$$E(z,t) = A_0 \sin(kz - \omega t), \tag{21}$$

which is described by its amplitude A_0, wavenumber $k = 2\pi/\lambda$, which is inversely proportional to wavelength λ, and frequency ω which is inversely proportional to wave period. Quantity $(kz - \omega t)$ is called the wave phase. Superposition of harmonic waves is also a solution of the Helmholtz equation. However, summation of waves described by trigonometric functions yields complex mathematical expressions that are inconvenient to use, therefore complex wave notation is usually applied which is derived using Euler's

formula:

$$e^{i\alpha} = \cos\alpha + i\sin\alpha. \tag{22}$$

Using it and a derived expression:

$$\sin\alpha = \frac{1}{2i}[\exp(i\alpha) - \exp(-i\alpha), \tag{23}$$

it is convenient to express the harmonic wave in the complex form:

$$E(z,t) = \frac{1}{2}E_0 \exp[i(kz - \omega t)] + c.c., \tag{24}$$

where E_0 is the complex amplitude of the wave and c. c. marks the complex conjugate part of the equation: if we analyze wave propagation in free space, the complex conjugate part of expression in Eq.(24) can be ignored.

There are several approximations of the light waves that are important when analyzing propagation of a laser beam. Firstly, for the sake of simplicity, let us assume that light waves are infinite in time, i.e. they have only a single frequency component (are monochromatic). **Plane waves** are waves whose phase at fixed moments of time is stationary in the entire plane perpendicular to the propagation direction (Fig. 20 a).

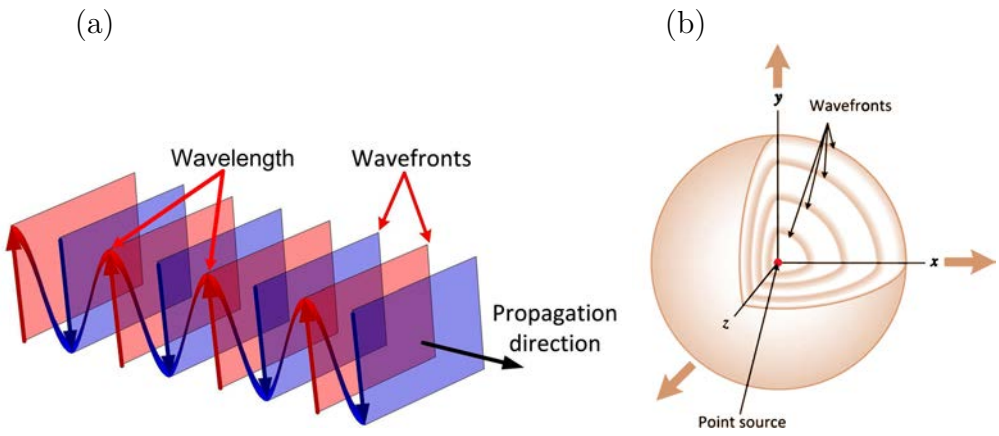


Figure 20: Illustration of plane (a) and spherical waves (b).

It is written as

$$E(z,t) = E_0 \exp[i(\vec{k}\vec{r} \pm \omega t)], \tag{25}$$

where $\vec{r} = (x, y, z)$, $\vec{k} = (k_x, k_y, k_z)$ and $\vec{k}\vec{r} = const$. Such wave repeats itself in the $\vec{k}$ direction after a distance equal to λ. This wave is infinite in space and its k vectors are unidirectional, so such wave does not diffract. In reality plane waves do not exist, but the concept of plane waves is often applied as a convenient approximation when diffraction of light can be neglected.

Another wave optics approximation is a **spherical wave** (Fig. 20 b). In spherical coordinate system these waves are expressed as

$$E(r,t) = \frac{E_0}{r} \exp\left[i(kz \pm \omega t)\right]. \tag{26}$$

The amplitude of a spherical wave decreases during propagation as a function of $1/r$ and when r becomes very large (approaches infinity), the spherical wave becomes a plane wave (see Fig. 21). Interestingly, spherical waves strongly diffract away from the source.

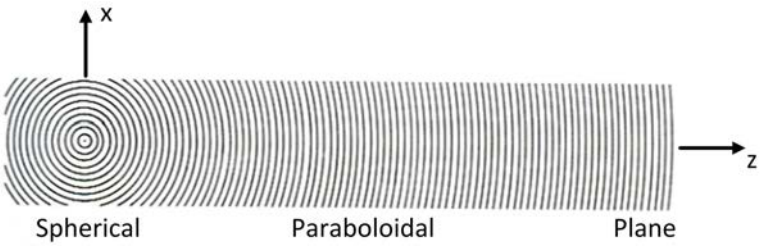


Figure 21: Principal view of spherical wave wavefront evolution during propagation in free space: far from the source the wavefront can be approximated by a paraboloidal function and as it propagates further the wavefront becomes very similar to plane.

Bearing in mind the above, it is clear that plane and spherical waves represent the two opposite cases considering angular and spatial energy distributions. A plane wave does not have angular spread and its energy is evenly distributed through the entire space, whereas a spherical wave comes from a point source and spreads at every direction with its amplitude rapidly decreasing at a given direction. The light beams represent an intermediate case between plane and spherical waves. They possess features of both, plane and spherical waves. The wave nature of light rules out the possibility for light to propagate in free space without angular spread, but in the case of light beams, the energy can be sufficiently well (as much as the uncertainty principle allows) localized along the propagation direction.

The light waves whose wavefront normal vectors make a small angle with the propagation axis are called **paraxial**. Such waves are also solutions of the Helmholtz equation. One of the most important approximations of paraxial waves is the **Gaussian beam** – the light beam whose spatial intensity distribution is expressed by a Gaussian function (Fig.22).

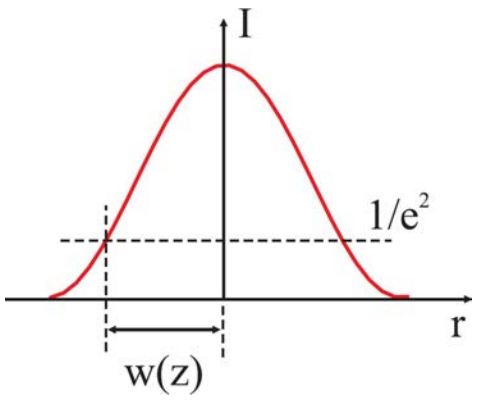


Figure 22: Intensity distribution of a Gaussian beam.

2.3 Gaussian beams

The Gaussian beam is of major importance in laser physics. First of all, as mentioned before, Gaussian beam has the smallest uncertainty of photon position and propagation direction, therefore, the energy in a Gaussian beam is localized very well along the cylinder surrounding the propagation axis. This implies that the angular spread of an ideal Gaussian beam is minimal for a given beam radius in the waist, so in other words, the Gaussian beam has very high quality and is **diffraction-limited**; it can be focused to a spot which size is limited only by diffraction. Secondly, the intensity distribution of a Gaussian beam in the plane perpendicular to the propagation axis is described by a Gaussian function, which has a maximum located on the propagation axis. Thirdly, all lasers generate beams with a Gaussian intensity distribution (the reason for this will be explained later).

We will now discuss the main parameters of the Gaussian beam and how they evolve during propagation in free space. The light intensity distribution of an ideal Gaussian beam (Fig.22) is described as

$$\boxed{I(r,z) = I_0 \left(\frac{w_0}{w(z)}\right)^2 \exp\left(-\frac{2r^2}{w^2(z)}\right)}, \tag{27}$$

where $I_0 = |E_0^2|$, E_0 is the amplitude of electric field, w_0 is the minimum beam radius at the waist (see Fig.23), $r = \sqrt{x^2 + y^2}$. For every propagation distance z the transverse (x-y plane) intensity distribution of the beam is a Gaussian function with a maximum intensity at $z = 0$ and monotonically decreasing when going away from beam symmetry axis. During propagation (when z increases) the beam intensity changes as:

$$I = \frac{I_0}{1 + \frac{z^2}{z_R^2}}, \tag{28}$$

where z_R is the **Rayleigh length**. Let us discuss the meaning of z_R. The minimum Gaussian beam radius w_0 is at the beam waist ($z = 0$) and the beam diameter is $2w_0$. During propagation (away form the waist) the beam radius continuously increases and at $z = z_R$ it becomes $w(z_R) = \sqrt{2}w_0$. Then we can define the Rayleigh length as:

$$\boxed{z_R = \frac{\pi w_0^2}{\lambda}}. \tag{29}$$

It defines the distance at which beam radius increases by $\sqrt{2}$ times. The minimum Gaussian beam wavefront radius of curvature is at $z = \pm z_R$, thus the Rayleigh length is often called the **diffraction length**. On the other hand, the beam radius can be expressed as $w_0 = \sqrt{\lambda z_R/\pi}$.

The entire power of a Gaussian beam can be determined by integrating the following expression:

$$P = \int_0^\infty I(r, z) 2\pi r dr = \frac{1}{2} I_0 \pi w_0^2. \tag{30}$$

The beam radius can be defined at different intensity levels (see Fig.12). Usually, the beam radius is defined at $1/e^2 \approx 0.135$ intensity level, as shown in Fig.22. Eq.(30) is valid for such beam radius definition. For the electric field amplitude, this would correspond to $1/e \approx 0.368$ level. If the integration limit in Eq.(30) is changed from ∞ to $w(z)$, such definition of $w(z)$ corresponds to concentration of 86% of the whole beam power, while 99% of the beam power fits into $1.5w(z)$ intensity level. In general, **an ideal Gaussian beam is infinite**, but only a negligible part of its power is contained in the far periphery. In practice, Gaussian beam radius can also be defined at a half maximum intensity level (HWHM – half width at a half maximum) and the beam diameter is determined at the same intensity level (FWHM – full width at a half maximum). Beam radii definitions at FWHM level and at $1/e^2$ level are related as:

$$\exp\left(-2\frac{w_{FWHM}^2}{w_{1/e^2}^2}\right) = 0.5. \tag{31}$$

Simplifying this expression yields the mathematical relation for different beam radii definitions:

$$w_{FWHM} = \sqrt{\frac{\ln 2}{2}} w_{1/e^2}, \tag{32}$$

which for the diameter of the beam is $2w_{FWHM} = \sqrt{2\ln 2}w_{1/e^2}$, thus a factor of $2\ln 2$ is necessary in the Gaussian beam power expression if the beam radius is defined at FWHM level.

The radius ($w(z)$) and wavefront curvature ($R(z)$) of a Gaussian beam change according to expressions (Fig.23):

$$w(z) = w_0\big(1 + \frac{z^2}{z_R^2}\big)^{1/2}, \tag{33}$$

and

$$R(z) = z\big(1 + \frac{z_R^2}{z^2}\big), \tag{34}$$

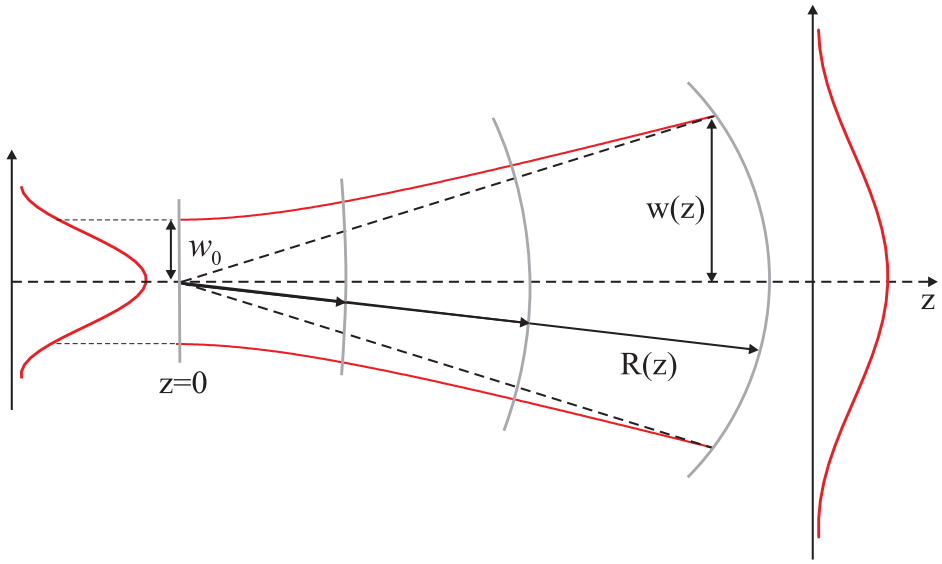


Figure 23: Gaussian beam parameters and their evolution during propagation.

Let us determine how radius of a Gaussian beam changes far from the beam waist. When $z \gg z_R$ in Eq.(33) "1" can be omitted and the beam radius increases linearly with z as depicted in Fig.24. This linear dependence can be expressed as

$$w(z) \approx \frac{w_0}{z_R} z = \theta z \tag{35}$$

where θ is the divergence angle:

$$\theta = \frac{w_0}{z_R} = \frac{\lambda}{\pi w_0}. \tag{36}$$

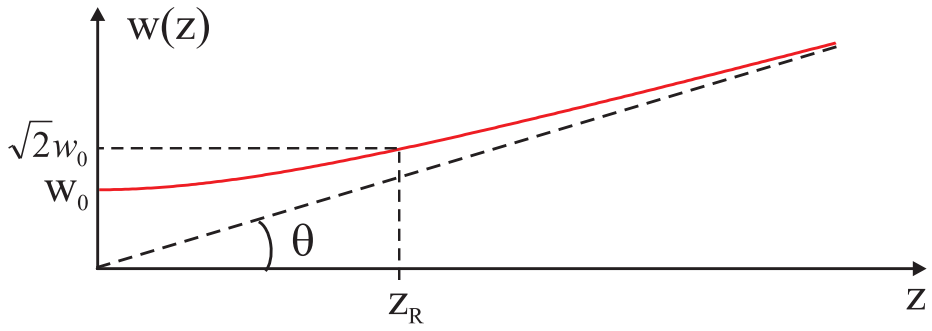


Figure 24: Definition of divergence angle of a Gaussian beam.

The divergence angle basically reflects the uncertainty principle which was discussed in the first section of this chapter. In addition, the divergence angle depends not only on the radius of beam waist (the divergence of beam with a small waist is large, and vice versa, so for the latter case a plane wave approximation could be applied) but also on the wavelength. In that regard, the ultraviolet Gaussian beams have lower divergence than infrared beams of the same waist size. For radio waves which could be considered as a boundary case, the concept of beams is meaningless since they have huge divergence and are essentially spherical waves.

Sometimes the Gaussian beam is characterized by a depth of focus (confocal parameter). The main idea is that the beam spreads equally on both sides with respect to its waist, so the confocal parameter is $b = 2z_R = 2\pi w_0^2/\lambda$. For example, for $\lambda = 633$ nm wavelength (He-Ne laser wavelength) and beam diameter of $2w_0 = 2$ cm, $b = 1$ km, whereas for the beam diameter of $2w_0 = 20$ μm, $b = 1$ mm.

Now let us discuss the phase factor of a Gaussian beam which is expressed as:

$$\phi(r, z) = kz - \xi(z) + \frac{kr^2}{2R(z)}. \tag{37}$$

When considering the phase of the beam we usually use the concept of wavefront. The wavefront is a surface of equal phase with respect to the beam propagation direction. The wavefront of a plane wave is flat (the plane perpendicular to the propagation direction) and the wavefront of a spherical wave is spherical. As follows from expression in Eq.(37), the Gaussian beam wavefront is neither flat, nor spherical. The phase of the beam on the propagation axis ($r = 0$) is expressed as $\phi(0, z) = kz - \xi(z)$. The first term describes a plane wave. So at the waist ($z = 0$) the Gaussian beam is a plane wave, i.e. its wavefront is flat and radius of curvature $R(z)$ is infinite. The

second term describes the phase shift of the on-axis beam part compared to either plane or the spherical wave. When $z = \pm z_R$, the phase shift is $\pm\pi/4$ and approaches $\pm\pi/2$ when $z \to \pm\infty$. This means that during propagation of the Gaussian beam through infinitely long distance the maximum accumulated axial phase shift is π. This is called the **Gouy effect**. The third term in Eq.(37) describes the phase shift for off-axis points compared with the phase of the on-axis points. In other words, it shows the degree of wavefront bending. It can be shown that a surface of equal phase is parabolic with a radius of curvature R. The evolution of the radius of curvature for a Gaussian beam during propagation is described by Eq.(34), and is graphically illustrated in Fig.25. It is clear that the minimum radius of curvature of a Gaussian beam is at $z = \pm z_R$: $R = \pm 2z_R$, i.e., the wavefront here is maximally curved. The radius of curvature both sides from z_R is only increasing with propagation.

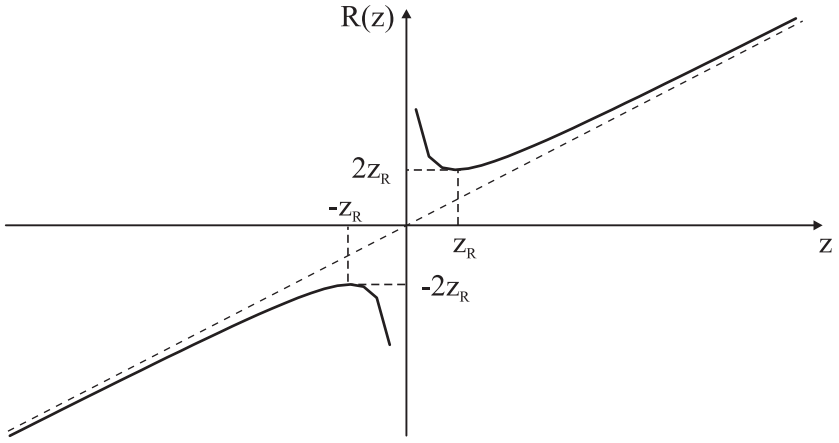


Figure 25: Evolution of a Gaussian beam radius of curvature during propagation.

At the beam waist $R = \infty$, so at this particular location the Gaussian beam can be approximated as a plane wave as illustrated in Fig.26. The same approximation can also formally be applied at $z = \infty$, however, the axial phase shift due to Gouy effect has to be taken into account. Far from the beam waist of a Gaussian beam wavefront is almost spherical.

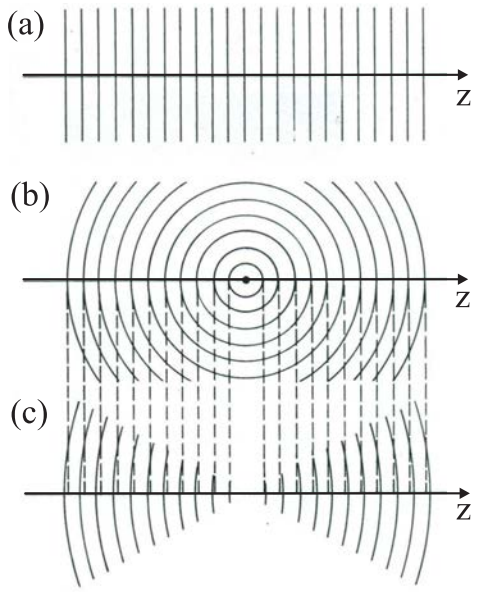


Figure 26: Comparison of wavefronts:(a) plane waves, (b) spherical waves and (c) Gaussian beam.

2.4 Coherence

Coherence is a distinctive feature of laser radiation. In general, the term **coherence** is used to describe **phase correlation** between two monochromatic waves. Waves with correlated phases are called coherent and waves with uncorrelated (random) phases are called **incoherent**. Two types of coherence can be distinguished:

- **Temporal (longitudinal) coherence** is the measure of the correlation of light wave's phase at different points in time along the direction of propagation and it describes how monochromatic the light source is (Fig.27).

- **Spatial (transverse) coherence** is the measure of the correlation of a light wave's phase at different points transverse in the plane perpendicular to the direction of propagation. It describes how uniform the phase of a wavefront is (Fig.27). In other words, it shows the ability for two points in space, y_1 and y_2, in the extent of a wave to interfere, when averaged over time.

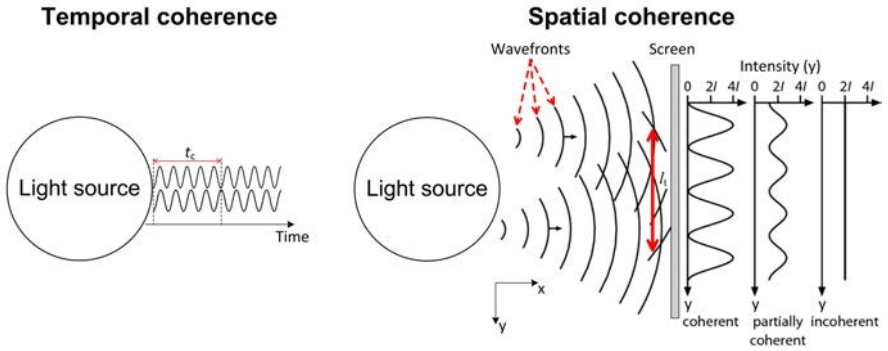


Figure 27: Principal illustration of temporal and spatial coherence of light waves. t_c and l_t mark temporal coherence time and spatial coherence length, respectively. In the spatial coherence part interference intensity pattern is depicted in the cases of perfectly coherent, partially coherent and incoherent light source.

Let us first discuss temporal coherence in more detail. From a practical point of view, temporal coherence characterizes how well a wave can interfere with itself at a different moments of time. When we observe interference pattern using an interferometer, greater temporal coherence corresponds to sharper contrast between fringe minima and maxima. If we increase the length difference between interferometer arms, the interference pattern blurs and eventually disappears (Fig.28). Let us discuss why.

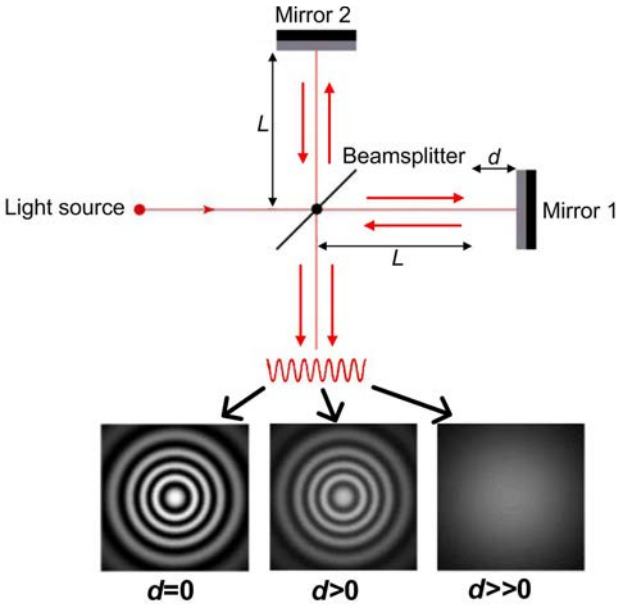


Figure 28: Example of interference fringes observed from a monochromatic light source using a Michelson interferometer. d is the optical path difference between Michelson interferometer arms.

An ideal monochromatic light source is temporally coherent, but it is clear that in the reality strictly monochromatic light sources do not exist. Light sources that we are called "monochromatic" in reality radiate a sequence of harmonic waves with finite lengths Fig.(29), which are interrupted by phase jumps due to various processes occurring in a system of excited atoms.

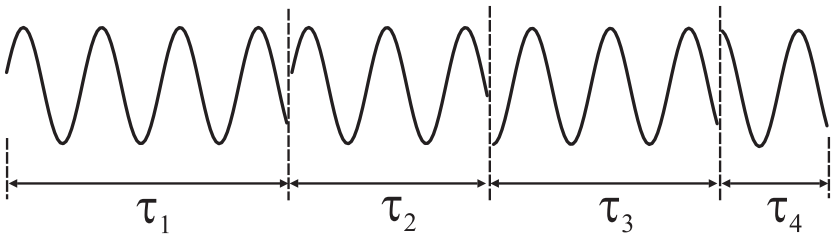


Figure 29: A sequence of harmonic waves with random length and phase.

Such a light source is described by an average duration of a harmonic wave t_c (with a constant phase), which is an average of $\tau_1, \tau_2, \tau_3, \tau_4, ...$ Consequently, any light source can be described with the duration t_c which is called the **coherence time**. The main parameter which determines the coherence time of a given source is the natural (intrinsic) spectral linewidth, i.e., $\Delta\omega = 2\pi/t_c$ or $\Delta\nu = 1/t_c$. It is obvious that a continuous phase and infinitely monochromatic wave would have an infinite coherence time, and vice versa: a radiation with infinitely broad spectral bandwidth would have an infinitely short coherence time, approaching the delta function. The same applies to the coherence length which is defined as $l_c = ct_c$:

$$\boxed{l_c = \frac{c}{\Delta\nu} \approx \frac{\lambda^2}{\Delta\lambda}}, \tag{38}$$

which essentially reflects the boundary case of the uncertainty principle; here $\Delta\lambda$ is the natural spectral linewidth of a given light source. Having this in mind, we can now explain what is depicted in Fig.28. When optical path difference between interferometer arms (d) approaches the temporal coherence length (l_c), the contrast of interference fringes drops and when d exceeds temporal coherence length, the interference pattern disappears. This simply reflects the fact that there are no ideal monochromatic light sources, therefore any real light source has a finite coherence time and temporal coherence length.

The temporal coherence of several representative light sources is compared in Table 1. For any CW light source (including CW lasers) the temporal coherence in this case describes the monochromaticity of such light

Table 1: Comparison of spectral linewidths, coherence times and coherence lengths of some representative light sources.

Source	Spectral linewidth, Hz	t_c, s	l_c, mm
The Sun	6×10^{14}	10^{-15}	3×10^{-4}
Blue LED	6×10^{13}	3×10^{-14}	9×10^{-3}
Filter	1.2×10^{12}	5.2×10^{-13}	0.2
Flashlamp	10^9	6.2×10^{-10}	190
Monochromator	10^8	6.2×10^{-9}	1900
CW laser	10^5	6.2×10^{-6}	10^6
fs laser	5×10^{13}	3.5×10^{-14}	10^{-2}

source. It is clear that CW lasers in regard to their monochromaticity by far surpass any non-laser light source.

For pulsed lasers, the coherence time is interpreted differently. Such lasers emit light in bursts or portions, which are called pulses, whose durations depend on the methods how they are generated. The spectrum of a light pulse represents a collection of many monochromatic waves with different frequencies and if all the waves are in phase, the pulse is bandwidth limited and its duration equals to a coherence time, which is inversely proportional to its spectral width. However, if the pulse is not bandwidth-limited (chirped), its duration its always longer than its coherence time.

For example, a femtosecond laser emitting 100 fs bandwidth-limited pulse has temporal coherence time of t_c=100 fs and temporal coherence length $l_c = ct_c$=30 μm. When observing interference of such laser pulse with a Michelson interfereometer (Fig.28) optical path difference between interferometer arms cannot exceed 30 μm.

2.5 Spatial coherence

In the previous section we considered the phase correlation at different moments of time along the propagation direction. Therefore, temporal coherence is also called the longitudinal coherence. Now let us discuss the coherence between two points of the same light source (or beam) separated in space. As already mentioned, the measure of the correlation of a light wave's phase at different points transverse to the direction of propagation is called **spatial coherence** (Fig.27 right). The **spatial (transverse) coherence length** is usually defined as a maximum distance l_t between the two points where the interference can still be observed, as schematically depicted in Fig.30.

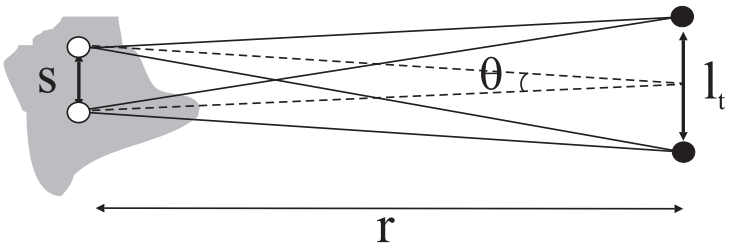


Figure 30: Definition of spatial coherence of light.

Then the coherence length is defined as

$$\boxed{l_t = \frac{r\lambda}{s} = \frac{\lambda}{\theta}}, \tag{39}$$

where r is the distance from the light source, s is the distance between the points of the source and θ is the angle at which these points are visible. According to this definition, an ideal plane waves and point light sources are fully spatially coherent. Any other incoherent light source can be made coherent by turning it into a point light source, however, this is possible only at the cost of its radiation power..

Let us discuss the meaning of spatial coherence in the case of a Gaussian beam. Comparing expressions in Eq.(36) and Eq.(39), it is clear that they match with precision of a constant, so the radius of a Gaussian beam waist corresponds to the length of spatial coherence.

To characterize a real Gaussian beam, another parameter called $\boldsymbol{M^2}$ is also introduced. This parameter indicates the difference between the divergence angles of real and ideal beams with the same beam waist. Moreover, this parameter is often used as a beam "quality" parameter. The term "quality" is more often defined as a measure for how well the beam can be focused (Fig.31) and it is quantified as either $\boldsymbol{M^2}$ parameter:

$$\boxed{M^2 = \frac{\pi w_0 \theta}{\lambda}}, \tag{40}$$

or **beam parameter product (BPP)**:

$$\boxed{BPP = w_0 \cdot \theta}. \tag{41}$$

When we calculate diameter of a focused beam (d_f) we have to take into account the M^2 parameter:

$$\boxed{d_f = M^2 \frac{4\lambda f}{\pi d_0}}, \tag{42}$$

here f is lens focal length, d_0 is the initial (unfocused) beam diameter. M^2=1 for diffraction-limited Gaussian beam, so the M^2 factor is a quantitative measure how much the beam is different from a diffraction-limited Gaussian beam.

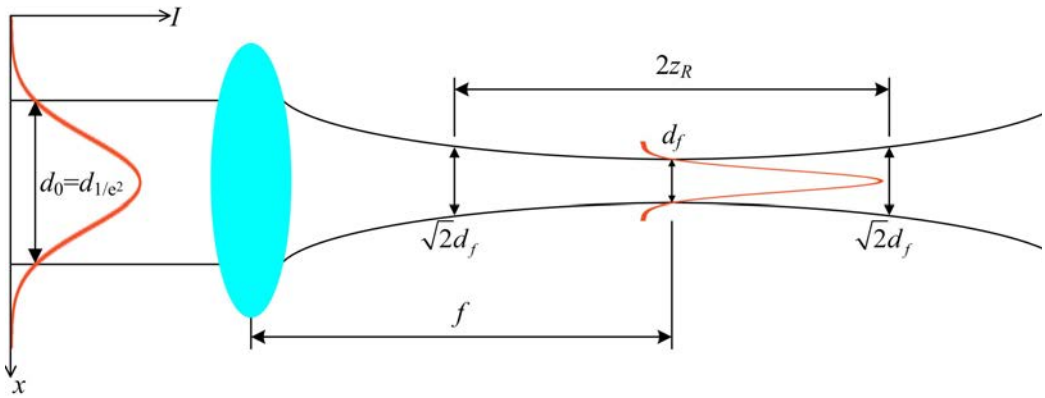


Figure 31: Gaussian beam size evolution when it is focused with a lens. The red contours depict laser beam intensity distribution and d_{1/e^2} denotes laser beam diameter at $1/e^2$ of maximum intensity level.

Using the light interference terminology, the Gaussian beam can be viewed as a result of interfering multiple plane waves propagating at different directions. Since these waves are phased in a certain way, due to interference they create a specific spatial modulation which is observed as the light beam. Thus, laser light sources are also unique with regard to spatial coherence as laser beams are spatially coherent and carry all the power emitted by the laser source.

Concluding the coherence issues, the concepts of temporal and spatial coherence could be merged by defining the volume (three-dimensional) coherence:

$$l_v = l_c l_t^2, \tag{43}$$

which describes the spatio-temporal coherence and directly indicates the volume where the energy of light is localized. The concept of volume (spatio-temporal) coherence applies to pulsed lasers.

Summary of Chapter 2

- Light is described either as electromagnetic waves or as particles (photons). Each concept attributes to different properties of laser radiation.
- Photons are quantum particles and are subject to uncertainty principle, which poses fundamental limitations to relevant parameters of laser radiation.
- Diffraction is related to the photon position-momentum uncertainty which in turn is the fundamental the reason for laser beam divergence (spreading in space during propagation). Laser beams with minimum radius for a given divergence angle are called diffraction-limited.
- The time-energy uncertainty relates the laser pulse duration with its spectral bandwidth. The shorter is laser pulse, the broader is its frequency spectrum, and vice versa. Laser pulses that have minimum duration for a given bandwidth are called bandwidth-limited (transform-limited).
- Laser produces a low divergence light beam which has an intensity distribution (in the plane perpendicular to propagation direction) that is described by Gaussian function.
- Laser emits spatially and temporally coherent light.

3 Optical resonators

An optical resonator (optical cavity) is an extremely important component of any laser. The simplest optical resonator consists of two reflecting mirrors, which play several relevant roles in laser operation:

- **Energy accumulation and feedback**: the laser resonator accumulates electromagnetic energy by confining radiation within a certain volume and provides a feedback function by returning radiation to the active element, where the process of stimulated emission takes place.
- **Formation of transverse modes**: the laser resonator forms a directional radiation in the shape of a light beam, which has with high spatial quality and high spatial coherence and which has a particular intensity distribution in the transverse plane that is called transverse mode.
- **Formation of longitudinal modes**: the laser resonator has its intrinsic temporal frequencies which satisfy the standing wave condition (the longitudinal modes) and which determine the spectral quality of the radiation and ensure high temporal coherence.

In this chapter we will discuss the role of resonators in lasers: light propagation analysis in optical resonators, resonator stability criteria, transverse and longitudinal modes and energy loss mechanisms in optical resonators.

3.1 Light rays and ray (or ABCD) matrices

Light propagation in various optical elements and resonators in the paraxial approximation can be successfully described using an approximation of geometrical optics: light rays. A **light ray** is defined as a straight line perpendicular to the wavefront. Therefore, if we know how rays behave when propagating through a certain optical medium or element, we also know how light waves and beams behave in such medium or element. Furthermore, light ray approximation can be applied not only for light propagation in isotropic optical media but also in optical elements (lenses, mirrors, etc.) and for light propagation in media with refractive index gradient, absorption, amplification or scattering.

Propagation of light rays (as well as their refraction or reflection) in optical elements or media can be described using simple 2×2 matrices which are called **ray** or **ABCD matrices**. A light ray can be completely described by two parameters: position x and inclination angle θ with respect to the

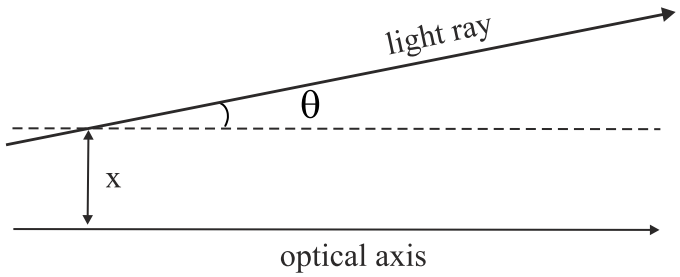


Figure 32: The main parameters defining propagation of a light ray.

optical axis, as depicted in Fig.32. These two parameters define the ray vector (x, θ). An effect of any optical element is equivalent to multiplication of the ray matrix defining the optical element by ray vector, resulting in new values of the parameters characterizing the ray:

$$\boxed{\begin{bmatrix} x_{\text{out}} \\ \theta_{\text{out}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_{\text{in}} \\ \theta_{\text{in}} \end{bmatrix}} \tag{44}$$

where indices *in* and *out* mark vector parameters of the initial (incident) and final (outgoing) rays respectively. We will now discuss the principle how ray matrices for various optical elements are constructed and how their A, B, C and D values are determined. Assuming that inclination angle of a ray to the optical axis is small (which is essentially the condition of paraxial approximation as was already stated), the ray parameter changes are described by a two equation system:

$$\begin{aligned} x_{\text{out}} &= \frac{\partial x_{\text{out}}}{\partial x_{\text{in}}}\, x_{\text{in}} + \frac{\partial x_{\text{out}}}{\partial \theta_{\text{in}}}\, \theta_{\text{in}} \\ \theta_{\text{out}} &= \frac{\partial \theta_{\text{out}}}{\partial x_{\text{in}}}\, x_{\text{in}} + \frac{\partial \theta_{\text{out}}}{\partial \theta_{\text{in}}}\, \theta_{\text{in}}. \end{aligned} \tag{45}$$

These equations are equivalent to expression in Eq.(44), so matrix elements can be expressed as partial derivatives where

$$A = \frac{\partial x_{\text{out}}}{\partial x_{\text{in}}}, \quad B = \frac{\partial x_{\text{out}}}{\partial \theta_{\text{in}}}, \quad C = \frac{\partial \theta_{\text{out}}}{\partial x_{\text{in}}}, \quad D = \frac{\partial \theta_{\text{out}}}{\partial \theta_{\text{in}}}. \tag{46}$$

Matrix elements A and D can be understood as spatial and angular magnification, respectively. With accordance to these expressions, let us write ABCD matrices for several optical media and elements which are important for optical resonators. To start with, let us determine the ABCD matrix for

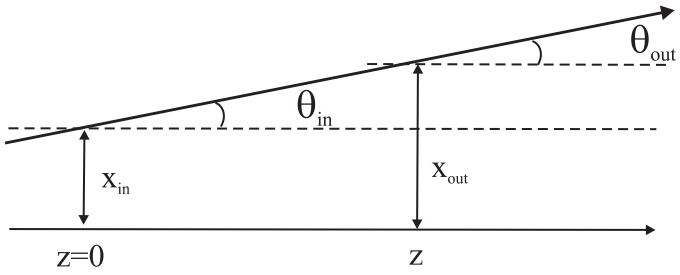


Figure 33: Propagation of a light ray in free space.

free space propagation M_{S}. Light propagation in free space is depicted in Fig.33.

