Script "Fundamentals of Laser Physics" – Chapter 2

2. Fundamentals of Optics

The goal of this chapter is to provide terms and definitions which are essential for Laser Physics. This is a brief discussion, for more details please study standard textbooks on optics.

<u>The electro-magnetic spectrum</u>

The sensitivity of the human eye defines the visible spectral range, called henceforth VIS.

VIS: ∼ 400 .. 750 nm wavelength
        ∼ $7.5 \cdot 10^{14}$ .. $4.3 \cdot 10^{14}$ Hz frequency
        ∼ 3.0 .. 1.5 eV photon energy

The following picture gives a schematic overview of the spectral regions covered by electro-magnetic radiation and their main applications.
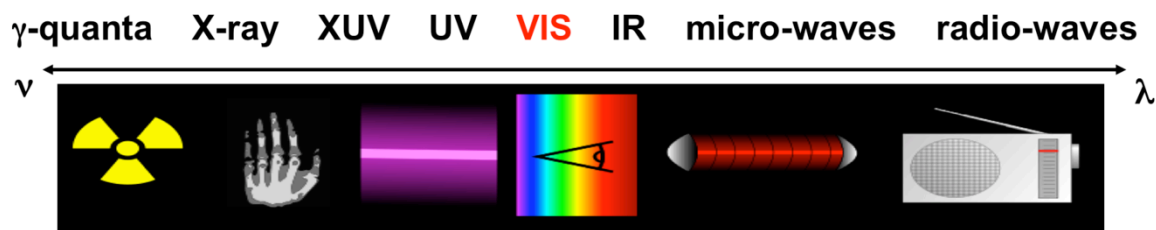
Over the whole electro-magnetic spectrum the wavelength λ and frequency ν are connected by:

$$c = \upsilon \cdot \lambda \tag{2.1}$$

c: speed of light in vacuum c = 299.792.458 m/s ≈ $3 \cdot 10^8$ m/s

Propagation effects of electro-magnetic waves are described by Maxwell's equations. Although they have been derived from experiments with static or slowly varying fields, it has been shown that Maxwell's equations are also valid at optical frequencies. Hence, we can use Maxwell's equations in the optical spectral range to understand the propagation of light including all interference and diffraction effects.

Nevertheless, in that classical picture, the interaction of electro-magnetic field with matter remains unexplained. In fact, for a long time it was not possible to explain the radiation emitted by hot bodies (so called black-body radiation) or the photoelectric effect, i.e. the emission of electrons from a metal surface under illumination of light. These shortcomings of the classical theory were solved by Planck and Einstein. A quantization of the radiation field (i.e. light quanta known as photons) was introduced for the description of light. This approach established the beginning of Quantum Physics.

According to Quantum Physics, in a field of frequency ν energy only exists in a quantized form. In this context, the smallest energy fraction (quantum) is the photon energy given by:

$$E = h \cdot \upsilon \qquad\qquad (2.2)$$

h: Planck's constant $h \approx 6.6 \cdot 10^{-34}\, Ws^2$

Hereby it is important to know that: *Radiation can only be taken from or given to the radiation field by integer values of the photon energy*.

Additionally, a photon is a particle with a momentum which magnitude is given by:

$$p = \frac{h}{\lambda}.$$

Further properties of photons can be summarized as follows:
  – They propagate with the speed of light
  – According to the theory of relativity their rest mass is zero and photons possess a spin equal to +/- 1 (Bosons).

In summary, electro-magnetic radiation can be described as a wave or as a particle. This two-fold aspect is known as wave-particle dualism. Within the scope of the lecture we will keep it as simple as possible: Whenever it makes more sense to highlight the particle characteristics to explain certain phenomena we will do so; on the contrary, whenever effects are easier explained with the wave nature of light, then wave characteristics will be used.

It is worth mentioning that quantum electrodynamics (QED) is the only theory which provides an exact description of light-matter interaction at the atomic level. QED is based on a quantization of the electro-magnetic field, it covers the wave and the particle aspects of the radiation and contains classical electro-dynamics as a boundary case. Unfortunately, the QED formalism is extremely difficult to handle and hardly applicable to real optical systems / problems. Hence, there is the need for a trade-off, which is provided by the semi-classical theory. This treatment is based on the fact that often very useful descriptions of optical effects can be obtained by a classical treatment of the radiation field and a quantum-mechanical treatment of matter and light-matter interaction.

2.1. Emission and Absorption of Radiation

What are the fundamental processes happening when light interacts with matter?

In the simplest view these processes are: *emission (spontaneous or stimulated) and absorption of photons*, which cause electronic transitions between the energy levels in a particle (such as atoms or molecules) under observation of the selection rules. Thereby the energy of the complete system (particle + radiation field) is conserved.

Each atom or molecule has an infinite number of energy levels. But often it is sufficient to consider only 2 of those levels, i.e. for the time being we reduce the description of absorption and emission processes to a two-level system.
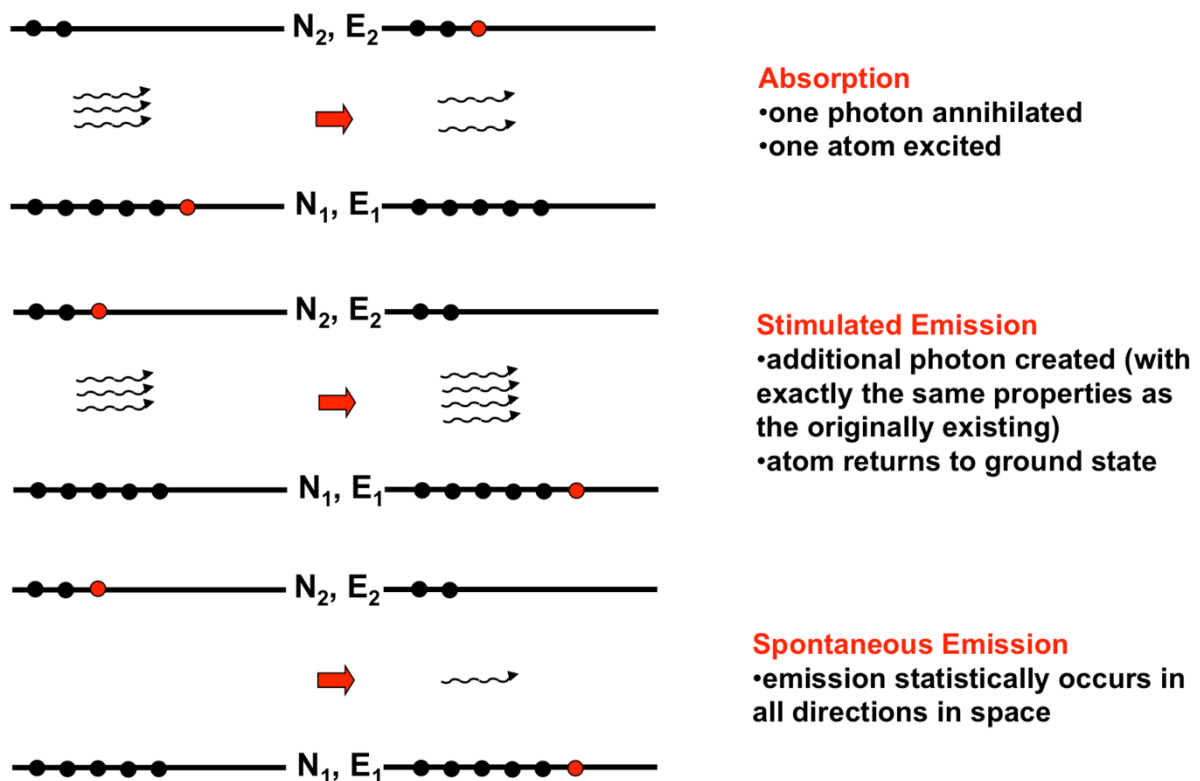
N_2, E_2 ——————————— $|2>$ (excited state)

