Astronomical units and useful constants

Astronomical unit 1 AU $= 149\,597\,870\,700$ m $\simeq 1.5 \times 10^{11}$m. The median separation between Earth and Sun. Often used when dealing with distances within the Solar System.
Parsec 1pc $= 648\,000/\pi$ au $\simeq 3.0857 \times 10^{16}$m $\simeq 3.3$ ly. It is the distance at which an object would show a parallax displacement of one arcsec.
Angstrom 1Å $= 10^{-10}$m $= 0.1$ nm.
Jansky 1 Jy $= 10^{-23}$ erg s^{-1} cm^{-2} Hz^{-1} $= 10^{-26}$ W m^{-2} Hz^{-1}. Unit for flux density (indicated as S_ν, F_ν, or f_ν).
Solar luminosity $L_\odot = 3.828 \times 10^{26}$ W $= 3.828 \times 10^{33}$ erg s^{-1}.
Solar mass $M_\odot = 1.989 \times 10^{30}$ kg.
Speed of light c $= 299\,792.458$ km S^{-1}.
Boltzmann's constant $k = 8.617 \times 10^{-5} eV/K = 1.380649 \times 10^{-16} erg/K$.
Electron volt $1eV = 1.60218 \times 10^{-19} J = 1.60218 \times 10^{-12} erg$. Widely used in astronomy as a unit of frequency or wavelength. It corresponds to the energy gained by an electron accelerating from rest through an electric potential of one volt in vacuum.

Conventions and units

γ and χ rays $\rightarrow$ energy units [keV, MeV, TeV].
Visible and IR $\rightarrow$ use wavelength units [Å, nm, μm].
Sub(mm) and radio $\rightarrow$ Use both wavelength [μm, mm, cm] and frequency units [THz, GHz, MHz].

Cosmological redshift

Cosmological redshift is due to the expansion of space and the consequent stretch of wavelengths, which makes wavelengths longer (and so redder).

$$z = \frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} \rightarrow (1+z) = \frac{\lambda_{obs}}{\lambda} = \frac{\nu_{em}}{\nu_{obs}}$$

Doppler shift

$$\frac{\lambda_{obs} - \lambda_{em}}{\lambda_{em}} = \frac{v_r}{v_s}$$

Specific intensity or brightness I_ν

(Often referred to as angular surface brightness or just brightness). The quantity I_ν corresponds to the rate of energy transported along a particular direction, per unit area, per unit solid angle, and per unit frequency.

$$I_\nu = \frac{dE}{dt \cdot dA \cdot d\Omega \cdot d\nu} [ergs^{-1}cm^{-2}sr^{-1}Hz^{-1}] \tag{1}$$

Defined for a direction oriented perpendicularly $dA(\cos\theta = 1)$. I_ν is the flux density per unit solid angle in the sky. By multiplying I_ν by ν we obtain the reduced brightness

$$\nu \cdot I_\nu = \lambda I_\nu \tag{2}$$

Étendue or throughput

Etendué or throughput, also referred to as the "area-solid angleproduct is defined as defined as the product $dA \cdot d\Omega$ (or product of aperture by field of view) and is a conserved measurement.

$$dA_2 \cdot d\Omega_2 = dA_1 \cdot d\Omega_1 \rightarrow dA \cdot d\Omega = const. \tag{3}$$

> Important: I_ν is independent of the distance from the astronomical source and hence it is a property of the source itself. Brightness is the same at the source and at the telescope (or detector).

The mean intensity is derived by removing the dependency on Ω (averaging across all solid angles)

$$J_\nu = \frac{1}{4\pi} \int I_\nu d\Omega \tag{4}$$

Flux density

Flux density (indicated wtih F_ν, or often in radio astronomy with S_ν) is the energy flowing through dA in the time dt:

$$dF_\nu = \frac{dE}{dt \cdot dA \cdot d\nu} = I_\nu \cdot \cos\theta \cdot d\Omega \tag{5}$$

dF_ν is defined for $\theta \neq 0$, and so we need to add the dependency on $\cos\theta$. If we are measuring the flux density of a source in the sky we need to integrate eq. 5 over the angle subtended by the source to get the flux density

$$F_\nu = \int_{\Omega_{source}} I_\nu \cdot \cos\theta \cdot d\Omega [ergs^{-1}cm^{-2}Hz^{-1}] \tag{6}$$

In (sub-)mm/radio astronomy, the unit typically used for flux density is Jansky. The concept of flux density is very convenient to measure the strength of the line or continuum emission of an unresolved or, in any case compact astronomical source (Ω is small). The flux density is much less convenient for extended sources (Ω is large), hence in these cases astronomers prefer to quote the specific intensity (I_ν).

Optical depth (photon penetration)

For a ray passing through matter, the variation of I_ν, through a distance ds is

$$dI_\nu = -\alpha_\nu I_\nu ds \rightarrow \frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu \tag{7}$$

where α_ν $[cm^{-2}]$ is the (monochromatic, because it depends on ν) absorption coefficient and j_ν $[ergs^{-1}sr^{-1}cm^{-3}Hz^{-1}]$ is the spontaneous emission coefficient. The absorption coefficient depends on the incident field I_ν (if there is no incident field, there is nothing to absorb!), but the spontaneous emission term does not depend on the incident radiation field, I_ν. The material can emit spontaneously also without incident radiation.
From this, we define the (monochromatic) optical depth as

$$d\tau_\nu = \alpha_\nu ds \rightarrow \tau_\nu(s) = \int_{s_0}^{s} \alpha_\nu(s')ds' \tag{8}$$

Integration is performed along the path of the traveling ray, from an arbitrary point s_0 (which sets the zero point for the optical depth scale) to the observer's position s.
$\tau_\nu > 1$ Optically thick (or opaque medium).
$\tau_\nu < 1$ Optically thin (or transparent medium)