The system of equations (Eq.(45)) for light ray propagation in free space can be written as:

$$\begin{aligned} x_{\mathrm{out}} &= x_{\mathrm{in}} + z\ \theta_{\mathrm{in}} \\ \theta_{\mathrm{out}} &= \theta_{\mathrm{in}}. \end{aligned} \tag{47}$$

In matrix form the system of equations transforms into:

$$\begin{bmatrix} x_{out} \\ \theta_{out} \end{bmatrix} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{in} \\ \theta_{in} \end{bmatrix} \tag{48}$$

or

$$M_{\mathrm{S}} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}. \tag{49}$$

Now let us determine the ABCD matrix for intersection of two optical media with different refractive indices n_1 ir n_2. Geometrically the situation is depicted in Fig.34.

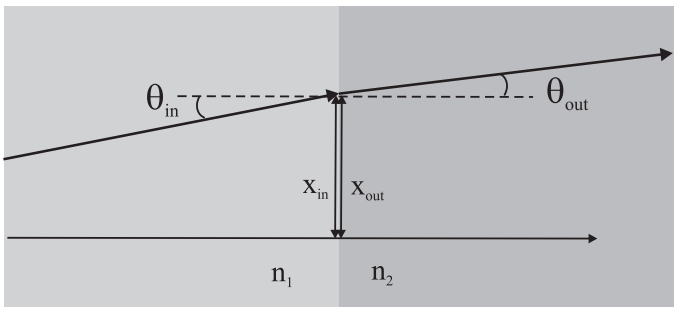


Figure 34: Propagation of a light ray through the intersection of two media.

At the intersection of two media, a distance of the light ray from the optical axis remains unchanged: $x_{\text{out}} = x_{\text{in}}$, whereas the change of the inclination angle can be calculated from the Snell's law: $\theta_{\text{out}} = \frac{n_1}{n_2}\theta_{in}$. Consequently, the ray matrix for the intersection of two media with different refractive indices is

$$M_{\text{I}} = \begin{bmatrix} 1 & 0 \\ 0 & n_1/n_2 \end{bmatrix}. \tag{50}$$

In analogy, the ray matrix for the intersection of two optical media with different refractive indices and curved (with R radius of curvature) intersection surface can be written as:

$$M_{\text{CI}} = \begin{bmatrix} 1 & 0 \\ \frac{n_1/n_2-1}{R} & n_1/n_2 \end{bmatrix}. \tag{51}$$

In that case, if we assume a lens as two curved surfaces, the ray matrix for a lens can be derived as a product of two ray matrices of curved surfaces:

$$M_{\text{L}} = M_{CI_1} M_{CI_2} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{bmatrix}. \tag{52}$$

Here we also assumed that the refractive index of the surrounding medium is $n_1 = 1$ and the refractive index of a lens material is $n_2 = n$. Then f is the focal distance of the lens and is expressed as:

$$\frac{1}{f} = (n-1) \quad \frac{1}{R_2} - \frac{1}{R_1} \quad , \tag{53}$$

where R_1 and R_2 are the radii of curvature of the lens' surfaces. This expression is also known as the **Lens Maker's formula**. Notice that if $f > 0$ the lens is convex and if $f < 0$ it is concave. Based on a similar consideration, we can determine ABCD matrix for a curved mirror, whose center of curvature is located on the optical axis:

$$M_{\text{CM}} = \begin{bmatrix} 1 & 0 \\ -\frac{2}{R} & 1 \end{bmatrix}, \tag{54}$$

where R is the mirror's radius of curvature. Comparing the expressions for M_{L} and M_{CM} it is clear that a curved mirror focuses rays as a lens with a focal length of $f_v = R/2$. Therefore a mirror with $R > 0$ is regarded as concave and a mirror with $R < 0$ as convex. In the case of a plane mirror, the same ray matrix can be used by setting $R = \infty$, i.e.:

$$M_{\text{PM}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{55}$$

Concluding the results of this section, we can note that an ABCD matrix can be written for any optical element no matter how complex it is. This means that light ray propagation can be modelled in sophisticated optical systems consisting of various optical elements whose ABCD matrices are known. Such method is called **ray tracing**. Moreover, ABCD matrix formalism can be directly applied to a Gaussian beam assuming that the optical element transforms its complex propagation parameter in the following way:

$$q_{out} = \frac{Aq_{in} + B}{Cq_{in} + D}, \tag{56}$$

where q_{in} and q_{out} are incident and passing (through the optical element) complex propagation parameters of the Gaussian beam, respectively. The complex propagation parameter is expressed via two real functions: beam radius $w(z)$ and wavefront curvature $R(z)$, which fully ascribe the Gaussian beam:

$$\frac{1}{q(z)} = \frac{1}{R(z)} - i\frac{\lambda}{\pi w^2(z)}. \tag{57}$$

From this we can determine how relevant parameters (i.e., $w(z)$ and $R(z)$) of the Gaussian beam change during beam propagation in various optical systems.

3.2 Laser resonator stability

Let us discuss light propagation in an optical resonator consisting of two curved mirrors with radii of curvature R_1 and R_2 separated by distance l as shown in Fig.35. The readers may notice that the depicted resonator does not have any gain medium inside: such resonators are called "cold" or empty. We can say in advance that such simplification does not alter light propagation analysis or any important results that we will arrive at in this section.

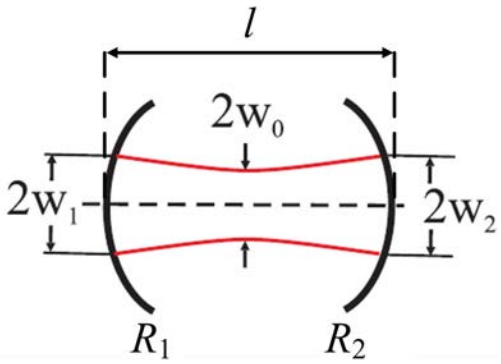


Figure 35: Laser beam in an empty optical resonator consisting of two curved mirrors.

It is obvious that in order for the laser to work, light should not escape the resonator: after each reflection the beam must maintain a stable size fitting into the resonator mirror aperture which is finite. Light will be confined in such optical system only if it fully reproduces its spatial distribution after each round-trip between the resonator mirrors. This is called **beam convergence condition**, which implies that only the laser beams with a specific spatial distributions of electric field (intensity) are able to do so. These are called **spatial (transverse) modes** of the resonator. On the other hand, beam convergence condition poses strict requirements on the resonator parameters and geometry: mirrors radii of curvature R_1 and R_2 and resonator length l. By applying ABCD law for a Gaussian beam that fully reproduces itself after each resonator round-trip ($q_{in} = q_{out}$):

$$q = \frac{Aq + B}{Cq + D}, \tag{58}$$

where A, B, C, D values are derived from $M_{resonator} = M_{CM2}M_S M_{CM1}M_S$ and skipping mathematics for the sake of simplicity we can derive **laser resonator stability condition**:

$$0 \leq \quad 1 - \frac{l}{R_1} \qquad 1 - \frac{l}{R_2} \quad \leq 1. \tag{59}$$

Here R_1 and R_2 are resoantor mirros radii of curvature and l is the length of a resonator. $g_i = 1 - \frac{l}{R_i}$ are called g parameters of the particular resonator mirror. This inequality is depicted graphically in Fig.36 which is called **resonator stability diagram**. Blue area in the diagram shows the space of g parameter values, where resonator is stable.

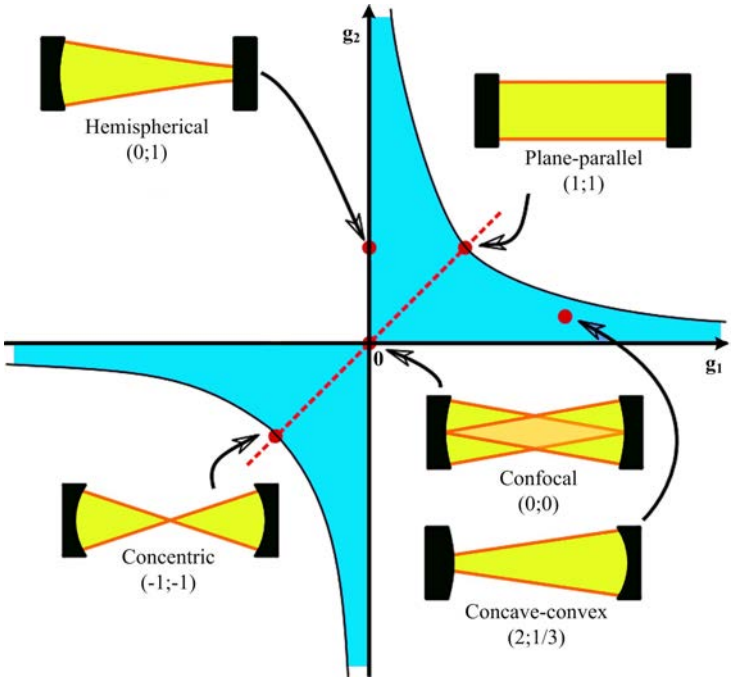


Figure 36: Resonator stability diagram, where stability zone is shown by blue area, and several resonator types corresponding to different points in the stability diagram. Image adapted from Wikimedia under CC BY-SA 3.0 license.

The stability diagram also shows that distinct resonator arrangements (which have different l and R ratio – Fig.37) correspond to different g_1 and g_2 parameter values in the diagram. It is important to understand that when g parameter values are at the boundary of stability zone, such resonator is very sensitive to any perturbations due to mechanical vibrations, temperature instabilities, etc. Therefore, concentric, confocal and plane resonators are very sensitive to external perturbations. Even very small changes of the environment make the resonator unstable, i.e., its diffraction losses increase significantly. Such resonator configurations are also critically sensitive to mirror alignment with respect to the optical axis. Therefore, in practice, asymmetrical resonators are usually used, e.g. consisting of plane and concave mirrors (hemispherical resonator). Such resonators are in the middle of the stability zone in the stability diagram, therefore, they are much less sensitive to external perturbations. On the other hand, we considered only "cold" (empty) resonators which do not have any other optical elements inside. In a real laser, apart from resonator mirrors there is also an active element whose operation in some cases can significantly change resonator stability conditions. For example, active elements of solid state lasers often have thermal lens effect, which must be taken into account when designing

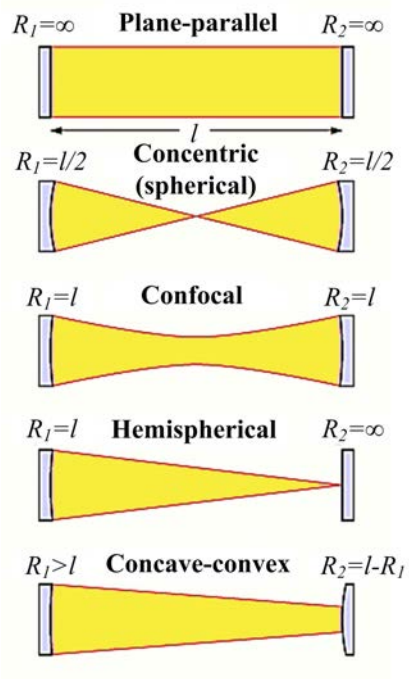


Figure 37: Types of laser resonators.

resonator geometry.

Therefore, summarizing both facts, we can claim that *Gaussian beam is a spatial mode of a spherical mirror resonator*. Generally, any resonator can be regarded as an approximation of a spherical resonator, so the Gaussian beam is a spatial mode of any optical resonator, and consequently, **all lasers produce radiation in the form of a Gaussian beam**. Furthermore, the M^2 factor introduced in the previous chapter quantitatively indicates how much a beam generated in a certain resonator differs from an ideal Gaussian beam generated in a spherical mirror resonator.

The current discussion may raise a question whether a laser with an unstable resonator can operate at all? It is possible indeed, however, lasers operating with unstable resonators typically have high gain which balances large diffraction losses. In an unstable resonator light will be ejected after a certain number of round-trips and output coupling is such lasers is also different: light leakage due to diffraction is often taken as the useful laser output (Fig,38). Unstable resonators also have transverse modes, but their intensity distributions are very complex and beyond the scope of this textbook. Interestingly, unstable resonators are less sensitive to alignment instabilities.

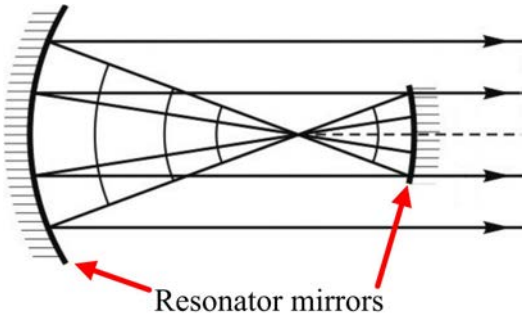


Figure 38: Light propagation in an unstable resonator. Output coupling in this case is at a hard-edge mirror.

3.3 Hermite-Gaussian beams and Laguerre Gaussian beams

When analyzing resonator stability conditions we have showed that Gaussian beam is an intrinsic spatial mode of a spherical mirror resonator (and in general, of any other resonator). The complex amplitude of electric field of a Gaussian beam can be expressed in a very general form as:

$$\begin{aligned} E(x,y,z) = E_0 \times \tfrac{w_0}{w(z)} \exp\left[-\tfrac{x^2+y^2}{w^2(z)}\right] \times \\ \exp\left[-i(kz)+i\xi(z)\right] \times \exp\left[-i\tfrac{k\left(x^2+y^2\right)}{2R(z)}\right] \end{aligned} , \tag{60}$$

where $\xi(z) = tan^{-1}(z/z_R)$ and $R(z)$ is the wavefront radius of curvature, E_0 is the amplitude of electric field, w_0 is the minimum beam radius (see Fig.23). The first line in this expression describes the amplitude evolution during propagation in the z direction, whereas the second line in the expression describes the phase evolution.

Lasers with very stable resonators may produce beams with more complex electric field (intensity) distributions, which reproduce themselves after each round-trip, and satisfy the convergence condition. These beams are called **higher order spatial (transverse) resonator modes**.

Two cases of higher order modes can be distinguished. In stable optical resonators **with no cylindrical symmetry** higher order transverse modes can be ascribed as **Hermite-Gaussian beams**, whose their electric field amplitude is expressed as

$$E(x,y,z) = E_0 H_m\left[\frac{\sqrt{2}x}{w(z)}\right] H_n\left[\frac{\sqrt{2}y}{w(z)}\right] \times \frac{w_0}{w(z)} \exp\left[-\frac{x^2+y^2}{w^2(z)} \quad \times \right. \tag{61}$$
$$\left. \exp\left[-i\left(kz - (1+m+n)\xi(z)\right)\right] \times \exp\left[-i\frac{k(x^2+y^2)}{2R(z)}\right] \quad .$$

where quantities E_0, w_0, $w(z)$, $R(z)$ and $\xi(z)$ are defined in the same manner as in the case of a Gaussian beam, while functions H_m and H_n are called Hermite polynomials, whose indices m and n mark the order of the polynomial. Hermite polynomials are defined as

$$H_0(u) = 1 \tag{62}$$
$$H_1(u) = u \tag{63}$$
$$H_2(u) = 2u^2 - 1 \quad \text{etc.} \tag{64}$$

In the case of the lowest order Hermite polynomial $m, n = 0$, we have an expression for the usual Gaussian beam Eq.(60), so the lowest order Hermite-Gaussian beam is simply a Gaussian beam. Comparing the expressions in Eq.(60) and in Eq.(56) we can notice that Hermite-Gaussian beam has all the relevant requisites of a Gaussian beam, but there are several key differences. The first multiplier of the product in expression Eq.(56) shows that the Hermite-Gaussian beam amplitude at $z = 0$ has interesting variations in the transverse plane (i.e., the plane perpendicular to the beam propagation direction) and values of m and n indicate how many times the field amplitude is equal to 0 on x and y axes, respectively (see Fig.39). The second multiplier of the product is identical to that of a Gaussian beam describing the evolution of field amplitude along the z axis: radius of the beam increases further away from the waist. The third multiplier describes the evolution of longitudinal phase: it is clear that the wavefront of Hermite-Gaussian beam is delayed on the axis with respect to spherical or plane wave. The delay is greater for higher order of the Hermite polynomials. The last multiplier describes the transverse phase (wavefront radius of curvature) and is identical to the term describing Gaussian beam wavefront.

Examples of Hermite-Gaussian beam intensity distributions are illustrated in Fig.39. Hermite-Gaussian beams are also the intrinsic spatial modes of a resonator. The spatial modes of a resonator are often called **Transverse Electromagnetic Modes and abbreviated** as $\mathbf{TEM_{nm}}$. The lowest index TEM mode is TEM_{00} and represents the Gaussian beam. The number of the amplitude peaks in higher order spatial resonator modes can be empirically evaluated as $(m+1)(n+1)$.

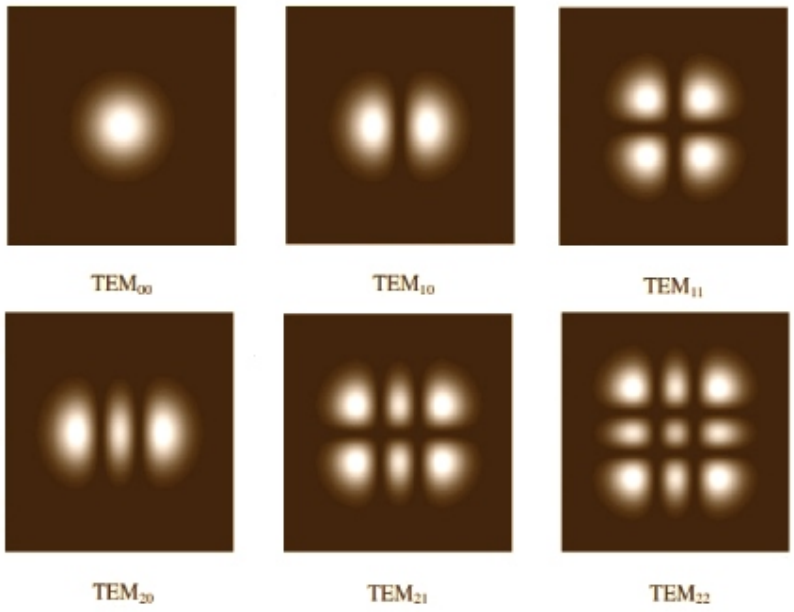


Figure 39: Intensity distributions of various Hermite-Gaussian modes.

It is useful to compare Gaussian and higher order spatial resonator mode parameters: beam radius and divergence. These are expressed as

$$\begin{aligned} w_{\mathrm{mn}} &= \sqrt{2m+n+1}\; w_{00} \\ \theta_{\mathrm{mn}} &= \sqrt{2m+n+1}\; \theta_{00}, \end{aligned} \tag{65}$$

where index 00 stands for the Gaussian beam (TEM_{00} mode). Higher order spatial resonator modes, compared to a Gaussian beam, have larger beam radius and larger angular divergence. This tells us that such laser beam quality is lower than that of a Gaussian beam and that higher order modes can only be excited in a very stable optical resonator where diffraction losses are small. On the other hand, the laser output with such a complex intensity distribution is rarely used in practice, thus usually higher order mode generation is suppressed: mitigated by specifically inducing losses (e.g., diffraction losses) to force laser resonator generate the TEM_{00} mode.

In stable optical resonators **with cylindrical symmetry** higher order transverse modes can be ascribed as **Laguerre-Gaussian beams** and their

electric field amplitude is expressed as

$$E(\rho,\varphi,z) = E_0 \frac{w_0}{w(z)} \left[\frac{\rho}{w(z)} \right]^l L_m^l \left[\frac{2\rho^2}{w^2(z)} \right] \times \exp\left[-\frac{\rho^2}{w^2(z)} \right] \times \\ \times \exp\left[-ikz - il\varphi + i(l+2m+1)\xi(z) \right] \times \exp\left[-i\frac{k\rho^2}{2R(z)} \right], \tag{66}$$

where ρ, φ and z are the coordinates (in cylindrical coordinate system) and L_m^l is the generalized Laguerre polynomial with indices m and l. These polynomials are the solutions of the Laguerre differential equation. They are defined via recurrence relation with the first ones being:

$$L_0^l(x) = 1 \tag{67}$$

$$L_1^l(x) = 1 + l - x \tag{68}$$

Fig.40 depicts the intensity distributions of Laguerre-Gaussian modes in resonators with cylindrical symmetry. As in the previous case, the lowest index TEM_{00} and represents the Gaussian beam. As in the case of Hermite-Gaussian beams, the Laguerre-Gaussian beam has lower quality than that of a Gaussian beam and such higher order Laguerre-Gaussian modes can only be excited in a very stable optical resonator.

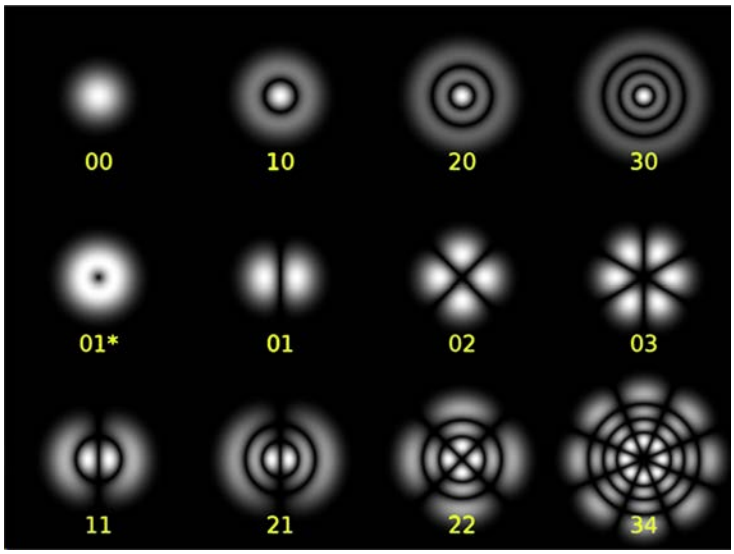


Figure 40: Intensity distributions of Laguerre-Gaussian modes. Indices mark the TEM_{ml} mode index in cylindrical coordinate system.

A specific and particularly interesting case seen in Fig.40 is when TEM_{ml} mode indices are 01*. This is a special case of Laguerre-Gaussian beam when

m=0 and rotational phase factor ($exp(+i\varphi l)$) is complex conjugate, which is called an **optical vortex**:

$$E(\rho,\varphi,z) = E_0 \frac{w_0}{w(z)} \left[\frac{\rho}{w(z)}\right]^l L_0^l \left[\frac{2\rho^2}{w^2(z)}\right] \times \exp\left[-\frac{\rho^2}{w^2(z)}\right] \times \\ \times \exp\left[-ikz + il\varphi + i(l+1)\xi(z)\right] \times \exp\left[-i\frac{k\rho^2}{2R(z)}\right] \tag{69}$$

Its phase front has a helical trajectory with an undefined phase and intensity minimum at the center of the beam (Fig.41).

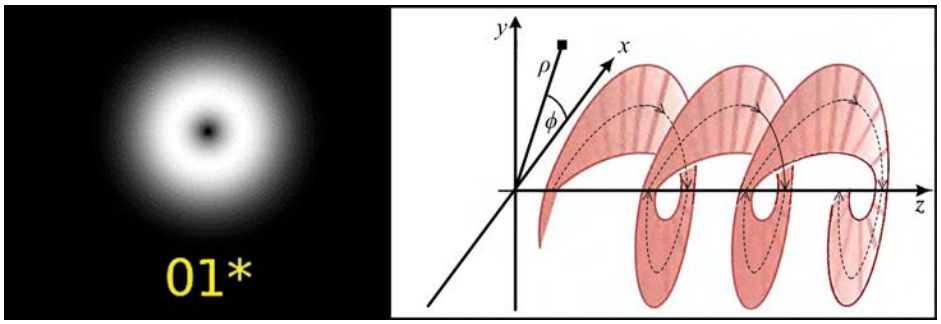


Figure 41: Intensity distribution (left) and wavefront (right) of an optical vortex. Dashed black line marks phase front rotation pattern.

Mode conversion in optical resonators is one way to obtain such peculiar laser beams which have several specific and unique applications. However, in depth analysis of optical vortices is beyond the scope of this textbook.

3.4 Resonance frequencies

Up to now we discussed spatial resonator modes which define the intensity distribution of radiation in the plane perpendicular to beam propagation direction. It was also assumed that radiation is monochromatic. However, laser resonator can sustain many temporal frequencies (Fig.42) which are defined by the requirement that phase change $\Delta\varphi$ of a light wave after one round-trip must be multiple of 2π:

$$\boxed{\Delta\varphi = q\pi}, \tag{70}$$

where q is an integer number called longitudinal mode index.

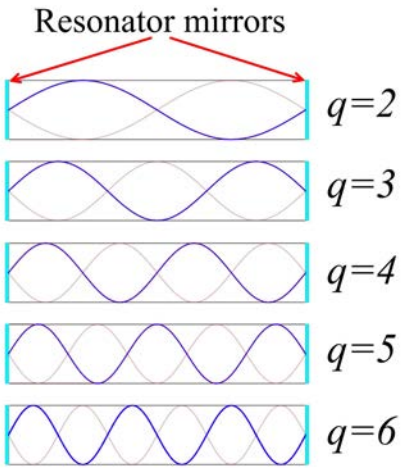


Figure 42: Standing waves in plane parallel resonator and the corresponding longitudinal mode index q.

In other words, Eq.(70) represents a standing wave condition and frequencies that satisfy this condition are called resonance frequencies or **longitudinal resonator modes**.

Let us determine how different transverse and longitudinal modes fit together, i.e., determine the resonance frequencies of a given spatial resonator mode. Assume a spherical mirror resonator where mirrors are placed at z_1 and z_2. The resonance condition for spatial mode with mn indices can be written as a difference of longitudinal phases provided that during reflection from a mirror the phase does not change:

$$\phi_{\mathrm{mn}}(z_2) - \phi_{\mathrm{mn}}(z_1) = q\pi, \tag{71}$$

where $\phi_{\mathrm{mn}}(z)$ is the longitudinal phase of Hermite-Gaussian beam:

$$\phi_{\mathrm{mn}}(z) = kz - (m+n+1)\arctan\frac{z}{z_R}. \tag{72}$$

Inserting the latter expression into Eq.(71) and defining the resonator length as $l = z_2 - z_1$, we get

$$k_q l - (m+n+1)\quad \arctan\frac{z_2}{z_R} - \arctan\frac{z_1}{z_R}\quad = q\pi. \tag{73}$$

Setting m and n constant, the former expression can be rewritten as

$$k_{q+1} - k_q = \frac{\pi}{l}, \tag{74}$$

and expressing $k = 2\pi\nu n_0/c$, where n_0 is the refractive index of the medium filling the resonator, and transferring into frequency domain we get

$$\boxed{\nu_{q+1} - \nu_q = \frac{c}{2n_0 l}.} \tag{75}$$

This frequency difference is called the **intermodal distance** (Fig.43), which is inversely proportional to resonator length. Assuming that m and n are equal to 0 (the case of a Gaussian beam), the intermodal distance is independent of resonator configuration.

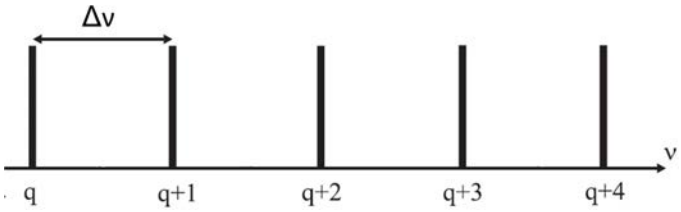


Figure 43: Intermodal distance in the case of a TEM_{00} mode (Gaussian beam).

Without going into mathematical details, it can be summarized that in a more complex case, i.e. for higher order transverse modes, the intermodal distance changes and the amplitudes of distinct longitudinal modes are not equal and the characteristics depend on the type (or configuration) of resonator. This can be understood from very simple assumptions: due to complex field amplitude distribution such modes can be expanded as a certain set of plane waves, where individual plane waves travel a slightly different distances between the resonator mirrors compared to the TEM_{00} mode (Gaussian beam), so standing waves in the resonator occur for slightly different frequencies. Furthermore, it is important to note that *a resonator sustains only a single transverse mode, i.e., electromagnetic field amplitude or intensity distribution in a plane perpendicular to propagation axis, which in turn is composed of many longitudinal modes (resonance frequencies).*

A set of longitudinal modes comprises spectrum of the laser radiation. However, it is important to note that here we analyzed an ideal "empty" resonator which can sustain an infinite number of longitudinal modes (the index q can be from 0 to infinity and their width is infinitively narrow, as depicted in Fig.43. In fact, *any real laser supports (amplifies) only a finite number of longitudinal modes, which fit under gain contour (the laser bandwidth) of the active medium.* For example, a fine structure of He-Ne laser spectrum which consists of 6 discrete longitudinal modes is shown in Fig.44. The overal spectral intensity distribution is described by a Gaussian function, therefore such spectrum is called **Gaussian spectrum** despite the fact

that it actually consists of only 6 discrete frequencies (longitudinal modes).

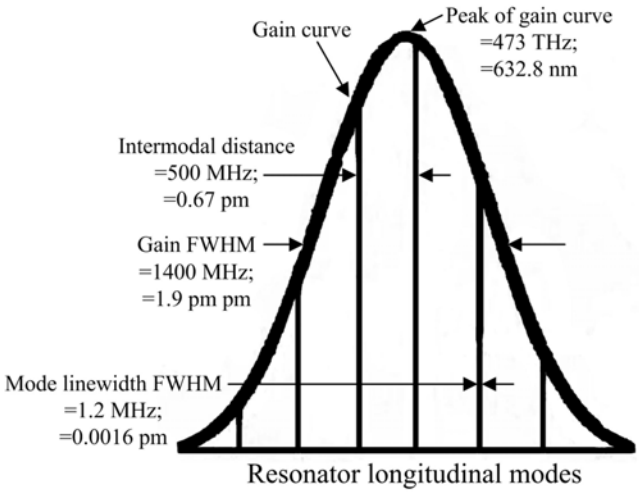


Figure 44: Example of multimode He-Ne laser spectral structure with 500 MHz intermodal distance.

The intensity peaks corresponding to intrinsic resonator frequencies (longitudinal modes) are rarely detected during laser radiation measurements. Since the intermodal distance is inversely proportional to the resonator length, longitudinal modes can be resolved only in the case of short resonators, e.g., with lengths of a few mm. In the case of long resonator with a meter length or more, the intermodal distances are very small, the adjacent modes are very close, so very often the resolution of the spectral device is insufficient to resolve them. To do this, interferometers with very high spectral resolution have to be used.

3.5 Fabry-Perot resonator

So far, the longitudinal modes were considered as monochromatic. In reality, these modes have a certain spectral bandwidth (linewidth). Let us find out what is the spectral bandwidth of a longitudinal mode and what determines its width. For this purpose let us analyze wave propagation in a Fabry-Perot etalon, which is a prototype of an optical resonator. Here we note that the first ruby laser had exactly this type of resonator, and currently Fabry-Perot type resonators are commonly used in solid-state microlasers and semiconductor lasers..

The Fabry-Perot etalon (resonator) is a piece of transparent material with parallel surfaces, whose thickness is l and refractive index is n_0. Let us see how such etalon transmits and reflects light waves. We mark A_i as

the complex amplitude of the incident wave, while $B_1, B_2, ...$ and $A_1, A_2, ...$ as complex amplitudes of reflected and transmitted waves, respectively, as shown in Fig.45.

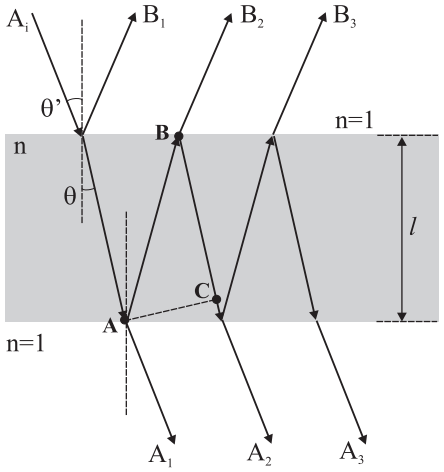


Figure 45: Light wave propagation in Fabry-Perot resonator.

The amplitudes of reflected waves can be written as

$$\begin{aligned} B_1 &= rA_i \\ B_2 &= tt'r'A_i e^{i\delta} \\ B_3 &= tt'r'^3 A_i e^{2i\delta} \ ... \end{aligned} \tag{76}$$

etc, where r and r' are external and internal reflection coefficients for wave amplitudes, t and t' are respective transmittance coefficients for wave amplitudes when the wave propagates from the etalon and vice versa. δ is the phase shift, which can be expressed from simple geometric assumptions through the optical path difference $(\delta l) = |AB| + |BC|$:

$$\delta = \frac{2\pi n_0 (\delta l)}{\lambda} = \frac{4\pi n_0 l \cos\theta}{\lambda}. \tag{77}$$

The full reflected wave complex amplitude is $A_r = B_1 + B_2 + B_3 + ...$:

$$A_r = \left[r + tt'r'e^{i\delta} \ \ 1 + r'^2 e^{i\delta} + r'^4 e^{i2\delta} + ... \right) \Big] A_i. \tag{78}$$

In the analogy, we can write transmitted wave complex amplitudes:

$$\begin{aligned} A_1 &= tt'A_i \\ A_2 &= tt'r'^2 A_i e^{i\delta} \\ A_3 &= tt'r'^4 A_i e^{2i\delta} \ ... \end{aligned} \tag{79}$$

etc, and the full transmitted wave complex amplitude is $A_t = A_1 + A_2 + A_3 + ...$:

$$A_t = tt' \ 1 + r'^2 e^{i\delta} + r'^4 e^{i2\delta} + ... \ A_i. \tag{80}$$

Assuming no energy losses (due to absorption or scattering) in reflecting surfaces (which essentially are resonator mirrors) and in the resonator itself, we can claim that $r = r'$ and at the same time energy conservation law is also satisfied, which in this case can be expressed as $r^2 + tt' = 1$. Introducing reflection and transmittance coefficients for wave intensities $R = r^2$, $T = tt'$ and $R + T = 1$ and summing infinite series, we get

$$\begin{aligned} A_r &= \frac{1 - e^{i\delta} \ \sqrt{R}}{1 - Re^{i\delta}} A_i \\ A_t &= \frac{T}{1 - Re^{i\delta}} A_i. \end{aligned} \tag{81}$$

Let us rewrite the latter expressions as wave intensity ratio assuming that $I = AA^*$. The ratio of reflected and incident wave intensities then is

$$\frac{I_r}{I_i} = \frac{A_r A_r^*}{A_i A_i^*} = \frac{4R \sin^2 \frac{\delta}{2}}{(1-R)^2 + 4R \sin^2 \frac{\delta}{2}}. \tag{82}$$

In a similar manner, the ratio of transmitted and incident wave intensities is:

$$\frac{I_t}{I_i} = \frac{A_t A_t^*}{A_i A_i^*} = \frac{(1-R)^2}{(1-R)^2 + 4R \sin^2 \frac{\delta}{2}}. \tag{83}$$

Since there are no losses in the resonator, it is easy to determine that $I_t + I_r = I_i$. From expression in Eq.(83) it follows that a maximum transmittance of the Fabry-Perot resonator is when $\delta = 2\pi m$, where m is any integer number. Then $I_t/I_i = 1$ and does not depend on the reflection coefficient R. On the other hand, we can notice that maximum Fabry-Perot resonator transmittance coincides with a standing wave condition (Eq.70) and intrinsic frequencies can be expressed as

$$\nu_m = m \frac{c}{2n_0 l \cos\theta}, \tag{84}$$

and the intermodal distance is

$$\boxed{\Delta\nu_m = \nu_{m+1} - \nu_m = \frac{c}{2n_0 l \cos\theta}}. \tag{85}$$

Assuming normal incidence ($\cos\theta = 1$), we get the same expression as in Eq.(75). Consequently, the maximum transmittance of the Fabry-Perot etalon (resonator) is for intrinsic longitudinal modes. In other words, these modes experience no losses inside the resonator and have infinitely long lifetime there. On the other hand, the minimum transmittance depends on R. Fig.46 illustrates the transmittance of Fabry-Perot resonator as a function of mirror reflection coefficient.

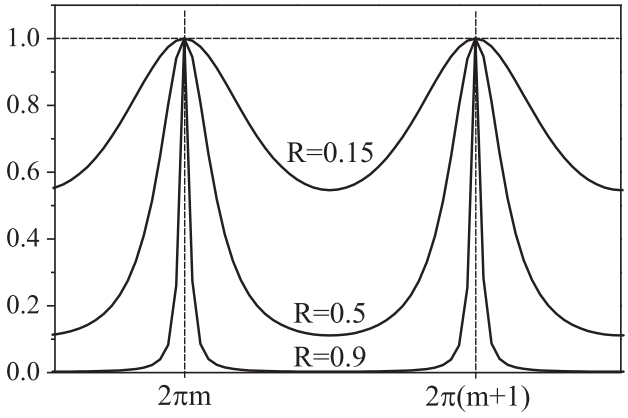


Figure 46: Fabry-Perot resonator transmittance as a function of mirror reflection coefficient.

In the case of large reflection coefficient, the transmittance maxima are sharp and minima are deep. When mirror reflectance decreases, the maxima spread and the amplitudes of minima increase. In that regard it is convenient to define the spectral width of a longitudinal mode which depends on the resonator length and mirror reflectance coefficient:

$$\boxed{\delta\nu = \frac{c}{2n_0 l\cos\theta}\,\frac{1-R}{\pi\sqrt{R}}}. \tag{86}$$

Fabry-Perot etalon can be also used as a spectral device: the intermodal distance can be varied by adjusting l and the spectral linewidth can be varied by changing R. In practice, Fabry-Perot etalon is inserted into a resonator of a CW laser when single longitudinal mode operation and very narrow spectral bandwidth is desired. Fig.47 shows an example of He-Ne laser spectral structure and a Fabry-Perot etalon modal structure. The infinite set of resonator longitudinal modes is limited by the width of the lasing medium gain curve and the Fabry-Perot etalon transmittance: without the Fabry-Perot etalon there would be a number of longitudinal modes in the emitted He-Ne

laser spectrum, but the Fabry-Perot etalon filters most of them. Leaving just a single longitudinal mode that falls under the gain curve.