$$h \cdot v_{21} = E_2 - E_1$$

N_1, E_1 ——————————— $|1>$ (ground state)
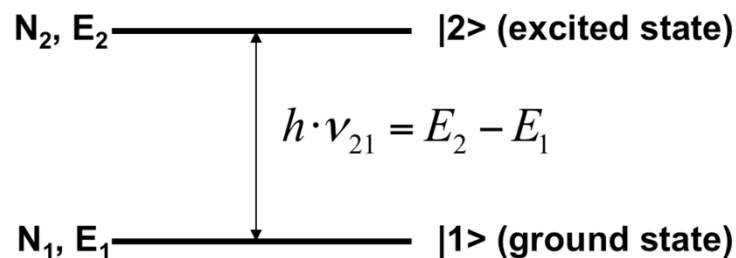
v_{21}: resonance frequency
$N_{1,2}$: population density in level 1,2
$E_{1,2}$: energy eigenvalue of level 1, 2

The following illustration is a schematic view of the possible absorption and emission processes taking place during the interaction of light with matter. The left-hand side column shows the situation before the interaction and the right-hand side shows the situation after the interaction.

Absorption
•one photon annihilated
•one atom excited

Stimulated Emission
•additional photon created (with exactly the same properties as the originally existing)
•atom returns to ground state

Spontaneous Emission
•emission statistically occurs in all directions in space

In this depiction an ensemble of particles (here 8 particles represented as dots) is considered to be distributed between the ground state and the excited state of a two-level system. The radiation field consists of a certain number of photons, illustrated in the picture by the wavy arrows.

In the case of absorption, each process annihilates one photon and excites one particle (one particle, marked in red, moves from the ground state to the excited state).

On the other hand, a stimulated emission process creates one photon with exactly the same properties as the originally existing photons (i.e. it is coherent to them) as one excited particle (marked in red) returns to the ground state. This transition is stimulated by the incoming photons.

Finally, the spontaneous emission occurs statistically in all directions in space, because no photons are required to trigger the process. Hereby, one particle in the excited state (marked in red) spontaneously relaxes to the ground state liberating a photon without correlation to any existing light field in the process (i.e. it is incoherent). Spontaneous emission is a quantum effect and the theory of quantum electrodynamics is necessary to fully describe it.

In the following we will take a closer, quantitative look at each one of these processes.

Absorption

It can be observed that the occurrence of absorption processes is proportional to both the number of particles available for absorption (i.e. in the lower energy state) and the energy density of the radiation field around the frequency of the resonance. Hence, we can write the following equation to describe the decrease of population density in level 1 (i.e. loss of particles in level 1) due to the absorption processes:

$$\left.\frac{dN_1}{dt}\right|_{Absorption} = -B_{12} \cdot \rho(\nu) \cdot N_1$$

(2.3a)

N_1:   population density in level 1
$\rho(\nu)$:   spectral energy density of the incoming radiation
            (energy content per frequency interval and unit volume)
B_{12}:   constant of proportionality (property of the transition)

(2.3a) describes the number of transitions (i.e. absorption processes) per unit volume and unit time.

Stimulated Emission

The existence of a reverse process to absorption can be expected, which is referred to as stimulated emission (or induced emission). In this case, the decrease of the population density in level 2 through stimulated emission is given by:

$$\left.\frac{dN_2}{dt}\right|_{Induced\ Emission} = -B_{21} \cdot \rho(\nu) \cdot N_2$$

(2.3b)

N_2:   population density in level 2
$\rho(\nu)$:   spectral energy density of the incoming radiation (energy content per frequency
            interval and unit volume)
B_{21}:   constant of proportionality (property of the transition)

Spontaneous Emission

For both of the processes above, i.e. stimulated emission and absorption, there would be no transition/occurrence observable if there is no radiation field present ($\rho(v) = 0$). However, this contradicts experimental observations, where it can be seen that an excited particle can spontaneously return to the ground state liberating a photon in the process. Hence, there must be a further emission process, which is the spontaneous emission.

One can observe that the rate of transitions/spontaneously generated photons is proportional to the population density in level 2 (i.e. the excited state). Thus, the decrease of population density in level 2 through spontaneous emission can be written as:

$$\left.\frac{dN_2}{dt}\right|_{Spontaneous\ Emission} = -A_{21} \cdot N_2$$

(2.3c)

N_2: population density in level 2
A_{21}: rate constant, defined by transition properties

Equations (2.3a-c) have introduced three constants of proportionality for the different modalities of light-matter interaction (i.e. B_{12}, B_{21}, A_{21}). Now the question is: Are there any relations between these constants of proportionality in equations (2.3a-c)?

As mentioned in the historical overview, Einstein published the work "On the Quantum Theory of Radiation" in 1917. In that work, based on statistical considerations, he derived the following relations between the A/B constants (later labelled Einstein coefficients):

$$\frac{A_{21}}{B_{12}} = 8\pi \cdot h \cdot \frac{v^3}{c^3}$$

(2.4a)

$$\frac{B_{21}}{B_{12}} = 1$$

(for non-degenerated states)

(2.4b)

The Einstein coefficients are a measure of the probability of each particular process and, hence, they are determined by the properties of the considered transition.

Using equations (2.3b-c) together with (2.4a) it is possible to derive the ratio of spontaneous to stimulated emission processes:

$$\frac{dN_2/dt|_{spontaneous}}{dN_2/dt|_{induced}} = \frac{8\pi \cdot h \cdot v^3/c^3}{\rho(v)} \sim v^3$$

It is interesting to see that, neglecting the frequency dependence of the spectral energy density, the ratio of spontaneous emission processes to induced emissions scales with the transition frequency to the power of 3. Moreover, this ratio is independent of the considered atomic or molecular transition. But, what are the consequences of this strong frequency dependence? For one, the probability of spontaneous emission strongly increases with shorter wavelengths. Hence, the generation of short-wavelength

(coherent) laser radiation is difficult. In fact there are special demands on the pump process the shorter the wavelength gets.