Radiative transfer transport equation

$$I_\nu(\tau'_\nu) = I_\nu(0)e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(\tau'_\nu)e^{-(\tau_\nu - \tau_\nu')}d\tau_\nu' \tag{9}$$

The right term of the equation can be interpreted as the sum of (I) the initial intensity diminished by absorption; (ii) the integrated source diminished by absorption. The entire equation translated into words sounds something like this: brightness at a certain position = initial brightness at $\tau = 0$ where the original light is diminished by $e^{-\tau}$ + (source function) how this slab of material is emitting at different optical depth, absorption happens in the slab of material.
Homogenous medium (S_ν does not depend on τ_ν)

$$I_\nu(\tau_\nu) = S_\nu + e^{-\tau_\nu}(I_\nu(0) - S_\nu) \tag{10}$$

In the homogenous medium limit, one can see that for large optical depths $\tau_\nu \rightarrow \infty$, we have $I_\nu \rightarrow S_\nu$, hence the specific intensity that emerges out of the medium approaches the source function, meaning that none of the incident radiation field ($I_\nu(0)$) manages to escape from the medium, and the observer only observed the radiation produced within the medium itself. In the opposite case, $\tau_\nu \rightarrow 0$, then $I_\nu \rightarrow I_\nu(0)$, hence the emergent brightness is equal to the incident one. The simplified limit of a homogenous medium is of particular interest in astronomy since it is the case that applies also to thermal emission, where the source function S_ν is the Planck function $B_\nu(T)$ which depends only on T.
The mean free optical path The average optical path a photon can travel before being absorbed. The probability of a photon to travel through one optical path τ_ν without being absorbed is $e^{-\tau_\nu}$.

$$\langle\tau_\nu(s)\rangle \equiv \frac{\int_0^\infty T_\nu(s) \cdot e^{-\tau_\nu(s)} d\tau_\nu(s)}{\int_0^\infty e^{-\tau_\nu(s)} d\tau_\nu(s)} = 0 \tag{11}$$

So for a homogenous material (α_ν does not depend on $\tau_\nu(s)$), $\langle\tau_\nu\rangle = 1 = \alpha_\nu s_\nu\rangle = \alpha_\nu l_\nu$ where we define l_ν the mean free geometrical path, $l_\nu = 1/\alpha_\nu$.

Thermal emission

The function describing I_ν for BB radiation is the Planck function

$$I_\nu^{BB} = \text{universal function of T and } \nu \equiv B_\nu(T) \tag{12}$$

From the radiative transfer transport equation (eq. 9) we have

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu = -B_\nu(T) + S_\nu \tag{13}$$

$dI_\nu/d\tau_\nu$ can not be different from zero, because the new configuration (BB cavity with the thermal material inside) is also a BB. So we have obtained that the source function of the thermal material must be equal to the Planck function

$$\text{Thermal emitter} \rightarrow S_\nu = B_\nu(T) \rightarrow j_\nu = \alpha_\nu \cdot B_\nu(T) \tag{14}$$

which is Kirchhoff's law of thermal radiation. Hence the radiative transfer transport equation for thermal material is:

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + B_\nu(T) \tag{15}$$

The Planck function

Black body: Emitter in thermal equilibrium (uniform $\neq$ constant temperature) that emits radiation that is itself in thermal equilibrium.
Many astronomical sources behave very similarly to a blackbody, allowing us to model their emission (at least in first approximation) using the Planck function). The Planck function (or Planck's law) can be expressed either as a function of frequency

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \tag{16}$$

or as a function of wavelength

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1} \tag{17}$$

Stefan-Boltzmann law

From Planck's law we calculate the flux density at the surface of a BB emitter (either an infinite surface or very close to it)

$$F_\nu = \pi B_\nu(T) = \frac{2\pi h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \tag{18}$$

From this, we obtain the total bolometric flux by integrating it across the whole spectrum

$$F = \int_0^\infty F_\nu d\nu = \left(\frac{2\pi^5 k^4}{15c^2h^3}\right) T^4 = \sigma T^4 \tag{19}$$

Wien's law

(High frequency domain approximation of Planck's law, $h\nu \gg k_B T$)

$$I_\nu^W(T) = \frac{2h\nu^3}{c^2} e^{-h\nu/kT} \tag{20}$$

Rayleigh-Jeans approximation

(Low frequency domain approximation of Planck's law, $h\nu \ll k_B T$)

$$e^{h\nu/kT} \simeq 1 + \frac{h\nu}{kT} \text{ (From Taylor expansion)} \tag{21}$$

From which we obtain

$$I_\nu^{RJ}(T) = \frac{2kT\nu^2}{c^2} \tag{22}$$

or, in units of wavelength

$$I_\nu^{RJ}(T) = \frac{2kTc}{\lambda^4} \tag{23}$$

RJ approximation can be used for all thermal radio/mm/submm sources except for low temperatures in the (sub-)mm range.

Brightness temperature

(The part where Claudia introduces something we will hate and something that will make us hate her) The I_ν of a source (any, source, not necessarily a BB or a thermal emitter) at a certain frequency ν can be expressed in terms of the temperature of a black body that has the same brightness as that source at that specific frequency. The brightness temperature is therefore defined as

$$I_\nu \equiv B_\nu(T_b(\nu)) \tag{24}$$

(Not a real temperature).

Effective temperature

The effective temperature T_{eff} is not the same as the brightness temperature. It is derived from the total amount of flux radiated, integrated over all frequencies (i.e. bolometric flux), and not from the specific intensity. It is defined as the temperature of a black body with the same bolometric flux as the source.

$$F_{Bol} = \int_0^\infty F_\nu d\nu = \int_0^\infty d\nu \int d\Omega I_\nu \cos(\theta) \equiv \sigma T_{eff}^4 \tag{25}$$

The effective temperature of the Sun is $\sim 5780K$.