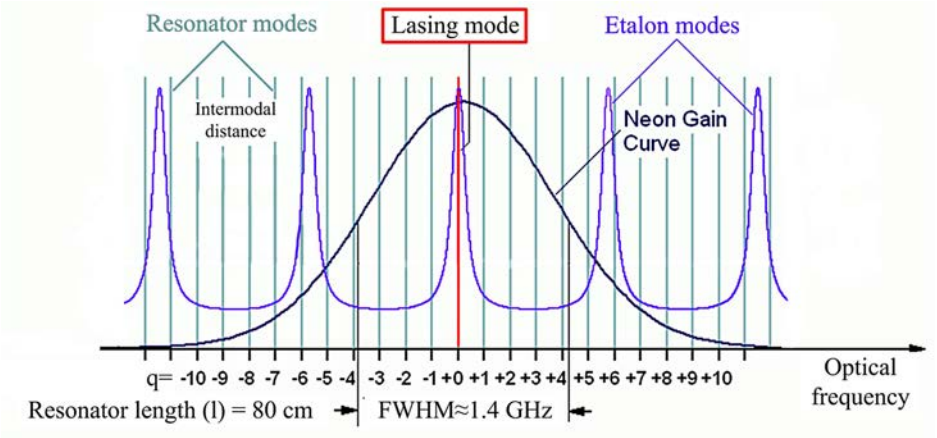


Figure 47: He-Ne laser spectral structure and Fabry-Petor etalon mode structure. The etalon filters longitudinal modes effectively forcing He-Ne laser to emit only a single longitudinal mode.

For ultrashort pulse (broad bandwidth) lasers Fabry-Perot etalon is not recommended since beatings between the intrinsic etalon and laser resonator modes occur, resulting in modulation of radiation spectrum and deterioration of the pulse temporal shape. This phenomenon is called the **etalon effect**. More generally, the etalon effect can be caused by any other optical element in the resonator which has parallel surfaces perpendicular to the optical axis of resonator. For this reason all surfaces of optical elements in the resonator (except for resonator mirror surfaces) are adjusted or cut at a slight angle with respect to the optical axis of resonator (e.g., a few degrees).

3.6 Losses in resonators

In order to understand optical resonator operation, it is important to evaluate sources of energy loss, i.e., find out where the electromagnetic wave energy dissipates. Resonator losses can be determined using various parameters: losses per round-trip, resonator quality or photon lifetime in a resonator. In this section we will provide simple evaluations of these parameters and we will also show how they are mutually related.

Firstly, let us determine the lifetime of a photon in a resonator. Assume that initially we have a number of photons circulating in the resonator N_{p0}. If the reflection coefficients of resonator mirrors are r_1 and r_2, respectively, after one round-trip the number of photons decreases to $r_1 r_2 N_{p0}$ and the lost

photon number is $(1 - r_1 r_2)N_{p0}$. The full round-trip time in a resonator is $2n_0 l/c$, so the evolution of photon number in time can be expressed as

$$\frac{dN_p}{dt} = -\frac{1 - r_1 r_2}{2n_0 l/c} N_p. \tag{87}$$

The solution of this equation is

$$N_p(t) = N_{p0} \exp \quad -\frac{t}{t_p} \quad , \tag{88}$$

where t_p is the photon lifetime in a resonator:

$$\boxed{t_p = \frac{2n_0 l/c}{1 - r_1 r_2}}. \tag{89}$$

The former expression can be interpreted in the following way: the **photon lifetime** in a resonator equals to a resonator round-trip time divided by round-trip losses. On the other hand, photon lifetime can be related with characteristic resonator mode extinction time:

$$\frac{dE}{dt} = -\frac{E}{t_p}, \tag{90}$$

where E is the energy of a resonator mode. **Resonator quality** is usually defined as a ratio of accumulated energy and dissipated power:

$$\boxed{Q = \frac{\omega E}{P} = -\frac{\omega E}{dE/dt}}, \tag{91}$$

where ω is the frequency of radiation. Comparing both expressions we get that $Q = \omega t_p$. Now let us discuss the causes of optical losses in resonators. These are produced by several mechanisms:

- **Losses due to reflections**. These losses are unavoidable since mirror reflections are not perfect, e.g., aluminum mirrors reflect of around 96% incident energy in the visible and near-infrared (an even less in the ultraviolet). The mirrors with dielectric coatings provide reflection coefficients as high as 99.9%. On the other hand, lower reflection of one of the resonator mirrors is set by purpose to enable the radiation to come out of the resonator. Such semitransparent mirrors are called output couplers. This is especially important for CW lasers which do not have any other elements for radiation control in the resonator. A certain amount losses are introduced by reflections from the surfaces of

optical elements inside the resonator. For example, Fresnel reflection from a single uncoated fused silica lens surface is roughly 4% and if there are multiple such elements, considerable losses are accumulated during a single round-trip. Therefore, surfaces of all elements inside the resonator are coated with anti-reflection dielectric coatings which reduce losses related to reflections by several or dozen of times. Moreover, reduction of reflection losses is also achieved by positioning elements at a Brewster angle when laser radiation is polarized.

- **Losses due to absorption and scattering**. Firstly, if the quantum defect (the difference between pump and laser photon energies) of a laser is small, i.e., the energy levels from which lasing occurs and the energy levels to which pump is absorbed are close, a fraction of generated radiation is absorbed in the active element itself. Secondly, optical elements in the laser resonator have to be transparent to laser radiation, but sometimes absorption losses due to impurity-related energy levels or due to being close to absorption band edge are notable. For ultraviolet lasers, the durability of resonator optical element maintaining the same high quality is of paramount importance. At long UV exposure times, high energy UV photons in some transparent materials (e.g. fused silica, etc.) can create long living color centers, which have large absorption. Thirdly, there is always scattering of light form dust, dirt or imperfections of optical elements. In that case light is scattered at large angles and does not return to the resonator. Losses due to scattering increase significantly when surface or volume of the optical element is damaged.

- **Diffraction losses** play a double role. On the one hand, when diffraction losses can be decreased or increased by selecting a certain resonator configuration: mirror radius of curvature and resonator length. Physical apertures of elements inside the resonator are often smaller than apertures of the resonator mirrors, so these diffraction losses due to the former must be taken into account. On the other hand, diffraction losses help to suppress higher order TEM modes whose generation is usually not desirable. Finally, diffraction losses, as a rule, are the largest in long resonators, but it is the diffraction-related spatial filtration of transverse modes which enables formation of light beams with very high spatial quality (in simple words, all beam irregularities due to not ideal optical surfaces, dust, etc are filtered out due to diffraction). Lasers with short resonators and low diffraction losses (e.g., semiconductor lasers and laser diodes) generate low spatial quality beams, their

M^2 is large and such beams are difficult to focus.

In conclusion, resonator losses are important not only when constructing high quality resonators: without them it is impossible to set the control on the output parameters of laser radiation. Diffraction losses enable generation of high quality light beams. By modulating losses in the resonator, pulsed laser operation and ultrashort pulse generation are performed. This last issue will be discussed in the coming chapters.

Summary of Chapter 3

- Optical resonator is an indispensable component of any laser, that is responsible for energy accumulation, feedback and formation of spatial and spectral characteristics of the laser radiation.
- The resonator stability condition is equivalent to convergence condition of light rays and beams, which states that a light field is confined in a resonator after infinite number of round-trips.
- Transverse (spatial) mode of a resonator is a certain distribution of light electric field that reproduces itself after each round-trip.
- The lowest order transverse mode is Gaussian beam. Higher-order transverse mode are Hermite-Gaussian and Laguerre-Gaussian beams, which possess complex intensity distributions.
- Longitudinal mode (resonance frequency) of a resonator is the light wave which satisfies standing wave condition.
- At a time laser resonator sustains many longitudinal modes, but only a single transverse mode.
- Any real resonator experiences energy losses due to a variety of reasons: reflections, diffraction, absorption and scattering. The resonator quality could be expressed by the photon lifetime in the resonator.

4 Light interaction with atomic systems

This chapter addresses the fundamental light-matter interaction processes that govern laser operation. Various general aspects of laser media, such as laser gain, amplification bandwidth, laser energy level systems and pump sources are discussed.

4.1 Einstein's treatment of spontaneous and induced transitions

In 1900 M. Planck postulated that light can be absorbed or emitted in the form of certain energy portions, called quanta. In addition, N. Bohr proved that electrons in atoms and molecules can occupy energy levels which can have only quantized energy states. At that time it was believed that light and matter interaction can by fully ascribed by the two processes: **absorption** and **emission**. In 1916 A. Einstein investigated the fundamental aspects of electromagnetic radiation interaction with matter and discovered that a steady state of electromagnetic field and matter is possible only when yet another previously unknown process is taken into account, which was called **stimulated (induced) emission**. Let us discuss these processes in more detail.

If the atom is in an excited state (the electron occupies a higher energy level), it spontaneously returns from the higher energy level E_2 to the lower energy level E_1, which for the sake of simplicity is regarded as the ground level. This transition is accompanied by emission of a photon, whose energy is equal to the difference between the energy levels (as schematically depicted in Fig.48):

$$\boxed{E_{ph} = h\nu = E_2 - E_1}. \tag{92}$$

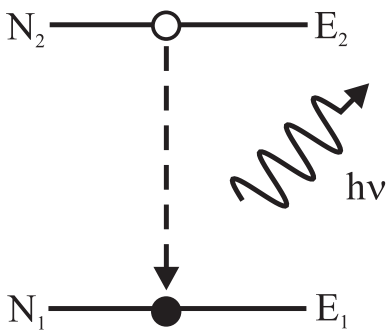


Figure 48: Spontaneous emission of radiation.

This process is spontaneous and is called spontaneous emission. The photon is emitted in a random direction and the corresponding light wave has random phase and polarization. The probability of such transition is represented by **Einstein's coefficient** A_{21}, which determines the probability of transition from higher to lower energy level per unit time.

Provided that the number of atoms in excited level per unit volume of material is N_2 (this is called **population**), the number of atoms returning to the ground energy level is $N_2 A_{21}$. The equation describing evolution of the excited level population can be written as

$$\frac{dN_2}{dt} = -A_{21} N_2, \tag{93}$$

and its solution is

$$N_2(t) = N_{20} \exp(-A_{21} t), \tag{94}$$

where N_{20} is the initial ($t = 0$) population of the excited level. So the population of excited level N_2 decreases in time with a characteristic time constant $t_{21} = t_{\text{spont}} = 1/A_{21}$. It is obvious that in a two-level system, which is illustrated in Fig.48, the population of ground level N_1 increases with the same time constant. The time t_{spont} is called **lifetime** of the excited level, which is inversely proportional to the probability of the corresponding electron transition. This probability is very different for energy levels of various atoms and molecules. Typically, probabilities for transitions allowed by the selection rules are of the order of $10^6 \div 10^8$ s^{-1}, or, in other words, lifetimes of such excited energy levels are in the range of $10^{-6} \div 10^{-8}$ s. If the electron transitions are forbidden (there are no absolutely forbidden transitions in nature; the probability of such transition is simply much smaller), the lifetime of such energy level is much longer and may vary from 10^{-4} s to even a few seconds. Such energy levels are called **metastable**. Their lifetime depends

on the physical state of the matter (e.g., in rarified gases metastable levels can exist for tens of seconds and even longer). Note that t_{spont} is a material-related characteristic. In general, lifetime of an energy level depends on the frequency of radiated photon: the greater is the frequency (also meaning greater energy difference between the levels), the shorter is the corresponding lifetime. This is a logical result since electrons in atoms always try to occupy the state with a lowest energy.

According to the uncertainty principle, the real energy levels are broadened. This means that these energy levels emit not infinitely narrow (monochromatic) light, but a spectrum with finite bandwidth, which is characterized by a certain linewidth (the so-called natural linewidth) which is described by a spectral function $g(\nu)$:

$$\int_0^\infty g(\nu)d\nu = 1. \tag{95}$$

The quantity $g(\nu)d\nu$ indicates the probability that the frequency of emitted photon is in the frequency interval $\nu + d\nu$, and the spectral intensity of radiation is defined as $I(\nu) = I_0 g(\nu)$.

Einstein discovered that besides the spontaneous emission there also exists another form of emission of radiation, which was called the stimulated emission. This means that excited atoms can return to the ground (unexcited) state not only spontaneously but also can be forced to do that. Consider an excited atom (or an ensemble of atoms) which is in the field of external radiation that has spectral energy density $\rho(\nu)$. If there are frequencies in the external field that match the frequency of transition from excited to the ground energy level, stimulated (induced) transitions occur, as shown in Fig.49. An important property of stimulated emission is that **emitted photon has the same direction and energy as the incident photon of the external radiation field**, and the corresponding electromagnetic wave has the same polarization and phase as the incident one.

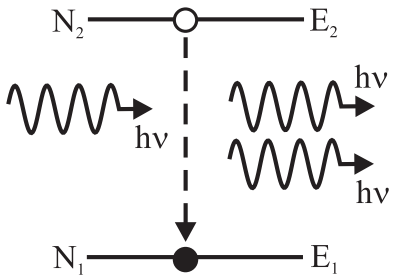


Figure 49: Stimulated emission of radiation.

The rate of stimulated transitions can be expressed as $B'_{21}(\nu)N_2$, where $B'_{21}(\nu)$ is a specific function describing the probability of such transition, and is expressed as $B'_{21}(\nu) = B_{21}g(\nu)$. Here B_{21} denotes the **Eintein's coefficient for stimulated emission**. In this case, the evolution of excited energy level population can be expressed by the following equation

$$\frac{dN_2}{dt} = -B_{21}N_2 \int_0^\infty g(\nu)\rho(\nu)d\nu = -N_2B_{21}\rho(\nu), \tag{96}$$

for simplicity, assume that $\rho(\nu)$ is a narrow line and $\rho(\nu) = \rho(\nu_0)$. Quantity $B_{21}\rho(\nu)$ is called rate of stimulated emission.

Now let us consider absorption. Since absorption is a process which occurs only in the presence of external field, it is a stimulated (induced) process, which is depicted schematically in Fig.50.

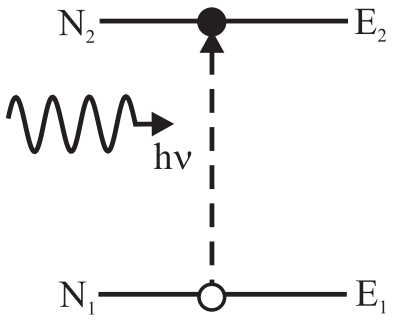


Figure 50: Stimulated absorption of radiation.

During stimulated absorption, a photon, whose energy is equal to the energy difference between the energy levels of interest, is absorbed and the atom becomes excited. Then the evolution of the ground energy level population N_1 can be described as:

$$\frac{dN_1}{dt} = -B_{12}N_1 \int_0^\infty g(\nu)\rho(\nu)d\nu = -N_1B_{12}\rho(\nu), \tag{97}$$

where B_{12} is the **Einstein's coefficient for (stimulated) absorption** and has a meaning of absorption probability.

Since all processes in nature must obey energy conservation laws, the discussed light matter interaction processes, which essentially represent energy transfer via certain mechanism, are related. Let us find the relation between the Einstein's coefficients. To make the task simpler, consider a two-level system, as illustrated in Fig.51.

Einstein's theory is based on several important assumptions:

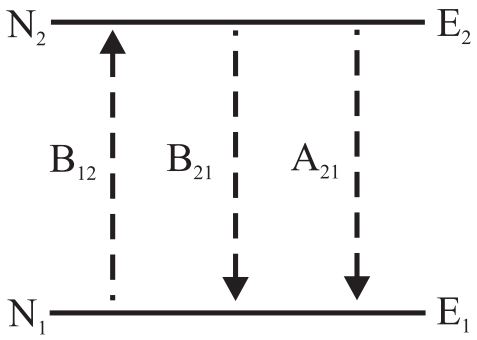


Figure 51: Equilibrium between stimulated absorption, stimulated emission and spontaneuos emission in a two-level system.

- Spectral density distribution of the external radiation corresponds to a black body radiation with temperature T:

$$\rho(\nu) = \frac{8\pi h\nu^3}{c^3} \quad \frac{1}{e^{\frac{h\nu}{kT}} - 1} \quad = \frac{8\pi\nu^2}{c^3} \times h\nu \times \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \tag{98}$$

  Here the first multiplier denotes the number of modes within a frequency interval, the second denotes the photon energy and the third denotes the modal population factor, i.e., number of photons per mode;

- Population distribution of the atom system obeys the Boltzman distribution:

$$\frac{N_2}{N_1} = \exp \quad -\frac{h\nu}{kT} \quad ; \tag{99}$$

- There is a thermodynamic equilibrium between the atom system and radiation, which means that the population of energy levels does not change when temperature is stable.

In the case of thermodynamic equilibrium we have

$$\frac{dN_2}{dt} = \frac{dN_1}{dt} = 0. \tag{100}$$

Let us also assume that radiation energy density of a black body has very small variations in the frequency domain of interest, which is determined by the electron transitions between excited and ground energy levels. Then the resulting population of an excited level is described as an equilibrium of stimulated absorption, stimulated emission and spontaneous emission:

$$\frac{dN_2}{dt} = N_1 B_{12}\rho(\nu) - N_2 B_{21}\rho(\nu) - N_2 A_{21} = 0. \tag{101}$$

Inserting expression of $\rho(\nu)$ we get:

$$N_2\left[B_{21}\frac{8\pi h\nu^3}{c^3\ e^{h\nu/kT}-1} + A_{21}\right] = N_1 B_{12}\left[\frac{8\pi h\nu^3}{c^3\ e^{h\nu/kT}-1}\right]. \tag{102}$$

With a reference to Eq.(99), we can rewrite the latter equation as

$$\frac{8\pi h\nu^3}{c^3\ e^{h\nu/kT}-1} = \frac{A_{21}}{B_{12}e^{h\nu/kT} - B_{21}}. \tag{103}$$

In this way we get one equation with three unknowns. It is satisfied when:

$$\boxed{B_{12} = B_{21}, \quad \frac{A_{21}}{B_{21}} = \frac{8\pi h\nu^3}{c^3}}. \tag{104}$$

These are **Einstein relations**. Let us discuss their meaning. From expressions in Eq.(104) it is clear that Einstein's coefficients A_{21}, B_{21} and B_{12} are mutually related: if we know at least one of them, we can find the other two. Another important thing to note is that coefficients for stimulated emission and stimulated absorption are equal (in the case of non-degenerate energy levels). This implies that stimulated emission and stimulated absorption are the opposite processes, and not spontaneous emission and stimulated absorption, as might seem at the first sight. If energy level population is $N_1 > N_2$, then external radiation field is absorbed and attenuated, while if $N_2 > N_1$, photons that appear as a result of stimulated emission add to the external radiation field which is thus amplified. A situation when ($N_2 > N_1$) is called **population inversion**. It is clear that in any laser the condition of population inversion must be satisfied.

Now let us discuss the role of spontaneous emission. The rate of stimulated emission depends on the energy density of external radiation field: $W_{21} = B_{21}\rho(\nu)$, while the rate of spontaneous emission does not. So if we ignore transitions related to spontaneous emission when describing the thermodynamic equilibrium between two energy levels, we get $N_2 B_{21} = N_1 B_{12}$ which, knowing that $B_{21} = B_{12}$, would yield $N_2 = N_1$; this is obviously impossible in the case of thermodynamic equilibrium.

An important aspect is the ratio between the rates of spontaneous and stimulated emissions, which can be defined as

$$R = \frac{A_{21}}{B_{21}\rho(\nu)} \simeq \exp \quad \frac{h\nu}{kT} - 1 \quad . \tag{105}$$

Let us compare this ratio in the case of microwave and optical range at equal ambient temperature, $T = 300$ K. Microwave frequency is roughly $\nu \approx 10^{10}$ Hz, then $h\nu/kT \approx 1.6 \times 10^{-3}$ and $R \approx 0.0016$. This indicates that in the microwave range the rate of stimulated emission significantly exceeds the rate of spontaneous emission. The opposite result is achieved in the optical range, assuming $\nu \approx 10^{15}$ Hz. Then we get $R \approx 160$. This means that the rate of the stimulated emission in the optical range is very low compared to that of the spontaneous emission, therefore, it is much harder to achieve laser amplification in this spectral range. With this in mind, it is easy to understand why masers were invented much earlier than lasers. It is also obvious that laser generation and amplification is much harder to realize in the blue and ultraviolet spectral range. Of course, emission rate ratio for metastable energy levels may be substantially different in the optical range but the general tendency is the same.

Although Einstein's relations were derived in the case of thermodynamic equilibrium, they are also valid in diverse conditions as transition probabilities are specific characteristics for a given material. In reality, a working laser is hardly in the thermodynamic equilibrium, but despite this, the meaning of Einstein relations remains the same.

4.2 Gain coefficient

Assume that a monochromatic wave with a frequency ν propagates in a two-level atom ensemble and the unit volume population of atom ensemble is as defined earlier: N_1 (the ground energy level) and N_2 (the excited energy level) which are measured in atoms per cm^3. For the sake of simplicity, assume that no spontaneous transitions occur in the atom ensemble (or their rate is very low compared to that of stimulated emission). In fact, spontaneous emission can be ignored if we analyze propagation of radiation at small angles with respect to the optical axis, e.g., in an optical resonator. Since the spontaneous emission occurs in random directions, its power is distributed evenly throughout the whole spatial angle and power density at a given direction (in this case, along the optical axis of the resonator) is very low. In contrast, the stimulated emission is directional. Expressing spectral intensity of the light field as $I_\nu = \rho_\nu/c$, let us find the change of intensity versus propagation distance z:

$$\frac{dI_\nu}{dz} = \ N_2 B_{21} g(\nu)\rho_\nu - N_1 B_{12} g(\nu)\rho_\nu \ h\nu. \tag{106}$$

Using Einstein relations Eq.(104) and definition of intensity we get

$$\frac{dI_\nu}{dz} = \ N_2 - N_1 \ \frac{c^2 A_{21}}{8\pi\nu^2} g(\nu) I_\nu. \tag{107}$$

The solution of this equation is

$$\boxed{I_\nu = I_\nu(0) \exp \ \gamma(\nu) z}, \tag{108}$$

where $I_\nu(0)$ is the initial intensity of radiation and $\gamma(\nu)$ is the **gain coefficient**, which is expressed as:

$$\boxed{\gamma(\nu) = \ N_2 - N_1 \ \frac{c^2 A_{21}}{8\pi\nu^2} g(\nu) = \Delta N \frac{\lambda^2}{8\pi t_{\text{spont}}} g(\nu)}. \tag{109}$$

The amplification of light is achieved when $\gamma(\nu) > 0$, while attenuation occurs when $\gamma(\nu) < 0$. The sign of gain coefficient depends only on the difference of energy level population ΔN, since all other coefficients are positive. If the ensemble of atoms is in a thermodynamic equilibrium, population of the energy levels obeys Boltzmann's distribution, then the following is always true: $N_2 < N_1$ and $\Delta N < 0$ when $T > 0$. Consequently, in a two-level system we always have only attenuation (absorption) of the external field and in this case $\gamma(\nu) < 0$ is simply equivalent to the absorption coefficient $\gamma(\nu) \equiv \alpha(\nu)$. On the other hand, even if the system is far from thermodynamic equilibrium and external energy source is present, the population inversion is still impossible because the rates of stimulated emission and stimulated absorption are equal. In such a two-level system of atoms, amplification is only possible when $T < 0$, which has no physical sense according to the classical (Boltzmann) definition of the absolute temperature. To this end, in the early studies, the population inversion was related to an effective negative absolute temperature.

It is clear that in order to achieve amplification of light (laser generation), a two-level system is insufficient to create the population inversion ($\Delta N > 0$). Leaving the issues how the population inversion is achieved in practice, we currently only note that population inversion is achieved not only by setting the system out of thermodynamic equilibrium (with the use of an external energy source) but also by involving additional energy levels.

4.3 Homogeneous and inhomogeneous spectral linewidth broadening

Although the spectrum of resonator longitudinal modes is infinitely broad, the expression of gain coefficient in Eq.(109) has the same frequency dependence as $g(\nu)$, which means that the laser has its own (intrinsic) frequency (amplification) bandwidth. Indeed, amplification bandwidth determines the frequency spectrum of laser radiation. Let us discuss how broad the gain bandwidth could be and which physical mechanisms define its spectral width.

Due to uncertainty principle, both emission and absorption (in that regard these processes could be considered identical) spectral linewidths are not infinitely narrow, they are broadened. The linewidth broadening can be of two types: **homogeneous** and **inhomogeneous**.

In the case of homogeneous broadening, excited atoms (or molecules) in the ensemble are treated as identical (undistinguishable). This means that spectrum of emitted radiation is identical for each atom and we cannot distinguish which atom exactly emitted one photon or another photon. The fundamental reason for homogeneous spectral broadening is the time-energy uncertainty. Moreover, spectral lines of emission are broadened due to elastic collisions between excited atoms with other atoms or phonons, all of which produce phase jumps of the emitted monochromatic waves. As a result of superposition of monochromatic waves with different phases, we get a certain temporal modulation which in turn makes emission spectrum not monochromatic, but broadened. There are several other reasons due to which homogeneous spectral line broadening occurs, or in other words, due to which additional spectral line broadening is induced. Firstly, the radiative and non-radiative transitions to other energy levels, which shorten the lifetime of the excited level and hence broaden the emission spectrum. Secondly, spectral line can additionally homogeneously broaden when an ensemble of atoms interacts with external field of radiation. This effect comes into play when gain saturation in the laser operation occurs.

A special type of homogeneous broadening occurs in certain laser media when transition metal ions are embedded into the structure of crystal lattice. Then the spectral line broadens due to crystal lattice oscillations (various phonon interactions) which reduce the lifetime of the excited level. Such oscillations occur in all crystalline media, but a considerable broadening of ion emission spectrum occurs only when the emitting energy levels are weakly shielded from the net crystal lattice field. This occurs exactly in the case of transition metal ions, e.g., Ti, Cr. Such lasers are often called vibronic and the most prominent example is the Ti ion doped sapphire which has a very wide homogeneously broadened emission band.

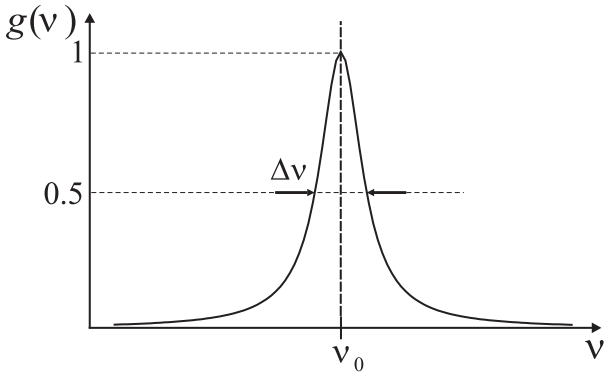


Figure 52: Shape and width of spectral line in the case of homogeneous broadening of energy levels.

In an ideal case the shape of **homogeneously broadened spectral line is described by the Lorentz function**:

$$\boxed{g_H(\nu) = \frac{\Delta\nu}{2\pi\ (\nu - \nu_0)^2 + (\Delta\nu/2)^2},} \tag{110}$$

where $\Delta\nu$ is the spectral linewidth at half intensity level as shown in Fig.52 and equals to

$$\Delta\nu = \frac{A_{21}}{2\pi} = \frac{1}{2\pi t_{\text{spont}}}. \tag{111}$$

Such spectral line broadening is identical for any atom in the ensemble and since we cannot distinguish between the radiating atoms, it is called homogeneous.

Inhomogeneous spectral line broadening is caused by a certain distribution of features of the individual atoms (and consequently, their energy levels). In this case we can distinguish between radiating atoms or their groups, as each atom or a group of atoms emit radiation with slightly different frequency. There are several physical reasons of inhomogeneous spectral line broadening, which are different for various laser materials. Here we will discuss several of them which are the most important and common.

In the case of solid-state lasers, there are two main reasons for inhomogeneous spectral line broadening. The first is attributed to amorphous structure of laser host material. Various glasses doped with rare-earth metal ions (Nd, Yb, etc) represent the most illustrative example. Glasses are amorphous materials, which from a macroscopic view are isotropic. However, from

the microscopic view of structure, they are not strictly isotropic but composed of many small ordered structures which are "frozen" in the isotropic matrix of the glass. Orientation and properties of distinct structures (such as mechanical stress, concentration of impurities and their energy levels) is slightly different, which in turn slightly modify the energy levels of ions that are embedded in the glass matrix. Then the overall radiation spectrum is a sum of all spectra emitted by these structures which consequently makes the spectrum relatively broad compared to that of homogeneously broadened line. For example, the spectral line of Nd ions embedded in a YAG crystal is homogeneously broadened, whereas the spectral line of the same Nd ions embedded in a glass is inhomogeneously broadened and its width is roughly 50 times broader.

Inhomogeneous spectral line broadening in crystalline materials occurs due to various defects and impurities in the crystal lattice, so each radiating ion is surrounded by a slightly different environment. However, in most cases such inhomogeneous broadening mechanism is relatively weak, as modern technologies enable production of laser materials with very high optical and chemical quality.

Several mechanisms of inhomogeneous spectral line broadening are specifically attributed to gas lasers. The first occurs in gas mixtures due to isotope impurities, which emit radiation with slightly shifted frequencies. The second mechanism is attributed to spectral line broadening due to Doppler effect. Random velocity distribution of gas atoms or molecules suggests that a stationary observer due to Doppler effect sees slightly shifted frequency of radiation, which depends on the direction of the moving particle with respect to the observer. Distribution of atom or molecule velocities depends on the mass and temperature and is described by the Maxwell-Boltzmann distribution function, which can be converted to the distribution function of radiation frequency. For an inhomogeneously broadened spectral line the function is

$$\boxed{g_I(\nu) = \left(\frac{4\ln 2}{\pi}\right)^{1/2} \frac{1}{\Delta\nu} \exp\left[-4\ln 2 \left(\frac{\nu-\nu_0}{\Delta\nu}\right)^2\right]}. \tag{112}$$

In the ideal case, the shape of the inhomogeneously broadened spectral line is a **Gaussian function** whose width is defined as

$$\Delta\nu = \nu_0 \left(\frac{8kT\ln 2}{Mc^2}\right)^{1/2} \tag{113}$$

and is a function of atomic mass M and temperature T. In the case of

inhomogeneous broadening, the spectral line can be viewed a superposition of multiple homogeneously broadened lines as shown in Fig.53.

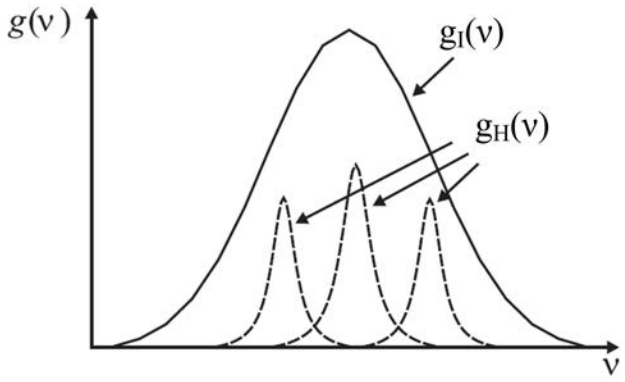


Figure 53: The shape of the spectral line in the case of inhomogeneous broadening which consists of multiple homogeneously broadened lines shown by dashed curves.

In many real laser systems both, homogeneous and inhomogeneous line broadenings are not symmetrical with respect to the central frequency and the emission spectrum of each laser material has a specific shape. The spectral line of any laser material is often broadened both homogeneously and inhomogeneously, but one of the broadening types is usually dominant. Spectral linewidth and its broadening strongly depends on the temperature, so lasers generating a very narrow spectrum of radiation operate only at low temperatures, e.g., cryogenic temperature. In conclusion, notice that in the case of homogeneous spectral line broadening, the linewidth of emitted spectrum is inversely proportional to the lifetime of the excited level, while for inhomogeneous spectral line broadening this assumption does not apply.

4.4 Three-level and four-level lasers

Now let us discuss possible energy level systems in the laser medium in which the population inversion can be achieved. In principle, all laser energy level systems in laser media, no matter how complex they are, can be approximated as either **three-level** or **four-level** systems regardless of the type of spectral line broadening.

To achieve laser operation, the energy levels must obey certain requirements. Let us first consider a **four-level laser**, whose energy level diagram is schematically illustrated in Fig.54. In a four-level laser, atoms by absorbing pump energy are excited from the ground level (0) to the level (3), which may consist of an entire system of energy levels. The lifetime of level (3)

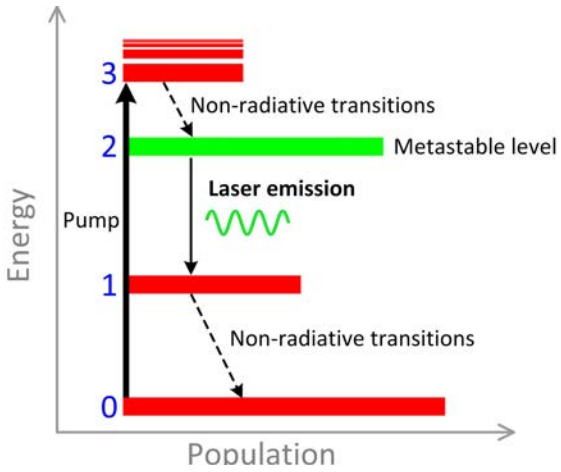


Figure 54: Principal diagram of a four-level laser system.

has to be short, it is also desirable that any radiative transitions from level (3) do not occur or their rate is very low. In this way, excitations of level (3) are rapidly transferred by non-radiative transitions to level (2), which is metastable, i.e. has long lifetime. Levels (2) and (1) are the upper and lower laser levels, respectively. To achieve population inversion between laser levels (2) and (1), lifetime of energy level (1) has to be short. i.e. its excitations rapidly relax to the ground state (0) by radiative and non-radiative transitions. Therefore, in a four-level laser, any excited atom in level (3) quickly transfers its excitation to level (2) and since level (1) is empty (very quickly relaxing), the population inversion is achieved by very simple means. Examples of four-level laser systems include Nd-doped solid-state lasers (e.g. Nd:YAG, Nd:glass), Ti:sapphire, excimer and most of gas lasers.

A different situation occurs in a **three-level laser** (Fig.55), where the lower energy level (1) matches the ground level. In this case, to achieve population inversion it is necessary to excite at least half of the atoms from the ground level to level (3). Obviously, this can be obtained only using sufficiently high pumping rate. The most prominent example of three-level laser is ruby laser.

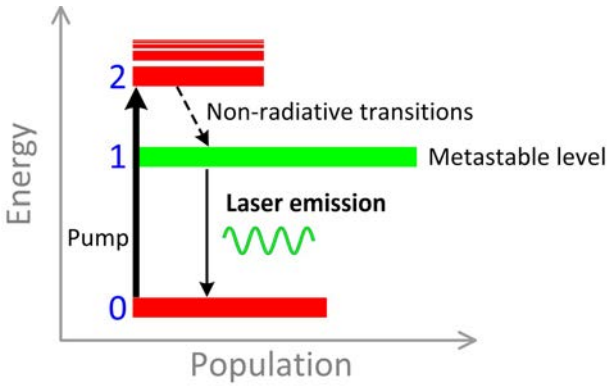


Figure 55: Principal diagram of a three-level laser system.

Compared to a three-level laser, four-level laser offers a great advantage that population inversion is ideally achieved when just one atom is excited to level (3) that consequently quickly decays to level (2). Therefore four-level lasers are much more used whenever possible.

More recently, the so called **quasi-three-level lasers** became an important laser category. In these lasers the energy level scheme is similar to that of four-level lasers (Fig.56), with the difference that energy levels (0) and (1) are very closely spaced (even partially overlapped). The advantage of such energy level configuration is small quantum defect; that is the energy difference between absorbed and emitted energy. Small quantum defect means that only a small amount of pump energy is converted to heat. On the other hand, since ground (0) and laser emission end (1) energy levels are closely spaced, pumping level (3) can absorb a fraction of laser radiation. This effect is called reabsorption. Examples of quasi-three level lasers are Yb-doped solid-state (Yb:YAG, Yb:KGW) and Yb-doped and Er-doped fiber lasers.

The most distinguished difference between four-level and quasi-three-level systems is that an unpumped four-level system is transparent to the laser light, while a quasi-three-level system has significant laser light absorption in that state: such laser system can absorb part of the radiation that it emits (reabsorption). Therefore, some level of pump intensity is required to obtain transparency of such system (when most electrons are excited in the upper energy levels, absorption of such medium significantly reduces effectively making it transparent), and laser gain is achieved only for intensities higher than that.

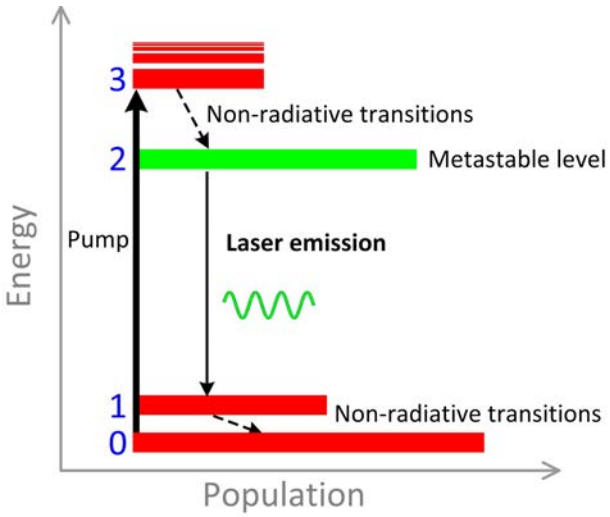


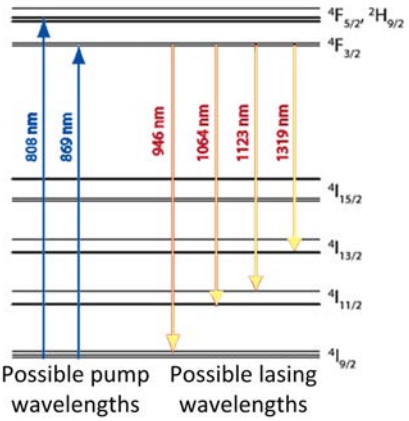
Figure 56: Principal diagram of a quasi-three-level laser system.