In this context stimulated emissions appear to be rather rare processes. Nevertheless, it constructively adds photons to a radiation field just by interaction of the light with excited particles in the active medium (i.e. it is the process that leads to optical amplification). This is possible because the stimulated photons have the same properties as the originally exiting ones. In the wave picture this statement reads: *The frequency, the propagation direction, the polarization and the phase of the incoming wave is preserved during stimulated emission.*

However, in order to be able to use these properties one has to ensure that induced emissions dominate over the spontaneous ones, despite the frequency scaling behavior. On top of that the absorption processes should not occur.

2.2. Homogeneous and Inhomogeneous Broadening of Transitions

We need to consider line broadening mechanisms before we can continue with the working principle of lasers. So far we have assumed that the energy levels E_1, E_2 are sharp and, therefore, a well-defined resonance frequency ν_{21} exists. However, in reality all levels and, hence, the transitions between them, have an uncertainty that broadens them. In this context, it is possible to distinguish between homogeneous and inhomogeneous broadening mechanisms.

A *homogeneous line broadening* is characterized by the fact that the effect responsible for the broadening *hits all emitters/absorbers in the same way*. Hence, *it is impossible to define different classes of particles* (emitters and absorbers), i.e. all particles in the medium behave identically. Examples of mechanisms leading to homogeneous line broadening are:
- spontaneous emission – natural line width
- collision broadening (gases, liquids)
- lattice vibrations (solids)

In contrast, *inhomogeneous broadening* is present if *the effect which causes the broadening acts differently upon the different particles of the considered ensemble*. Consequently, *individual particles* may have different transition frequencies and *can be divided into classes* attending to their emission/absorption frequency. Examples of mechanisms leading to inhomogeneous broadening are:
- Doppler-effect (gases)
- Stark-effect (solids)

In the following the most important broadening mechanisms are described in more detail.

Homogeneous Broadening

Spontaneous Emission – Natural Linewidth

In the absence of photons (ρ = 0) the population of the excited state is only changed by spontaneous emission since no induced emission or absorption processes will occur. In that case, the rate equation for the evolution of the population density in level 2 (excited state) was given by:

$$\left.\frac{dN_2}{dt}\right|_{Spontaneous\,Emission} = -A_{21} \cdot N_2$$

Where A_{21} is the rate constant for spontaneous emission. This simple first order differential equation with constant coefficients can be solved with the following ansatz:

$$N_2 = C \cdot \exp\left(-t/\tau_2\right)$$

Hence, an exponential decay of the population density in the excited state N_2 with a characteristic time τ_2 can be expected. The ansatz becomes the solution of the differential equation if $\tau_2 = 1/A_{21}$, i.e. τ_2 is the average lifetime of the exited state. On a different account, C is a constant given by the initial conditions, C = N_2(t=0).

Since the population density of level 2 (N_2) decays exponentially with a time constant τ_2, the intensity I of the spontaneously emitted light will also decay exponentially with the same time constant (since the presence of less and less particles in the excited state, which are the ones that can potentially emit light, will lead to a decaying emitted intensity):

$$I \sim N_2 \sim \exp(-A_{21}t)$$

Hence, the electric field of the emitted light can be viewed as a damped oscillation at the resonance/transition frequency ν_{21}, which intensity at a fixed position decays exponentially with a time constant τ_2. Thus, the amplitude of the electric field of the emitted light is given by:

$$E(t) \sim \exp\left(-t/2\tau_2\right) \cdot \sin(2\pi\nu_{21}t + \varphi)$$

with φ being a phase constant. This electric field represents, as mentioned above, a damped wave at frequency ν_{21}.

As it is known, a damped wave cannot be monochromatic, it has to possess a spectral distribution to account for the temporal changes of its amplitude. This spectral distribution can be obtained by a Fourier transformation, which, after some calculations, will lead to the so-called power spectrum:

$$P(\nu) \sim \frac{\left(\frac{\Delta\nu}{2}\right)^2}{(\nu-\nu_{21})^2+\left(\frac{\Delta\nu}{2}\right)^2}$$

(2.6a)

The power spectrum of a damped wave is a Lorentzian function, which generic shape normalized to 1 is given by Equation (2.6a) and is plotted in the following figure:
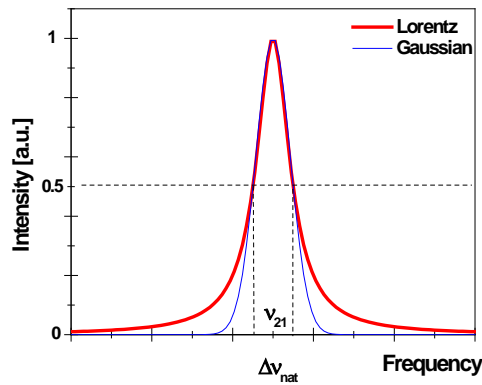
The full width at half maximum (FWHM) of that function is given by:

$$\Delta\nu_{nat} = \left(2\pi\tau_2\right)^{-1},$$

(2.6b)

Which is known as the natural linewidth and is a fundamental property of the transition. Taking into account the definition of photon energy, a variation of frequency (i.e. a frequency bandwidth $\Delta\nu_{nat}$) results in a variation of the photon energy (i.e. in an energy band ΔE):

$$\Delta E = h \cdot \Delta\nu_{nat} = \frac{\hbar}{\tau_2}$$

(2.6c)

Hence, *if an ensemble of atoms emits* such a damped wave, e.g. *through spontaneous emission, it seems as if level 2 would have an energy distribution with its FWHM given by equation (2.6c)*. This is also known as the Heisenberg uncertainty relation.

Example: The lifetime of an excited atomic state is typically in the order of some nanoseconds (for an allowed transition). Let's assume that it is τ_2 = 5 ns, then the natural linewidth is 30 MHz.

If the particles in an ensemble interact with each other or with their proximity, then other broadening mechanisms come into play and need to be considered. Such as:

Collision Broadening / Pressure Broadening

In gases, collisions can lead to an additional broadening, which is homogeneous because the probability of a collision is the same for all particles. As the rate of collisions scales with the pressure, collision broadening is also known as pressure broadening.