Apparent magnitude, m

(The one you measure from your instrument on earth)
Given F_ν or F_λ the monochromatic flux (or flux density) of a source at a given frequency/ wavelength, its apparent magnitude, non-corrected for atmospheric absorptions, is:

$$m_\nu = -2.5\log_{10}\frac{F_\nu}{F_{\nu 0}} \tag{26}$$

where F_ν is the flux of the source and $F_{\nu 0}$ is the flux density of your reference. This can also be written as

$$m_\nu = -2.5\log_{10} F_\nu + q_{\nu 0} \tag{27}$$

where the constant $q_{\nu 0}$ defines the zero point. This can again be rewritten as

$$\frac{F_1}{F_2} = 10^{-0.4(m_1 - m_2)} \tag{28}$$

AB magnitude

Defined such as a source with $F_\nu = 3.63 \times 10^{-20} ergcm^{-2}s^{-1}Hz^{-1}$ has AB mag = 0, hence:

$$m_{\nu,AB} = -2.5\log_{10} F_\nu - 48.6 \tag{29}$$

where F_ν is given in $[ergcm^{-2}s^{-1}Hz^{-1}]$.

Space telescope magnitude (STMAG)

A source with $F_\nu = 3.63 \times 1o^{-9} ergcm^{-2}s^{-1-1}$ has magnitude ST = 0 in every filter, hence:

$$m_\lambda = -2.5\log_{10} F_\lambda - 21.1 \tag{30}$$

where F_ν is given in $[ergcm^{-2}s^{-1-1}]$.

Bolometric magnitude

The (apparent) bolometric magnitude is the integral of the monochromatic flux over all wavelengths (usually optical spectrum)

$$m_{bol} = -2.5\log_{10}\left[\frac{\int_0^\infty F_\lambda d\lambda}{e_b}\right] \tag{31}$$

where $e_b = 2.52 \times 10^{-8} Wm^{-2}$.

Absolute magnitude, M

The absolute magnitude M is defined as the magnitude that the source would show if placed at a standard distance of 10 pc.

$$m_1 - m_2 = 5\log_{10}\frac{d_1}{d_2} \tag{32}$$

The relation between apparent and absolute magnitude

$$m - M = 5\log_{10}(d[pc]) - 5 \tag{33}$$

Basic optics

Telescopes divide into refractions (using lenses), and reflectors (using mirrors). When a beam of light travels from one medium to another, part of the beam is scattered forward and so transmitted"(or refracted), while a part of it is scattered backward (reflectedbeam). *The processes of refraction, reflection, and diffraction are macroscopic manifestations of scattering of electromagnetic radiation by atoms and molecules in a medium.* Scattering can be seen as the redirection of radiation out of the original direction of propagation, usually due to interactions of the radiation with charged particles.

Rayleigh, Mie, and geometric scattering

Scattering of electromagnetic radiation involving particles smaller than the wavelength, i.e. $d < \lambda/10$ is referred to as **Rayleigh scattering**. Lord Rayleigh (1871) understood scattered sunlight in terms of molecular oscillators and derived that the intensity of the scattered light is proportional to $\propto \lambda^{-4} \propto \nu^4$. If we compare with $\lambda \sim 5000$nm, then both atoms and ordinary molecules (whose diameters measure a few 10s of nm) satisfy the condition for Rayleigh scattering.
When instead the dimension of the particle is comparable to λ, i.e. $d \sim \lambda$, the theory to adopt is that of the **Mie scattering**, which becomes independent of λ when the particle size $d > \lambda$ (i.e. white light in, white light out). Mie scattering is stronger in the forward direction. The theory requires the scattering particles to be spherical.
When $d > 10\lambda$, the ordinary laws of geometric optics are sufficient to provide a macroscopic explanation for the interaction of light and matter and we refer to the process as **geometric scattering**.

Refraction

Quoting Hecht: The properties of the material determine whether the absorption and emission process (scattering) of the photons retard or advances the phases of the scattered radiation, even as they travel at speed c". The path of light is also bent at the interface between the two media, and the bending depends on the ratio of the so-called indices of refraction of the two media, which also depend on wavelength.
Refractive index is a dimensionless number that describes the speed of light through a material, and it is defined as

$$n_\lambda = \frac{c}{v_\lambda} \tag{34}$$

where c is the speed of light, and v_λ is the speed of light within the specific medium ($v_\lambda = c$ in vacuum, independent of wavelength).
Snell's law

$$n_i \sin(\theta_i) = n_t \sin(\theta_t) \tag{35}$$

Note that θ_i and θ_t are respectively equal to the angle of incidence and to the angle of refraction with respect to the normal to the interface.
(I) A ray entering a higher-index medium bends towards the normal to the interface, while if the medium has a lower index, light bends away from the normal ($n_t > n_i \to \theta_t < \theta_i$).
(II) Relative index of refraction: $n_{ti} = \sin(\theta_i)/\sin(\theta_t)$.

Reflection

When a beam of light impinges on a surface that is at the interface between two media, some light is scattered backward, a phenomenon called reflection.

$$\sin(\theta_i) = \sin(\theta_r) \to \theta_i = \theta_r \tag{36}$$

The angle of incidence is equal to the angle of reflection.

Fermat's principle

Proposed by Pierre de Fermat in 1657, the Principle of Least Time, which encompasses both refraction and reflection, states that "The actual path between two points taken by a beam of light is the one that is traversed in the least time". From this principle, it is possible to derive both Snell's law (law of refraction) and the law of reflection.