4.5 Lasing at multiple wavelengths

Due to complexity of the real laser medium energy level system, the same lasing medium may have several possible three-level or four-level combinations that enable lasing at multiple wavelengths. Nd ion energy level structure and possible lasing transitions are depicted in Fig.57. Nd:YAG has multiple possible laser transitions. The most important fact is that different transitions have different cross-sections (probabilities) with transition emitting at wavelength of 1064.1 nm being the most probable. In practice, the main lasing wavelength (which is usually associated with a particular lasing medium and offered in commercial lasers) is chosen to be that of the highest cross-section (probability): this corresponds to lowest laser generation threshold and highest gain. This is done by using resonator mirrors with high reflectivity for that wavelength.

The pump wavelength is usually chosen such that its absorption would be maximum. It is 808 nm is the case of Nd:YAG, but pumping at wavelength of 869 nm is also possible. Interestingly, as seen from Fig.57, in case of 869 nm pump wavelength Nd:YAG medium operates as a three-level system.

Another example of multiple lasing wavelengths is the He-Ne laser. It has several possible emission wavelengths at 543.36 nm, 593.93 nm, 604.61 nm, 611.8 nm, 632.8 nm, 1.15 μm, 3.39 μm, some of which are shown in Fig.58. Interestingly, the first demonstrated He-Ne laser operated at 1.15 μm wavelength, but later a more efficient transition was discovered at 632.8 nm, which is the common commercially offered He-Ne laser wavelength.



Wavelength, nm	Laser transition	Cross-section (10^{-19} cm^2)
1064.1	$^4F_{3/2} \rightarrow {}^4I_{11/2}, (R_2 \rightarrow Y_3)$	2.8
1338.1	$^4F_{3/2} \rightarrow {}^4I_{13/2}, (R_2 \rightarrow X_3)$	1.0
1318.7	$^4F_{3/2} \rightarrow {}^4I_{13/2}, (R_2 \rightarrow X_1)$	0.95
946	$^4F_{3/2} \rightarrow {}^4I_{9/2}, (R_1 \rightarrow Z_5)$	0.5

Figure 57: Energy level structure of Nd ion embedded in YAG crystal, pump and laser transitions.

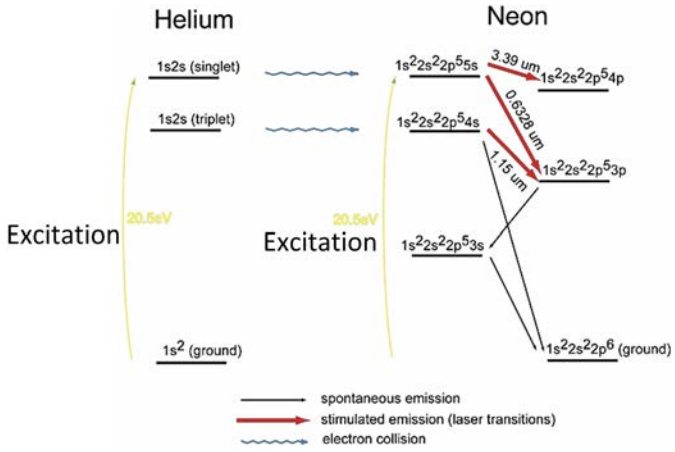


Figure 58: Simplified view of excited states of He and Ne atoms and possible transitions.

There are also some unique laser media, e.g., dye lasers, which possess an extremely broad emission spectrum. This is related to the energy structure of dyes. They are usually in liquid phase and their electronic states have multiple vibrational energy sub-levels. Lasing can occur between any (or all) of these vibrational energy levels, therefore, lasing is broadband (see Fig.59). This unique feature is important in femtosecond lasers: the shorter is the

pulse duration, the broader laser gain bandwidth is required to generate such pulse (see Eq.13). Dye lasers very historically the first femtosecond lasers.

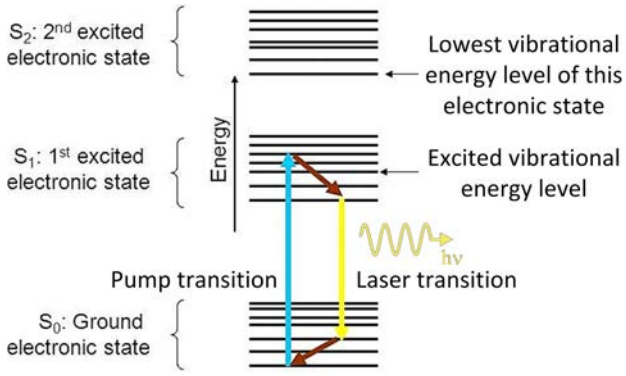


Figure 59: Simplified principal view of dye laser energy levels.

The relevant parameters of several widely used laser media are listed in Table 2. Gain bandwidth for various laser media differs by more than 6 orders of magnitude. Rhodamine 6G dye and Ti:sapphire crystal are two unique media with very large gain bandwidth.

Table 2: The main parameters of several laser media. λ_0 is the laser wavelength, σ is the laser transition cross-section, t_{spont} is is the lifetime of the upper laser level, $\Delta\nu$ is the gain bandwidth, I and H stand for inhomogeneous and homogeneous line broadening, respectively.

Laser medium	λ_0,μm	σ,cm^2	t_{spont},μs	$\Delta\nu$	Broadening
He-Ne	0.6328	5.8×10^{-13}	0.03	1.7 GHz	I
Ar^+	0.5145	2.5×10^{-13}	0.01	3.5 GHz	I
CO_2	10.6	3×10^{-18}	2.9×10^6	60 MHz	I
Nd:YAG	1.064	2.8×10^{-19}	230	135 GHz	H
Nd:glass	1.054	4.1×10^{-20}	350	8 THz	I
Ti:sapphire	0.8	3.8×10^{-19}	3.9	100 THz	H
Rhodamine 6G	0.57	2×10^{-16}	0.005	45 THz	H/I

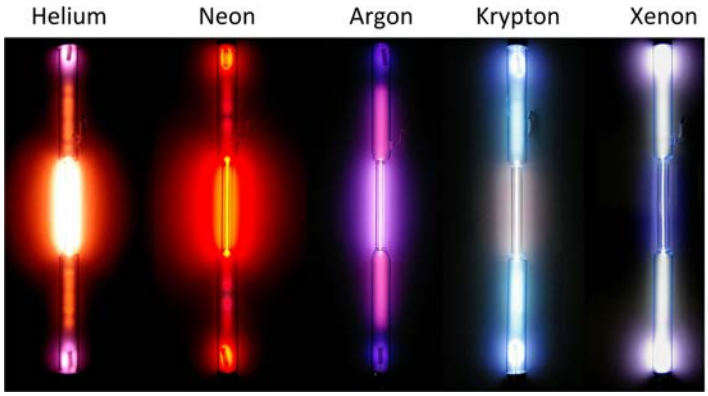
4.6 Pump sources

External energy sources of lasers are called **pump sources**. They are very different and their choice depends on the laser operation mode, properties of the laser medium and lifetime of laser energy levels. All pump sources, according to the physical mechanism of their operation, can be classified into three types:

- **Optical pumping**. The lasing medium is pumped by electromagnetic radiation (light). All solid-state lasers are pumped optically. A source of optical pumping can be:

  1) **A low pressure flashlamp** filled with noble gasses (Fig.60).

(a)



(b)

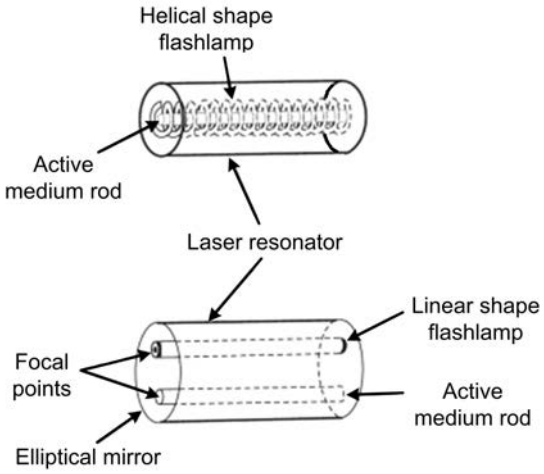


Figure 60: (a): noble gas-filled flashlamps used for optical laser pumping. Image adapted from Wikimedia under CC BY-SA 2.0 license. (b): solid-state laser resonator pumping geometries using helical or linear shape flashlamps.

Principle of operation of flashlamps is based on electric discharge. An electric discharge produces an electric current that travels through the tube and ionizes the gas inside. As ions recombine with electrons, photons are emitted. A flashlamp emits broadband (white light) radiation, hence multiple energy levels of the active laser medium are excited at the same time. This is not energetically efficient since, as

a rule, radiative and non-radiative transitions between various energy levels cause significant losses and unwanted thermal effects. Flashlamps are used to pump various Nd-doped lasers, as Nd ion has multiplicity of absorption lines in the UV, visible and near-infrared.

2) **A laser diode or another laser** (usually gas or solid-state). Laser diodes are semiconductor lasers with electrical pumping. It is important to note that the term "semiconductor laser" also includes optically pumped semiconductor lasers which are not to be confused with laser diodes. As compared with flashlamp pumping, diode pumping is much more efficient, since emission wavelength of a laser diode may be chosen to exactly match a single absorption line of the laser medium. Furthermore, laser diodes have better quality (focusability) of the beam compared to flashlamps, which enables advanced pumping setups with almost 100% pump consumption efficiency (Fig.61).

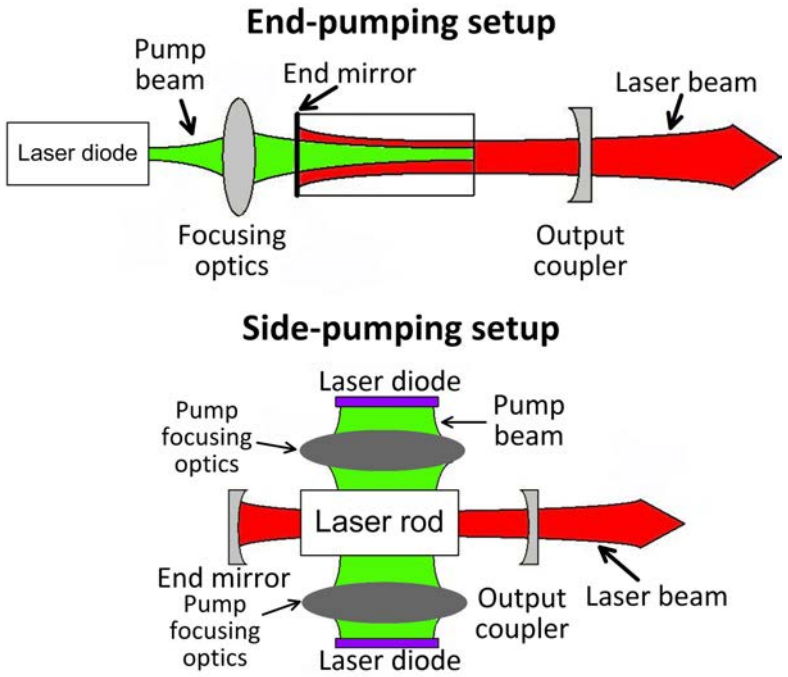


Figure 61: End- and side-pumping arrangements using laser diodes.

Most laser diodes emit in the near-IR and are of particular importance for Yb-doped laser pumping, as Yb ion has absorption lines only in the near-IR. Laser pumping is used only when the absorption lines or bands of the active laser medium are in the visible spectral range and excited laser energy levels have short lifetimes. This is the case for the

Ti:sapphire laser, which is pumped by the second harmonic of solid-state lasers operating in CW or Q-switch regime or by CW argon ion laser.

3) **The Sun**. Sun may be considered as an exotic continuous-wave white-light pump source, a kind of natural flashlamp analogue. Interestingly, the first laser pumped by the sunlight (Nd:YAG) was demonstrated back in 1966[11]. Currently Sun-pumped iodine, Nd:Cr:YAG and several semiconductor lasers have been demonstrated.

4) **Microwaves** can be used to pump certain gas lasers, e.g., the CO_2 laser. Along with the Sun, this is also a rather and rarely used pump source.

- **Electrical pumping**. This pumping type uses electrical particles to excite the lasing medium and achieve population inversion. Electrical pumping includes:

  1) **Electrical discharge**. This type of electrical pumping is used in gas lasers (e.g., He-Ne, Kr, etc), where high voltage electric discharge produces and accelerates free electrons, which transfer their energy to atoms or molecules of the gas via collisions, creating population inversion.

  2) **Electric current**. This type of electrical pumping is usually used in laser diodes. An electric current flowing through the p-n junction which injects electrons/holes into the depletion region. When an electron and a hole are present in the same region, they may recombine producing spontaneous emission (photons) which is used as initial light (seed) for amplification via stimulated emission.

  3) **Electron-ion beams**. Free electron lasers and some excimer lasers are pumped by electron/ion beams, which excite atoms or molecules in a similar fashion as electric discharge.

- **Chemical pumping**. Chemical lasers use this type of pumping when population inversion is achieved due to energy released during exothermal (emitting heat) chemical reactions.

4.7 Laser oscillation

From what we discussed earlier, we can easily formulate conditions for laser generation. In this context it is important to distinguish between the terms

[11] C. G. Young, A Sun-pumped CW one-watt laser, Applied Optics **5**, 993–997 (1966)

"amplifier" and "generator". If the population inversion ($\Delta N > 0$) and $\gamma(\nu) > 0$ are achieved, the medium acts as a **laser amplifier**: radiation that passes the medium is amplified (middle image of Fig.62). The resonator for a laser amplifier is not necessary. For the **laser generation** to occur, the laser medium needs to be inserted into a resonator (or made to work as a resonator) and laser oscillation condition needs to be fulfilled – this setup is called **a laser oscillator or simply a laser** (Fig.62 top).

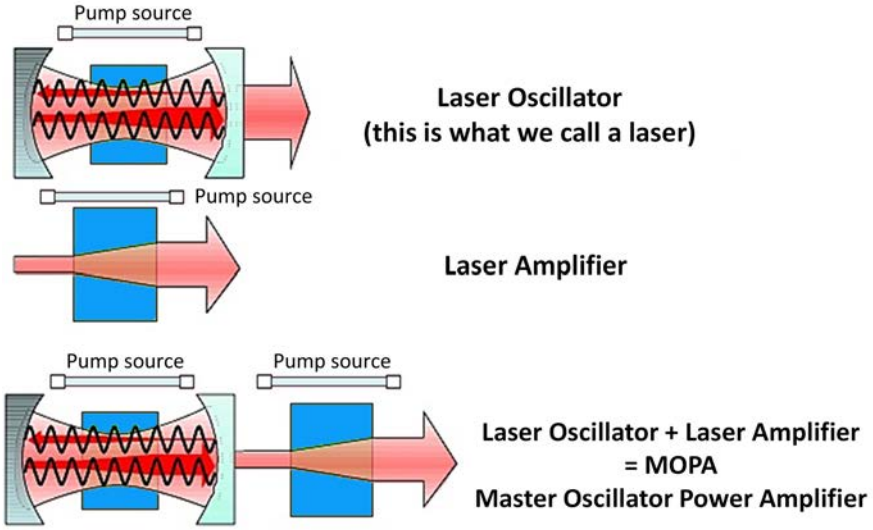


Figure 62: Principal setup of laser oscillator (generator), amplifier and MOPA system.

For laser amplification it is enough to create the population inversion and then the amplifier is able to amplify the launched signal. However, it is clear that achieving population inversion $\Delta N > 0$ and positive gain coefficient is not enough for laser oscillation in a resonator. Taking into account losses in the resonator, laser oscillation occurs only when amplification per round-trip exceeds losses per round-trip:

$$r_1 r_2 \exp\ 2\gamma_0(\nu) l_l\ \geq 1, \tag{114}$$

where l_l is the length of the laser medium and we also assume that the resonator losses due to diffraction, absorption and scattering are very small; r_1 and r_2 are coefficients of reflection for resonator mirrors. Then the gain coefficient necessary to achieve laser oscillation is

$$\boxed{\gamma_0(\nu) \geq \frac{1}{2l_l} \ln\ \frac{1}{r_1 r_2}\ = \alpha}. \tag{115}$$

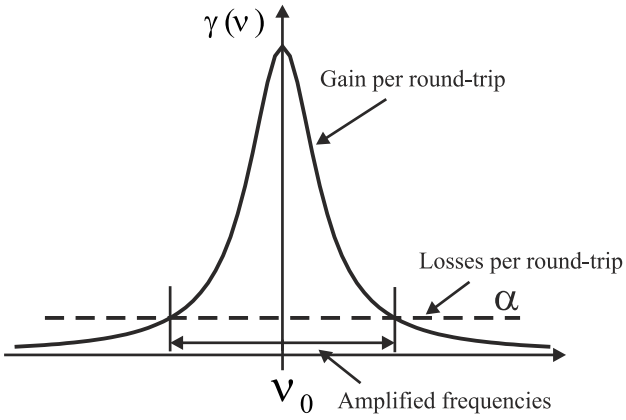


Figure 63: Frequency range of laser generation.

Having depicted the former expression in Fig.63, we notice that from all longitudinal modes supported by the resonator laser radiation consists only from those longitudinal modes that fall under the frequency interval in which expression in Eq.(115) is satisfied. In the case of equality, expression in Eq.(115) defines the laser oscillation threshold and the threshold pump rate as well.

Now let us discuss the key difference between the laser oscillator and the laser amplifier. Since the purpose of the laser amplifier is only to amplify a certain external (injected) signal, the laser amplifier does not require a resonator and laser amplification takes place as long as the population inversion $\Delta N > 0$ is achieved. In contrast, there is no such external signal in the laser oscillator. Therefore the essential question is from where the initial signal comes from? In a laser oscillator, once the generation threshold is exceeded, the oscillation starts from spontaneous emission, a small fraction of which occasionally propagates along the optical axis of a resonator. The resonator mirrors return the spontaneous emission back to the laser medium, where it is amplified and the further oscillation process is dominated by the stimulated emission. In this sense, the resonator performs the function of positive feedback amplifying only the radiation that propagates close to the optical axis and forming a high quality beam.

A combination of laser oscillator and amplifier (or multiple amplifiers) is called **master oscillator power amplifier – MOPA** (Fig.62 bottom). Such systems produce laser pulses with high energy.

Summary of Chapter 4

- The stimulated emission is the fundamental effect that produces laser radiation.
- A necessary condition for laser operation is population inversion: a state when population of electrons in a higher energy level exceeds population of electrons in a lower energy level, which cannot be achieved in a two-level system.
- All energy level systems in laser media, no matter how complex they are, can be approximated as either three-level or four-level systems.
- Laser gain bandwidth is broadened due to various physical mechanisms or fundamental limitations. The broadening can be homogeneous or inhomogeneous.
- Laser generation starts when laser gain exceeds losses per round-trip in the resonator. This is the laser oscillation condition.
- The purpose of laser pump source is to excite the medium in order to create population inversion. There are three types of laser pumping: optical, electrical and chemical.

5 Pulsed laser operation

Compared to all other (non-laser) light sources, a unique feature of lasers is that they emit light not only in the form of a coherent beam but also they can emit light in certain portions, called **light pulses**. Pulsed laser operation impleas that energy accumulated in the laser medium is emitted within a very short time interval, which is called pulse duration, or pulse width. With a reference to the fact that laser radiation is localized in space in the form of a light beam, pulsed laser operation allows achieving high instantaneous power (termed peak power) and high intensity (instantaneous power per unit of area).

There are three distinct types of pulsed laser operation (regimes): free-running, Q-switching and mode-locking. These operation regimes differ in terms of achievable pulse durations and their technical realization. Free-running is the simplest laser operation regime, which starts when the laser oscillation threshold is exceeded. If the pump is steady (continuous), free-running regime is also continuous (with several exceptions) and if the pump is pulsed, free-running becomes pulsed as well. Achievable pulse durations in the free-running regime are of the order of several microseconds (1 μs=10^{-6} s) and the peak power of such pulses can reach several kW. Q-switching and mode-locking regimes are always pulsed regardless of the pumping type. In Q-switching regime the laser generates light pulses with durations ranging from several to several tens of nanoseconds (1 ns=10^{-9} s), with peak powers reaching hundreds of MW. The shortest light pulses with durations ranging form 100 picoseconds (1 ps=10^{-12} s) to a few femtoseconds (1 fs=10^{-15} s) are generated in mode-locking operation regime. In this regime, the pulse duration depends on the properties of the laser medium (achievable gain bandwidth) and the methods of mode-locking, and the peak power of such pulses may reach several hundreds of MW as well. Light pulses with durations shorter than several picoseconds are called **ultrashort**. Both Q-switching and mode-locking regimes are usually employed in solid-state lasers since these lasers have low laser transition cross-section (see Table 2) and broad gain bandwidth, i.e., can accumulate large population inversion and generate broadband laser radiation. Other lasers, such as CO_2, excimer, dye and semiconductor lasers, can also operate in Q-switching or mode-locking regimes, but these are rare exceptions. The most powerful light pulses with peak powers reaching PW (10^{15} W) and higher are achieved by amplifying ultrashort pulses in laser amplifiers. Their principles and configurations will be discussed at the end of this chapter.

5.1 Free-running regime and relaxation oscillations

Free-running is the simplest laser generation regime which starts simply when laser generation threshold is exceeded. Even in the case of continuous pump, for a certain transient time interval periodic oscillations of laser intensity may occur. These oscillations are called **relaxation oscillations**. They are observed in all lasers and their period is significantly longer than the resonator round-trip time and even than the photon lifetime in the resonator. The reason of relaxation oscillations is the dynamic interaction between the light field in the resonator and variation of population of the laser medium. When the light field is amplified (the rate of stimulated transitions increases), this inevitably causes a decrease of the upper laser level population, which in turn leads to a decrease of gain coefficient. As a result, the intensity of the light field itself starts decreasing. Since an external pump source is involved, whose pumping rate is steady (assuming continuous pump operation), the temporal evolution of intensity develops a certain oscillating pattern which depends on the relation of these parameters.

To characterize relaxation oscillations quantitatively, let us assume an ideal four-level laser with a homogeneously broadened spectral line which generates a transverse mode with the lowest index (TEM_{00}) and a single longitudinal mode. For the sake of simplicity, we assume that the lower laser level is depleted very quickly, so $N_1 = 0$ and $\Delta N = N_2$ (this is also part of our idealization), and the lifetime of the upper laser level is t_{spont}. Assuming that the rate of stimulated emission is $W_i = \sigma(\nu)\phi$ and the intensity is defined as $I = h\nu\phi$, let us write the rate equation for population of the upper laser level, which is analogous to the rate of population difference:

$$\frac{d\Delta N}{dt} = R - W_i \Delta N - \frac{\Delta N}{t_{\mathrm{spont}}}. \tag{116}$$

Since the rate of stimulated transitions W_i is proportional to the intensity of radiation, the middle term on the right hand side of the expression can be rewritten as $W_i \Delta N = \phi B \Delta N$, where ϕ is the number of photons per unit volume and B is a certain proportionality constant (a modified Einstein coefficient). This term describes the rate at which photons occurr. Consequently, we can write an equation for the evolution of the photon flux density as:

$$\frac{d\phi}{dt} = \phi B \Delta N - \frac{\phi}{t_p}, \tag{117}$$

where t_p is the photon lifetime in the resonator. When the equilibrium is established, i.e.,

$$\frac{d\phi}{dt} = \frac{d\Delta N}{dt} = 0, \tag{118}$$

we can find the equilibrium values for population and photon flux densities:

$$\begin{aligned} \Delta N_0 &= \frac{1}{Bt_p} \\ \phi_0 &= \frac{RBt_p - 1/t_{\text{spont}}}{B}. \end{aligned} \tag{119}$$

From the above expressions it follows that when $R_s = (Bt_p t_{\text{spont}})^{-1}$, $\phi_0 = 0$. Here R_s marks the threshold pump rate. When it is exceeded, the laser starts to generate and photon flux density increases. We define a ratio $r = R/R_s$ indicating how many times the threshold pump rate is exceeded. Then the second equation in the system of equations (119) can be written as

$$\phi_0 = \frac{r-1}{Bt_{\text{spont}}}. \tag{120}$$

Let us examine what happens when we have small deviations of population and photon flux densities from their equilibrium values:

$$\begin{aligned} \Delta N(t) &= \Delta N_0 + N'(t) \\ \phi(t) &= \phi_0 + \phi'(t), \end{aligned} \tag{121}$$

where $N'(t) \ll \Delta N_0$ and $\phi'(t) \ll \phi_0$ are small quantities. Inserting these expressions into the initial equations and using relations in Eq.(119), we get a system of equations for small variations of population and photon flux densities:

$$\begin{aligned} \frac{dN'}{dt} &= -RBt_p N' - \frac{\phi'}{t_p} \\ \frac{d\phi'}{dt} &= \left(RBt_p - \frac{1}{t_{\text{spont}}}\right) N'. \end{aligned} \tag{122}$$

We will search for solutions of these equations in the form $\phi'(t) = e^{At}\sin(\omega_m t)$, i.e., assuming that a perturbation of the photon flux density has the shape of a damped sine function with a damping parameter (a characteristic time constant) A and oscillation frequency ω_m. Solving equations (without focusing on mathematical subtleties) we get:

$$\boxed{A = \frac{r}{t_{\text{spont}}}}, \qquad \boxed{\omega_m \approx \sqrt{\frac{r-1}{t_p t_{\text{spont}}}}}. \tag{123}$$

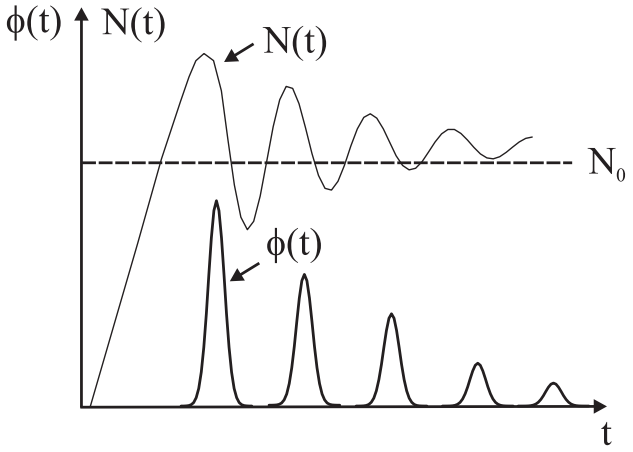


Figure 64: Relaxation oscillations in a laser and evolution of energy level population.

An example of relaxation oscillations is provided in Fig.64. The period and amplitude of relaxation oscillations depend on the pump rate, namely, how many times the threshold pump value is exceeded. On the other hand, the period of relaxation oscillations also depends on the photon lifetime in the resonator, i.e., the quality of the resonator, which in turn is determined by the losses in the resonator.

Continuous or pulsed pumping produce different dynamics of relaxation oscillations. In the case of a continuous pump, the relaxation oscillations are observed only when the laser is turned on and cease when a steady state of laser operation is established. In the case of pulsed pumping, the relaxation oscillations repeat with every pump pulse.

Here we derived parameters of relaxation oscillations in the case when the laser generates a lowest-order transverse and a single longitudinal mode. If there are more longitudinal modes, the situation does not change qualitatively, the relaxation oscillations still have a damped sinusoidal form with exponentially decreasing intensity. An interesting situation occurs when the laser is able to generate higher order transverse modes, i.e., the resonator

quality is very high. Then the relaxation oscillations become chaotic and emerge as a sequence of non-periodic intensity spikes. The reason for this behavior is explained as follows. The lowest order mode TEM_{00} is generated first and strongly depletes population of the laser medium within its volume. Since the volume of higher order modes is larger, they are easily excited because the population of laser medium in the peripheral part remains large and exceeds the threshold value. Each of the transverse mode produces a sequence of its own relaxation oscillations, which overlap producing non-periodic intensity spikes. Thereafter the TEM_{00} mode is generated again and such mode "hopping" can continue for quite a long time. The shape of relaxation oscillations essentially indicates the ability of the laser resonator to support only the lowest index transverse mode and such simple observations of relaxation oscillations in free-running mode are readily used for laser adjustment.

5.2 Q-switching

Q-switching regime is used in order to concentrate laser radiation in a form of a single pulse and so to increase the peak power of laser output. Q-switched lasers produce nanosecond pulses with a peak power up to several hundreds of MW. The first Q-switched laser was demonstrated back in 1962. At present, Q-switched lasers are used for many scientific and technological applications such as precision materials processing, laser ranging and detection and pumping ultrashort pulse lasers and amplifiers, to mention a few. **The idea of Q-switching technique is based on the control of losses in the resonator and consequently the resonator quality.** This is performed as follows: after the pump is turned on, resonator losses are artificially increased. Therefore, the laser generation is prevented, but the active laser medium at that time efficiently accumulates population acting as an energy storage. When the population inversion reaches its maximum value, the resonator quality is rapidly increased (returned to its initial high value state). Then the gain coefficient rapidly becomes very high (exceeding the threshold value by many times) and radiation intensity inside the resonator rapidly (during just several round-trips) increases, "emptying" the whole accumulated population in the upper laser energy level. Such laser pulse is often called giant (gigantic) pulse.

Before going into technical details, how Q-switching is performed in practice, we first analyze the basic features of operation of such laser. The principal scheme of a Q-switched laser is depicted in Fig.65. Initially, let us estimate the parameters of such laser radiation and factors from which they depend on. Since the entire radiation (laser generation) process usually oc-

curs faster than in 20 ns, we can ignore slow processes, such as change of the pump rate in time and consequent change of the population.

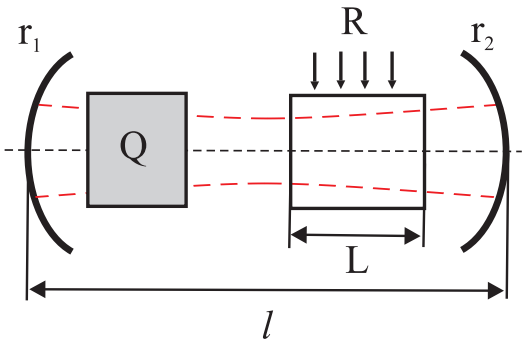


Figure 65: Principal scheme of a Q-switched laser. r_1 and r_2 are the mirror reflection coefficients, l is the length of the resonator, L is the length of laser medium, R denotes the pump source and Q denotes the resonator Q-switching unit.

We also assume that resonator quality is switched on instantly. The laser emission, the laser medium and the resonator itself are described by the following parameters: photon number in a resonator ϕ, mode volume V, population difference in the mode volume $n = \Delta N\ V$ and photon lifetime in passive resonator t_p. Since the gain coefficient γ is proportional to population difference n, the radiation intensity during propagation in a resonator grows exponentially:

$$I(z) = I_0 \exp(\gamma z), \tag{124}$$

or $dI/dz = \gamma I$. An observer traveling with the light wave sees the following intensity evolution in time:

$$\frac{dI}{dt} = \frac{dI}{dz}\frac{dz}{dt} = \gamma I \frac{c}{n_0}, \tag{125}$$

where c is the speed of light and n_0 is the refractive index of laser medium. It is clear that temporal growth constant of the intensity is $\gamma c/n_0$. Since the intensity increases only during propagation in the laser medium with length L, the growth of photon number after propagation from one resonator mirror to the other is only L/l, as the rest of the resonator length is just a free space. The growth constant of photon number is then $\phi\gamma c/n_0(L/l)$. On the other hand, the number of photons in the resonator is reduced due to losses, which are described by the photon lifetime t_p, so the constant of photon number loss is ϕ/t_p. When both of these factors are added together, evolution of photon number in time can be expressed as

$$\frac{d\phi}{dt} = \phi \quad \frac{\gamma cL}{n_0 l} - \frac{1}{t_p} \quad . \tag{126}$$

Let us introduce a dimensionless time $\tau = t/t_p$, which is normalized to the photon lifetime in a resonator. Then after multiplying both sides of Eq.(126) by t_p, we get

$$\frac{d\phi}{d\tau} = \phi \quad \frac{\gamma}{n_0 l/cLt_p} - 1 \quad = \phi \quad \frac{\gamma}{\gamma_t} - 1 \quad , \tag{127}$$

where $\gamma_t = n_0 l/cLt_p$ is the threshold gain coefficient when number of photons in the resonator is constant: $d\phi/d\tau = 0$. As the gain coefficient is directly proportional to population inversion, the former expression can be rewritten in terms of population inversion:

$$\frac{d\phi}{d\tau} = \phi \quad \frac{n}{n_t} - 1 \quad , \tag{128}$$

where $n_t = \Delta N_t V$ is the threshold population inversion. The term $\phi(n/n_t)$ indicates the number of photons generated per unit of normalized time due to stimulated emission. Because each generated photon is a result of electron transition between the laser energy levels, after emission of a single photon the difference of population is reduced by a factor of two: population of the upper energy level is decreased by one and population of the lower energy level is increased by one. This means that a single optical transition corresponds to change of population $\Delta n = -2$ and then its evolution is:

$$\frac{dn}{d\tau} = -2\phi \frac{n}{n_t}. \tag{129}$$

Eq.(128) and Eq.(129) describe the time evolutions of photon number and population, respectively. Solving these equations yields the radiation parameters. Dividing Eq.(128) by Eq.(129), we get photon number evolution which is related to the change of population:

$$\frac{d\phi}{dn} = \frac{n_t}{2n} - \frac{1}{2}. \tag{130}$$

After integration we get:

$$\phi - \phi_i = \frac{1}{2} \left[n_t \ln \frac{n}{n_i} - (n - n_i) \right] \quad , \tag{131}$$

where ϕ_i and n_i are photon number and population, respectively, at the initial moment of time. Assuming that the initial photon number in the resonator

is infinitely small ($\phi_i \approx 0$ and this is essentially the case, because at the very beginning only photons produced by spontaneous emission exist), we get expression of photon number in the resonator:

$$\phi = \frac{1}{2}\left[n_t \ln \frac{n}{n_i} - (n - n_i) \right] . \tag{132}$$

The former expression describes the relation between the photon number and population at any moment of time. Let us take time $t \gg t_p$, when we know that there are no photons left in the resonator ($\phi = 0$), then we can evaluate the final population n_f:

$$\boxed{\frac{n_f}{n_i} = \exp \frac{n_f - n_i}{n_t}} . \tag{133}$$

The quantity $\eta = \frac{n_i - n_f}{n_i}$ indicates which portion of the accumulated population is converted into radiation and η approaches unity when $n_i/n_t \to \infty$, i.e., **the greater is the initial accumulated population at the moment of switching to high quality, the more efficient is its conversion to radiation**. The time evolutions of photon number and population, as described by Eq.(128) and Eq.(129), are plotted in Fig.66.

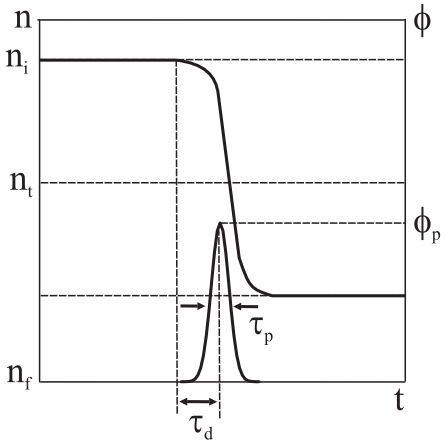


Figure 66: Time evolutions of photon number and population.

It can be assumed that the generation of a gigantic pulse essentially starts from a single photon ($\phi_i = 1$) which appears due to spontaneous emission and whose propagation direction coincides with the optical axis of resonator. The giant pulse develops in a few resonator round-trips. The time delay τ_d between switching the resonator to high quality and temporal peak of the pulse is expressed as

$$\tau_d = \frac{t_p}{n_i/n_t - 1} \tag{134}$$

and is inversely proportional to the ratio of initial and threshold populations. The greater is this ratio, the faster is pulse evolution when Q-switching is turned on. The pulse duration is also a function of this parameter and is expressed as

$$\boxed{\tau_p = t_p \frac{\eta\ n_i/n_t}{n_i/n_t - \ln(n_i/n_t) - 1}}. \tag{135}$$

Note that duration of the gigantic pulse is directly proportional to photon lifetime in the resonator: **in order to get shorter pulses, some losses in the resonator should be introduced**, e.g., choosing an output coupler with greater transmission.

Another important parameter of the laser output is the **peak power of the pulse** which is defined as $P = \phi h\nu/t_p$. According to Eq.(132), the peak power is expressed as:

$$P = \frac{h\nu}{2t_p}\left[n_t \ln\frac{n}{n_i} - (n - n_i)\right. \quad . \tag{136}$$

The maximum peak power is achieved at the moment of time when $n = n_t$ (it can be derived from the condition $\partial P/\partial n = 0$). Inserting $n = n_t$ to expression in Eq.(136) we get

$$P_{\max} = \frac{h\nu}{2t_p}\left[n_t \ln\frac{n_t}{n_i} - (n_t - n_i)\right. \quad . \tag{137}$$

Assuming that the initial population significantly exceeds the threshold value, i.e., $n_i \gg n_t$ (this is essentially the case), the expression for peak power simplifies to

$$\boxed{P_{\max} = \frac{n_i}{2}\frac{h\nu}{t_p}}. \tag{138}$$

In practical setups Q-switching mode is optimized for either maximum peak power (minimum pulse duration) or energy, depending on the planned applications. An important feature of the Q-switched regime: the longer is the photon lifetime in the resonator t_p (the greater resonator quality), the lower is the peak power.