What happens during a collision?
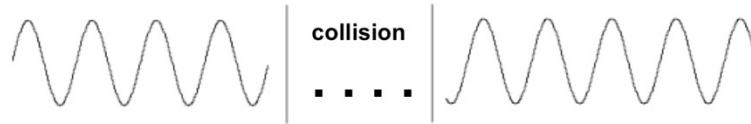
For this explanation we assume elastic collisions between the particles. If a collision happens during an emission process, the phase of the emitted wave is changed (i.e. there is a loss of phase information). This essentially results in the emission of finite wave trains with constant phase. An intuitive explanation for that loss of phase relation between the wave packets is that, during the collision, the energy levels of the particles are slightly shifted due to the particle interaction, which leads to a transitory change of the emission frequency (just during the time of collision) and, hence, to a phase shift.

A Fourier analysis of the emission can, again, reveal its spectral distribution, which is also given by a Lorentz profile. The FWHM is determined by the average time between collisions and given by:

$$\Delta \nu_{Coll} = \left(2\pi \cdot \tau_{coll}\right)^{-1} = \sqrt{\frac{3}{4mkT}} \cdot d^2 \cdot p$$

(2.6d)

with m being the particle mass, d the diameter of the particle, k the Boltzmann constant k = $1.39 \cdot 10^{-23}$ J/K, T the absolute temperature, p the gas pressure and τ_{coll} the average time between collisions (which can be derived from text-book equations of thermo-dynamics).

Example: Assuming an ensemble of CO_2 molecules, a gas pressure of 1 bar and an absolute temperature of 300 K, then the linewidth of the collision broadened transition is 4 GHz. *It is worth remembering that the pressure/collision broadening is typically orders of magnitude larger than the natural linewidth.*

Inhomogeneous Broadening

In contrast to the homogenous broadening case discussed so far, *there is also the case when different particles of the ensemble emit/absorb at different frequencies.* This case is known as *inhomogeneous broadening*. One example of this is caused by the static Stark effect. Hereby, the electrical field acting upon the various particles that can interact with the laser radiation is different because it is created by their neighborhood in the matrix of the solid body (host), which changes from particle to particle. This creates a so-called Stark splitting / shifting in the energy levels, which is not identical for all particles. This notwithstanding, *the most prominent example of mechanism leading to inhomogeneous broadening is the Doppler Effect.*

Doppler broadening

In gases, particles can move in all direction in space (with a certain statistical particle velocity distribution). If we assume that particles with a transition frequency v_{21} have a velocity component in the direction of the detector (z-direction), they will appear to have a shifted transition frequency due to the Doppler Effect.
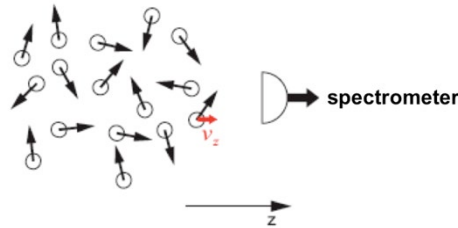
If that velocity component in the z-direction ($\pm v_z$) fulfills the condition $|v_z| \ll c$, then the maximum emission / absorption will occur at the shifted frequency:

$$v_D = v_{21} \cdot \left(1 \pm \frac{v_z}{c}\right),$$

which is a manifestation of the Doppler frequency shift. Obviously, the velocity component v_z is not the same for all particles and, hence, we can expect that each particle will emit / absorb at a slightly different frequency, thus leading to an inhomogeneous line broadening for the ensemble. If the considered gas (or in general the particles of the ensemble) is in thermal equilibrium, then the velocity component v_z can be described by Maxwell's velocity distribution, which states that the average number of particles per unit volume in a certain level i with velocities in the interval v_z to $v_z + dv_z$ is given by:

$$g_i(v_z) \cdot dv_z = N_i \cdot \left(\frac{1}{v_m \cdot \sqrt{\pi}}\right) \cdot exp\left(-\frac{v_z}{v_m}\right)^2 \cdot dv_z \qquad (2.7a)$$

with $\qquad v_m = \left(\frac{2kT}{m}\right)^{1/2} \qquad$ and $\qquad \int_{-\infty}^{\infty} g_i(v_z) \cdot dv_z = N_i$

In the equation above, v_m denotes the so-called most probable velocity, m is the particle mass, k is the Boltzmann constant, T is the absolute temperature and N_i is the particle population density in level i.

Example: Assuming room temperature, the most probable velocity is ~500 m/s. At a transition wavelength of 500 nm and a velocity component v_z equal to the most probable velocity, the resulting Doppler frequency shift is:

$$v_{21} \cdot \frac{v_z}{c} \approx 1 \, GHz.$$

To simplify things, we assume a gas medium at low pressure so that, consequently, pressure broadening can be neglected. As a result, the homogeneous linewidth is determined by the lifetime of the excited state (i.e. natural linewidth), which is about 2 orders of magnitude narrower than the Doppler shift. Hence, in this Doppler-broadened medium we can neglect the homogeneous component of the broadening and assume that the power spectrum of the ensemble is determined by the Maxwell's velocity distribution.

As the Maxwell's velocity distribution is a Gaussian function, the power spectrum will also have a Gaussian profile (here normalized to 1):

$$P(v) \sim \exp\left(-4 \cdot \ln 2\left\{\frac{(v - v_{21})}{\Delta v_D}\right\}^2\right)$$

(2.7b)

The linewidth of this power spectrum is usually referred to as Doppler width and it is given by:

$$\Delta v_D = 2 \cdot (\ln 2)^{1/2} \cdot v_{21} \cdot \frac{v_m}{c} = \frac{2v_{21}}{c} \cdot \sqrt{\frac{2kT \ln 2}{m}}$$

(2.7c)

Note that the Doppler width should be derived by you from the Maxwell's velocity distribution and just the final result has been provided here.

A more accurate description of inhomogeneous broadening via the Doppler effect can be done attending to a classification of the emitters/absorbers (i.e. particles that can interact with the laser radiation). Accordingly, all atoms with their velocity components in the z-direction (v_z) falling in the interval between v_z and $v_z + dv_z$ will be assumed to belong to one class. This way, the different classes can be illustrated in a histogram. If the classification is fine enough (i.e. the chosen velocity interval is narrow enough), then Doppler broadening within one class can be neglected, i.e. only one Doppler-shifted, homogeneously broadened line profile belongs to each class.