Telescope specifications

A telescope is an instrument that produces an image of a portion of the sky, corresponding to our astronomical target. What we get from Earth is actually the two-dimensional projection of our astronomical target on the celestial sphere. A telescope has two main functions:

1. T0 collect light emitted by your astronomical target (or a portion of the sky).
2. To transport that light to the focus (by refracting the light using lenses, or by reflecting the light using mirrors).

The main technical specifications of a telescope (independent of the type and wavelength):
Primary mirror or lens The first optical element encountered by light from the astronomical target when hitting the telescope. It usually has a paraboloid surface, which is the best at producing anastigmatic images. However, sometimes spherical mirrors are used because they are less expensive to polish.
Aperture (D) Usually given as the diameter of the primary mirror. The aperture determines the collecting area, and the maximum energy (or light) collecting capacity of the telescope.
Focal plane It is the plane passing through the focal point and oriented perpendicular to the optical axis of the system, where the astronomical image is formed.
Focal length (f_λ, or sometimes indicated simply as f) Distance between the first optical component of the telescope (e.g. the primary mirror) and the point of focus. The focal length of a spherical mirror is $f \simeq R/2$ where R is the radius of curvature of the mirror. The focal length F_λ of a thin lens can be calculated directly from its index of refraction n_λ and geometry of the lens (i.e. its radius of curvature R, which is positive for converging lenses and negative for diverging lenses). If both surfaces are spheroidal, the lensmaker's formula gives

$$\frac{1}{f_\lambda} = (n_\lambda - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \tag{37}$$

Field of view (FoV) Portion of the celestial sphere (sky) visible through the telescope. It is usually given in terms of an angular diameter measured in degrees or arcmin or arcsec. Sometimes the FoV is indicated by its area (hence a solid angle, e.g. given in square degrees).
Focal ratio or f-number(f or f/N where N is the f-number) It is the ratio between focal length f and diameter of the optical aperture.

$$N = f/D \to D = f/N \tag{38}$$

For example: if the focal length is $f = 10$ mm and the optical aperture has $D = 5$ mm, the f-number N is 2, often expressed using this notation: $f/2$.
Focal scale (s) It is the scale of the image formed in the focal plane, which depends on the equivalent focal length f. An object seen at the angle u, forms an image at height s

$$s = f\tan(u) \approx f \cdot u[mmm/"] \tag{39}$$

This leads to the definition of the plate scale, usually indicated with a differential equation du/ds

$$p = \frac{du}{ds} = \frac{1}{f} \tag{40}$$

The plate scale p of a telescope connects the angular separation of two targets du on the sky with the linear separation ds of their images on the focal plane. **Magnification**

$$\omega = u'/u \approx f/f' \tag{41}$$

Throughput or étendue ($A \cdot \Omega$ product) For a telescope it is the product of its collection area A and the solid angle subtended by the observed potion of the sky. Given the aperture diameter D, and Φ_{FoV} the (angular) diameter of a circular FoV, then the $A \cdot \Omega$ can be expressed as:

$$A \cdot \Omega = \eta \cdot (\pi D^2/4) \cdot (\pi \phi^2_{FoV}/4) = \eta \pi^2/16(D \cdot \phi_{FoV})^2 \tag{42}$$

where η is a factor < 1 that accounts for losses, blocking, etc.

A perfect optical system produces an image with the same etendue as the source.

Diffraction and angular resolution

Francesco Grimaldi: "the effect is a general characteristic of wave phenomena occurring whenever a portion of a wavefront, be it sound, a matter wave or light, is obstructed in some way". Huygen's principle states that each point in a propagating wavefront (=surface of the constant phase of waves) can be seen as a source of secondary spherical waves with the same frequency ν as the primary wave. The relevant theory to describe the effect of diffraction on astronomical observations is the **Fraunhofer diffraction** equation, which applies in the **far field** case, i.e. when the diffraction pattern is seen at a large distance from the diffracting object. Here we simply report the Fresnel equation, which provides a solution to the Diffraction problem, and its simplified version that hold in the far-field approximation, which is the Fraunhofer equation. From this we obtain

$$\rho_{spot} = 1.22\frac{f}{\lambda D} \to \Delta\theta = \frac{\rho_{spot}}{f} = 1.22\frac{\lambda}{D} \tag{43}$$

where $\Delta\theta$ is the annuglar spread, λ is the wavleength of the incident radiation, D is the aperture diamenter, and θ is given in radians. The Airy pattern is therefore the Fraunhofer diffraction pattern from a ciruclar aperture. Most of the flux (83.3 %) will be included in the first minimum. This is often used to define the theoretical **angular resolution** of a telescope. The angular resolution (resolving power) of a telescope is the minimum angular separation of two points on the celestial sphere (or two targets) that can be identified by the telescope. In other words, it is the telescope's ability to separate two objects in the sky (or, alternatively, to distinguish the small details of an astronomical source).
Accordingg to the **Rayleigh criterion**, *two sources are resolved when the peak of the Airy pattern of one source coincide with the first minimum of the Airy pattern of the other one."*
The angualr resolution of an instrument (especially in the optical/IR and X-ray regimes) is referred to as its **Point Spread Function (PSF)**, which is the responce function defined as the intensity distribution of the diffraction pattern produced by a point-like source normalized to its maximum value.

$$\text{Sthrel ratio (S)} = \frac{\text{observed PSF peak}}{\text{theoretical PSF peak}} = e^{-(\Delta\phi)^2} \tag{44}$$

≤ 1, the closer to 1, the better. Telescope is well built, otherwise the telescope is affected by optical defects. T he theoretical angular resolution is set only by the diffraction of light, the actual angular resolution of a telescope depends also by optical aberrations and by the Earth's atmosphere: the 'seeing' for optical/IR observations. Radio/millimeter telescopes are instead diffraction-limited, meaning that the angular resolution of radio/mm single-dish telescopes corresponds to their theoretical diffraction limit.