Up to now we have analyzed the case when the quality of resonator is switched instantaneously. In practice, this means that Q-switching time is

shorter than the development time of the pulse τ_d. What happens when resonator quality is switched slowly (the resonator losses α are reduced gradually), is depicted in Fig.67. Then the gain coefficient γ can exceed losses several times during Q-switching phase (moments t_1 and t_2) and as a result several pulses are generated, which is often not a desired result.

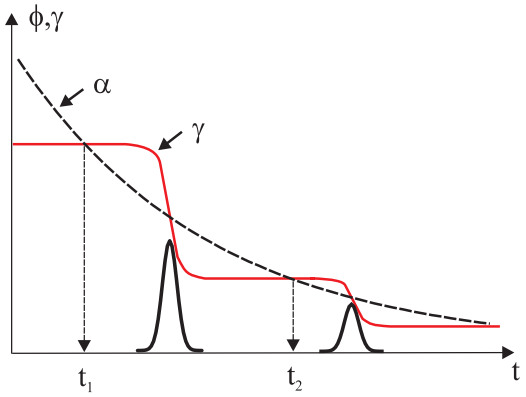


Figure 67: Dynamics of photon number and gain coefficient in the case of slow Q-switching. α marks the level of losses in the resonator.

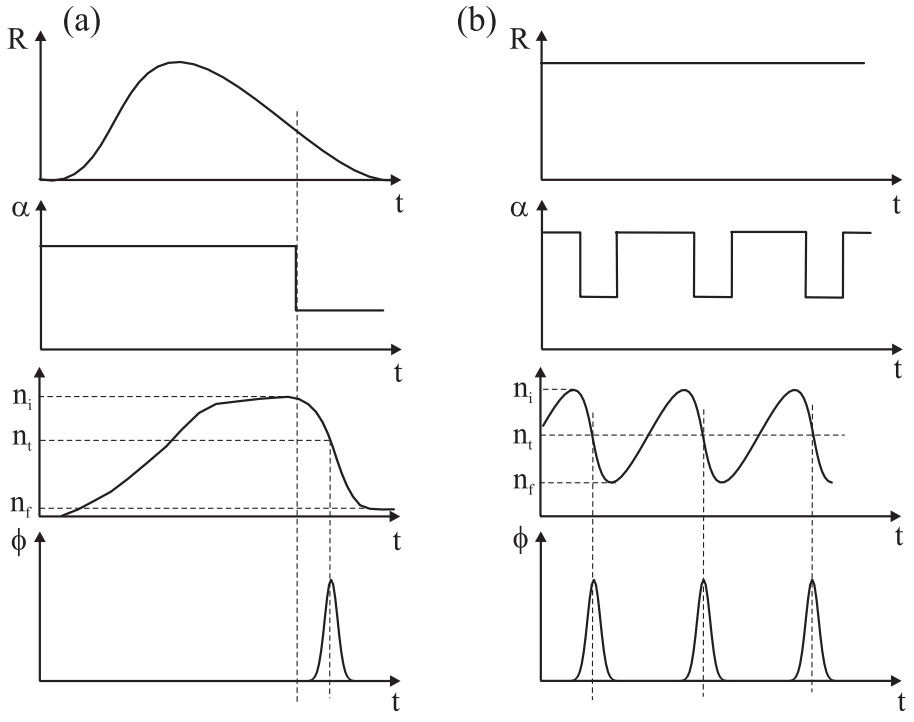


Figure 68: Dynamics of the Q-switching operation in the cases of (a) pulsed pump and (b) continuous pump.

Q-switching can be implemented with both continuous and pulsed pumping, these operation regimes are compared in (Fig.68). In the case of continuous pumping, the resonator quality can be switched at kHz and greater repetition rates, which depend on the time required for the population to grow from n_f to $n_i \gg n_t$. The corresponding repetition rate of the gigantic pulses will be the same. In the case of pulsed pumping, the repetition rate of gigantic pulses will be determined by the repetition rate of the pump pulse, which is typically of the order of several or several hundred of Hz.

5.3 Methods of Q-switching

Up to now the Q-switching unit in the Q-switched laser setup (Fig.65) was depicted as a "black box". Now let us discuss the physical principles and methods of Q-switching in practical setups. The Q-switching methods are be classified into **active** (the resonator quality is controlled by an external signal) and **passive** (the resonator quality changes by itself due to nonlinear losses introduced by a certain optical element).

Active Q-switching methods include:

1. **Rotating prism or mirror**. This is a very simple Q-switching method, when the resonator quality is adjusted by temporarily aligning or misaligning one of the resonator mirrors. A rotating right-angle prism instead of a resonator mirror could be used as well. This Q-switching method is schematically depicted in Fig.69.

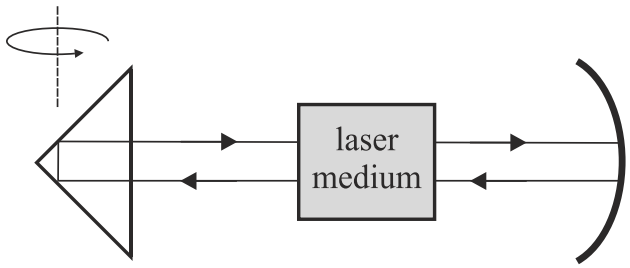


Figure 69: Q-switching employing a rotating prism.

When reflection from the prism or mirror aligns with the optical axis of the resonator, the quality of the resonator suddenly increases and a giant pulse is generated at this moment. This is a very simple and cheap Q-switching method, which performs equally well at any laser wavelength. The first Q-switched ruby laser operated exactly according to this principle. However, this method has one key drawback: the speed of prism rotation is limited (in real conditions it is up to several hundred rotations per second), so Q-switching is slow and, as a consequence, several light pulses are generated.

Therefore in practice this method is no longer used and has only a historic value as a simple idea of Q-switching .

2. **Electro-optic switch**. The electro-optic switch is the most commonly used method for Q-switching. This method is based on controlling the losses by controlling the polarization of light in the resonator. An electro-optic switch consists of an electro-optical modulator (Pockels cell) and a polarization analyzer. Operation of an electro-optical modulator is based on electro-optical effect. A birefringent crystal is oriented in such a way that light propagates along its optical axis. When no external electric field is applied, the refractive indices for ordinary (o) and extraordinary (e) waves in the direction of the optical axis are equal, the crystal shows no birefringence and has no effect on light propagation. When a constant external electric field is applied, the refractive index ellipsoid of the birefringent crystal is modified along the light propagation direction; because of that a certain phase difference between the ordinary and extraordinary waves is induced. The induced phase difference depends linearly on the strength of the applied electric field (the phenomenon is called the Pockels effect):

$$\Delta\phi = \frac{2\pi l n_o^3 r_{nm} E}{\lambda} = \frac{\omega n_o^3 r_{nm} V}{c}, \tag{139}$$

where E is the strength of external electric field, l is the crystal length, $V = El$ is the applied voltage voltage, ω is the frequency of light, n_o is the refractive index for the ordinary wave and r_{nm} is a certain element of crystal electro-optic tensor (electro-optic coefficient). In other words, when voltage is applied, the birefringent crystal, which is oriented along the optical axis, acts as a voltage-controlled wave-plate. The voltage value when $\Delta\phi = \pi$ is called half-wave voltage. When a half-wave voltage is applied, the crystal acts as a half-wave ($\lambda/2$) plate: it rotates the initial linear polarization by 90°. When twice less voltage is applied, the phase difference is $\Delta\phi = \pi/2$ and the crystal acts as a quarter-wave ($\lambda/4$) plate; as a result linearly polarized light becomes circularly polarized. The polarization of light then is analyzed by means of a polarization analyzing element (e.g., Glan prism or a thin-film polarizer) which is inserted in the resonator.

The simplified layout of a Q-switched laser that uses an electro-optical modulator is depicted in Fig.70. The laser operation is controlled by applying the voltage to the Pockels cell. When no voltage is applied, Pockels cell does not rotate polarization, so the electro-optic switch is open and the quality of resonator is high. When any voltage is applied, Pockels cell rotates the polarization of light circulating in the resonator and the analyzer reflects part of it out of the resonator axis, so introducing losses and reducing the resonator

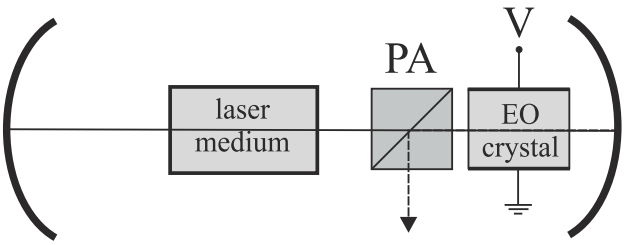


Figure 70: The schematic of a Q-switched laser that uses an electro-optic modulator. V is the voltage applied to the electro-optical crystal (Pockels cell), PA is the polarization analyzing element.

quality. If the half-wave voltage to the Pockels cell is applied, introducing a phase difference of π, the resonator becomes completely blocked, the laser generation does not start and the active element can accumulate population inversion. As the maximum value of population inversion is achieved, the voltage is switched off, the resonator opens and a giant pulse is emitted.

The size of electro-optic coefficient is determined by physical parameters of the crystal and its symmetry class. One of the most common crystals used in these devices is KD_2PO_4 (potassium dideuterium phosphate, KD*P) which has one of the largest electro-optical coefficients (r_{63}). A half-wave voltage for KD*P crystal is ≈ 8 kV for a laser wavelength of 1 μm. In practice, twice less (4 kV) voltage is applied, since light in the resonator passes the Pockels cell twice after bouncing the resonator mirror, so the same polarization rotation effect as per single pass applying a half-wave voltage is achieved. Because of that, the Pockels cell is placed close to one of the resonator mirrors. Switching on and off 4 kV voltage with a high speed is not a technically simple task, however, devices and methods enabling control of high voltage electrical pulses with high repetition rate (kHz) are developed, and high voltage electrical pulse fronts (which essentially determine the speed of Q-switching) are steeper than 10 ns. So electro-optical Q-switching is essentially performed by changing the polarization of light: losses are increased or decreased by analyzing polarization of light. The Q-switching moment is matched with the time moment when the maximum population inversion is accumulated, yielding the highest gain coefficient.

There are two types of Pockels effect: longitudinal and transverse. In the former case voltage is applied along the light propagation direction (as in the case just discussed), while the in the latter case, voltage is applied transverse to light propagation direction. The geometries of electro-optical modulators corresponding to these two types of Pockels effect are illustrated in Fig.71.

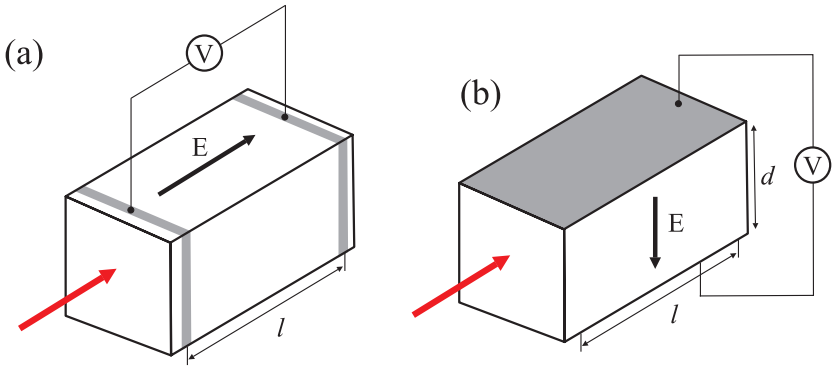


Figure 71: Electro-optical modulator based on (a) longitudinal and (b) transverse Pockels effect.

The phase difference due to transverse Pockels effect is:

$$\Delta\phi = \frac{\omega l}{c}\left[(n_o - n_e) - \frac{n_o^3}{2} r_{63} \frac{V}{d}\right. \quad , \tag{140}$$

where n_e is the refractive index for the extraordinary wave, l is the length of crystal, d is the width of crystal. Unlike the longitudinal Pockels effect, the polarization rotation due to transverse Pockels effect depends on the crystal length. It may seem convenient, since the half-wave voltage can be reduced significantly, however, excitation of strong high frequency acoustic waves can considerably modulate the transmission of such electro-optic switch in time.

3. **Acousto-optic switch**. The method of acousto-optic switch relies on the control of resonator quality by changing propagation direction of light in a resonator that is equivalent to a misalignment of a resonator. An acousto-optic switch (modulator) is based on photoelastic effect. The photoelastic effect is a change of the material density induced by propagating acoustic wave, which in turn causes modulation of the material refractive index. A light wave propagating in such medium is diffracted; the situation is equivalent to light diffraction from a transparent diffraction grating. In the present case, the period of the grating corresponds to the wavelength of the acoustic wave.

The operation principle of an acousto-optic modulator is schematically depicted in Fig.72. The most commonly used media for acousto-optic modulation are fused silica (SiO_2) and tellurium oxide (TeO_2). These materials have excellent optical and mechanical properties such as broad transmission range and sufficiently large photoelastic coefficient. High frequency traveling acoustic waves are generated by applying an alternating voltage at radio

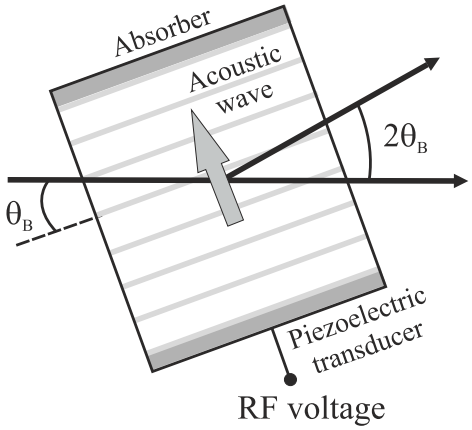


Figure 72: Principle of operation of an acousto-optic modulator.

frequency to a piezoelectric transducer which is attached at one end of the crystal. An acoustic wave absorber is attached at the other end of the crystal. Diffraction of a light wave from an acoustic wave is described by Bragg diffraction condition:

$$\sin\theta_B = \frac{m\lambda}{2\lambda_a}, \tag{141}$$

which for small angles and first diffraction order ($m = 1$) can be rewritten as $\theta_B = \lambda/2\lambda_a$, where θ_B is the angle between the light and acoustic waves and λ_a is the wavelength of acoustic wave. In that case the diffraction angle of a light wave is $2\theta_B$.

Q-switching is performed by simply placing an acousto-optic modulator into the resonator. The control of resonator quality is simple: when alternating voltage at radio frequency is switched on, the light beam changes its propagation direction due to diffraction and is deflected from the optical axis of the resonator. Then the resonator becomes misaligned, its quality reduces to 0 and such laser cannot generate. However, at this time the active element accumulates population inversion. When the maximum population inversion is achieved, the radio frequency signal is turned off and the resonator quality suddenly increases and a gigantic pulse is generated. Acousto-optic modulators can operate at high repetition rates (tens of kHz) and their switching times are very short, e.g., 5 ns for TeO_2 crystal.

To achieve **passive Q-switching**, losses in the resonator are modulated using a **saturable absorber** which is inserted into the resonator. It is a passive Q-switch since it does not require any source of external control. A

saturable absorber is essentially a two-level system which absorbs laser radiation very well and has low saturation intensity. The absorption coefficient of a saturable absorber is a function of intensity:

$$\alpha = \frac{\alpha_0}{1 + I/I_s}, \tag{142}$$

where α_0 is the absorption coefficient for low intensity (when no saturation is present) and I_s is the saturation intensity. When intensity is low, the saturable absorber is opaque: it simply absorbs all laser radiation and so induces significant losses in the resonator. Therefore, the laser cannot generate but just accumulates the population inversion. When the gain coefficient exceeds the absorption coefficient, intensity in the resonator rapidly increases, while the absorber starts to become transparent and losses in the resonator rapidly decrease. When the absorber saturates, the maximum resonator quality is achieved. In order to have an efficient and timely switching of the resonator quality, it is important to match parameters of the pump and saturable absorber. The lifetime of saturable absorber excited level has to be sufficiently long, longer than duration of generated pulses, i.e., several tens of ns, so that after generation has started it would relax slowly and would not introduce any further unnecessary losses.

Historically, cyanine dyes dissolved in certain solvents (usually methyl or ethyl alcohols) were widely used as saturable absorbers. Their advantage is that by adjusting dye molecule concentration, it is easy to establish the necessary saturation intensity of the absorber. On the other hand, choice of dyes is very wide, so one can always choose a dye with the desired absorption and relaxation properties. However, dye solutions have several of drawbacks: they are toxic, chemically unstable and when irradiated by intense laser light, they degrade quite fast (in a few months or faster). Another drawback is strong thermal effects which manifest themselves in the form of a thermal lens, which dynamically changes resonator stability conditions. Due to this fact, dye solutions have to be permanently mixed. The next generation saturable absorbers are based on chromium-doped crystalline media (e.g., Cr:YAG) which do not have the aforementioned problems.

Typically, the most efficient in the Q-switching operation are solid-state lasers, especially those having long lifetimes of their upper laser energy levels, such Nd;glass, Nd:YAG, Nd:YLF, etc. However, efficient Q-switching was also demonstrated with gas lasers: CO_2 and iodine.

5.4 Mode-locking

Another laser operation regime is **mode-locking**. Mode-locking represents a technique that allows phasing of all frequencies (**longitudinal modes**) circulating in the laser resonator. Mode-locking produces the shortest and the most intense laser pulses with durations varying from several hundred of ps to < 10 fs (depending on the gain bandwidth of the laser medium and mode-locking method). The first mode-locked laser was demonstrated in 1966[12]. Absolutely all ultrashort pulse lasers employ a certain type of mode-locking technique. Let us discuss the differences between unphased and phased (locked) mode laser operation.

Consider a laser medium with inhomogeneously broadened spectral line. Only longitudinal modes for which the gain is greater than losses are amplified. Assuming sufficiently long resonator, a number of such modes can be very large and the intermodal distance (the difference of adjacent mode frequencies) which is defined as

$$\Delta\omega = \omega_{q+1} - \omega_q = \frac{\pi c}{n_0 l} \tag{143}$$

is much smaller than the laser amplification bandwidth $\Delta\omega_L$, as schematically illustrated in Fig.73. Only in this case the discussion on mode-locking is meaningful.

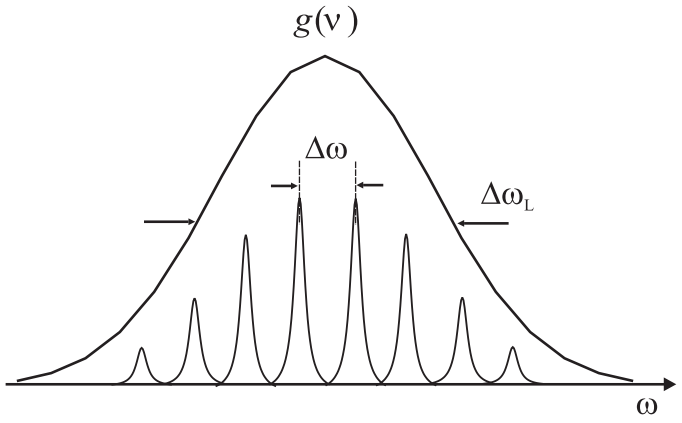


Figure 73: Intermodal distance and laser amplification bandwidth.

Without any additional means, such laser generates N independent longitudinal modes whose phases are uncorrelated. The complex amplitude of the n-th mode can be written as

[12]A. J. De Maria, D. A. Stentser, and H. Heynau, Self-mode locking of lasers with saturable absorbers, Applied Physics Letters **8**, 174–176 (1966)

$$E_n(t) = E_n e^{i(\omega_n t + \phi_n)}, \tag{144}$$

where E_n is amplitude of the mode electric field, ω_n is the mode frequency, ϕ_n is its phase. The complex amplitude of the total radiation field in the resonator is a superposition of all N modes with random phases:

$$E(t) = E_0 \sum_{n=0}^{N-1} e^{i(\omega_n t + \phi_n)}, \tag{145}$$

and the intensity is

$$\boxed{I(t) = |E(t)^2| = EE^* = E_0^2 \sum_{n=0}^{N-1} e^{i(\omega_n t + \phi_n)} e^{-i(\omega_n t + \phi_n)} = NE_0^2}. \tag{146}$$

This result shows that the total radiation intensity is a sum of intensities of the individual modes: $I = NI_{\text{mode}}$. This is the average value because a certain number of modes may accidentally have the same phase, but the average result is exactly as specified. Such (unphased mode) radiation is generated by free-running and Q-switched lasers.

Now let us consider the situation when all longitudinal resonator modes are phased, i.e., have the same phase $\phi_n = \phi_0$. Then the the complex field amplitude is

$$E(t) = E_0 \sum_{n=0}^{N-1} e^{i(\omega_n t + \phi_0)} = E_0 e^{i\phi_0} \sum_{n=0}^{N-1} e^{i\omega_n t}. \tag{147}$$

After performing certain mathematical operations, the intensity can be expressed as

$$I(t) = |E(t)^2| = EE^* = E_0^2 \left| \frac{1 - e^{-iN\Delta\omega t}}{1 - e^{-i\Delta\omega t}} \right|^2 = E_0^2 \frac{\sin^2(N\Delta\omega t/2)}{\sin^2(\Delta\omega t/2)}. \tag{148}$$

The latter result suggests that phasing of longitudinal modes induces certain intensity modulation in time with intensity maxima at

$$\frac{\Delta\omega t}{2} = 0, \pi, 2\pi, ..., n\pi. \tag{149}$$

It is then easy to determine the temporal separation (period) between the adjacent intensity maxima:

$$\Delta t = t_{n+1} - t_n = \frac{2(n+1)\pi}{\Delta\omega} - \frac{2\pi n}{\Delta\omega} = \frac{2\pi}{\Delta\omega} = \frac{2n_0 l}{c} = t_r, \tag{150}$$

where t_r is the resonator round-trip time.

This means that the light pulses in the resonator are generated and their repetition period coincides with the resonator round-trip time t_r. Let us estimate the maximum intensity of such a pulse. Let us rewrite expression in Eq.(148) assuming that $\Delta\omega t/2 = 0$:

$$\boxed{I(t)_{\max} = \lim_{\Delta\omega t/2 \to 0} E_0^2 \frac{\sin^2(N\Delta\omega t/2)}{\sin^2(\Delta\omega t/2)} = \lim_{\Delta\omega t/2 \to 0} E_0^2 \frac{N^2(\Delta\omega t/2)^2}{(\Delta\omega t/2)^2} = N^2 E_0^2}. \tag{151}$$

In the case of mode-locking, intensity of the light pulse is $I = N^2 I_{\text{mode}}$, i.e., significantly greater than in the case of modes with random phases. Therefore, mode-locking produces light pulses with very high intensity. A comparison of the intensity evolutions $I(t)$ in the cases of unsynchronized and synchronized modes is shown in Fig.74.

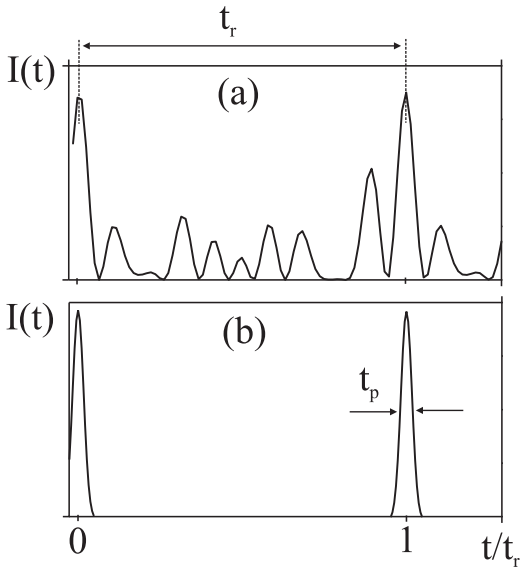


Figure 74: Evolution of light intensity in a resonator in the cases of (a) unsynchronized and (b) synchronized mode (mode-locked) cases. The intensities in (a) and (b) are not to scale.

Let us now estimate the minimum duration of such pulses:

Table 3: Relevant parameters of several mode-locked lasers. τ_i is the minimum theoretical pulse duration that is inversely proportional to amplification bandwidth $\Delta\nu_L$ and τ_p is the experimentally achieved minimum pulse duration.

Laser medium	λ, μm	$\Delta\nu_L$, THz	τ_p	τ_i
Nd:YAG	1.064	0.135	5 ps	3.3 ps
Nd:YLF	1.047	0.390	2 ps	1.1 ps
Nd:glass	1.054	8	60 fs	55 fs
Cr:LiSAF	0.85	57	18 fs	8 fs
Ti:sapphire	0.8	100	6 fs	4.4 fs
Rhodamine 6G	0.57	45	27 fs	10 fs

$$\boxed{\tau_p = \frac{2\pi}{\Delta\omega N} = \frac{t_r}{N} = \frac{2n_0 l}{Nc} = \frac{1}{\Delta\omega_L}}. \tag{152}$$

This result suggests that the more modes are phased, the shorter is the generated light pulse. Assuming that all the longitudinal modes are phased, the minimum pulse duration is inversely proportional to the width of laser medium amplification band. In a real laser, it is possible to synchronize a major fraction of longitudinal modes, the relevant examples are presented in Table 3. Note that with several exceptions (e.g., dye or semicoductor lasers) only solid-state lasers operate in mode-locking regime as only these lasers provide sufficiently broad laser amplification bandwidth.

5.5 Methods of mode-locking

Up to now we considered mode-locking in the spectral domain, without paying attention how the mode-locking is performed in a real laser. For that reason, it is convenient to understand the process of mode-locking in the time domain. In practice, all the mode-locking methods rely on the control of resonator losses. The principal layout of a mode-locked laser is very similar to that of the Q-switched laser, which is schematically is depicted in Fig.65. However in the case of mode-locking, the modulation depth of the resonator losses is considerably smaller and the modulation speed is much faster, matching the resonator round-trip time. Formally, the methods of mode-locking are classified into active and passive, depending on the way how the resonator losses are controlled.

Active mode-locking. The term active mode-locking itself means that mode-locking is performed by active interference into the laser operation using a certain external signal. Active mode-locking can be performed in several

ways and their choice depends on specific properties of the laser medium, i.e., the gain coefficient and lifetime of the upper laser energy level. According to the principle of operation, active mode-locking is performed via amplitude or phase modulation. The amplitude mode-locking relies on periodic modulation of resonator losses, while the phase mode-locking is based on periodic adjustment of the resonator length. In lasers which have short upper level lifetime (e.g., dye lasers) active mode-locking is implemented by modulation of gain coefficient (e.g., by pumping with short pulses from another laser). Such mode-locking method is called synchronous pumping.

We will now discuss the basic principle of only the most commonly used active amplitude mode-locking method.

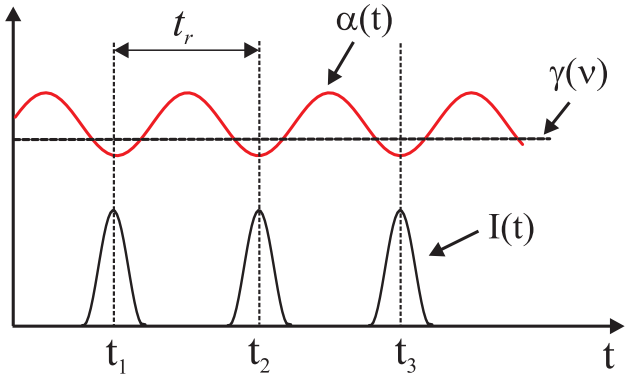


Figure 75: Principle of active amplitude mode-locking.

Active amplitude mode-locking is performed by placing either an electro-optical or acousto-optical switches into the resonator, which are controlled by an external signal, whose period exactly matches the resonator round-trip time (a typical signal frequency is 100 MHz for a 1.5 m long resonator). In the case of electro-optical modulator, it is impossible to switch its half-wave voltage at such high repetition rate, so the actual modulation depth of the introduced losses is small and the desired effect is achieved by introducing additional steady losses at a certain constant level by applying a bias voltage. The periodic control signal basically creates a periodic loss modulation and the light pulse is formed only within the time interval when laser gain exceeds the resonator losses, as shown in Fig.75. The amplitude modulator is usually placed very close to one of the resonator mirrors to avoid possible generation of several pulses at a time and their subsequent competition.

How the introduction of losses makes longitudinal modes synchronized and how can this process be envisaged in the time domain? If the modulation period of losses matches the resonator round-trip time, the light pulse generated at the moment of minimum resonator losses returns back to

the modulator exactly at the time when resonator losses are reduced again. Moreover, with each round-trip the pulse shortens due to uneven amplification: its peak is amplified more efficiently than its fronts. Mode-locking in the time domain can be understood in a quite straightforward way: when the pulse gets shorter, its frequency spectrum broadens and this automatically implies that more resonator longitudinal modes are made to oscillate in phase. When such process repeats many times, in an ideal case of inhomogeneous linewidth broadening the minimum pulse duration is defined as

$$t_p \approx \frac{0.44}{\Delta\omega_L}. \tag{153}$$

In the reality, a complete locking of all the longitudinal modes is never achieved, and the final pulse duration which establishes after many round-trips is determined by the modulation depth or frequency. In the case of homogeneous amplification linewidth broadening the minimum pulse duration can be defined as

$$t_p \approx \frac{0.45}{\sqrt{\Delta\omega_L f_m}}, \tag{154}$$

where f_m is the modulation frequency of resonator losses. The active mode-locking critically depends on how accurately the modulation frequency matches the resonator length (or its round-trip time): even a small mismatch between these two, e.g., length change due to environmental temperature variations, mechanical vibration, etc, can cause large instabilities in the mode-locking process, which eventually can be lost.

Using active mode-locking methods, light pulses with durations ranging between 1 ns and 50 ps are usually generated. This pulse duration interval is quite unique, as there are no other methods to generate bandwidth-limited pulses with such duration. Usually, electro-optical modulators are used in lasers with pulsed pumping, where the gain coefficient is large, while acousto-optical modulators are used in lasers with continuous wave pumping where the gain coefficient is small.

Passive mode-locking. Passive mode-locking is implemented by placing into the resonator a specific element, which operation is based on intensity-dependent transmission or reflection. One of such elements is a saturable absorber, or any other element whose operation is based on a similar principle and which has fast relaxation time (much faster that the round-trip time of a resonator). Let us discuss operation of these elements and how they perform mode-locking in more detail.

The operation principle of a saturable absorber was explained when we described Q-switching. The same principle holds also in the case of passive mode-locking. However, in this case, the saturable absorber has very fast relaxation time (several ps and shorter, 10^{-12} s). Fast absorber relaxation ensures that the light pulse propagating through the absorber gets shorter with each round-trip: the pulse peak which has the highest intensity, propagates without any losses, while the pulse slopes with lower intensity are partially absorbed, consequently making pulse fronts steeper. This situation is schematically depicted in Fig.76. Eventually, the established pulse duration approximately equals to

$$t_p \cong \frac{0.79}{\Delta\nu_L}\sqrt{\frac{\gamma}{\alpha}\frac{I_s}{I_p}}, \tag{155}$$

and is determined by saturable absorber parameters, i.e., losses due to absorption α, gain coefficient γ, saturation intensity I_s, pulse peak intensity I_p, and gain bandwidth of the laser medium $\Delta\omega_L$.

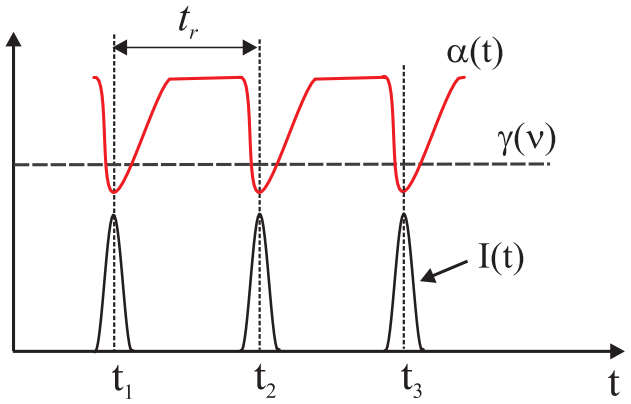


Figure 76: Principle of passive mode-locking. t_r is the resonator round-trip time.

It is important to understand how passive mode-locking starts, since no external signal that initiates the process is applied. The laser oscillation starts from a certain noise background (an example is presented in Fig.74(a)) when many longitudinal modes (falling under the laser medium amplification contour) having random phases are excited at once. The saturable absorber absorbs all this "mode noise" and in the laser medium new "mode noise" with random phases is being constantly generated. Sooner or later (this time interval is of the order of μs) there emerges a fluctuation with sufficiently high intensity (random combination of modes with the same phase) which turns the absorber partially transparent and then returns to the laser medium. Here this fluctuation is further amplified and eventually overgrows all

the remaining noise background. In this way, a single light pulse remains circulating in the resonator and becomes shorter and shorter with every pass through the absorber. As a result, its frequency spectrum broadens that is equivalent to adding more and more longitudinal modes in phase.

We have already mentioned that organic dye saturable absorbers posses several drawbacks related to their chemical instability (lack of durability). Furthermore, they are liquid and have to be constantly mixed to avoid undesired thermal effects. Solid-state saturable absorbers (such as Cr:YAG) also have drawbacks related to relatively long relaxation time and limited operating wavelength range, so they are practically used only for Q-switching. In 1992 qualitatively new type of saturable absorbers based on semiconductor media was proposed [13]. These saturable absorbers were called **SESAM (Semiconductor Saturable Absorber Mirror)**. Unlike organic dye absorbers which are based on saturable transmission, semiconductor absorbers are based on saturable reflection (Fig.77).

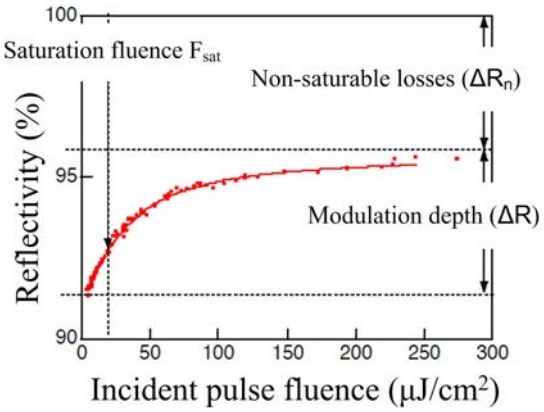


Figure 77: Examples of SESAM mirror reflectivity as a function of laser pulse fluence.

SESAM consists of a semiconductor layer (e.g., AlAs/AlGaAs) with an intensity-dependent reflection coefficient and serves as one of the resonator mirrors. By choosing various combinations of semiconductor materials, such absorber can operate in a relatively broad spectral range (limited by the near IR range, however) and its relaxation time is very fast, in the order of 100 fs, which makes it possible to generate very short light pulses. Although SESAM is chemically stable, its main drawbacks are relatively small modulation depth (the intensity-dependent reflection coefficient may vary only within $\sim$ 5%

[13] U. Keller et al, Semiconductor saturable absorber mirrors (SESAMs) for femtosecond to nanosecond pulse generation in solid state lasers, IEEE Journal of Selected Topics in Quantum Electronics **2**, 435–453 (1996)

range) and low saturation intensity. Therefore, the use of SESAM is limited to very low pulse energies; such pulses are generated by pulsed lasers that employ continuous wave pumping (quasi-cw lasers). With high energy pulses, the coating of the absorber degrades in time as is eventually optically damaged.

Another method of passive mode-locking is based on the **optical Kerr effect** in transparent materials, which results in the intensity-dependent refractive index, which is expressed as

$$\boxed{n = n_0 + n_2 I}, \tag{156}$$

where n_2 is the nonlinear index of refraction. For all solid-state media that are transparent in the visible and IR range, this quantity is in the order of 10^{-16} cm^2/W. The typical intensities present in laser resonators induce a refractive index change of the order of $10^{-5} - 10^{-6}$, and this is enough that an intense light beam starts to self-focus when propagating in such a medium. Self-focusing occurs because beam in the resonator has intensity distribution $I(r)$ described by the Gaussian function: beam parts with different intensities acquire different phase shift, which is the largest at the center of the beam:

$$\delta\phi(r) = \frac{2\pi n_2 l_n}{\lambda} I(r), \tag{157}$$

where l_n is the length of nonlinear medium. Due to that intensity-dependent phase shift, the initially flat phase front of the beam bends as if the beam passes through a convex lens and the beam starts to self-focus, as shown in Fig.78.

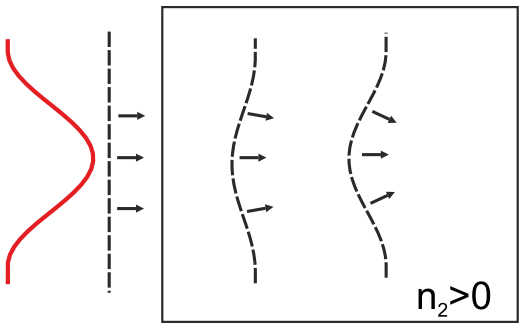


Figure 78: Bending of the phase front of intense Gaussian beam due to intensity-dependent phase shift which causes self-focusing.

The lens induced by nonlinear medium is called **Kerr lens** and its focal length

$$\frac{1}{f_n} = \frac{4n_2 l_n}{n_0 w^2} I_p \tag{158}$$

is a function of the peak intensity I_p. Here w is the radius of a Gaussian beam.

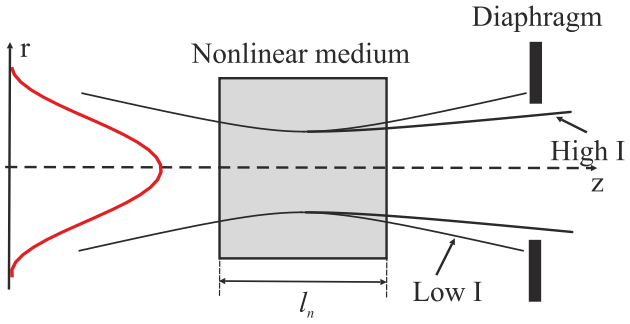


Figure 79: Principle of Kerr lens operation.