The line shape of the entire ensemble is then given by the sum over the homogeneously broadened line profiles of all classes of velocities, i.e. it will be a superposition of all homogeneous line shapes.
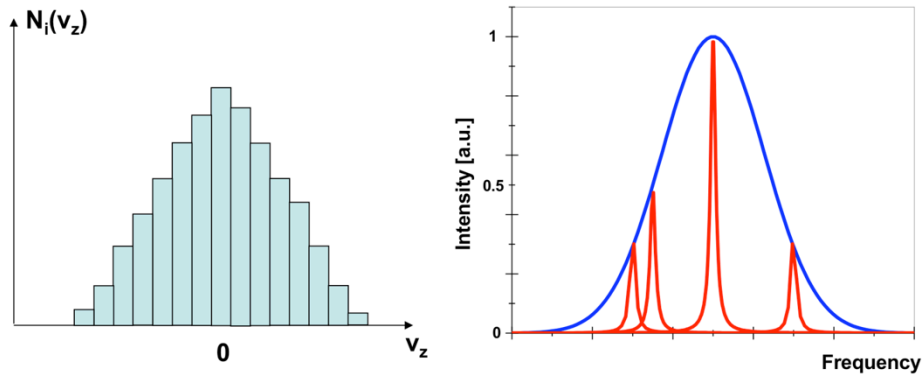
Typical broadening mechanisms in other types of matter

In solids:
A further inhomogeneous broadening in solids (e.g. dopants in a crystalline host) happens as a consequence of locally inhomogeneous crystal fields which result in a frequency shift through the static Stark-effect, as already mentioned before. In addition, homogeneous broadening can be observed in solids as well due to the interaction with lattice vibrations, the so-called phonons. The mechanism of line broadening due to lattice vibrations is very similar to that of pressure broadening in gases. In reality, almost always a superposition of homogeneous and inhomogeneous broadening is present.

In semiconductors:
The typical *energy bands* of holes and electrons (known as conduction and valence bands) have their origin in the interaction of the carriers with the periodic potential of the crystal structure. These energy bands are a manifestation of line broadening.

In dyes:
Organic dyes in liquid solutions can be used as an active laser medium offering an extremely broad amplification bandwidth. This extreme broadening is the consequence of the interaction of molecules with each other, which leads to an overlap of adjacent rotational and vibrational levels.

The following table provides a list of examples of transition linewidths in some prominent lasers. Please note that the stated linewidth is not necessarily the laser emission bandwidth, it just represents the range of frequencies around the resonance frequency in which emission can take place.

Type of Laser	Wavelength	linewidth	Mechanism
HeNe gas temp. 300 K	0.633 µm	1.5 GHz	Doppler / inhomogeneous
Argon-Ion gas temp. 2000 K	0.488 µm	4 GHz	Doppler / inhomogeneous
Excimer KrF	0.248 µm	10 THz	Overlapping vibrational levels
CO_2, 300 K 10 mbar	10.6 µm	60 MHz	Doppler / inhomogeneous
1 bar	10.6 µm	4 GHz	Collisions / homogeneous
10 bar	10.6 µm	150 GHz	Overlapping rotational levels
Dye Rh6G	0.6 µm	80 THz	Overlapping vibrational levels
Ruby laser	0.69 µm	330 GHz	Lattice vibrations / homogeneous
Nd:YAG	1.06 µm	120 GHz	Lattice vibrations / homogeneous
Nd:glass	1.06 µm	7.5 THz	Stark-effect / inhomogeneous
Semiconductor GaAs	0.8 µm	10 THz	Energy bands of electrons in periodic crystal field

2.3. Optical Amplification and Cross-sections

At first we need to derive the link between the absorption coefficient and the cross-sections. *Cross-sections are a very important term and their understanding is essential for laser physics.*

Absorption Coefficients and Cross-sections

From the point of view of a simple phenomenological consideration of light-matter interaction, we have learnt that the intensity of the light gets attenuated along its propagation through a medium.
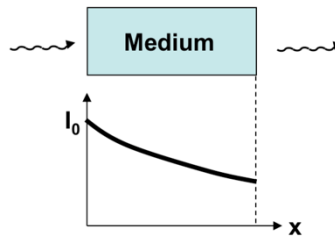
But, what is intensity? A common definition is: *Intensity I is the energy flux density of the radiation field, i.e. the energy per unit time (= power) through a surface element dA perpendicular to the pointing direction and can be written as*:

$$P = \int \mathbf{I} \cdot d\mathbf{A}$$

.

Consequently, the unit of intensity is W/m^2. Since each radiation field contains a certain spectral distribution, it is possible to describe the intensity by means of its spectral characteristics using:

$$I = \int I_\nu d\nu$$

with I_ν being the intensity per frequency interval at frequency ν, unit: $W/(m^2 \cdot Hz)$.

If the spectral intensity I_ν propagating through a medium is small, the change of the population density distribution within the energy levels of the medium can be neglected. In other words, in this case no saturation effects can be observed, what is commonly referred to as small signal absorption case. Under this assumption, the absorbed (lost/attenuated) intensity in a layer of the medium is proportional to the launched intensity and the layer thickness dx. Hence, we can write for the attenuation:

$$dI_\nu = -\alpha_\nu \cdot I_\nu \cdot dx$$

(2.8).

If absorption is the only physical process leading to attenuation of the light intensity, then the constant of proportionality is the absorption coefficient α_ν, which typically is frequency dependent.

The integration of (2.8) leads to

$$\frac{I_\nu(x)}{I_\nu(0)} = \exp(-\alpha_\nu \cdot x)$$

,

which describes the exponential decrease of I_ν with increasing penetration depth x. This is the well-known Lambert-Beer's law of absorption.

But... what is the relation between the intensity per frequency interval I_ν and the spectral energy density $\rho(\nu)$, which was introduced in lecture 1?

Since the radiation energy is transported with the speed of light, we can write:

$$\rho(\nu) = \frac{I_\nu}{c}$$

(2.9a)

Another important characteristics of the light radiation is the photon flux (number of photons per unit time) per frequency interval and surface element, which is defined as:

$$\Phi_\nu = \frac{I_\nu}{h \cdot \nu}$$

.

(2.9b)

Hence, instead of (2.8) we can also describe the attenuation of the photon flux, by:

$$d\Phi_\nu = -\alpha_\nu \cdot \Phi_\nu \cdot dx$$

Which has the straightforward solution:

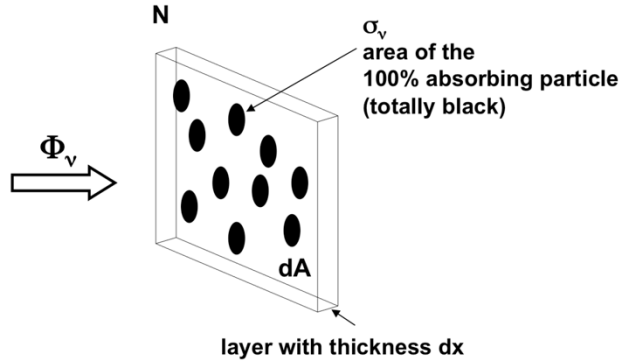$$\frac{\Phi_\nu(x)}{\Phi_\nu(0)} = \exp(-\alpha_\nu \cdot x)$$

.

It is easy to see that the absorption coefficient is proportional to the density of absorbing particles N (the more absorbing particles present in the medium, the stronger the attenuation will be). Thus, we can write:

$$\alpha_\nu = N \cdot \sigma_\nu$$

with σ_ν being a constant of proportionality known as *cross-section*. The cross-section encompasses all the properties of the considered transition, i.e. its strength, shape and resonance frequency.