Refracting telescopes

Refracting telescopes have two problems:
1) big lenses are heavy and so they tend to deform under gravity
2) since the refractive index of a lens depends on wavelength, different colors are brought to different focal points, a problem known as chromatic aberration.
This problem was overcome by the introduction of "achromatic"and "apochromaticlenses. Problem 1) instead has never been solved.

Reflecting telescopes

Reflecting telescopes use mirrors instead of lenses, coated with some layer of other material such as Aluminium. Reflectors do not suffer from chromatic aberrations, but they still suffer from diffraction and various aberrations such as coma.
The main telescope designs, which are often the basis used for modifications are listed below:

1. **Newtonian:** A paraboloid primary mirror plus a flat secondary that reflects to a focal plane at the side.
2. **Cassegrain:** A two-mirror on-axis design, with a concave (paraboloid) primary mirror, and a convex (hyperboloid) secondary that reflects the light back to a hole in the primary mirror (Cassegrain focus).
3. **Nasmyth:** Similar to Newtonian, but with the addition of a tertiary mirror that reflects the light to the side of the telescope to allow for the mounting of heavy instruments. A telescope can have both Cassegrain and a Nasmyth focus. The Nasmyth focus can also be called "Coudéfocus if further optics are added to a Nasmyth-style telescope to deliver light to a fixed focus point.
4. **Schmidt:** The spherical aberration due to the use of spherical primary mirrors is corrected by a lens placed along the optical path.

Active optics is a technique that corrects for mirror deformations (due to e.g. gravity or temperature fluctuations).

Optical aberrations of reflecting telescopes

As a compromise between image quality and field of view, the image of a geometric point in the sky is never just a point, due to one or more optical aberrations, which corresponds to departures from idealized optical systems. The five primary aberrations are

1. **Spherical aberration:** In a spherical primary mirror (cheaper and easier to produce compared to a parabolic one), the light rays hitting the center and edges of the mirror converge at different points
2. **Coma:** The main optical aberration of parabolic primary mirrors is coma. Coma makes point sources appear comet-like.
3. **Astigmatism:** After Coma, Astigmatism is the main optical aberration of parabolic mirrors. The focus in the radial direction is in a different place from that in the circumferential direction. Astigmatism makes point sources appear eye shape-like.
4. **Field curvature:** The curvature of the field of view is a real problem for many reflector designs. The instruments placed at the focus need to compensate for this.
5. **Distortion:** It destroys the geometrical similarity between the target and the image on the focal plane while maintaining a one-to-one correspondence between points of the targets and points of the image.

Telescope mounts

In the **equatorial mount**, one of the axes (the so-called polar or hour axis) is directed toward the celestial pole. The other axis, the declination axis, is perpendicular to it. The polar alignment allows the telescope to naturally track astronomical targets on the sky while Earth rotates.
The **azimuthal or alt-aximuth** mount has a vertical and horizontal axis. this mounting is easier to construct and more stable for large telescopes. However, with this mount, in order to track a source on the sky (and so compensate for Earth's rotation), the telescope must be turned around both axes with changing velocities. This requires the use of computers.

Earth's atmosphere

The Earth's atmosphere affects astronomical observations in many ways. The physical processes responsible for these effects are strongly dependent on the geographical location (including the altitude of the site), the time of the day and period of the year, and the wavelength of the observed radiation.

1. **Absorption** of light by atmospheric constituents (e.g. mostly molecules), which diminishes the observed brightness (I_ν) of astronomical radiation and so it is at the origin of the atmospheric **opacity** at certain wavelengths.
2. Scattering of light with air molecules (e.g. Rayleigh scattering, see lecture note 3, sect. 1) and aerosols (pollution, volcanic ashes, etc.), which prevent daytime observations in the optical and cause light pollution at night.
3. Refraction, dispersion of light: deviate the apparent direction of a star in a wavelength-dependent fashion.
4. Emission of the atmosphere: perturbs IR and sub-mm observations (both day and night time). The emission from the atmosphere has two components: fluorescence (or airglow)and thermal emission.
5. Atmospheric turbulence: degrades the angular resolution of optical and IR observations, and it determined the astronomical seeing.
6. Phase instability: fluctuations of atmospheric constituents such as water vapor are the dominant source of interferometric phase fluctuations.
7. Ionization of the upper atmosphere: above 60 km the atmosphere becomes increasingly ionized because of UV photons from the Sun. This created a plasma of electrons and ions that interacts with low-frequency radio waves.

The relative contribution of different molecules to absorption depends on λ:
(sub-)mm Rotational transitions of H_2O.
IR Rotational and rotational-vibrational transitions of H_2O and CO_2 (mainly mid-IR).
Near-UV Electronic transitions of oxygen O_2 and ozone O_3, which is responsible for the complete absorption of UV radiation below $\lambda < 300$ nm.
Far-UV Absorption continuum of N_2 at $\lambda < 20$ nm.