Now let us discuss how the self-focusing effect can be exploited for mode-locking. The method is called **Kerr lens mode-locking** (KLM) and its principle of operation is schematically depicted in Fig.79. A nonlinear medium is placed in the laser resonator together with a diaphragm. The diameter of the diaphragm is chosen such that only a high intensity beam that carries a very short light pulse passes the diaphragm without any losses. Only a very short (mode-locked) light pulse has sufficient intensity to induce a nonlinear lens, while a noise-like radiation with low intensity experiences large losses. Hence a Kerr lens operates similarly to a saturable absorber with very fast relaxation time, which is only few fs (10^{-15} s), and which is determined by the response time of electron cloud to the electric field of light. Kerr lens can be induced in any transparent material, since all media possess electronic nonlinearity. Moreover, this method is suitable to any wavelength and is non-resonant: the losses are created without absorption. A small problem is that often the initial fluctuations appearing in the laser resonator have insufficient intensity to initialize Kerr lens mode-locking. Then the laser generation is triggered by an external perturbation e.g., by gently hitting the resonator mirror mount. Of course, KLM can only be applied in the case of quasi-cw laser and would be difficult to implement in pulsed pumping regime.

Kerr lens mode-locking was first demonstrated in 1991 by generating femtosecond pulses from Ti:sapphire oscillator[14]. Interestingly, in that configuration, sapphire crystal served itself as a nonlinear medium, while the

[14]D. E. Spence, P. N. Kean, and W. Sibbett, 60-fsec pulse generation from a self-mode-locked Ti:sapphire laser, Optics Letters **16**, 42–44 (1991)

focused pump from an argon ion laser served as a diaphragm. Mode-locking operation was achieved by simply aligning the laser resonator in a certain way. Such mode-locking method was termed self-mode-locking since no specific element directly modulating losses was used. Thanks to this discovery, Ti:sapphire quickly became the most reliable and widely used femtosecond laser, which outperformed the whole generation of dye lasers, which were the only available femtosecond lasers at that time. At present, Kerr lens mode-locking is used in many modern solid-state laser oscillators.

In conclusion, we note that since passive mode-locking starts from random fluctuations, the time moment when the sufficient fluctuation appears in the resonator and a pulse is formed is quite undefined. For quasi-cw lasers the exact start time of mode-locking is not important: when the laser is turned on and once mode-locking starts, it will continue as long as the laser keeps running. However, for lasers with pulsed pumping (such as Nd:glass and Nd:YAG) the start time of mode-locking is important, as mode-locking has to restart with every pump pulse. For this purpose, along passive mode-locking, active mode-locking is employed as well, which forms the initial light pulse (although quite long), which is capable to initiate the process of passive mode-locking at rather well-defined moment of time. Such mode-locking method is called **hybrid mode-locking**.

Summary of Chapter 5

- A unique feature distinguishing lasers from other light sources is that they can emit light not only in the form of a low divergence beam but also in the form of very short pulses.
- There are three types of pulsed laser operation regimes: free-running, Q-switching and mode-locking.
- Free-running is the simplest laser operation regime, which starts when the laser oscillation threshold is exceeded and produces μs pulses with kW peak power.
- Q-switching regime is achieved via control of resonator losses and produces ns pulses with a peak power up to hundreds of MW.
- Mode-locking regime is achieved when longitudinal modes in the laser resonator are made to oscillate (locked) in phase. Mode-locking produces ultrashort pulses with durations ranging form 100 picoseconds to a few femtoseconds and peak powers of up to several hundreds of MW.

6 Generation and amplification of ultrashort light pulses

The shortest (the so-called, ultrashort) light pulses are generated only by employing passive mode-locking. Although passive mode-locking could be implemented in many solid-state lasers, the **shortest light pulses are generated only in laser media which possess the broadest amplification (gain) bandwidth**. In that regard, there are two unique femtosecond lasers to note: dye lasers and Ti:sapphire lasers, which have the broadest gain bandwidth and which served as main femtosecond laser sources at different decades[15]. The minimum pulse duration for both of these lasers has been reduced by more than two orders of magnitude (from several ps by the end of 1960s to tens of fs in 1980s) as a result of development of mode-locking techniques. It is important to mention that the shortest pulse durations are achieved by additionally compressing the pulses outside the resonator – the method is generally termed the **extracavity compression**.

In this chapter we will discuss the main problems related to ultrashort pulse generation, amplification and techniques how such ultrashort pulse duration is measured.

6.1 Dispersion control in the resonator

One of the main problems facing the generation of ultrashort light pulses is their dispersive broadening in a resonator due to very broad frequency spectrum. There is always at least one optical element in the resonator (i.e., the laser medium) in which dispersive effects occur: longitudinal modes with different frequencies propagate at different phase velocities in the dispersive medium, which causes uneven phase shifts between adjacent modes, so the pulse broadens in time and becomes chirped (phase modulated). This problem is important for all ultrashort pulses as **dispersion is a cumulative effect**: the phase shifts accumulate with each resonator round-trip. **The shorter is the pulse, the broader is its frequency spectrum and the more pronounced is the dispersion effect**. For example, it is estimated that in Ti:sapphire laser resonator a light pulse with a duration of 10 fs can broaden up to several times during just a single round-trip. Naturally, due to temporal broadening, the intensity of a pulse decreases, and as a result, the process of mode-locking could be disturbed or completely lost. To avoid this, dispersion must be compensated by placing a specific element in the

[15]U. Keller, Recent developments in compact ultrafast lasers, Nature **424**, 831–838 (2003)

resonator. Let us discuss what these elements are and how they work.

The group velocity dispersion of transparent dielectric media in the visible and near-IR spectral range is normal: the long-wavelength ("red") spectral components propagate faster than the short-wavelength ("blue") spectral components and therefore the pulse becomes **chirped (phase modulated)**. A chirped pulse is "colored": its leading front is "red" and its trailing front is "blue", and the frequency across the pulse changes linearly (the phase changes quadratically). The pulse duration during propagation changes according to

$$\boxed{t_p = t_0\sqrt{1+\frac{z^2}{L_d^2}}}, \tag{159}$$

where t_0 is the initial pulse duration, z is the propagation distance and L_d is the dispersive broadening length, which is expressed as $L_d = t_0^2/|2g_0|$, where $g_0 = \frac{\partial^2 k}{\partial \omega^2}$ is the **group velocity dispersion coefficient**. $g_0 > 0$ denotes the **normal** group velocity dispersion, whereas $g_0 < 0$ denotes the **anomalous** group velocity dispersion. The dispersive broadening length is proportional to the pulse duration squared, and its physical meaning could be deduced from the case when $z = L_d$: it is the propagation distance when the pulse duration increases by a factor of $\sqrt{2}$. The phase modulation occurring due to dispersive broadening is expressed as

$$\frac{d^2\phi}{d\omega^2} = \frac{\lambda^3 L}{2\pi c^2}\frac{\partial^2 n}{\partial \lambda^2}, \tag{160}$$

where L is the length of the dispersive medium and $n(\lambda)$ is the dispersion law of the refractive index.

Compensation of dispersion requires delaying the spectral components of the light pulse that propagate faster with respect to spectral components that propagate slower. There are several ways to do this.

It would be very simple if one could introduce into the resonator an **optical element, which has anomalous group velocity dispersion** ($g_0 < 0$), where the short-wavelength ("blue") spectral components propagate faster than long-wavelength ("red") spectral components. Unfortunately, the range of anomalous group velocity dispersion in transparent dielectric materials lies close to the mid-IR region, e.g., in glass or sapphire, the dispersion is anomalous for wavelengths longer than 1.3 μm, while most of the solid-state lasers emit at shorter wavelengths.

A solution was found by using **a pair of prisms**, where anomalous group velocity dispersion occurs due to optical path difference between distinct

spectral components, as illustrated in Fig.80. In this configuration, the optical path of the long-wavelength spectral components is always longer than that of the short-wavelength spectral components, so propagation through a prism pair is equivalent to propagation through a medium featuring anomalous group velocity dispersion. The prism pair yields the following phase modulation:

$$\frac{d^2\phi}{d\omega^2} \approx -\frac{l\omega_0}{c}\left(\frac{d\alpha}{d\omega}\right)_{\omega_0}^2, \tag{161}$$

where l is the distance between prisms and the last term of the product describes the angular dispersion of a prism.

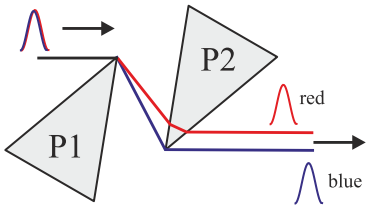


Figure 80: Anomalous group velocity dispersion created in a pair of prisms.

The amount of introduced anomalous dispersion can be easily adjusted by either varying the distance between prisms or by simply translating one of the prisms perpendicularly with respect to the beam propagation direction. In such a way dispersive broadening of the pulse in other resonator elements can be fully compensated. In fact, all lasers producing light pulses shorter than 0.5 ps employ dispersion compensating element in the resonator. A typical layout of Ti:sapphire laser with dispersion compensation is depicted in Fig.81

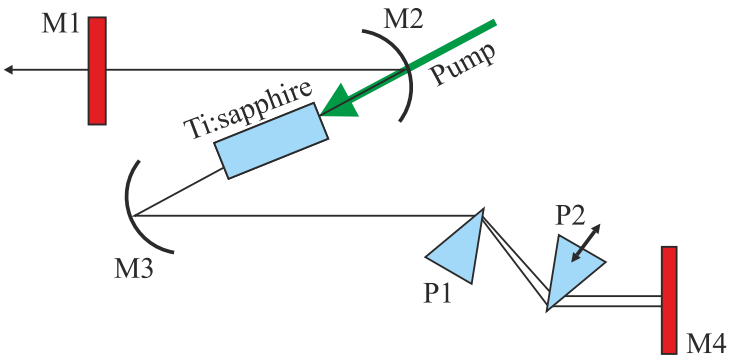


Figure 81: Layout of Ti:sapphire laser resonator with dispersion compensation.

Note that after passing the prism pair, the laser beam becomes "colored", as its colors become separated in space due to different refraction angles. The coloration of the beam is cancelled after back reflection from the mirror (providing yet another pass), which also doubles the temporal dispersion of the setup.

If the pulse duration is very short (consequently, its frequency spectrum is very broad), dispersion cannot be fully compensated using a pair of prisms because of higher-order dispersion effects (related to higher order derivatives: $d\phi^3/d\omega^3$, etc.) that become significant. The effects related to higher order dispersion manifest themselves as nonlinear distortions of the pulse chirp preventing optimal compression of the pulse. In that case, optimal dispersion compensation (pulse compression) is achieved using the so-called **chirped mirrors**[16]. Chirped mirror is a mirror with a specific dielectric coating which consists of e.g. alternating SiO_2 and TiO_2 layers of variable thickness, deposed in such a way that light penetration depth (where it is reflected) depends on its wavelength, as illustrated in Fig.82.

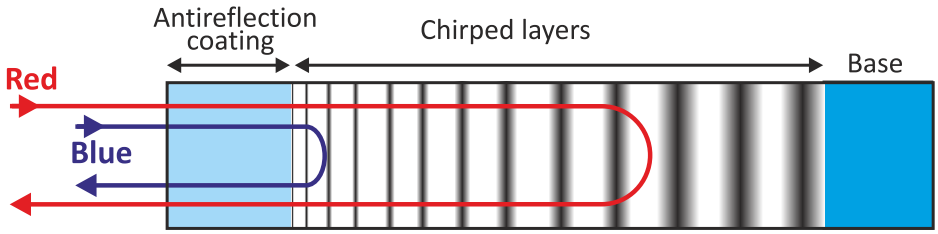


Figure 82: Structure and operation principle of a chirped mirror.

The long-wavelength ("red") spectral components of the pulse (which are faster) are reflected from a deeper layer of the coating, while the short-wavelength ("blue") spectral components are reflected from a layer close to the surface, so the group velocity dispersion is compensated. By selecting an appropriate layer thickness and positioning even a very complex law of group velocity dispersion can be fully compensated. Chirped mirrors have only one drawback: when the dielectric layers are deposited their dispersion and its sign cannot be adjusted.

One of the most flexible dispersion compensation techniques employs the **combination of a prism pair and chirped mirrors**: the prism pair is used to compensate the second order (group velocity) dispersion and the chirped mirrors are used to compensate higher order dispersion.

[16]N. Matuschek, F. X. Kartner, and U. Keller, Theory of double-chirped mirrors, IEEE Journal of Selected Topics in Quantum Electronics **4**, 197–210 (1998)

Kerr lens mode-locking enables generation of the shortest possible pulses whose duration is comparable to a single **optical cycle**[17]. An optical cycle is a time interval corresponding to a period of a single oscillation of the electric field. The duration of an optical cycle is defined as

$$\tau_c = \frac{\lambda_0}{c}, \tag{162}$$

where λ_0 is the central (carrier) wavelength. It is easy to calculate that one optical cycle at $\lambda_0 = 300$ nm equals to $\tau_c = 1$ fs, at $\lambda_0 = 600$ nm, $\tau_c = 2$ fs and so on. For the central wavelength of Ti:sapphire laser, $\lambda_0 = 800$ nm, one optical cycle is $\tau_c \simeq 2.7$ fs. Spectral bandwidth which corresponds to a bandwidth-limited single optical cycle pulse can also be estimated:

$$\Delta\lambda = \frac{c}{\nu_0 \pm \Delta\nu/2}. \tag{163}$$

In the visible range (e.g., $\lambda_0 = 550$ nm) the spectrum of a single optical cycle pulse will extend through the entire range and could be regarded as white light. In that regard we can view such laser sources as **white light lasers**.

In some cases, laser pulses can be compressed outside the resonator by applying the above discussed dispersion compensation techniques. The shortest optical pulses to date were generated by employing the **extracavity compression technique**. It is based on increasing the spectral bandwidth via self-phase modulation (SPM) in a nonlinear medium and subsequent removal of the frequency modulation by using an appropriate dispersive delay line. This constitutes a simple and robust method for obtaining few optical cycle pulses whose spectral widths extend well beyond the gain bandwidth supported by the ultrashort pulse lasers. Pulse compression based on SPM-induced spectral broadening in a bulk solid-state medium offers the advantages of a wide variety of suitable nonlinear materials, technical simplicity, low cost, and easy implementation to virtually any existing ultrashort pulse laser system. For instance, this pulse compression technique attracts an increased practical interest, in particular due to its applicability to novel Yb-based lasers, which produce either relatively long (few hundreds of fs) femtosecond or sub-picosecond pulses with very high average power.

[17] G. Steimeyer, D. H. Suter, L. Gallmann, N. Matouschek, and U. Keller, Frontiers in ultrashort pulse generation: Pushing the limits in linear and nonlinear optics, Science **286**, 1507–1512 (1999)

6.2 Measurement of ultrashort laser pulses

Measurement of the laser pulse duration is a simple task if the pulse duration is long (several ns or longer). The duration of free-running and Q-switched laser pulses is easy to measure using a sufficiently fast photodetector and an oscilloscope. However, this method is unsuitable in the case of ultrashort laser pulses because of insufficient temporal resolution of electronic detection components, i.e., the detector response time is significantly longer than the pulse duration. Therefore, ultrashort pulse duration measurement techniques are essentially different and are based on the **measurement of correlation functions**. One of the most common used correlation techniques is measurement of the second order (intensity) correlation function:

$$I_{AC}(\tau) = \int_{-\infty}^{\infty} I(t)I(t-\tau)dt, \tag{164}$$

where $I(t)$ is a function describing temporal envelope (profile) of the pulse, $I(t-\tau)$ is the same temporal envelope (its replica) shifted in time by τ. By changing the temporal interval τ (i.e., the delay between the original pulse and its replica), the entire correlation function, which is called the autocorrelation function, is measured. It is obvious that $I_{AC}(\tau) = I_{AC}(-\tau)$, i.e., the autocorrelation function has a maximum at $\tau = 0$ and is always symmetrical regardless of the shape of the measured pulse.

In practice, the autocorrelation function of the light pulse is measured using an optical setup which is in principle identical to Michelson interferometer setup, as depicted in Fig.83. The incident light beam is divided into two identical beams (carrying identical light pulses) by means of a beamsplitter. One arm of the interferometer is kept fixed, while the other arm is translated (by means of mechanical translation stage) and in this way temporal delay between the pulse and its replica is changed. Thereafter both beams are focused by a lens into a nonlinear crystal where second harmonic is generated. Since the intensity of the second harmonic radiation is proportional to the product of intensities of the pulse and its replica:

$$I_{2\omega}(t,\tau) \propto I(t)I(t-\tau) = I_{AC}(t,\tau), \tag{165}$$

by changing the delay τ between them and measuring the intensity of the second harmonic, we essentially record the autocorrelation function of the pulse. Moreover, the photodetector performs integration in time, since its response time is significantly longer than the duration of the second harmonic pulse, so the time variable t is eliminated and the measured function is described by the expression in Eq.(164).

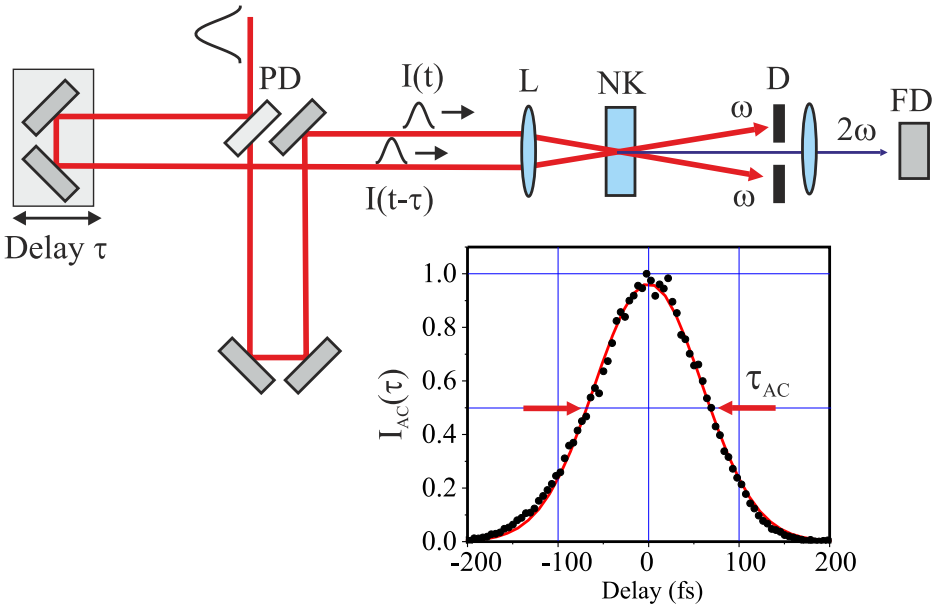


Figure 83: Setup for measurement of the autocorrelation function. PD is the beam splitter, L is the lens, NK is the nonlinear crystal for second harmonic generation, D is the diaphragm and FD is the photodetector. The inset depicts measured autocorrelation function of an ultrashort pulse.

An example of the autocorrelation function is shown in the inset of Fig.83. In order to increase the accuracy of measurement, tens to hundreds autocorrelation function values are measured at each delay, which thereafter are averaged. To further increase the measurement accuracy, each point in the autocorrelation function is normalized to the intensity of the incident pulse, so eliminating pulse-to-pulse intensity fluctuations from the laser source. The step of time delay variation is chosen to provide sufficient time resolution for the measurement. Modern computer-controlled mechanical translation stages ensure high accuracy, so the delay between the pulse and its replica can be varied with an accuracy of 1 μm, i.e., 3 fs. Once the autocorrelation function is measured, it is important to know how the measured duration of the autocorrelation function τ_{AC} is related to the duration of the pulse τ_p. This relation depends on the shape of the measured pulse. The autocorrelation function of a Gaussian pulse also has a Gaussian shape, so $\tau_{AC}/\tau_p = \sqrt{2} = 1.414$. For optical pulses whose shapes are different from Gaussian, the ratio is different as well, e.g., for pulses with squared hyperbolic secant (sech^2) shape this ratio is $\tau_{AC}/\tau_p = 1.543$.

The above discussed method is well suited for measurement of autocorrelation functions of pulses at high repetition rate. However, if the pulse

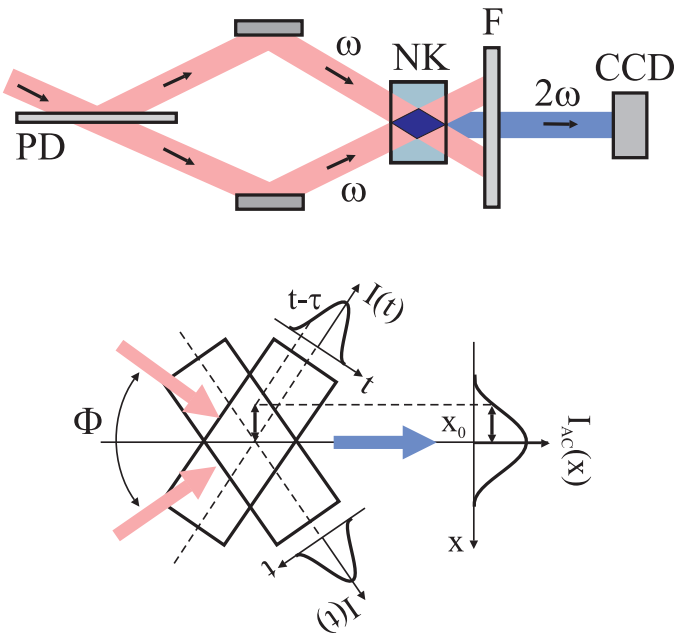


Figure 84: Setup for measurement of the autocorrelation function in a single-shot. PD is the beam splitter, NK is the nonlinear crystal for second harmonic generation and CCD is the CCD camera. The bottom picture depicts how the overlap zone of beams carrying the ultrashort pulses is transformed into the autocorrelation function.

repetition rate is low, measurement of the autocorrelation function (data accumulation and averaging) can take a relatively long time. Therefore, in that case a slightly different method is used which is based on the measurement of autocorrelation function in a single laser shot (i.e., single pulse). The layout of a single-shot autocorrelator is depicted in Fig.84. The light beams in a nonlinear crystal are crossed at a certain angle Φ and their longitudinal dimensions correspond to the temporal intensity distribution of the pulse $I(t)$. The beam and pulse intersection zone in the nonlinear crystal has the shape of a rhombus and only in this zone the second harmonic is generated. The second harmonic signal has a certain spatial intensity distribution which is recorded by the CCD camera. From the intersection geometry it is easy to estimate that spatial and temporal coordinates are related as follows

$$\tau = \frac{n_0 x_0}{c} \sin \frac{\Phi}{2}, \tag{166}$$

where n_0 is the refractive index of the nonlinear crystal. The spatial intensity distribution of the second harmonic corresponds to the autocorrelation function

$$I_{AC}(x_0) = \int_{-\infty}^{\infty} I(x)I(x - x_0)dx. \tag{167}$$

The temporal delay in such setup is calibrated very simply: when one of the pulses is delayed, the image of autocorrelation function on the x axis shifts accordingly. Since the temporal autocorrelation function is converted into spatial intensity distribution, the measurement accuracy also depends on the beam diameter. In this case it is chosen such that it would be much greater than longitudinal pulse parameters (its duration).

The intensity autocorrelation function has one significant advantage. This method background free: if the pulses do not overlap in time, the measured signal is zero. On the other hand, since the autocorrelation function is always symmetrical even when the intensity profile of the measured pulse is not, or the pulse shape is complex (i.e. containing several intensity peaks, subpulses, etc.), this kind of information is not retrieved. This shortcoming is circumvented by the measurement of cross-correlation function. The cross-correlation function is produced in a similar manner as the autocorrelation function, just instead of the pulse replica, a pulse with a well-known known shape (usually much shorter as well), which is called the probe pulse, is used:

$$I_{CC}(\tau) = \int_{-\infty}^{\infty} I(t)I_z(t - \tau)dt, \tag{168}$$

where $I_z(t)$ is the intensity of probe pulse. It is obvious that the cross-correlation function is not symmetrical and the cross-correlation technique is often applied for the measurement of complex pulses.

The temporal information can also be extracted by measuring higher-order, e.g., **third order correlation function**:

$$I^{(3)}(\tau) = \int_{-\infty}^{\infty} I(t)^2 I(t - \tau)dt. \tag{169}$$

Third-order correlation function is obtained by making use of third-order nonlinear processes, which take place in media with cubic nonlinearity. These media are isotropic materials, e.g., fused silica. It is easy to notice that in the case of third-order autocorrelation function $I^{(3)}(\tau) \neq I^{(3)}(-\tau)$, the information on the direction of the time axis is maintained. It is important that the wavelength of the third-order autocorrelation signal is the same as that of the incident radiation. Third-order autocorrelation functions are measured when wavelength of radiation is short (e.g., in the UV spectral region) and where usual second-order correlation methods cannot be applied for some reason.

Very often, besides the duration and shape of the ultrashort pulse, the knowledge of its spectral and temporal phase is required, especially concerning optimization of pulse compression setups. Measurement methods that enable phase retrieval are based on the measurement of instantaneous frequency of the autocorrelation function. The measurement setup is basically identical to that shown in Fig.83, only in the present case instead of a photodetector a spectrometer is used. These methods are called **FROG**; the abbreviation stands for frequency-resolved optical gating. The obtained results (the so-called frogograms) are processed using mathematical algorithms which are able to retrieve the actual shape and instantaneous phase of the pulse.

6.3 Amplification of ultrashort light pulses

Amplification of ultrashort light pulses is one of the most important tasks in laser physics. Mode-locking technique enables the generation of very short light pulses, but their energy is low (varies from several nJ to several μJ) and the peak power does not exceed tens of MW. To increase the peak power up to GW, TW or even PW level, laser amplifiers are used.

All modern laser systems which produce ultrashort pulses with high peak power consist of: 1) a laser oscillator which operates in the mode-locking regime and generates very short light pulses with low energy; 2) a laser amplifier, whose purpose is to increase the pulse energy (and the peak power) as much as possible without distorting its temporal and spatial characteristics. Here we also note that there is a principal difference between the pulse repetition rate of the oscillator and the amplifier. In a quasi-cw oscillator, the pulse circulates in a resonator and repeats itself after each round-trip, so at the output of the oscillator a train of pulses with a period inversely proportional to resonator round-trip time (a repetition rate of tens on MHz) is obtained. The amplifier usually amplifies only one pulse from the sequence (e.g., every 100th or every 1000th), so the pulse repetition rate at the output of the amplifier is determined by the repetition rate of amplifier pump source, thermal effects in the amplifier medium and the ability to electronically control operation of the amplifier.

The purpose of the amplifier is to increase energy of the ultrashort pulse without introducing any distortion to the pulse and beam. The main factors limiting efficient amplification of ultrashort pulses are attributed to self-action phenomena, which may occur in the optical elements of the amplifier or in the amplifying medium itself. Self-action phenomena arise from the optical Kerr effect, i.e. the intensity-dependent refractive index of transparent materials: self-phase modulation, which causes

spectral broadening of the pulse and distortion of its phase characteristics (bandwidth-limited pulse becomes chirped) and self-focusing of the beam, which causes uncontrolled decrease of the beam diameter and increase of the intensity, which eventually leads to damage of amplifier components.

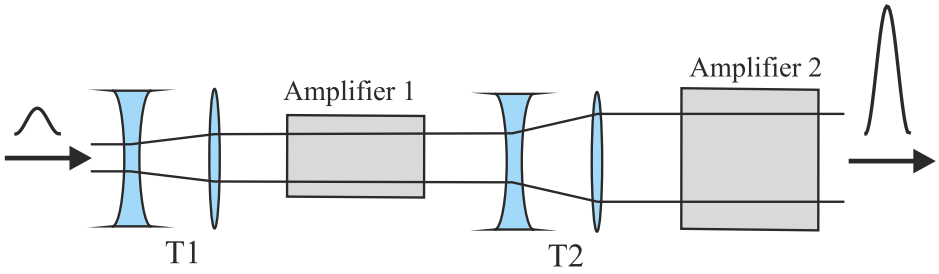


Figure 85: The layout of a two-stage single-pass laser amplifier. T1 and T2 are telescopes which increase the size of the beam.

The simplest amplifier is a single-pass (sometimes called linear) amplifier which is schematically depicted in Fig.85. The gain per single pass of such amplifier is not large, and varies from several to several tens of times, depending on the amplifier active medium. Since the intensity is inversely proportional to the beam area, the beam dimensions must be enlarged before each amplification stage as to keep the intensity below the threshold for self-action effects. It is obvious that in order to effectively increase pulse energy, the size of the amplifier medium in each successive stage increases very quickly. This causes new problems: how to achieve homogeneous population inversion within the entire volume of the amplifier medium, at the same time avoiding thermal effects due to its limited thermal conductivity. As a result, the repetition rate of the amplifier drops significantly: such amplifier systems can only operate at very low (from several to several tens of Hz) repetition rates.

The physical limits of such amplifier setups were reached around 1985, and further increase of the amplified pulse energy became hardly possible. These fundamental difficulties in short pulse amplification were circumvented by the invention of **Chirped Pulse Amplification (CPA)** technique in 1985 by D. Strickland and G. Mourou in 1985[18]. The principle of CPA is depicted in Fig.86. The idea of the CPA concept is to boost the energy of an ultrashort pulse, while avoiding very high fluence in the laser amplifier. In contrast to the above discussed approach, this is done in an elegant way: the ultrashort, broadband pulse is first stretched (chirped) in time without the

[18]G. Mourou and D. Strickland, Compression of amplified chirped pulses, Optics Communications **56**, 219–221 (1985)

loss of its spectral content by the use of dispersive elements. The stretching factor may vary from hundreds to thousands of times, in this way significantly reducing the pulse intensity. The long chirped pulse is then amplified in a laser amplifier by a factor of $10^4 - 10^6$ and then recompressed to the original duration by dispersive elements, which introduce an opposite dispersion as the stretched does. The CPA technique solved the long-standing problem of safe and efficient amplification of ultrashort optical pulses without the onset of optical damage of the amplifier material and other optical components, enabling a tremendous leap in the peak power and intensity of laser pulses, boosting an exciting progress in ultrafast laser technology. At present, all modern ultrashort pulse laser systems are based on this principle[19], while the inventors of this technique were awarded the Nobel Prize in Physics in 2018.

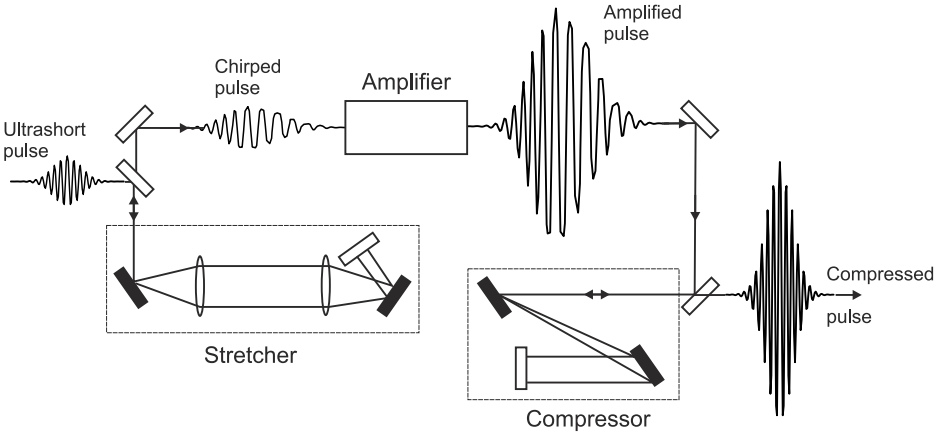


Figure 86: The principal layout of chirped pulse amplification.

After the advent of the CPA technique, construction of the amplifiers themselves was also improved since stretching of the pulse in time and the reduction of its intensity eliminated the need to enlarge beam dimensions each amplification stage. Currently, two configurations of laser amplifiers are in use: multi-pass and regenerative.

Multi-pass amplifier. A multi-pass amplifier is essentially a "folded" single-pass amplifier, its layout is schematically depicted in Fig. 87. The multi-pass amplification is very efficient when the gain coefficient of the laser medium is large and the beam area is relatively small. This setup perfectly fits Ti:sapphire, which provides large amplification per single-pass and which makes laser pump of the active medium easy to implement in a small volume.

[19]S. Backus, C. G. Durfee III, M. M. Murnane, and H. C. Kapteyn, High power ultrafast lasers, Review of Scientific Instruments **69**, 1207–1223 (1998)

The number of passes through the active medium is adjusted with the mirrors and the main advantage of such amplifier is simplicity of the setup which does not require any complex control elements. However, the mechanical stability of such amplifier system is low: even a small misalignment of mirrors significantly alters direction the light beam and small beam distortions occurring after each pass only increase.

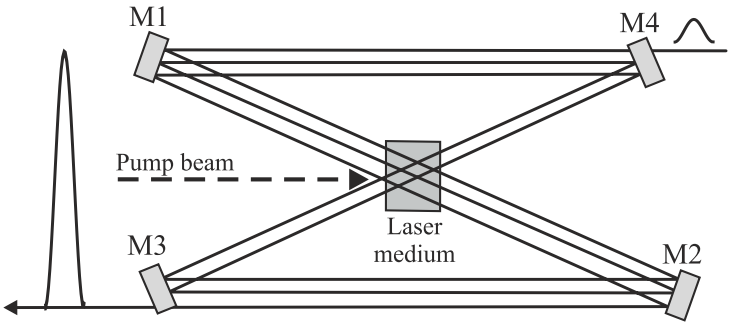


Figure 87: Layout of the multi-pass amplifier.

Regenerative amplifier.

The setup of a regenerative amplifier (Fig. 88) is very similar to that of a Q-switched laser. Here the electro-optic switch (the Pockels cell and polarization analyzer) performs functions of coupling the seed pulse (the stretched oscillator pulse) into the resonator and dumping the pulse after amplification.

The principle of operation of the regenerative amplifier is as follows. Let us assume that initially the resonator is closed ($\lambda/4$ voltage is applied to the Pockels cell). Then the amplifier medium accumulates population inversion. When the seed pulse arrives, it is reflected from the polarizer and passes the Pockels cell which turns its polarization into circular. After reflection from the mirror and second pass through the Pockels cell, pulse polarization becomes perpendicular with respect to the initial polarization, therefore, it passes the polarizer and is coupled into the resonator.

At this moment the voltage applied to the Pockels cell is turned off and the resonator quality suddenly increases, making the pulse "trapped" and amplified in the resonator. The pulse makes several tens to several hundreds of round-trips in the resonator to reach the gain saturation. When the gain saturates, the $\lambda/4$ voltage is again applied to the Pockels cell which rotates the polarization by 90° (note a double pass) and the amplified pulse is coupled out (dumped). It is obvious that the regenerative amplifier operates as a Q-switched laser when there is no seed pulse, but the resonator is opened.

The main advantage of the regenerative amplifier is its resonator which ensures very high spatial quality of the amplified beam due to diffraction losses. Any spatial distortions occurring during amplification or present in the seed signal are filtered out. Precise timing between the pulse seeding and dumping moments and the pump pulse is adjusted by the electronic control system.

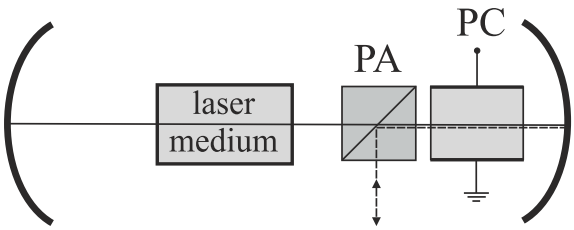


Figure 88: Layout of the regenerative amplifier. PE is the Pockels cell, PA is the polarization analyzer.

6.4 Pulse stretchers and compressors

Pulse stretchers and compressors are very important parts of the CPA-based laser system. Their purpose is to stretch and compress the light pulse in time without changing its spectral bandwidth. As discussed in the previous chapter, these manipulations could be performed using dispersive elements (prism pair and chirped mirrors). CPA requires very large pulse stretching and compression factors (300÷10000), so for this purpose a diffraction grating pair is used which yields very high group velocity dispersion.

Let us briefly discuss the performance of pulse stretchers and compressors. Let us first discuss operation of a **pulse compressor**. A pair of two parallel identical diffraction gratings, like a pair of dispersive prisms, introduces anomalous group velocity dispersion through the difference of optical paths. Due to fundamental diffraction law, angles for the short-wavelength radiation are smaller, so the "blue" spectral components travel a shorter distance than the "red" spectral components, as shown in Fig. 89, so the grating pair compressor introduces large anomalous group velocity dispersion.

The difference in optical paths between distinct frequency components can be estimated according to the diffraction grating equation:

$$\sin\gamma + \sin\theta = \frac{\lambda}{d}, \tag{170}$$

where γ and θ are the incidence and diffraction angles, respectively and d is the period of the diffraction grating. The optical path of any frequency

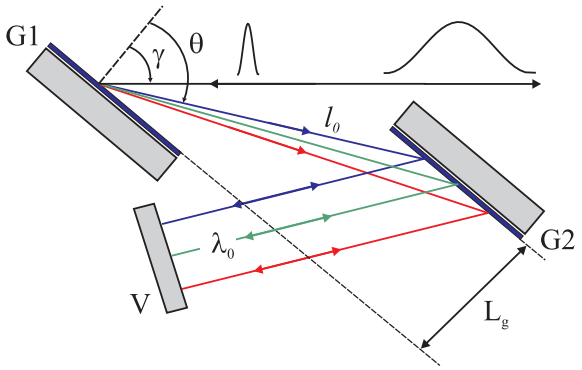


Figure 89: Principal setup of a diffraction grating pulse compressor. G1, G2 are the diffraction gratings, V is the mirror.

component can be estimated as

$$P = \frac{L_g}{\cos\theta}\ 1 + \cos(\gamma - \theta)\ , \tag{171}$$

where L_g is the distance between the gratings. The optical path difference between the marginal spectral components of the pulse determines the temporal delay between them $\Delta t = \Delta P/c$ and so the compression factor. In analogy with a prism pair compressor, the light beam is reflected back by a plane mirror (see Fig. 89) to pass the grating compressor for a second time to eliminate the spatial chirp (coloration of the beam). As a result, the time delay between the marginal spectral components (the compression factor) is doubled ($2\Delta t$). Here we note that the compression factor of grating compressor depends on the spectral bandwidth of the pulse, and not on its initial duration.