For an intuitive understanding of the cross-sections we can consider an incoming photon flux Φ_ν arriving at a unit volume dV of a material. This unit volume of the material is

characterized by the layer thickness dx and the surface element dA, and it contains a certain density N (or a certain number n) of absorbing particles.

$$d\Phi_\nu = -\alpha_\nu \cdot \Phi_\nu \cdot dx$$
$$= -N \cdot \sigma_\nu \cdot \Phi_\nu \cdot dx$$
$$= -\frac{n}{dV} \cdot \sigma_\nu \cdot \Phi_\nu \cdot dx$$
$$= -\frac{n}{dA} \cdot \sigma_\nu \cdot \Phi_\nu$$

Using the above stated equations, we can derive a ratio of the absorbed photon flux $d\Phi_\nu$ to the incoming photon flux Φ_ν, which turns out to be equal to the number of absorbing particles (n) times the cross-section (σ_ν) over the surface element dA. Hence, this becomes an area ratio. Thus, *the cross-sections* represent an "effective" interaction area. In other words, they *represent the probability of the considered process (here absorption) projected to an area*. In this context, σ_ν is the cross-section (area) a particle would have if it would be 100% absorbing (represented by black disks in the figure above).

What are the properties of cross sections when line broadening mechanisms act on the considered transition?

- *For homogeneous broadening:* σ_ν is the cross-section of one particle for the absorption of a photon at the frequency ν. Since all particles behave the same way, they will all have the same cross-section.
- *For inhomogeneous broadening:* all particles absorb differently and, consequently, σ_ν is an average cross-section of the ensemble.

In the following we will derive the relation between the Einstein coefficients and the cross-sections. We start by considering two different ways of calculating the power consumption of the medium which is absorbing light.

Possibility 1:

$$dI_\nu = -\alpha_\nu \cdot I_\nu \cdot dx$$
$$dP = dV \int \alpha_\nu \cdot I_\nu \, d\nu = dV \cdot N \int \sigma_\nu I_\nu d\nu \quad , \tag{2.10a}$$

with dP being the absorbed / consumed power in a unit volume dV.

Possibility 2:

We assume a simple two level system and that all particles are in the ground state (available for absorption, i.e. $N_1=N$) and, then, we use equation (2.3a):

$$\frac{dN_1}{dt}\bigg|_{Absorption} = -B_{12} \cdot \rho(\nu) \cdot N_1$$

,

which provides the number of absorption processes per unit volume and unit time. With equation (2.9a) and the knowledge that each absorption process adds an energy $h\nu_{21}$ to the medium, we can write the power consumption (energy load per unit time) as:

$$dP = dV \cdot B_{12} \cdot \frac{I_\nu}{c} \cdot N \cdot h\nu_{21}.$$

(2.10b)

By comparing (2.10a) and (2.10b) (under the assumption that the spectral intensity I_ν is constant over the absorption line) it can be obtained:

$$\sigma = \int \sigma_\nu d\nu = \frac{B_{12} \cdot h\nu_{21}}{c}.$$

(2.11a)

σ is typically referred to as transition line strength.

In general, the cross-section σ_ν can be expressed as a line shape function $F(\nu)$ normalized to one and multiplied by the peak value of the cross-section σ_0. With equation (2.11a) we can write:

$$\sigma = \int \sigma_\nu d\nu = \int \sigma_0 F(\nu) d\nu = \frac{B_{12} \cdot h\nu_{21}}{c}$$

$$\Rightarrow \quad \sigma_0 \approx \frac{B_{12} \cdot h\nu_{21}}{c \cdot \Delta\nu}.$$

(2.11b)

$\Delta\nu$ is the width of the line shape function. Obviously, the broader the line gets (due to broadening mechanisms), the weaker the peak value of the cross-section will be.

In any case the exact calculation of the cross-sections is difficult. The Einstein coefficients (or the related dipole matrix elements) can only be calculated for simple atoms and molecules. In the case of dopants in complex environments such calculations fail. Hence, the values of cross-sections given in the literature have been experimentally obtained by means of spectroscopy.

Stimulated Emission

So far we have assumed that all particles are in the ground state. We will now consider the opposite extreme: all particles are in the excited state. In analogy to the absorption case, a cross-section for stimulated emission can be defined. Following the Einstein relation (2.4b), we can conclude that the cross-section for absorption and stimulated emission should be identical (since stimulated emission is just the reversal process of absorption with identical probability):

$$\sigma_\nu\big|_{Absorption} = \sigma_\nu\big|_{Stimulated\ Emission}$$

Assuming that all particles are excited, then the change of photon flux is described by

$$d\Phi_v = +N \cdot \sigma_v \cdot \Phi_v \cdot dx$$ (2.12)

resulting in an exponential growth of the photon flux with increasing propagation length through an excited medium:

$$\frac{\Phi_v(x)}{\Phi_v(0)} = \exp(N \cdot \sigma_v \cdot x)$$

Of course, in reality we will have a certain particle density N_1 in level 1 (available for absorption) and a certain particle density N_2 in level 2 (available for stimulated emission). Therefore, we can replace equations (2.8) and (2.12) by:

$$d\Phi_v = (N_2 - N_1) \cdot \sigma_v \cdot \Phi_v \cdot dx$$ (2.13)

leading to

$$\frac{\Phi_v(x)}{\Phi_v(0)} = \exp((N_2 - N_1) \cdot \sigma_v \cdot x)$$

Consequently, whether we can observe a growth or an attenuation of the photon flux depends on which of the two levels is more populated!

Thus, *in order to get amplification of the photon flux we need that $N_2 > N_1$. This condition is known as population inversion* and it will be discussed in detail later on.

The following table provides an overview of excited state lifetimes and stimulated emission cross-sections of some prominent laser transitions.

| Type of Laser	λ [μm]	τ_2 (upper level)	τ_1 (lower level)	σ_v [cm²]
HeNe	0.633	10-20 ns	12 ns	$3 \cdot 10^{-13}$
Ar-Ion	0.488	9 ns	0.4 ns	10^{-16}
Excimer	0.248	1-10 ns	<1ps	10^{-16}
CO₂ low pressure	10.6	1-10 ms	100 ns	10^{-16}
CO₂ high pressure	10.6	1-10 ms	1 ns	10^{-16}
Dye	0.6	5 ns	<10 ps	$2 \cdot 10^{-18}$
Ruby	0.694	3 ms	infinity	$2 \cdot 10^{-20}$
Nd:YAG	1.06	230 μs	30 ns	$8 \cdot 10^{-19}$
Nd:glass	1.06	300 μs	50-100 ns	$4 \cdot 10^{-20}$
semiconductor GaAs	0.8	4 ns	infinity	10^{-16}

2.4 Temporal and Spatial Coherence

Before we can move on to the principle of operation of lasers we need to recap another important characteristics of a radiation field: *coherence*. When talking about this characteristic *it is possible to distinguish between temporal and spatial coherence.* Interestingly, both types of coherence can be defined by measurement methods: the temporal coherence by means of a Michelson Interferometer and the spatial coherence using Young's double-slit experiment.