Atmospheric absorption

The absorption of astronomical light is quantified by the optical depth τ_ν produced by the material distributed along the path traversed by the radiation before reaching the observer, eq. 8. Along the vertical line connecting the position of the telescope at the altitude of the observatory h_0, to $h = \infty$ (which in reality corresponds to the altitude of the top boundary of the absorbing medium, i.e. Earth's atmosphere).

$$\tau_\nu = \int_{h_0}^{h=\infty} \alpha_\nu(s')ds' \tag{45}$$

τ_ν is defined (and measured) for an astronomical target placed at zenith, i.e. directly above the observer and so vertical on the sky with respect to the telescope (zenith angle $\theta_z = 0^\circ$ and elevation angle $\alpha_{el} = 90^\circ$).
For sources that are not at zenith, but have a zenith angle $\theta_z > 0^\circ$, we need to account for the extra path that light has to travel through the atmosphere, and so we need to multiply τ_ν by the so-called airmass, defined as

$$Airmass = \frac{1}{\cos(\theta_z)} = sec(\theta_z) \tag{46}$$

Following the formal solution to the radiative transfer equation (eq. 9), and ignoring the spontaneous emission coefficient, the attenuated intensity of the radiation emitted by an astronomical source at zenith angle θ_z after traversing the atmospheric layer characterized by τ_ν is

$$I_\nu^{obs}(\tau_\nu) = I_\nu^{em}(\tau_\nu = 0)e^{-\tau_\nu/\cos(\theta_z)} \tag{47}$$

50% transmission, i.e. $I_\nu^{obs}/I_\nu^{em} = 0.5$, is equivalent to $\tau_\nu = 0.69$ (calculated at zenith). For $\tau = 2.3$, the atmospheric transmission is only 10%, and for $\tau = 4.6$, the transmission is 1% (the atmosphere is opaque).

Precipitable Water vapor (PWV)

Water in the troposphere exists in the form of vapor and it is responsible for the absorption of sub-mm radiation and for phase instabilities of sub-mm interferometric observations. The atmospheric quality of sub-mm observing sites is quantified by the quantity of Precipitable Water Vapor (PWV) above a given altitude h_0, measured in [mm]. This is the equivalent depth of water that would result if all the water vapor in the atmosphere were concentrated into a layer of liquid. PWV gives a good estimate of the opacity τ at sub-mm wavelengths.

$$h_{H_2O} = \frac{m_{H_2O} \cdot n_{H_2O}(h_0) \cdot H}{\rho_l} \equiv PWV \tag{48}$$

Telluric bands

When the absorption of radiation by atmospheric constituents is partial, the light will come in, but the observed data (e.g. near-IR spectra) will be affected by atmospheric features called telluric absorption bands which need to be corrected during the calibration of the data.
A common way to correct for telluric absorptions is to observe a telluric standard star, which could be either a rapidly rotating early-type star (whose advantage is to be almost featureless except for the hydrogen lines, so that the telluric features can be easily identified).

Astronomical seeing and adaptive optics

Turbulent mixing in the atmosphere causes inhomogeneities in the atmospheric density and temperature that produce variations in the index of refraction of light through air. The main effect is a degradation of the quality of astronomical images and in particular of their angular resolution (at optical and IR wavelengths), a degradation in what is commonly referred to as astronomical seeing. Because of this mixing due to atmospheric turbulence, ground-based telescopes do not usually achieve the diffraction-limited solution". This causes seeing, i.e. a random variation of the direction of the radiation entering the aperture that results in an overall decrease of the angular resolution, and scintillation, which is a random fluctuation of intensity that is at the origin of the twinkling of stellar lights (note this effect is only observed for point sources such as stars and not for extended objects such as planets).
Adaptive optics allows for corrections of seeing. This is done by introducing an equal but opposite distortion.

1. A **wavefront sensor** that measures with a high sampling rate the distortion of the incoming wavefront by observing a close star called a guide star". Either a bright real star near the target (natural guide star), or an artificial laser guide star.

2. A **control system** uses the information of the wavefront sensor to produce a deformation of the mirror for correcting the atmospheric effects.

3. A **deformable mirror** that produces the conjugated of the incoming distorted wavefront.

Extinction due to galactic interstellar medium

Before entering Earth's atmosphere, the radiation emitted by astronomical sources can be absorbed or scattered also by dust and gas in the interstellar medium (ISM) of the Milky Way (or even the ISM of other galaxies along the line of sight). The absorption and scattering due to the ISM is referred to as interstellar extinction (the colors of astronomical sources tend to become redder due to extinction, the so-called reddening). Taking this effect into consideration, an extinction term A_λ is added to the equation to eq. 33.

$$m - M = 5\log_{10}(d[pc]) - 5 + A_\lambda \tag{49}$$

A_λ is measured in magnitudes. Higher values of A_λ make sources appear much fainter, and it depends on observing bands and the position of the source on the celestial sphere with respect to the Galactic plane.

Introduction

In this section Claudia tries to include all the basic definitions, assumptions, and knowledge neede to understand the language"of sub-mm/radio observations, including the (sometimes very confusing) nomenclature and conventions.

What is a single-dish?

In the radio/(sub-)mm bands, the primary mirror (main rreflector") is called **(main) dish** or **antenna.** A radio/sub-mm telescope can be made of a **single dish**, i.e. one single antenna connected to one o rmore **receivers** (which is how we call the instruments in the radio/(sub)-mm bands), or it can be made of an array of $N \geq 2$ antennas working as an interferometer.

Design and components of a (sub-)mm/radio telescope

Telescope operations

Observations are conducted both during the day and the night, since the sky background (due to atmospheric emission) at these wavelengths is similar between day and night.

Antenna power pattern and beam

In radio/(sub-)mm astronomy, we commonly use the word **beam** as a synonymous of "angular resolution"(e.g. the point spread function (PSF)) of the telescope. However, to fully understand the concept of beam"we need to introduce some elements of antenna theory, since the beam is strictly related to the pattern of the power radiated or received by an antenna.
An **antenna** is a device that converts electromagnetic radiation in space into electric currents moving in metal conductors or vice-versa, depending on whether the antenna is used for receiving (e.g. a telescope) or transmitting (e.g. a radar). The **power pattern** of an antenna is easier to define and imagine for transmitting antenna: it is the angular distribution (i.e. as a function of a pair of angular coordinates (θ, ϕ) describing the position on the sky) of the radiation power. The same definition applies to a receiving antenna, the only difference is that for the alter we consider the pattern of the power received from the sky.