The same considerations apply to the **pulse stretcher**. However, the pulse stretcher must introduce the group velocity dispersion with an opposite sign as compared to pulse compressor, implying that the configuration of the pulse stretcher is different from that of the pulse compressor. Let us discuss how the pulse stretcher works. First of all, a pulse stretcher must provide large normal group velocity dispersion. This is determined also by the fact that the group velocity dispersion of the laser medium and other optical elements in the amplifier is also normal, so it adds to the present dispersion. The simplest way to stretch the pulse in this way is to make a pulse propagate in a long dispersive medium, e.g. optical fiber. Indeed this approach is used, however it is feasible only in the case of the pulses with very low peak power as to avoid accumulation of nonlinear effects during propagation in the optical fiber.

A considerably more flexible approach is to use a **pair of diffraction gratings with a telescope** in between, as shown in Fig.90.

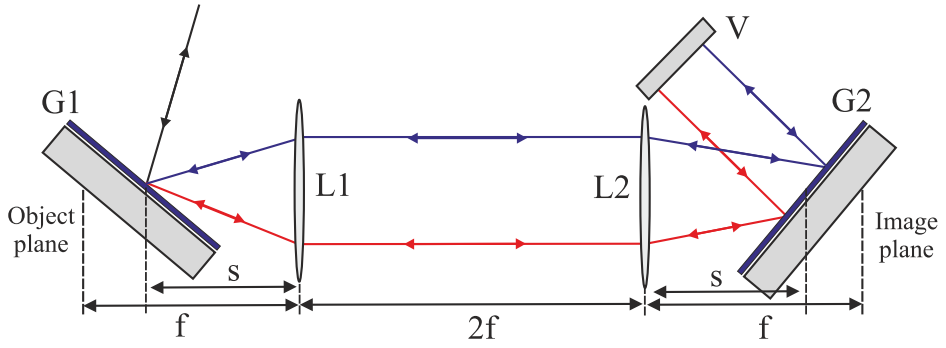


Figure 90: Layout of a grating pulse stretcher. G1, G2 are the diffraction gratings, L1, L2 are the telescope lenses, V is the mirror.

If both gratings are placed exactly at the focal planes of the lenses ($s = f$), group velocity dispersion of such system is zero, as the optical paths of all spectral components are equal, and pulse duration does not change when passing through such arrangement. Moving the gratings closer to the lenses (as depicted in Fig.90), the group velocity dispersion of the system becomes normal: the optical path of "red" spectral components becomes shorter than that of the "blue" spectral components. The closer to the lenses are the gratings, the larger is the introduced normal group velocity dispersion. The effective length of such pulse stretcher is expressed as:

$$L_{\text{eff}} = -2(f - s), \tag{172}$$

where s is the distance between the grating and the closest lens. A reflecting mirror plays the same role as in the compressor: the second pass through the pulse stretcher eliminates spatial distribution of the spectral components (spatial chirp) and doubles the dispersion. If the distance between the grating and lens is made longer than the focal length of the lens ($s > f$), the setup introduces anomalous group velocity dispersion, and the stretcher works as a simple grating compressor. The lengths of pulse stretcher and compressor are adjusted so that the latter also compensates the amount of normal group velocity dispersion accumulated by the propagation in the amplifier optics.

The optical path lengths of grating stretchers and compressors in ultrashort pulse laser systems are of the order of a meter and typical grating periods are $d = 1200 \div 1800$ mm^{-1}. The diffraction gratings are aligned in such a way that incidence and reflection angles for the central frequency (wavelength) of the pulse are equal: $\gamma = \theta(\lambda_0)$. This angle is termed the

autocollimation (or Littrow) angle. For this angle a specific profile of diffraction grating blaze is used to achieve the maximum diffraction efficiency (a fraction of the incident energy diffracted to the first diffraction maximum). In the near-IR spectral range the diffraction efficiency of aluminum-coated diffraction grating is up to 90%, for gold-coated grating it is up to 95% and gratings with special dielectric coatings have up to 98% diffraction efficiency. The diffraction efficiency of the compressor is very important, for example if the grating reflects 90% of the incident energy to the first diffraction maximum, then after four reflections in the compressor only 66% of amplified pulse energy remains ($0.9^4 \approx 0.66$), while the increase of reflection up to 98% yields the throughput increase up to 92%. For the pulse stretcher this parameter is of much less importance, since this kind of energy losses can be easily compensated during amplification.

6.5 Laser amplifier media

There are strict requirements for laser media which can be used to amplify ultrashort pulses:

- One of the most important requirements is high saturation fluence of the laser amplifier medium, which is defined as

$$\boxed{J_{\text{sat}} = \frac{h\nu}{\sigma}}. \tag{173}$$

  The saturation fluence for dye and most of gas lasers is ~ 2 mJ/cm^2 which means that in order to get, for example, 5 mJ energy amplified pulse, the diameter of the active medium should be around 2 cm. Therefore, due to low saturation fluence dye and gas lasers are very inefficient amplifiers. For solid-state lasers the saturation fluencies are much higher, e.g. 0.9 J/cm^2 for Ti:sapphire, 7 J/cm^2 for Nd:glass and 32 J/cm^2 for Yb:glass. In a 1 cm diameter Ti:sapphire amplifier it is possible to obtain the amplified pulse with an energy of 1 J.

- Another important requirement for laser amplifier medium is broad amplification bandwidth, which ensures that all spectral components of the chirped pulse are amplified; otherwise, the amplified pulse cannot be compressed to its initial duration.

- Long lifetime of the upper laser energy level, which determines the ability to accumulate large population inversion.

Summarizing these requirements we can define a certain figure of merit for the amplifier medium:

$$\boxed{M = \frac{t_{\text{spont}}}{t_p J_{\text{sat}}}}. \tag{174}$$

The evolution of achieved laser powers from the invention of the laser is illustrated in Fig.91.

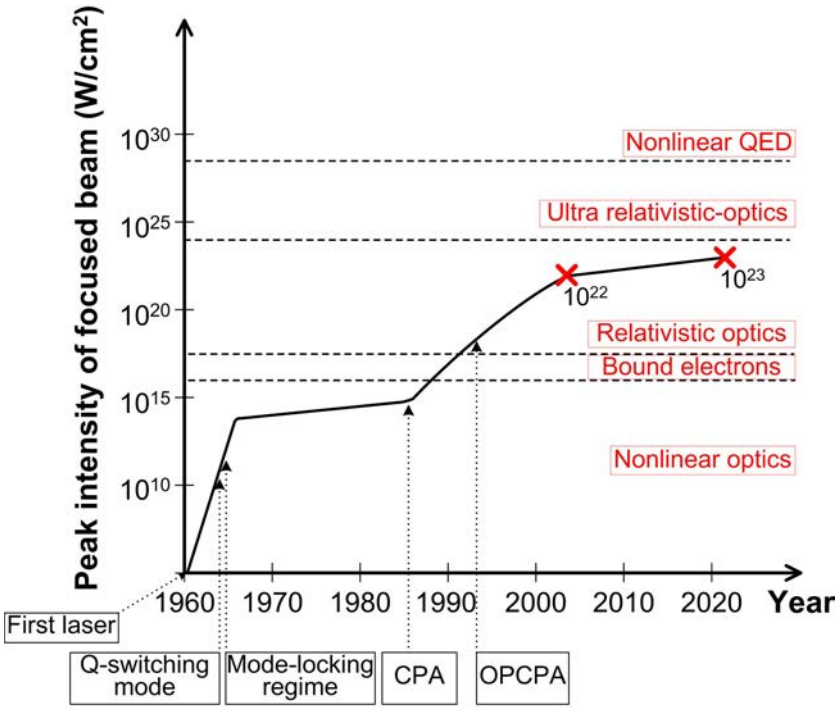


Figure 91: The maximum focused laser radiation intensity versus year.

The early rise of the focused intensity was achieved thanks to the invention of laser pulse generation techniques, Q-switching and mode-locking and subsequent amplification of mode-locked pulses. Thereafter, the progress has slowed down due to physical limits of short pulse amplification, which were imposed by the optical damage of the laser amplifier material. A major breakthrough came in 1985, when the chirped pulse amplification technique was invented, which currently lies on the basis of every modern ultrashort pulse laser system. Apart from tremendous impact on laser physics, invention of CPA facilitated rapid development of experimental sciences and opening new areas of physics, technology and multidisciplinary research. In particular, CPA-based lasers provided experimental access to strong-field physics

and relativistic nonlinear optics, eventually drawing the guidelines for probing vacuum nonlinearity and reaching the fundamental laser intensity limits. At present, the most powerful laser systems employ Nd:glass and/or Ti:sapphire as amplifier media. Fabrication technology of Nd:glass is well developed, there are many efficient pump sources for this amplifying medium and the wavelength of Ti:sapphire oscillator can be matched with the Nd:glass emission spectrum. However, Nd:glass is an amorphous medium, its thermal conductivity is low which significantly reduces repetition rate of the amplified pulses. Moreover, amplification bandwidth of Nd:glass can sustain only pulses with sub-picosecond duration. The highest peak power of 1.5 PW was achieved in a hybrid Ti:sapphire and Nd:glass laser system, which produced 0.44 ps pulses with an energy of 660 J, yielding the focused intensity of $> 7 \times 10^{20}$ W/cm^2 [20]. More recently, 2 PW femtosecond laser system based on sole Ti:sapphire laser medium was reported to produce the compressed pulses with a duration of 26 fs and an energy of 72.6 J[21]. Currently the highest reported focused beam peak intensity is 1.1×10^{23} W/cm^2 [22].

Let us determine what the maximum peak power can be achieved by amplifying ultrashort light pulses. Assuming that t_p is duration of the amplified pulse, which is inversely proportional to its spectral width ($t_p = 0.441/\Delta\nu$):

$$P_{\text{th}} = \frac{J_{\text{sat}}}{t_p} = \frac{h\nu}{0.441\sigma}\Delta\nu. \tag{175}$$

Assuming that we can focus such a beam into a spot size of the order of a wavelength, we get the maximum intensity:

$$I_{\text{th}} = \frac{P_{\text{th}}}{\lambda^2} = \frac{h\nu^3}{0.441\sigma}\frac{\Delta\nu}{c^2}. \tag{176}$$

It is easy to estimate that for known laser materials $I_{\text{th}} \approx 10^{24}$ W/cm^2.

[20] M. D. Perry et al, Petawatt laser pulses, Optics Letters **24**, 160–162 (1999)

[21] Y. Chu et al, High-contrast 2.0 Petawatt Ti:sapphire laser system, Optics Express **21**, 29231–29239 (2013)

[22] J. Woo Yoon et al, Realization of laser intensity over 10^{23} W/cm^2, Optica **8**, 630–635 (2021)

Summary of Chapter 6

- Dispersive broadening in laser resonator prevents the generation of ultrashort light pulses carrying very broad spectrum. Dispersion compensation is performed by introducing a delay of the fastest (long-wavelength) spectral components with respect to the slowest (short-wavelength) ones by using prism pairs and/or chirped mirrors
- Since mode-locked lasers generate ultrashort pulses with very low energy (from several nJ to several μJ), the energy and peak power scaling of these pulses is performed by amplification in single-pass, multi-pass and regenerative laser amplifiers.
- Chirped pulse amplification (CPA) is the technique that allows boosting the energy of an ultrashort pulse, while avoiding optical damage of the laser amplifier. In doing this, the ultrashort oscillator pulse is stretched in time, amplified in a laser amplifier (either multi-pass or regenerative) and then re-compressed to the original duration.
- CPA technique constitutes the standard conceptual basis of modern ultrafast solid-state lasers and laser systems.

7 Types of lasers

In this chapter we will briefly discuss the basic features, materials and principles of operation of solid-state, gas and semiconductor lasers.. The key difference between these three types of lasers is that the laser radiation is generated by very different means. In solid-state lasers, the laser radiation is produced by rare-earth or transition metal ions embedded in a solid-state host material. In gas lasers the laser radiation is produced by neutral atoms, ions or molecules, while in semiconductor lasers the laser radiation is generated at the p-n junction by recombination of electron-hole pairs, i.e. is produced by the semiconductor crystal itself. These key differences determine the construction principles, operation regimes, parameters of output radiation and the range practical applications of these laser types.

7.1 Solid-state lasers

Solid-state lasers are the main laser sources for the generation and amplification of ultrashort light pulses and therefore play a central role in the development of modern ultrashort pulse science and technology. Solid-state lasers can efficiently perform in all regimes of laser operation: free-running mode, Q-switching and mode-locking.

Fabrication of solid-state laser materials essentially relies on three criteria, which determine the basic features of such laser operation, construction and characteristics of the generated pulse:

- the activator ions, which have a specific configuration of energy levels;
- the laser host material having specific mechanical, optical and thermal properties;
- the optical pump source with its spectral and temporal properties and pump geometry.

Solid-state laser materials are crystals or glasses (laser hosts) doped with small amounts of activator ions which produce laser radiation. In general, solid-state laser materials have to have sufficient narrow emission lines (except those used for ultrashort pulse generation), strong absorption bands or lines and high quantum efficiency of the laser transition. The host material, where these activator ions are embedded, must be transparent to both pump and laser radiation, and the absorption of an ion must be in the spectral

region where optical pumping with flashlamps, semiconductor laser diodes or other lasers is possible[23].

Activator ions.

Ions of a certain group of chemical elements termed rare-earth metals, lanthanide and actinide, serve as activator ions in solid-state lasers. Rare-earth metal ions (see the two special rows positioned at the very bottom of the periodic table shown in Fig. 92) possess narrow emission linewidths (this essentially ensures that the lifetime of the upper laser level is long) and electronic transitions occur between the inner layers which are not fully occupied. Since the electronic transitions occur between the inner layers which are shielded by the outer layers, emission spectrum of such ion is almost identical to that of a free ion and is almost independent of the laser host the ion is embedded in.

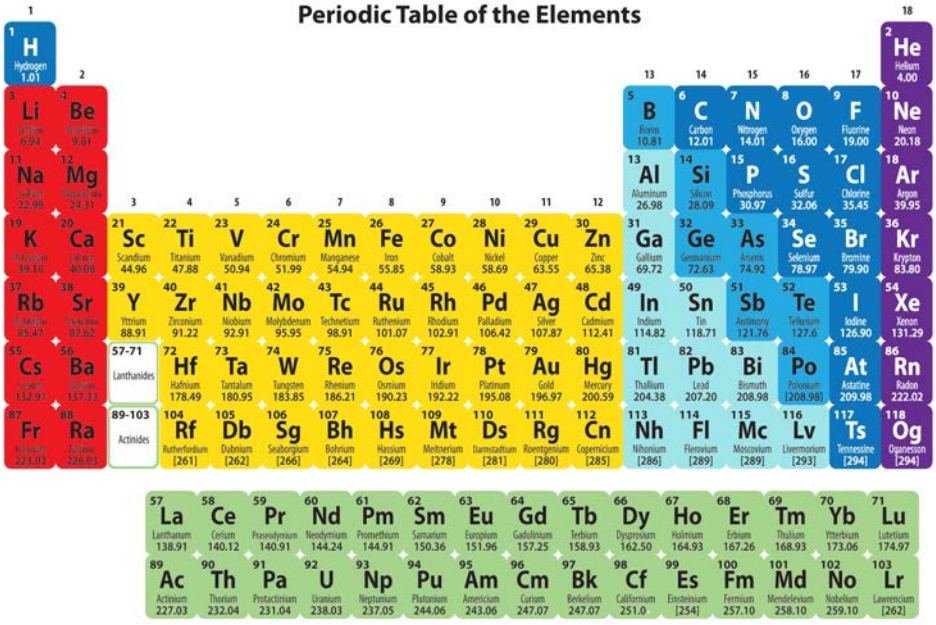


Figure 92: Periodic table of elements.

The configuration of electron layers in rare-earth metals is similar to Xe, in which layers with quantum numbers $n = 1, 2, 3$ are fully occupied, sublayers s, p, d of layer $n = 4$ are also fully occupied and $4f$ sublayer which can contain up to 14 electrons is empty, whereas sublayers $5s$ and $5p$ are fully occupied again:

[23]M. Eichhorn, Quasi-three-level solid-state lasers in the near and mid infrared based on trivalent rare earth ions, *Applied Physics B* **93**, 269–316 (2008)

$$\text{Xe}: \quad 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 5s^2 5p^6$$

In the case of rare-earth metals, electrons further fill the empty orbital of sublayer $4f$:

$$\begin{aligned} \text{Ce}: &\quad ...4s^2 4p^6 4d^{10} 4f^2 5s^2 5p^6 6s^2 \\ \text{Nd}: &\quad ...4s^2 4p^6 4d^{10} 4f^4 5s^2 5p^6 6s^2 \end{aligned} \tag{177}$$

As a rule, rare-earth metal ions are trivalent; the neutral atoms have lost three electrons from $6s$, $5d$ (if present) and $4f$ sublayers:

$$\begin{aligned} \text{Ce}^{3+}: &\quad ...4s^2 4p^6 4d^{10} 4f^1 5s^2 5p^6 \\ \text{Nd}^{3+}: &\quad ...4s^2 4p^6 4d^{10} 4f^3 5s^2 5p^6 \end{aligned} \tag{178}$$

After excitation of such ions, electronic transitions between $4f$ sublayer levels occur, which are fully shielded by $5s$ and $5p$ sublayers.

The lanthanide ions emit radiation in the near-IR spectral range, e.g., the wavelengths of strongest emission lines are 1.06 μm in Nd^{3+}, 1.03 μm in Yb^{3+}, 1.53 μm in Er^{3+}, 2.0 μm in Ho^{3+}, etc. Actinides also possess a shielded inner electron sublayer (which is shielded by $6s$ and $6p$ sublayers), however all these elements and their ions are radioactive and therefore are not used.

Another metal ion group consists of trivalent and bivalent transition metal ions: Cr^{3+}, Ti^{3+}, Ni^{2+}, etc. Laser levels of transition metal ions are not shielded, as the electronic transitions occur between outer $3d$ sublayer levels. Therefore embedding these ions into different materials, significantly alters their emission wavelength. Due to strong interaction between the electronic energy levels and the electric field of the crystal, the energy levels of transition metal ions may become significantly broadened. Such lasers have a broad emission band (although the linewidth broadening is homogeneous) and often are called vibronic.

Laser hosts.

Laser hosts are the materials into which the activator ions are embedded. Laser hosts have to comply with strict requirements which can be summarized as follows:

- good optical quality, which includes broad transmission range (transparency to both laser and pump radiation), homogeneity (smoothness of refractive index) and purity;
- good mechanical, thermal and chemical properties: large thermal conductivity, rigidity, chemical stability and resistance to color center formation;
- activator ions in laser host must maintain their energy level structure;
- crystal lattice structure of laser host must not be altered during doping and should preserve its optical, chemical, mechanical and thermal properties;
- fabrication technique of laser host material has to be as simple and cheap as possible.

Laser hosts, according to their structure are divided into two large groups: glasses and crystals. Glasses are essentially isotropic and amorphous materials produced by cooling of liquids and alloys. Among huge variety of glasses, the main types of glasses that serve as laser hosts are silicate (SiO_2), phosphate (P_2O_5) and fluoride (usually a mixture of various fluorides: ZrF_4, BaF_2, LaF_3, AlF_3 and NaF, the so-called ZBLAN). Glasses can be molded with high optical quality and in large dimensions; the optical fibers are also made from these materials. Glasses are easily processed and could be doped to a high degree. Since glasses are amorphous materials, their thermal conductivity is low, so imposing limits on the repetition rate of glass-based lasers and amplifiers. Nd^{3+}, Yb^{3+} and Er^{3+} ions serve as the main activators used for glass host materials. In general, the marking of laser material includes the names of a doping ion and laser host, e.g. Nd:glass, which reads as glass host material doped with Nd^{3+} ions. The charge of an ion is usually omitted.

Compared to glasses, crystalline hosts possess very good optical properties and thermal conductivity in particular. However, production technology of crystalline hosts is more complex, since during the crystal growth, the original ion in the crystal lattice is replaced with the activator ion. For this exchange to be successful and no residual defect to appear in the lattice, the charge of the activator ion has to be identical and its physical size has to be very close to the replaced element.

The main and most widely used crystalline hosts are oxides and fluorides, which have very broad transparency range. Basically, these are dielectric crystals with a large energy bandgap.

Sapphire (Al_2O_3) is an excellent optical material possessing high crystalline quality, and high optical damage threshold, its production technology

is relatively simple and well elaborated. While doping the sapphire crystal, Al^{3+} ions are replaced with trivalent activator ions. However, there is a slight problem: the size of Al^{3+} ion is very small, only ≈ 0.6 Å, while all rare-earth metal ions are almost twice larger. Therefore, A3+ ions in the sapphire crystal could be replaced only by transition metal ions: Cr^{3+} and Ti^{3+}, which have similar size (≈ 0.7 Å). Consequently, two doped sapphire materials are in use: Ti:sapphire and Cr:sapphire (ruby), the latter exists naturally in nature.

Apart sapphire, other gem class crystals could be doped with transition metal ions: Cr:$BeAl_2O_4$ (alexandrite), Cr:$Be_3Al_2(SiO_3)_6$ (emerald), etc. Some fluoride crystals, called colquirites, are also doped with transition metal ions (usually Cr^{3+}): $LiCaAlF_6$ (LiCAF), $LiSrAlF_6$ (LiSAF) and $LiSrGaF_6$ (LiSGAF). Compared to sapphire, their thermal properties are significantly poorer, but they have a very broad amplification bandwidth (they are so called vibronic materials) that affords either generation of femtosecond pulses or production of wavelength-tunable laser radiation.

Garnets comprise the most widely used oxide group. Out of these, the most popular and widespread is yttrium aluminium garnet, $Y_3Al_5O_{12}$ (YAG). YAG exhibits outstanding mechanical, thermal, and optical properties, which makes it a versatile optical material that is widely used in optoelectronics, laser physics, and nonlinear optics. YAG is one of the most important laser host materials, which could be doped with various rare earth metal ions, since radius of yttrium ions (≈ 1.1 Å) is practically identical to rare-earth metal ions radii. Neodymium doped YAG (Nd:YAG) is one of the most popular laser materials, which has low generation threshold, large gain and could be pumped with flashlamps and laser diodes. Currently, several other artificial garnets were produced: $Gd_3Ga_5O_{12}$ (GGG), $Ga_3Sc_2Al_3O_{12}$ (GSGG), etc. Garnets could also be doped with Yb^{3+} ions. For example, in the meantime a lot of research is carried out to develop Yb:YAG based laser systems that are capable of operating at very high repetition rates and producing ultrashort light pulses with sub-picosecond duration. Among the other commonly used laser hosts it is worth to mention yttrium-based crystals: yttrium vanadate (YVO_4) and yttrium lithium fluoride $YLiF_4$ (YLF) which all are doped with Nd^{3+} ions.

The search for new efficient laser materials continues nowadays, for example, a recently invented but already commercially used laser material is ytterbium ion doped potassium gadolinium tungstate $KGd(WO_4)_2$ (Yb:KGW). Another laser research line is directed toward the development of ultrashort pulse lasers operating in the mid-IR spectral range. To this end, novel laser sources based on Ho-doped gain media (Ho:YLF and Ho:YAG), operating around 2 μm and delivering multimilijoule, multigigawatt pulses with dur-

ation of a few picoseconds were developed more recently. A particularly promising and interesting research line concerns the development of transition metal doped chalcogenide lasers (Cr or Fe doped ZnS and ZnSe), which produce femtosecond pulses in the wavelenght range from 2.3 μm to 3 μm and are often termed as "Ti:sapphire of mid-infrared".

In conclusion, a huge variety of solid-state laser materials was invented and demonstrated in operation. However, only a limited cnumber of these have a commercial perspective, which is defined by ability to provide a specific wavelength of laser emission, pulse duration and energy, repetition rate, etc., which are essential for practical applications.

7.2 Gas lasers

Unlike in solid-state lasers, linewidth broadening in gaseous media is insignificant (practically the only important broadening mechanism is the Doppler effect), so gas lasers, as a rule, emit narrowband radiation. Most of gas lasers (except CO_2, metal vapor and excimer lasers) operate in only continuous wave (CW) regime; their ability to generate or amplify light pulses is poor due to narrow gain bandwidth and low saturation fluence. However, due to their diversity, gas lasers cover a very wide spectral range, from the vacuum ultraviolet to the mid-infrared. So the main applications of gas lasers are oriented to medicine and industry.

Gases are transparent over a very broad spectral range (almost throughout the entire optical range), and electronic energy levels in gas atoms and molecules are usually distributed quite widely. Therefore gas lasers cannot be pumped optically and due to large quantum defect gas lasers are less efficient than solid-state lasers. Almost all of gas lasers are pumped by an electric discharge which is created along the gas tube by applying high voltage. Gas atoms or molecules are excited by electrons, which are accelerated by the electric field and transfer their kinetic energy during collisions. In many cases low pressure gases are used as laser media, since it is much easier to create a homogeneous discharge at low pressure. Mixtures of gases are used to significantly increase the efficiency of electrical pumping: additive gas is selected in such a way that its energy levels are very similar to that of radiating gas. Some of the gas lasers are pumped chemically, where population inversion is achieved due to excess energy during chemical reactions; these lasers are called chemical lasers, or temporarily creating molecules which exist only in excited state, the so called excited dimers; these lasers are called excimer lasers.

In general, it is more difficult to achieve the population inversion in gases than in solid-state media. Although general requirements for laser energy

levels are the same for all laser media, gas media have certain specifics. It can be briefly described as factors reducing the lifetime of the upper laser level:

- collisions between gas atoms (molecules or ions) and electrons: the excitation energy can be converted into electron kinetic energy;
- collisions between excited and not excited atoms. In this way the excitations of the upper laser level are lost because they are become redistributed between the other energy levels;
- collisions of excited atoms with the gas tube walls;

The above factors suggest that in order to achieve population inversion in gases, the pump rate has to be sufficiently high. With a reference to the gases that are used, gas lasers can be classified into 5 types:

- **Neutral atom gas lasers.** The laser media are noble gases and their mixtures, also metal vapors. All noble gas lasers emit in the infrared, except the He-Ne laser which emits in the visible spectral range, at the wavelength of 632.8 nm. Metal vapor lasers emit visible light and operate only in a pulsed free-running regime. Continuous population inversion cannot be achieved in these media, because their lower laser energy level is depleted only through collisions with gas tube walls. Metal vapor lasers operate only at high ($> 1000°$C) temperature, since the metal has to be evaporated. The best-known example of metal vapor laser is copper vapor laser (CVL).
- **Ion lasers.** The energy levels of an ionized atom are shifted with respect to those of a neutral atom due to an unbalanced charge. As a result, the gaps between energy levels increase, and all ion (noble gas, metal vapor) lasers emit at shorter wavelengths, in the visible and ultraviolet spectral range. The most important ion gas lasers are argon ion Ar^+ laser and helium-cadmium He-Cd laser.
- **Molecular gas lasers.**

  In molecular gas lasers the laser radiation is generated during the transitions between vibrational levels. The energy levels in molecular gases split into vibrational and rotational, due to relative motions of the constituent atoms or their vibrations, respectively. The gaps between vibrational energy levels are much smaller than those between electronic energy levels, therefore the emitted wavelength is in the mid-IR

spectral range. Out of variety of molecular gas lasers, the most important laser is CO_2 laser emitting at wavelengths of around 10 μm. CO_2 laser is one of the most powerful continuous wave laser (the average power can exceed kW and more) and one of the most important lasers for industrial applications, such as laser cutting and welding.

- **Excimer lasers.** An excimer (termed after an excited dimer) is a diatomic molecule that exists only in an excited state. When excimer molecule emits a photon, the temporary bond between atoms breaks and the molecule disintegrates. The excimer laser is an ideal four-level laser, in which the lower laser level is always empty. Excimers are composed of noble gas and halogen atoms: ArCl, ArF, KrF, etc. To form an excimer molecule, a gas mixture has to be pumped with an electron beam or electric discharge. All excimer lasers emit light in the ultraviolet spectral range and operate only in the pulsed free-running regime. Due to broad gain bandwidth, the excimer lasers can be used as efficient amplifiers for ultrashort UV pulses, although they can not produce ultrashort pulses themselves. As the UV beams can be focused very sharply, the main applications of excimer lasers are related to biological research, medicine and lithography. However, halogens are very volatile, reactive and poisonous, which does not make them attractive materials.

- **Chemical lasers.** The chemical lasers are lasers in which population inversion is created during chemical reactions. All chemical lasers are oriented to military applications, e.g., long range (outside the atmosphere) shooting or destruction of enemy satellites or rockets. Chemical lasers emit in the IR, in wavelength range of $1.3 - 3$ μm, where the Earth's atmosphere is transparent. The power of chemical lasers can exceed MW in continuous wave regime and the beam diameter is up to 1.5 m, however, not much is known about these lasers, since most of the data is classified. Currently, several types of chemical lasers were developed: HF laser, AGIL (All Gas-phase Iodine Laser), COIL (Chemical Oxygen Iodine Laser) and deuterium fluoride (DF) laser MIRACL (Mid-Infrared Advanced Chemical Laser).

7.3 Semiconductor lasers

In a semiconductor laser, the radiation is produced by stimulated electron-hole recombination that is based on transitions between the conduction and the valence band of the semiconductor crystal (interband transitions). Since

the first demonstration in 1962, semiconductor lasers were recognized as a potential breakthrough in telecommunication technology. However, the first semiconductor lasers were based on simple p-n junction and operated only at cryogenic temperature. The development of double-heterostructure allowed to reduce the threshold current and achieve operation at room temperature. Further progress in semiconductor lasers was facilitated by the development of material processing technologies, such as metal-organic chemical vapor deposition (MOCVD) and molecular beam epitaxy (MBE), which allowed fabrication of semiconductor crystal structure with atomic layer accuracy.

At present, semiconductor lasers provide the output power up to several watts and their emission wavelengths cover very broad spectral range, from the UV to the mid-IR with. Besides numerous applications in various fields of modern technology, semiconductor lasers are used for pumping of fiber and solid-state lasers.

The most important class of semiconductor diode lasers is based on III-V semiconductors. An example for this is GaAs and $Ga_{1-x}Al_xAs$, where x indicates the molar fraction fraction of Ga in GaAs replaced by Al. The emission wavelength of such lasers depends on doping fraction and is in the 750–880 nm range. A generic double-heterojunction GaAs laser represents a thin (typically $0.1 - 0.2$ μm) GaAs region sandwiched between two regions of p- and n-doped $Ga_{1-x}Al_xAs$, as illustrated in Fig 93.

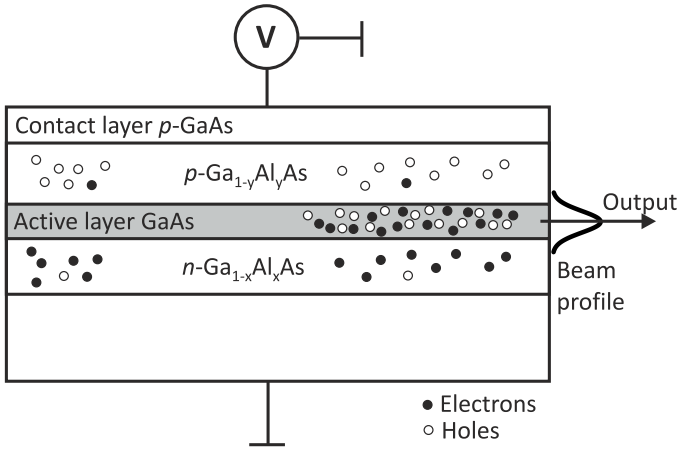


Figure 93: A typical double-heterostructure formed by GaAs and $Ga_{1-x}Al_xAs$ regions.

Since the bandgap energy of $Ga_{1-x}Al_xAs$ depends on the Al mole fraction, this double-heterostructure is fabricated so as the potential well for electrons of height ΔE_c coincides spatially with a well for holes of height ΔE_v. Applying the forward bias of $eV_a \sim E_g$, large densities of electrons are

injected from the n side and holes from the p side into the well, providing the population inversion condition, which is expressed as

$$E_{Fc} - E_{Fv} > \hbar\omega,$$

where E_{Fc} and E_{Fv} are the quasi-Fermi levels for the conduction and valence bands, respectively. The GaAs layer where stimulated emission takes place is called the active region, as shown in Fig. 94(a).

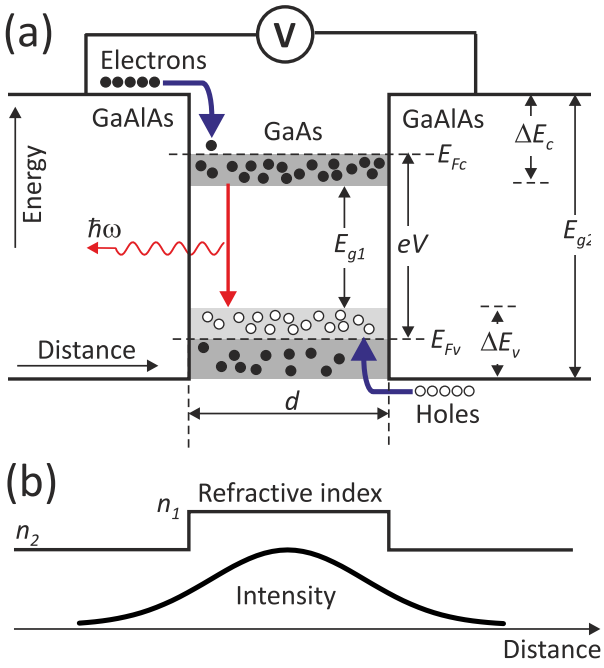


Figure 94: (a) Schematic energy diagram under positive (forward) bias in GaAlAs/GaAs/GaAlAs double-heterojunction laser diode. (b) The refractive index profile that confines the optical field.

For the maximum gain, the light has to be confined within an active region. This is achieved by a waveguiding effect in a GaAlAs/GaAs/GaAlAs sandwich, since reduction of the energy gap of a semiconductor by doping, causes increase of its refractive index. The refractive index distribution and the output profile of a typical double-heterojunction laser is shown in Fig. 94(b). In this case the gain scales as d^{-1}, where d is the height of the active region. The concept of double-heterojunction GaAlAs laser is transferred to other semoconductor lasers, e.g. InAlP.

Conventional heterostructures confine the gain into a small zone (the active region) and prevent electron and hole diffusion. Further miniaturization of the active region leads to quantized electron motion that is characterized

by discrete energy levels, giving rise to quantum film, quantum wire and quantum dot lasers. These lasers are not just simply smaller, but also show qualitatively novel radiation properties. Structures with reduced dimensionality offer lower threshold currents, larger gain and lower temperature sensitivity.

The most unwanted property of a laser diode is the mode hop that results in the change of laser wavelength and consequently the gain with temperature. To overcomes this, the simplest approach is based on an external cavity concept, which makes use of a diffraction grating at the Littrow configuration, which serves as a cavity mirror. Various modifications of semiconductor lasers were developed with regard of control the laser wavelength and coherence properties by integrating additional components during laser crystal manufacturing. In such a way, the so-called distributed feedback (DFB), distributed Bragg reflector (DBR) lasers are designed.

Summary of Chapter 7

- In solid-state lasers the laser radiation is produced by either rare-earth or transition metal ions embedded in a solid-state host material (transparent crystals and glasses). Solid-state lasers efficiently operate in all regimes: free-running, Q-switching and mode-locking.
- In gas lasers the laser radiation is produced by neutral atoms, ions or molecules. Gas lasers are efficient only in free-running regime (either CW or pulsed).
- In semiconductor lasers the laser radiation is produced by the crystal itself via stimulated electron-hole recombination in p-n junction.
- Electrically pumped semiconductor lasers are called laser diodes. They are the most common lasers.

8 Laser wavelength conversion by nonlinear optical methods

All solid-state lasers emit in the near infrared spectral range and the possibilities to tune their emission wavelengths are very limited or absent at all. These limitations come from laser energy level broadening, which in most cases is very small (Ti:sapphire and other vibronic lasers are the exceptions). **Nonlinear optics** offers simple methods how to efficiently change the wavelength of laser radiation. These methods can be formally divided into **frequency up-conversion** and **down-conversion**, which increase or decrease the frequency of laser radiation, respectively. Frequency up-conversion is based on the second harmonic and sum-frequency generation, while frequency down-conversion is based on the optical parametric generation and amplification. In the latter case the laser radiation at a fixed wavelength is converted into a broadly tunable radiation preserving original duration of laser pulses. All these phenomena take place in transparent birefringent crystals and rely on the second-order nonlinear optical susceptibility of these materials. In this chapter we will discuss the basic principles and practical approaches how laser frequency conversion can be performed.

8.1 Nonlinear optical susceptibility

Nonlinear optical phenomena are nonlinear in the sense that polarization response of the material depends on the electric field strength of a light wave in a nonlinear fashion. To understand the origin of the optical nonlinearity, it is useful to recall how material dipole moment per unit volume (polarization) depends on the strength of electric field. In linear optics, assuming that medium response is instant (non-inertial), this dependence is expressed simply as:

$$\boxed{P(t) = \epsilon_0 \chi^{(1)} E(t)}, \tag{179}$$

where ϵ_0 is the vacuum dielectric permittivity constant and $\chi^{(1)}$ is the linear optical susceptibility. In general, quantities $P(t)$ and $E(t)$ are vectors and rapidly oscillate (at the optical frequency) in time. The electric field of a light wave propagating in a transparent dielectric medium shifts the charges inducing a dipole moment $\mu(t)$, which oscillates at a frequency of the light wave. The induced polarization is then $P(t) = N < \mu(t) >$, where N is the number of dipoles and $<>$ denotes the averaging of all dipole moments. Note that all dipole moments which may naturally exist in the medium are not accounted for, since they do not emit electromagnetic waves.