Temporal coherence

What is a Michelson Interferometer?
In this device an electro-magnetic wave characterized by the field strength $E_{in}(t)$ illuminates a beam splitter BS (which typically has a 50% splitting ratio). The resulting two waves propagate until they reach two mirrors at the end of their paths and are reflected back. Consequently, they are superimpose at the beam splitter and create the out-going wave $E_{out}(t)$. The two waves need propagation times t_1 and t_2 to travel their respective paths from the BS to the mirror (lengths l_1 and l_2, respectively) and back.
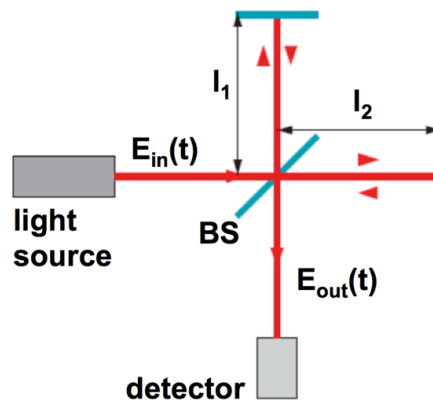
The general expression for the intensity of 2 superimposed waves with intensity I_1 and I_2 is:

$$I = I_1 + I_2 + 2 \cdot \sqrt{I_1 \cdot I_2} \cos(\phi)$$

(2.14)

where ϕ is the phase difference between the two waves.

If we assume that the incoming wave is perfectly monochromatic and that the BS splits the intensity (or power) exactly at 50%, then the intensity of the out-going wave is given by:

$$I_{out} = \frac{1}{2} \cdot I_{in}(1 + \cos\phi)$$

(2.15)

with

$$\phi = 2\pi \cdot v \cdot (t_2 - t_1)$$

being the phase difference and I_{in} being the intensity of the incoming wave.

Plotting equation (2.15), i.e. the intensity of the wave leaving the Michelson interferometer (normalized to the input intensity) as a function of the propagation time difference between the two branches, we can see that a perfect contrast (visibility) of the interference pattern is given for all time differences.
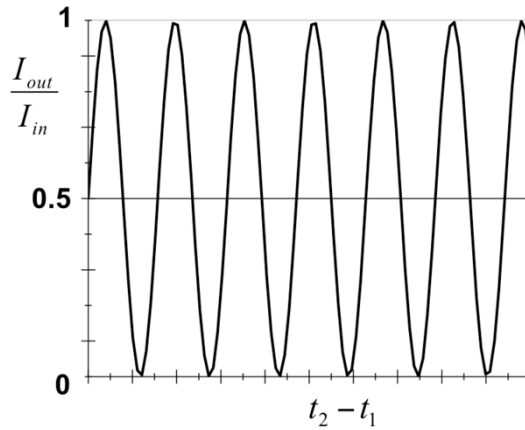
However, in reality what it is usually observed is:
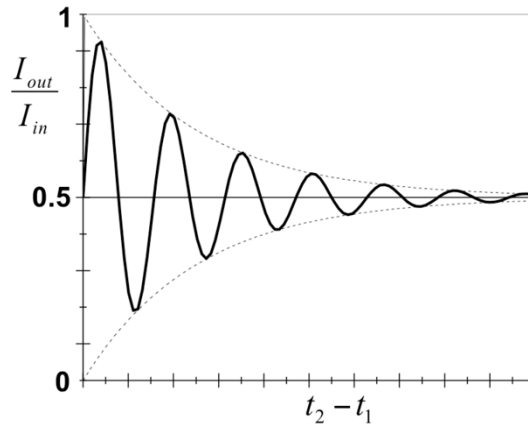
The spectral distribution of the incoming wave is the reason for the decaying contrast in the figure above. Assuming that the spectrum of the incoming wave has a spectral width Δv, then the variation of the phase differences between the two branches of the interferometer at a certain propagation time difference is given by:

$$\Delta\phi = 2\pi \cdot \Delta v \cdot (t_2 - t_1)$$

as soon as this variation gets in the order of 2π (by increasing the propagation time difference) some frequencies experience no phase difference and other a phase difference in the order of π. This implies that the interference pattern is washed out, because, constructive and destructive interference are roughly equally distributed among the different frequency components of the spectrum.

In other words: *for time differences larger than*

$$\tau_C = \frac{1}{\Delta \nu}$$

(2.16)

it is assumed that no correlation between the interference partners is given anymore. τ_c is known as *coherence time.* Associated to this time difference, there is a path difference l_c between the two branches that can be calculated by:

$$l_C = \tau_C \cdot c$$

(2.17)

l_c is known as *coherence length. A correlation between the two beams can only be found if the path difference in the interferometer is smaller than l_c.*

The following table compares a conventional spectral lamp with a stabilized laser attending to their temporal coherence. Lasers can emit a significantly narrower linewidth and, hence, they typically possess a temporal coherence that is orders of magnitude higher than broader light sources.

source	wavelength [µm]	rel. spectral bandwidth	l_c [m]	τ_c [s]
low press. Ne spectral lamp	0.633	$3 \cdot 10^{-6}$ (Doppler)	0.2	$7 \cdot 10^{-10}$
Stabilized HeNe Laser	0.633	$2 \cdot 10^{-12}$	$3 \cdot 10^5$	$1 \cdot 10^{-3}$

At this point it is important to note that, in principle, the temporal coherence of a source can be arbitrarily increased, if a sufficiently narrow spectral filtering is implemented. Of course, this happens at the cost of intensity.
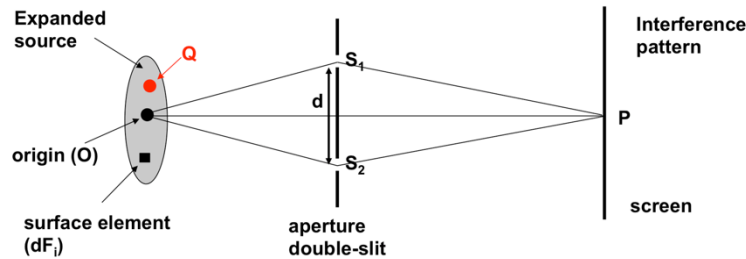
Spatial coherence