Beam area and beam solid angle

The definiton of **beam solid angle** is tied to the pattern of the *power* $P(\theta, \phi)$ received by the antenna, which described the sensitivity towards the given direction (described by a pair of spherical coordinates θ and ϕ). We introduce the normalized beam pattern, which is $P(\theta, \phi)$ normalized by the power in the pointing direction (the coordiantes (θ_0, ϕ_0) at which the power is maximum: $P_{max}(\theta_0, \phi_0)$):

$$P_n(\theta - \theta_0, \phi - \phi_0) = \frac{P(\theta - \theta_0, \phi - \phi_0)}{P_{max}(\theta_0, \phi_0)} \tag{50}$$

The power pattern has a **main beam lobe** (maximum radiation) in the pointing direction (θ_0, ϕ_0) (which usually corresponds to the direction of the astronomical target) with minor lobes (side and back) in other directions. An important parameter of the power pattern is the halfpower beam width (HPBW) and the beam width between first nulls (FNBW).
The **beam area** or beam solid angle Ω_A (it is solid angle so it is measured it steradian (sr) or deg^2) is defined as follows

$$\Omega_A = \int_{4\pi} P_n(\theta, \phi)d\Omega = \int_{\phi=0}^{\phi=2\pi}\int_{\theta=0}^{\theta=\pi} P_n(\theta, \phi)sin(\theta)d\theta\phi \tag{51}$$

The integration is extended over the full sphere (i.e. over 4π sr) such that the beam area Ω_A is a solid angle of an ideal antenna having $P_n = 1$ for all directions within Ω_A and $P_n = 0$ elsewhere, with total received power $P(\theta)\Omega_A$[Watts]. This is an idealized case, since all antennas will receiver $P_n \lesssim 1$ values for a small range of (θ, ϕ), and smaller values $P_n 1$ for a broader range of directions (side and back lobes). Shorter wavelengths tend to have larger side lobes.

Main beam solid angle, main beam efficiency η_B

For real antennas we define, in analogy with Ω_A, the **main beam solid angle** (Ω_{MB}), which is the solid angle between the first nulls, i.e. the fraction of

$$\Omega_{MB} = \int\int_{\text{main lobe}} P_n(\theta, \phi)d\Omega \tag{52}$$

hence the definition of Ω_{MB} is similar to Ω_A but the integration is limited to the main beam. Ω_{MB} contains the power pattern out to the first zero. The quality of an antenna depends on how well the power pattern is concentrated within the main beam (which is centered on the source). Quantitatively, this property is given by the **main beam efficiency** (often just beam efficiency"), indicated with η_B (or sometimes η_{MB}), which represents the fraction of the total beam solid angle Ω_A that lies inside the main beam, i.e.

$$\eta_B = \frac{\Omega_{MB}}{\Omega_A} \tag{53}$$

Note that η_B is not related to the size of the antenna: a small dish (with a wide beam) can still have a high beam efficiency.

Power gain

The directive gain $G(\theta, \phi)$ of an antenna is defined as

$$G(\theta, \phi) = \frac{4\pi P(\theta, \phi)}{\int_{4\pi} P(\theta, \phi)d\Omega} = \frac{P(\theta, \phi)}{\frac{1}{4\pi}\int_{4\pi} P(\theta, \phi)d\Omega} \tag{54}$$

hence it correponds to the power transmitted in the direction (θ, ϕ) per unit solid angle normalized by the mean power transmitted by lossless isotropic antenna (an isotropic antenna received equally from all directions of the solid angle 4π sr). This definition, as usual, is more intuitive for a transmitting antenna but it applies also to receiving antennas.
Notes:

1. The antenna gain is usually measured in decibels, using the logarithmic scale

$$G(dBi) = 10log_{10}(G) \tag{55}$$

2. In the direction (θ_0, ϕ_0) of maximum power $P_{max}(\theta_0, \phi_0)$, the **maximum gain** is given by

$$G(\theta, \phi) = \frac{P(\theta, \phi)}{\frac{1}{4\pi}\int_{4\pi} P(\theta, \phi)d\Omega} \tag{56}$$

$$= \frac{4\pi}{\frac{1}{4\pi}\int_{4\pi} P(\theta, \phi)d\Omega} \equiv \frac{4\pi}{\Omega_A} \tag{57}$$

where in the last sted we applied the definiton of beam area Ω_A. So higher gain corresponds to a smaller Ω_A.

Angular resolution of a single-dish observations

The angular extent of the main beam (or simply the beam size) is described by the **full width to half power (FWHP)** or half power beam width (HPBW), which is the angle (measured in deg/arcmin/arcsec) between the two points of the main beam for which $P_n = 0.5$. The FWHP (or HPBW) of the main beam can be identified with the **angular resolution** of a single-dish telescope. Another definition that can be sued to characterize the main beam size is the **beam widt between first nulls (BWFN)** (also known as first nulls beam width (FNBW)), which is the angle separating the first two nulls (points where $P_n = 0$). Follwing form the Fraunhofer diffraction equation (valid in far-field approximation, which is always satisfied at the focal plane of a telescope), the power pattern corresponds to the Fourier transform of the aperture function. Therefore, in analogy with the diffraction-limited Point Spread Function (PSF) of circular aperture, the intensity (or power, hence $\propto E^2$) pattern of a uniformly illuminated circular antenna is the well-known Airy disk. In this case, the main beam solid angle Ω_{MB} corresponds to the central component of the Airy disk. However, an Airy disk power pattern is an idealized case, since real antennas are not univormly illuminated. For real antennas, the aperture function is replaced by the **aperture field illumination**, which describes the illumination of the antenna. In real telescopes, where the illumination is **tapered** and so not uniform, the beam size needs to be calculated empirically, becuase it depends not only on the geometric shape of the antenna and wavelength observed, but also on the illumination of the antenna.