The nonlinear optical phenomena occur when the electric field of light is strong enough to induce the nonlinear polarization, which is however much smaller that the linear counterpart. This is called perturbative regime of light-matter interaction or perturbative approximation. In that case, the polarization can be expressed in power series of the electric field strength:

$$P(t) = \epsilon_0 \chi^{(1)} E(t) + \epsilon_0 \chi^{(2)} E^2(t) + \epsilon_0 \chi^{(3)} E^3(t) + ..., \tag{180}$$

where $\chi^{(2)}$ and $\chi^{(3)}$ are second and third order nonlinear optical susceptibilities, respectively. It is also assumed that the material is transparent, dispersionless and its nonlinear response is instantaneous. The former expression can be rewritten by separating linear and nonlinear polarizations:

$$\boxed{P(t) = P^{(1)}(t) + P^{(2)}(t) + P^{(3)}(t) + ... = P_L + P_{\mathrm{NL}}}. \tag{181}$$

Here $P^{(2)}(t) = \epsilon_0 \chi^{(2)} E^2(t)$ and $P^{(3)}(t) = \epsilon_0 \chi^{(3)} E^3(t)$ are called second order (quadratic) and third order (cubic) nonlinear polarizations, respectively. In what follows, we will further restrict to **second order nonlinear optical phenomena** which occur in a certain class of materials: **noncentrosymmetric dielectric crystals**, i.e. crystals that exhibit birefringence.

8.2 Second harmonic generation

The simplest nonlinear optical phenomenon is the **second harmonic generation**: a nonlinear process when frequency of light is doubled. It can be explained in general details as follows. Consider the incident monochromatic light wave whose electric field is expressed as

$$E(t) = Ee^{-i\omega t} + c.c., \tag{182}$$

where ω is the frequency of the light wave. The electric field induces the nonlinear polarization in the medium

$$P^{(2)}(t) = \epsilon_0 \chi^{(2)} E^2(t) = 2\epsilon_0 \chi^{(2)} EE^* + (\epsilon_0 \chi^{(2)} E^2 e^{-2i\omega t} + c.c.). \tag{183}$$

The induced quadratic polarization has two components. The first polarization component (the first term on the right hand side) has zero frequency and describes the nonlinear phenomenon called optical rectification, i.e. creates a static electric field in the crystal which exists as long as the optical field is present. The second polarization component (the second term on the right hand side) denotes the polarization which oscillates at the doubled frequency of the incident light wave and is responsible for second harmonic generation.

It is convenient to ascribe the nonlinear optical phenomena using **virtual energy level diagrams**. Second harmonic generation can be viewed as simultaneous absorption of two photons at frequency ω into a virtual energy level and subsequent emission of a single photon at frequency 2ω, as depicted in Fig. 95.

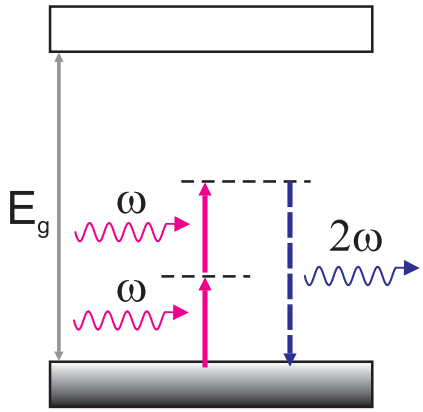


Figure 95: Energy level diagram depicting second harmonic generation. E_g is the energy bandgap of dielectric medium. Dashed lines denote virtual energy levels.

Note that virtual energy levels are not real, and they are associated with a small perturbation of bound electron quantum mechanical state by a light wave. Therefore, lifetime of virtual energy level can be formally estimated from the uncertainty principle: $\hbar/\delta E$, where δE is the energy difference between virtual and the closest real energy level. So the lifetime of virtual energy level is always much shorter than the lifetime of any real energy level, suggesting that processes involving virtual energy levels are very fast.

Figure 95 also represents the energy conservation law: $\hbar\omega_1 + \hbar\omega_1 = \hbar\omega_2$, where $2\omega_1 = \omega_2$. However, for efficient second harmonic generation, the momentum conservation law also has to be satisfied, i.e., $\hbar k_1 + \hbar k_1 = \hbar k_2$, where $k_{1,2} = n_{1,2}\omega_{1,2}/c$ denote fundamental (incident) and second harmonic wave vector lengths, respectively. Assuming that all interacting waves propagate at the same direction, we get that the momentum conservation is equivalent to the condition

$$n(\omega_1) = n(2\omega_1). \tag{184}$$

The above expression represents the **phase matching condition** as it imposes the phase velocities (c/n) of the interacting waves to be equal. In other words, the dipoles excited by the wave at fundamental frequency emit waves at doubled frequency (second harmonic) which are always in phase

with an exciting wave. This in turn guarantees that emitted second harmonic waves interfere constructively during propagation and the intensity of second harmonic radiation increases quadratically versus the propagation distance.

The fulfillment of the phase matching condition makes use of crystal birefringence and calls for setting appropriate polarizations of the interacting waves. For instance, in a negative uniaxial crystal ($n_e < n_o$, where n_e is the principal value of the extraordinary refractive index and n_o is the ordinary refractive index), the phase matching condition could be satisfied for the fundamental wave of an ordinary polarization, and the second harmonic wave of an extraordinary polarization. This is achieved only for certain propagation direction of the interacting waves with respect to the optical axis of the crystal, and which is called the "phase matching angle":

$$n_e(2\omega, \theta) = n_o(\omega). \tag{185}$$

The angle θ is computed from the refractive index ellipsoid equation.

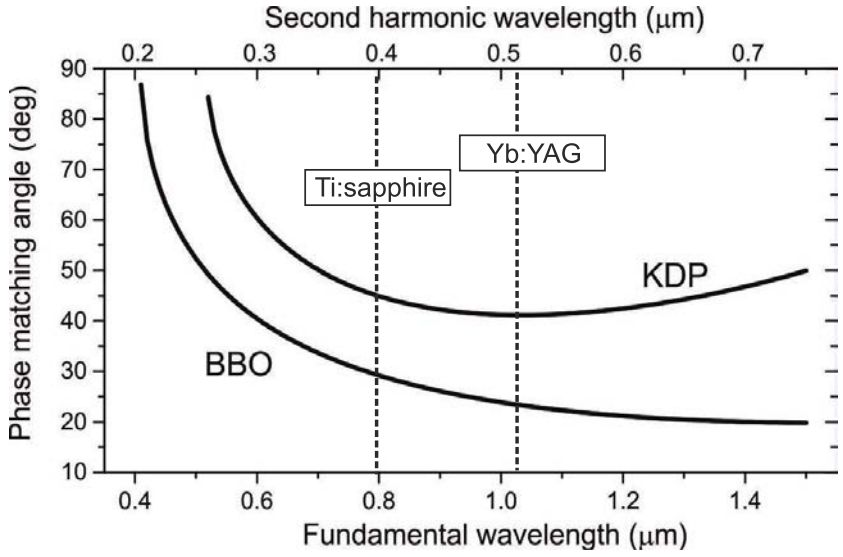


Figure 96: Phase-matching angle for the second harmonic generation in BBO and KDP crystals.

The phase matching angle for second harmonic generation changes by varying the fundamental wavelength, as illustrated in Fig.96. In practice, the nonlinear crystal is cut at the phase-matching angle for a given fundamental laser wavelength, i.e., so the phase-matching condition is satisfied for normal incidence.

The above considerations apply also to **sum-frequency generation**. In the case of sum-frequency generation, there are two incident waves with

frequencies ω_1 and ω_2, and the process is described as $\omega_1 + \omega_2 = \omega_3$, where ω_3 is the sum frequency. The incident waves with different frequencies can be produced by either two different lasers or the same laser. For example, the most common case of sum-frequency generation is generation of higher order (third, fourth, fifth) harmonics. Third harmonic generation is equivalent to sum-frequency generation between the fundamental and second harmonics: $\omega + 2\omega = 3\omega$. Fourth harmonic can be generated via $\omega + 3\omega = 4\omega$, and so on.

8.3 Optical parametric amplification

Although harmonic generation is a very simple and efficient method to change laser radiation wavelength and to extend the wavelength range accessible by a particular laser, many applications in physics, biology, medicine and chemistry demand continuous wavelength tuning over a broad frequency (wavelength) range. This could be achieved by making use of **optical parametric amplification**. The physical mechanism of the optical parametric amplification is essentially different from the stimulated emission, which lies on the basis of laser operation. At the photon level, the parametric amplification process does not require the population inversion between real energy levels or bands of the medium. Only **virtual** energy levels, which occur as a result of a weak perturbation of the electron states via nonlinear polarization induced by an intense pump wave, and which are located within a transparency band of the medium, are involved, highlighting instantaneous and non-resonant nature of the parametric amplification process. The principle of optical parametric amplification is schematically depicted in Fig. 97. The energy from an intense pump wave at frequency ω_p is transferred to a much weaker incident signal wave at frequency ω_s, which is thus amplified alongside the occurrence of a third wave at frequency ω_i, which is called the idler wave. A simplified physical picture of the process could be understood as follows. As a pump photon at frequency ω_p is absorbed by a *virtual* energy level, the presence of a photon at the signal frequency stimulates the emission of a photon pair: the signal photon with identical characteristics to the incident one and the idler photon at difference frequency $\omega_i = \omega_p - \omega_s$, as schematically shown in Fig. 97. Since the roles of the signal and idler photons are interchangeable, the idler photon stimulates the emission of another identical idler photon and the signal photon at difference frequency $\omega_s = \omega_p - \omega_i$. The process repeats a multiplicity of times consuming the energy of the pump. In the macroscopic scale, the amplification of the signal wave reinforces the generation of the idler wave, and vice versa, thus the amplitudes of the interacting waves (pump, signal and idler) become

coupled, and the intensities of the signal and idler waves grow exponentially with propagation distance, at the cost of the pump wave, whose energy is continuously depleted.

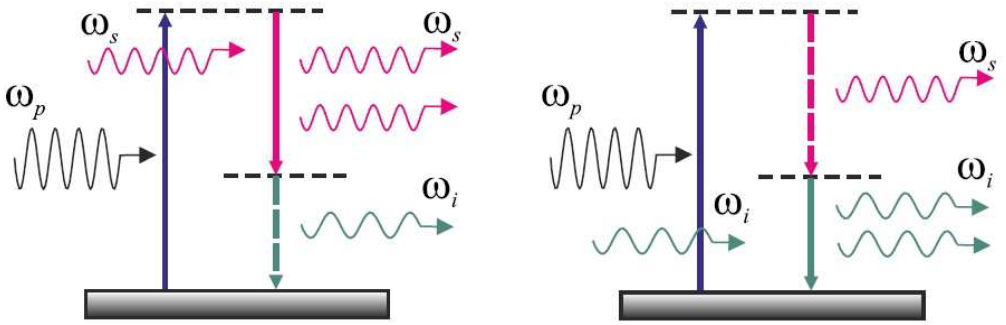


Figure 97: The principle of optical parametric amplification.

Optical parametric amplification is based on the same physical principles as the second harmonic or sum-frequency generation and requires momentum conservation (phase matching condition) to be fulfilled. In a collinear case (assuming that pump, signal and idler beams propagate in the same direction), for a negative uniaxial crystal ($n_o > n_e$) the phase-matching condition can be written as:

$$\omega_p n_e(\omega_p, \theta_{\mathrm{pm}}) = \omega_s n_o(\omega_s) + \omega_i n_o(\omega_i), \tag{186}$$

which implies that the pump wave has an extraordinary polarization, whereas the signal and idler waves have ordinary polarizations. The angle θ_{pm} between the propagation direction and optical axis of the crystal is called the phase-matching angle.

When propagation direction of the pump beam with respect to crystal optical axis is changed (i.e., the crystal is rotated), a different pair of signal and idler frequencies satisfies collinear phase-matching condition. Hence, continuous signal and idler wave frequency tunability can be achieved in a very broad spectral range as shown in Fig. 98. Note that when the signal wavelength decreases, the idler wavelength increases and vice versa, as defined by the energy conservation law. In the ideal case, the signal wavelength can be tuned from ω_p to $\omega_p/2$, whereas the idler wavelength can be tuned from 0 to $\omega_p/2$ respectively. However, in real crystals wavelength tuning range is limited by crystal absorption and/or phase-matching condition.

Optical parametric amplification is a process when energy of the pump wave is transmitted to signal and idler waves, therefore, their intensity increases exponentially from propagation distance in the crystal z:

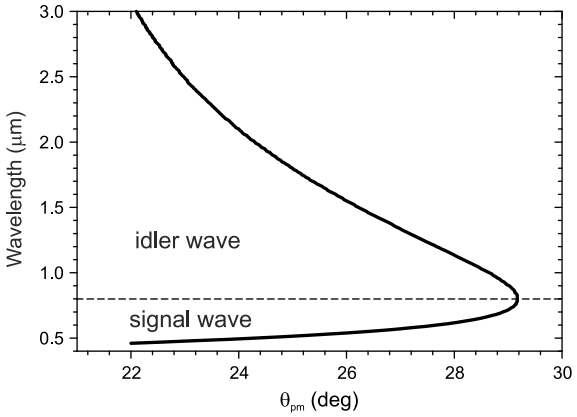


Figure 98: Phase-matching curve for BBO crystal. The pump wavelength is 400 nm, which corresponds to the second harmonic of Ti:sapphire laser.

$$I_s(z) = \frac{1}{4} I_s(0) \exp(2\Gamma z) \tag{187}$$

$$I_i(z) = \frac{\omega_i}{4\omega_s} I_s(0) \exp(2\Gamma z) \tag{188}$$

here

$$\boxed{\Gamma^2 = \frac{\omega_i \omega_s d_{\text{eff}}^2 I_p^2}{n_s n_i n_p \epsilon_0 c^3}}, \tag{189}$$

and quantity Γ is called the **amplification increment**, which depends on the effective nonlinearity of the crystal d_{eff}, proportional to certain matrix elements of the second order nonlinear susceptibility $\chi^{(2)}$, and pump wave intensity I_p. Here we notice that only radiation intensity of a laser source is high enough for amplification, while using non-laser sources makes optical parametric amplification practically impossible.

In this sense the nonlinear crystal performs as a certain catalyzer and gain coefficient is very high (conversion from pump energy to signal and idler waves can be more than 80%). The gain coefficient of optical parametric amplification is defined as:

$$\boxed{G = \frac{I_s(z)}{I_s(0)} = \frac{1}{4} \exp(2\Gamma z)}, \tag{190}$$

here $I_s(0)$ and $I_s(z)$ are the signal wave intensities at crystal input and output respectively. At GW/cm^2 level pump intensity, which is easy to achieve with

ultrashort pulse lasers, the gain coefficient can be $10^3 - 10^6$ for a crystal of a few mm length. So optical parametric amplification is a unique process allowing fixed wavelength radiation to be converted into tunable wavelength radiation with high efficiency and without loss of coherence properties.

8.4 Optical parametric amplifiers

There are two types of optical parametric amplifiers, whose design depends on the pump laser parameters: pulse duration, power and intensity.

Optical parametric oscillators

The first type of optical parametric amplifiers are made of nonlinear crystal which is placed in a resonator as shown in Fig.99. Such optical parametric amplifiers are called standing wave amplifiers or optical parametric oscillators (OPO). Here the pump beam is coupled through one of the resonator mirrors and the mirrors are chosen to reflect either signal or idler waves (singly resonant OPO, Fig.99(a)) or both (doubly resonant OPO, Fig.99(b)).

For OPO pumping either low power lasers which operate in CW or Q-switched modes or femtosecond laser oscillators generating ultrashort but low energy pulses at high repetition rate are usually used. Since the peak power of such pump laser radiation is low, single-pass amplification is relatively low as well (from several to a dozen of times). The resonator performs the same roles of energy accumulation, feedback and formation of high spatial quality beam (due to diffraction losses) as in a conventional laser. During the first round-trip parametric superfluorescence is generated in the nonlinear crystal; a part of which propagating on the resonator axis is amplified during the following round-trips through the nonlinear crystal. As in conventional laser, the oscillation condition is defined in the same way: parametric amplification after one resonator round-trip must exceed losses due to light scattering, absorption, diffraction and mirror transmission.

In the case of the simplest singly resonant OPO a small fraction (usually few percent) of resonating signal is coupled out. In addition, the entire idler wave is coupled out which after each pass is generated again. Notice that all resonator stability criteria discussed at the beginning of this textbook also apply to OPO resonator. Wavelength tuning of OPO is performed in a simple way – by rotating the nonlinear crystal, i.e., by changing pump beam direction with respect to crystal optical axis. In such case at each time collinear phase-matching is achieved for a different signal and idler frequency pair. Wavelength tuning range of OPO is usually limited by spectral characteristics of resonator mirrors because fabricating mirrors with uniform

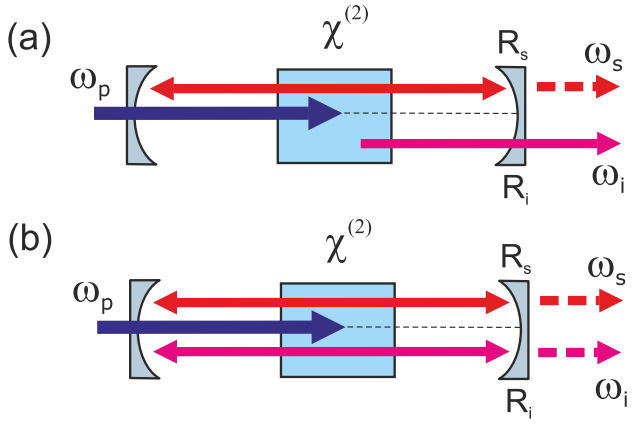


Figure 99: The main configurations of optical parametric oscillators: (a) singly resonant, (b) doubly resonant.

coefficient of reflection over a broad spectral range is very difficult or even impossible. Therefore, in such device the entire wavelength tuning range allowed by phase-matching conditions is not achievable.

Traveling wave optical parametric amplifiers

The second type of parametric amplifiers operate according to traveling wave principle, i.e., parametric amplification is performed at a straight setup using several nonlinear crystals. The main difference of traveling wave optical parametric amplifiers (simply termed OPA) from OPO is the absence of a resonator, a number of passes through the crystal is small but the gain factor is large. High power ultrashort pulses, which are generated by amplifying the laser oscillator pulses, are used to pump OPA. The peak power and intensity of the pump pulses are very high, so a single pass through the nonlinear crystal provides huge (several hundreds to thousands of times) amplification.

The simplest OPA consists of several nonlinear crystals. In the first crystal parametric superfluorescence is generated which is then amplified using one or more amplification stages. The only drawbacks of such amplification system is the requirement to form spatial and spectral characteristics of parametric superfluorescence and the fact that any distortions occurring in the beam are amplified at each successive stage. However, the main advantage of such device is large energy of output radiation and broad frequency tuning range which is only limited by phase-matching conditions.

The principal setup of traveling wave optical parametric amplifier is depicted in Fig.100. The pump laser beam is divided (by means of beam splitters BS) into several independent channels used to pump parametric

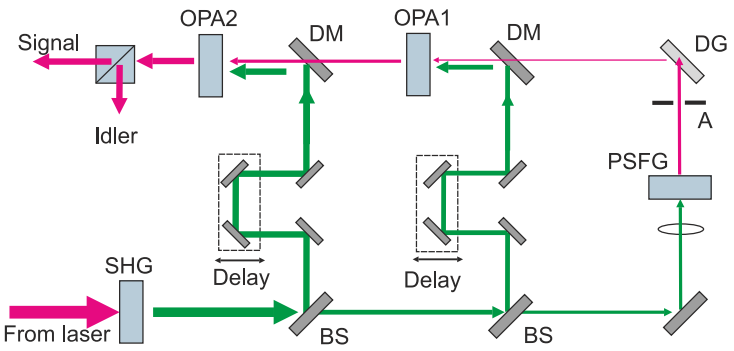


Figure 100: The principal setup of traveling wave optical parametric amplifier: SHG - second harmonic generator (optional), BS - beamsplitters, DM - dichroic mirrors, PSFG - parametric superfluorescence generator, OPA - optical parametric amplification stages.

superfluorescence generator (PSFG), first and second amplifiers OPA1 and OPA2. Spatial and temporal characteristics of the PSFG signal are formed using an aperture (A) and diffraction grating (DG). The signal and pump beams are directed to OPA1 and OPA2 using dichroic mirrors (DM) that have high transmittance for the signal and high reflectance for the pump. The pump energies for each amplification stage are chosen to achieve gain saturation. Since amplification of ultrashort light pulses is considered, precise temporal matching of pump and signal pulses is required, which is provided by mechanical delay lines. Wavelength tuning is performed in the same way as in OPO only in this case all three nonlinear crystals have to be rotated synchronously.

One of the most important and commonly used nonlinear crystals is beta barium borate (BaB_2O_4, BBO). This crystal has large nonlinearity and phase matching can be achieved in the whole transparency range (from 190 nm to 3.5 μm). Moreover, growth technology for this crystal is well developed and relatively cheap.

Fig.101 depicts wavelength tuning ranges of BBO crystal-based OPA when pumping with commonly used ultrashort pulse lasers and their harmonics. More than that, frequency tuning range of OPA can be additionally extended by employing second harmonic generation, sum-frequency and difference frequency generation. Consequently, wavelength tuning range of OPA may cover a considerable part of the optical spectrum: from the ultraviolet to the mid-infrared. Because of that, optical parametric amplifiers are often called multi-color lasers.

Finally, replacement of the laser amplifier by an optical parametric amp-

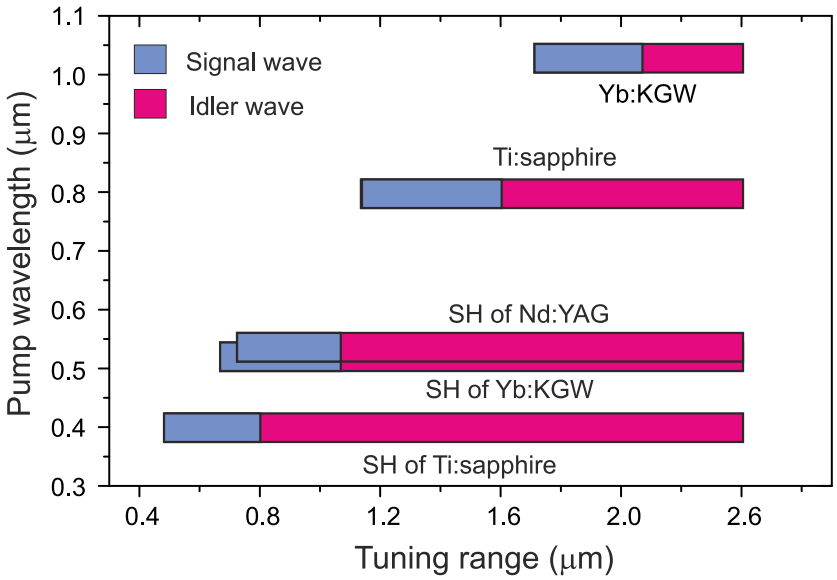


Figure 101: Tuning ranges of BBO crystal based optical parametric amplifiers pumped with commonly used ultrashort pulse lasers and their harmonics.

lifier[24] in CPA system has led to novel high power amplification technique, termed **Optical Parametric Chirped Pulse Amplification (OPCPA)**. OPCPA favorably combines the advantages of CPA with the advantages of optical parametric amplification, which offers very high gain, low thermal effects, great wavelength flexibility and intrinsically broad amplification bandwidth, extending well beyond that afforded by existing solid-state laser amplifiers

At present, compact optical parametric chirped pulse amplifiers comprise a unique class of ultrafast light sources, which currently amplify octave-spanning spectra and produce few optical cycle pulses with multi-gigawatt to multi-terawatt peak powers and multi-watt average powers, with carrier wavelengths spanning a considerable range of the optical spectrum.

[24]A. Dubietis, G. Jonušauskas, A. Piskarskas, Powerful femtosecond pulse generation by chirped and stretched pulse parametric amplification in BBO crystal, Optics Communications **88**, 437–440 (1992)

Summary of Chapter 8

- Nonlinear optics offers simple and efficient way to change the wavelength of laser radiation.
- Frequency conversion via nonlinear optical phenomena involves only virtual energy levels, so none of the interacting waves is really absorbed.
- Efficient second harmonic generation, as well as any other nonlinear optical phenomenon require energy and momentum conservation laws to be satisfied. The latter is called phase matching and makes use of crystal birefringence.
- Optical parametric amplification is a nonlinear process when energy from intense pump (higher frequency) wave is transferred to the seed (lower frequency) wave and a third wave with difference frequency is generated.
- OPOs and OPAs allow generation of broadly tunable radiation from a pump laser source operating at the fixed wavelength.

9 EN - LT Glossary

No.	English term	Lithuanian term (and additional remarks)
1.	Acoustooptic modulator	Akustooptinis moduliatorius
2.	Active Q-switch	Aktyvus kokybės moduliatorius
3.	Amplification	Stiprinimas
4.	Angle of incidence	Kritimo kampas (spindulio ar panašiai)
5.	Anti-reflective coatings	Skaidrinančios dangos
6.	Arc lamp	Elektros lanko lempa
7.	Attenuation	Silpimas (silpninimas)
8.	Autocorrelation	Autokoreliacija
9.	Bandwidth	Spektro plotis
10.	Bandwidth-limited pulse	Spektriškai ribotas impulsas
11.	Beam divergence	Pluošto skėstis (matuojama radianais/laipsniais)
12.	Beam parameter product (BPP)	Pluošto parametrų sandauga
13.	Beam waist	Pluošto sąsmauka (mažiausio diametro vieta)
14.	Birefringence	Dvejopalaužiškumas
15.	Broadening	Plitimas (išplitimas)
16.	Cavity	Rezonatorius (vartojamas turint omenyje tuščią rezonatorių)
17.	Chemical laser	Cheminis lazeris
18.	Chirped pulse	Čirpuotas (faziškai moduliuotas) impulsas
19.	Chirped pulse amplification (CPA)	Čirpuotų impulsų stiprinimas
20.	Coherence	Koherentiškumas
21.	Continuous wave (CW) laser	Nuolatinės veikos lazeris
22.	Convergence	Konvergavimas
23.	Correlation	Koreliacija
24.	Coupled amplitude wave equations	Susietų amplitudžių bangų lygtys
25.	Cross-correlation	Kryžminė koreliacija
26.	Cross-section	Skerspjūvis (atomo fizikoje tai reiškia tikimybę)
27.	Curved mirror	Kreivo paviršiaus veidrodis
28.	Delay	Vėlinimas [s]

No.	English term	Lithuanian term (and additional remarks)
29.	Dye laser	Dažų lazeris
30.	Dielectric permittivity	Dielektrinė skvarba
31.	Diffraction-limited beam	Difrakciškai ribotas pluoštas
32.	Dynamical equilibrium	Dinaminė pusiausvyra
33.	Dipole moment	Dipolinis momentas
34.	Dispersive prism	Dispersinė prizmė
35.	Doubly resonant oscillator	Dvibangis osciliatorius
36.	Electric discharge	Elektros išlydis
37.	Electrooptic modulator	Elektrooptinis moduliatorius
38.	Energy fluence	Energijos tankis [J/m^2]
39.	Excimer laser	Eksimerinis lazeris
40.	Excited	Sužadintas
41.	Experimental setup	Eksperimento schema
42.	Feedback	Grįžtamasis ryšys
43.	Fiber laser	Šviesolaidinis lazeris
44.	Figure of merit	Naudingumo skaičius
45.	Fission	Skilimas (pvz. branduolių)
46.	Flashlamp	Blykstinė lempa (blykstė)
47.	Fourier transform	Furjė transformacija
48.	Free space resonator	Laisvos erdvės rezonatorius
49.	Free-electron laser	Laisvų jų elektronų lazeris
50.	Free-running mode laser	Laisvos veikos lazeris
51.	Full width at half maximum (FWHM)	Pusės maksimalaus lygio plotis
52.	Fusion	Sintezė/susijungimas (pvz. branduolių)
53.	Gain bandwidth	Stiprinimo juosto plotis
54.	Gain coefficient	Stiprinimo koeficientas
55.	Gas laser	Dujų (dujinis) lazeris
56.	Gaussian beam	Gauso pluoštas
57.	Glass laser	Stiklo lazeris
58.	Group delay dispersion (GDD)	Grupinio vėlinimo dispersija [s^2]
59.	Group velocity	Grupinis greitis [m/s]
60.	Group velocity dispersion (GVD)	Grupinių greičių dispersija [s^2/m]
61.	Harmonic generation	Harmonikų generacija
62.	Heat induced birefringence	Šilumos indukuotas dvejopalaužiškumas
63.	Higher order mode	Aukštesnės eilės moda

No.	English term	Lithuanian term (and additional remarks)
64.	Highly reflective coatings	Didelio atspindžio dangos
65.	Yb doped medium	Yb jonais legiruota terpė
66.	Idler wave	Šalutinės banga
67.	Impurities	Priemaišos
68.	Intensity	Intensyvumas [W/m^2]
69.	Intensity autocorrelation	Intensyvumo autokoreliacija
70.	Kerr effect	Kero efektas
71.	Kerr lens mode-locking	Kero lęšio modų sinchronizmas
72.	Lase	Generuoti lazerio spinduliuotę
73.	Laser	Lazeris
74.	Laser amplifier	Lazerinis stiprintuvas (vartojama pabrėžiant, kad tai lazerinio stiprinimo terpė be rezonatoriaus)
75.	Laser beam	Lazerio pluoštas
76.	Laser diode	Lazerinis diodas
77.	Laser induced damage	1) Lazerio sukelta pažaida (reiškinys)
		2) Lazerio sukeltas pažeidimas (efektas)
78.	Laser oscillator	Lazerio osciliatorius (lazeris – vartojama pabrėžiant, kad tai lazerinio stiprinimo terpė su rezonatoriumi)
79.	Laser pulse	Lazerio impulsas
80.	Lasing	Lazerio generacija
81.	Lifetime	Gyvavimo trukmė
82.	LED (Light-Emitting Diode)	Šviestukas
83.	Longitudinal mode	Išilginė moda
84.	Losses	Nuostoliai
85.	Maser	Mazeris
86.	Mechanical stress	Mechanins į tempimas
87.	Metal vapor laser	Metalo garų lazeris
88.	Microlaser	Mikrolazeris
89.	Mode-locked laser	Sinchronizuotų modų veikos lazeris
90.	Multimode beam	Daugiamodis pluoštas
91.	Multi-pass	Daugelio praėjimų
92.	Multi-shot autocorrelator	Daugiašūvis autokoreliatorius
93.	Noncentrosymmetric media	Terpės be simetrijos centro
94.	Nonlinear coupling coefficient	Netiesinės sąveikos koeficientas
95.	Nonlinear medium	Netiesinė terpė

No.	English term	Lithuanian term (and additional remarks)
96.	Nonlinear response	Netiesinis atsakas
97.	Nonlinearity	Netiesiškumas
98.	Nonradiative transitions	Nespinduliniai šuoliai
99.	Optical density	Optinis tankis
100.	Optical fiber	Šviesolaidis
101.	Optical parametric amplification (OPA)	Parametrinis šviesos stiprinimas
102.	Optical parametric chirped pulse amplification (OPCPA)	Optinis parametrinis čirpuotų impulsų stiprinimas
103.	Optical parametric generation (OPG)	Parametrinė šviesos generacija
104.	Optical parametric oscillator (OPO)	Parametrinis šviesos osciliatorius (generatorius, turintis rezonatorių)
105.	Optical rectification	Optinis lyginimas (netiesinis efektas)
106.	Optical susceptibility	Optinis jautris
107.	Optical tweezer	Šviesos pincetas
108.	Optical vortex	Šviesos sukūrys
109.	Output coupler	Išvadinis veidrodis
110.	Output radiation	Išvadinė (išeinanti iš rezonatoriaus) spinduliuotė
111.	Parametric fluorescence	Parametrinė fluorescencija
112.	Parametric gain	Parametrinio stiprinimo koeficientas
113.	Parametric superfluorescence	Parametrinė superfluorescencija
114.	Paraxial approximation	Paraksialinis artinys
115.	Passive Q-switch	Pasyvus kokybės moduliatorius
116.	Peak power	Smailinė galia [W]
117.	Periodically poled crystal	Periodiškai poliuotas (orientuotas) kristalas
118.	Perturbation	Perturbacija (trikdis)
119.	Phase matching condition	Fazinio sinchronizmo sąlyga
120.	Phase mismatch	Fazinis nederinimas
121.	Phase velocity	Fazinis greitis [m/s]
122.	Photon	Fotonas
123.	Photon flux density	Fotonų srauto tankis
124.	Photon lifetime	Fotonų gyvavimo trukmė
125.	Plane wave	Plokščia banga
126.	Pockels cell	Pokelso elementas

No.	English term	Lithuanian term (and additional remarks)
127.	Population	Užpilda (elektronų tankis energijos lygmenyje)
128.	Population inversion	Užpildos apgrąžą
129.	Power	Galia [W]
130.	Propagation	Sklidimas
131.	Pulse compressor	Impulsų spaustuvas
132.	Pulse duration	Impulso trukmė [s]
133.	Pulse repetition period	Impulsų pasikartojimo periodas [s]
134.	Pulse repetition rate	Impulsų pasikartojimo dažnis [Hz]
135.	Pulse stretcher	Impulsų plėstuvas
136.	Pulsed beam	Impulsinis pluoštas (retai vartojamas terminas)
137.	Pulsed laser	Impulsinės veikos lazeris
138.	Pump	Kaupinimas
139.	Pump-enhanced singly resonant oscillator	Vienbangis osciliatorius su kaupinimo grąžinimu
140.	Pump rate	Kaupinimo sparta [Hz]
141.	Q-switched laser	Moduliuotos kokybės veikos lazeris
142.	Quantum	Kvantas (mažiausias į manomas dydis)
143.	Quantum defect	Kvantinis defektas
144.	Quasi-phase matching	Kvazifazinis sinchronizmas
145.	Radiation	Spinduliuotė (vienas iš galimų vertimų)
146.	Radiative transitions	Spinduliniai šuoliai
147.	Rayleigh range	Reilėjaus atstumas
148.	Regenerative amplifier	Regeneracinis stiprintuvas
149.	Relaxation oscillation	Relaksaciniai svyravimai
150.	Resonator	Rezonatorius
151.	Rod	Strypas (taip kartais vadina didelių matmenų lazerinius kristalus)
152.	Round-trip time	Lėkio (pilno apėjimo) trukmė (pvz. lazerio rezonatoriuje)
153.	Running wave	Bėganti banga
154.	Saturable absorber	Įsisotinantis sugėriklis
155.	Seed radiation	Užkrato spinduliuotė
156.	Semiconductor laser	Puslaidininkinis lazeris

No.	English term	Lithuanian term (and additional remarks)
157.	Semiconductor saturable absorber mirror (SESAM)	Puslaidininkinis įsisotinančios sugerties veidrodis (dažniausiai tiesiog vartojamas akronimas SESAM)
158.	Signal wave	Signalinė banga
159.	Single-pass	Vieno praėjimo
160.	Single-shot autocorrelator	Vienašūvis autokoreliatorius
161.	Singly resonant oscillator	Vienbangis osciliatorius
162.	Solid-state laser	Kietojo kūno lazeris
163.	Spatial	Erdvinis
164.	Spectral responsivity	Spektrinis atsakas
165.	Spherical wave	Sferinė banga
166.	Spontaneous emission	Savaiminė spinduliuotė
167.	Standing wave	Stovinti banga
168.	Stimulated emission	Priverstinė spinduliuotė
169.	Stimulated emission depletion (STED) microscopy	Priverstinio spinduliavimo nuskurdinimo mikroskopija
170.	Streak camera	Fotoelektronų kamera
171.	Temporal	Laikinis
172.	Thermal lens	Šiluminis lęšis
173.	Thermal stress	Šiluminis įtempimas
174.	Third order dispersion (TOD)	Trečios eilės dispersija [s^3]
175.	Threshold	Slenkstinis (slenkstis)
176.	Time-bandwidth product	Impulso trukmės ir spektro pločio sandauga
177.	Timing jitter	Laikinis tirtėjimas (nukrypimas nuo periodiškumo)
178.	Transform-limited pulse	Spektriškai ribotas impulsas
179.	Transverse (spatial) mode	Skersinė (erdvinė) moda
180.	Transverse mode suppression	Skersinių modų slopinimas
181.	Triply resonant oscillator	Tribangis osciliatorius
182.	Uncertainty principle	Neapibrėžtumo principas
183.	Wave vector	Bangos vektorius
184.	Wavefront	Bangos frontas
185.	Waveguide	Bangolaidis
186.	Wavelength	Bangos ilgis
187.	Wavelength tuning	Bangos ilgio derinimas

This page is intentionally left blank

Julius Vengelis, Audrius Dubietis
Laser Physics: Lecture notes. Vilnius University, Vilnius 2023.

Lasers and laser-related technologies are an integral part of modern science and are widely used in many areas of everyday life. This textbook provides the basic principles of laser operation, describes the underlying physical effects a nd gives an overview of relevant scientific and technological innovations that make possible generation and amplification o f u ltrashort l ight p ulses. T he m aterial presented here is intended for Vilnius University Faculty of Physics undergraduate students of *Physics* and *Light engineering* study programmes, but is also suitable to anyone with a background in natural sciences, interested in this subject.

knygynas.vu.lt
journals.vu.lt