For optics spatial coherence might be even more important than temporal coherence. Spatial coherence can be defined using Young's double-slit experiment. Hereby, as it is known, the illumination of a double-slit with a monochromatic plane or spherical wave results in a pronounced interference pattern on a screen placed behind the double-slit.
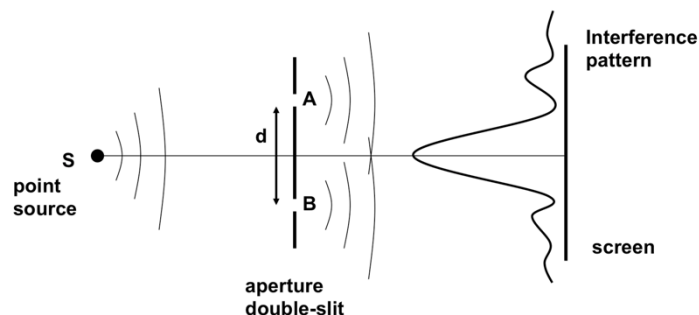
In contrast, if one places an independent thermal light source in front of each aperture no interference pattern can be observed. The reason for this is the missing (spatial) coherence between the two independent light sources. In other words, there is no correlation between the phase and amplitude of the two independent light sources.

Spatial coherence can be evaluated by illuminating a double-slit with a spatially extended light source. In this case, an interference pattern becomes visible, but the contrast of that pattern decreases with an increasing distance d between the two apertures / slits.

The understanding of that observation leads to an understanding of spatial coherence.

The slits S_1 and S_2 can be seen as starting points of new waves (according to the Huygens' principle each point of a wavefront is the source of a new spherical wave). The intensity at a certain point P on the screen I is determined by the amplitudes A_i and phases φ_i from all surface elements dF_i at S_1 and S_2 and by the path difference of each slit (1/2) to P:

$$\Delta s = S_1 P - S_2 P$$
.

If the individual surface elements dF_i emit independently, then the superposition of the emission from all surface elements dF_i of the expanded source at the slits S_1 and S_2 will statistically fluctuate over time. Consequently, the interference pattern at the screen is averaged out. However, if the emission from all surface elements is synchronized, then the contrast of the interference at the screen would not be influenced. This condition is indeed exactly fulfilled for individual points, e.g. light just from origin O of the extended source. In that special case there is no path difference (source to slit 1 and source to slit 2) and

$$OS_1 = OS_2$$
.

However, any other source points Q possess a path difference

$$\Delta s = QS_1 - QS_2$$
,

which grows towards the edges of the expanded source. The following simple geometric considerations aim at deriving an expression of that path difference:
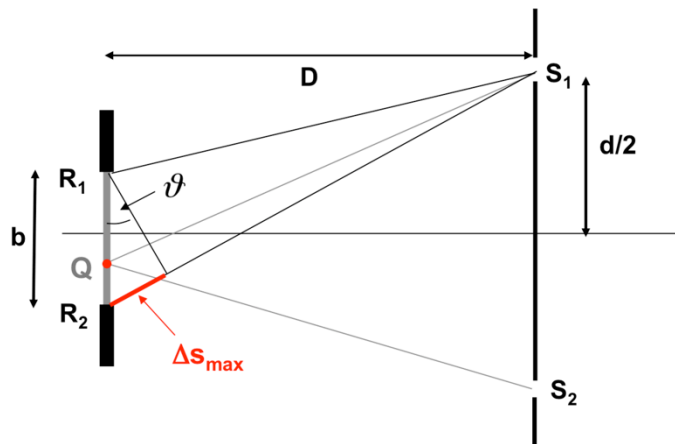
$$R_1 S_2 = R_2 S_1$$

$$\Delta s_{max} = R_1 S_2 - R_1 S_1 = R_2 S_1 - R_1 S_1$$

$$\approx b \cdot \sin \vartheta = \frac{b \cdot d}{2 \cdot D}$$

$$with \quad D \gg d$$

If Δs_{max} is larger than $\lambda/2$ and a statistical emission from all source points Q is assumed, then the variation in the phase difference at the slits:

$$\Delta \phi = \frac{2\pi}{\lambda} \cdot \Delta s,$$

fluctuates more than π. Consequently, the interference pattern is washed out with time. Thus, the condition for a coherent illumination of the double slit is

$$\Delta s_{max} \approx \frac{b \cdot d}{2 \cdot D} < \frac{\lambda}{2} \qquad \qquad \frac{d}{\lambda} < \frac{D}{b} \qquad \qquad d < \frac{\lambda}{\Theta_s}$$
$$\text{i.e.} \qquad \qquad \text{or}$$

with Θ_s being the aperture angle (under which the expanded source is seen from the double-slit).

In reality, the degree of spatial coherence of a source lays somewhere between that of statistical emission (what we have assumed above) and that of a plane wave (perfect spatial coherence). This fact motivates the introduction of a quantity f to quantify the degree of spatial coherence: the transversal coherence length l_t, which is defined as the distance d in Young's double-slit experiment, at which the interference disappears.

What is behind that transverse coherence length and what are its consequences?

Let's consider the aberration-free focusing of a plane wave (l_t = infinity) with a lens of limited diameter D_L.
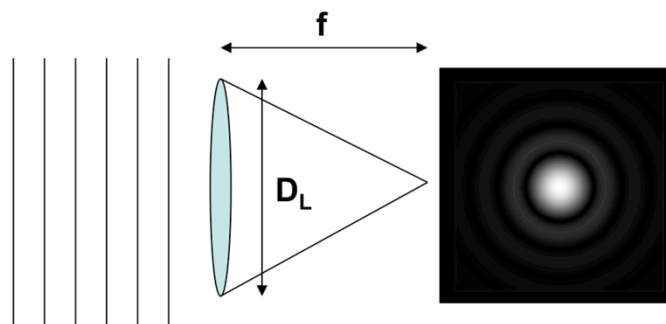
A diffraction pattern can be observed in the focal plane, which is known as the Airy diffraction pattern. The diameter of the central lobe of that pattern is given by:

$$q \approx \frac{f}{D_L} \lambda$$

where f is the focal length of the lens.

That spot size is the smallest possible one that can be obtained with that lens at that particular wavelength. The diffraction pattern is a result of interference of all the wavelets, which originate from the entire lens surface. This interference takes place because, over the entire lens surface, spatial coherence (correlation) exists ($l_t > D_L$).

In contrast, focusing a light beam with a transversal coherence length smaller than the lens diameter ($l_t < D_L$) leads to larger spot sizes. This is because in this case only wavelets in a surface area of diameter l_t can interfere with each other. Consequently, the intensity distribution in the focal point has a significantly larger diameter, given by

$$q \approx \frac{f}{l_t} \lambda > \frac{f}{D_L} \lambda$$

Hence, *the lack of spatial coherence* of thermal light sources *prevents the focusing of* their *radiation energy to small spatial areas.* However, the spatial coherence of any source can be arbitrarily enhanced by spatial filtering, of course, again, at the cost of power.