Position accuracy for a high S/N point source

Even though the FWHP of the main beam represents the angular resolution of a single-dish observation, meaning that structures with details $\Delta\theta < FWHM_{MB}$ cannot be distinguished, it is possible, for high signal-to-noise (S/N) isolated astronomical sources, to determine their central peak position with a precision $\Delta\theta_{peak}$ that is slightly better than the FWHP. For a high S/N source the **position accuracy** is then given by:

$$\Delta\theta_{peak} = \frac{FWHP}{S/N} \tag{58}$$

Effective area and aperture efficiency η_A

The effective collecting area or effective aperture of a receiving antenna is indicated with A_e and it describes the "cross-section"with respect to the radiation received from the source. The A_e (measured [m^2]) is defined as the ratio between the power collected by the antenna (P_e) and the total power carried by the radiation received from the source (i.e. the modulus of the Poynting vector, $|\langle \mathbf{S} \rangle|$):

$$A_e = \frac{P_e}{|\langle \mathbf{S} \rangle|} \tag{59}$$

The effective aperture is related to the geometric aperture of an antenna A_g (equal to $\pi D^2/4$ for a circular antenna with diameter D) through the so-called **aperture efficiency** (η_A):

$$A_e = \eta_A A_g \rightarrow \eta_A = \frac{A_e}{A_g} \tag{60}$$

For an ideal antenna with uniform illumination and no surface errors, A_e is simply equal to the geometrical aperture. However, in real antennas the illumination is *tapered* (= not uniform and reduced at the edges), so in general $A_e < A_g$.
Étendue of coherence
The maximum effective area A_e and the beam area are related to each other by this relation

$$A_e\Omega_A = \lambda^2 \tag{61}$$

The quantity λ^2 is the étendue of coherence of quasi-monochromatic radiation of wavelength λ.

Antenna temperature

We define the antenna temperature as the temperature T_A of an equivalent thermal source for which the power is equal to such power W_ν, so:

$$W_\nu = kT_A \tag{62}$$

The relation is the Nyquist theorem. W_ν is the power per unit bandwidth (the dependency on ν is included in $T_A = T_a(\nu)$ received at terminals of a resistor with temperature T_A. We can derive an expression for the antenna temperature as a function of brightness temperature:

$$T_A(\nu) = \frac{1}{\Omega_A}\int T_B(\nu)P_n(\theta, \phi)d\Omega \tag{63}$$

This relation tells us that the antenna temperature is equal to the brightness temperature of the source convolved with the normalized beam pattern of the antenna. The antenna temperature is the actual output of the single-dish observation, from which we have to retrieve T_B.
Point source $\Omega_S \ll \Omega_A$: In the case of a point unresolved source, i.e. a source that subtends a solid angle Ω_S that is much smaller than the beam solid angle $Omega_A$, some of the beam sees the cold sky that lies in the background of the astronomical source (this effect is called beam dilution)

$$T_A(\nu) \simeq \frac{\Omega_S}{\Omega_A}T_B(\nu) \tag{64}$$

Beam filling source $\Omega_S\Omega_A$: If the source fills the beam area (or it is bigger than that) then the relation becomes:

$$T_A(\nu) \simeq T_B(\nu) \tag{65}$$

Hence in this case the antenna temperature is a good approximation of the brightness temperature of the source. Note that this definition assumes that the source fills the beam area. This is equivalent to assuming that the telescope resolves the source, i.e. the angular size of the source exceeds or equals the beam of the telescope.

Antenna temperature corrected for atmospheric losses T_A^*

The antenna temperature measurement is corrected for blockage, ohmic losses, or absorption in the Earth's atmosphere: In this case, the data delivered to the observer are indicated with a different notation: T_A^* or $T_A^{'}$, to indicate the antenna temperature measurements have been corrected for atmospheric losses.

Relation between T_A and F_ν for a point souce

The Γ is factor called **point-source sensitivity of the telescope** or, more often, simply **K/Jy calibration factor**, and it is generally defined directly in terms of T_A^*, hence

$$T_A^* \equiv \Gamma F_\nu \rightarrow F_\nu = \Gamma^{-1} T_A^* \qquad (66)$$

All quantities depend on the observed frequency. In general, the Γ factor is provided in the telescope documentation.

The components of a superheterodyne receiver

The main type of receivers (=instruments) used in submillimeter and radio telescopes, especially when performing spectroscopy is called **super-heterodyne receivers**. HET receivers are **coherent** detectors, which means they can measure both the phase and polarization of the incoming EM wave. The main feature of HET receivers is that they operate by **mixing** the electric field of the source photons with that of a local source. By adding signals of different frequencies, the result is a beat pattern, which contains frequencies from the two original signals, but whose amplitude is modulated at the beat frequency.

1. **Antenna** the first stage is to convert the electric field of the incident EM wave, into an electric signal.

2. **RF amplifier** A radio frequency power amplifier (RF power amplifier) is a type of electronic amplifier that converts the input low-power radio-frequency signal (within a certain bandpass $\delta\nu$ which must be as narrow as possible and centered on the frequency of the astronomical signal) into a higher power signal. This must be high gain and low noise.

3. **Mixer** The mixer is a device that converts the RF signal into a lower frequency signal. The mixer accepts two input signals: the actual signal from the amplifier 20 (at the observed frequency ν_{RF},) and a sine wave signal produced by a local oscillator (LO) at a frequency ν_{LO}. The mixer adds these two signals and produces an output beat frequency signal11, at an intermediate frequency (IF) given by the difference between the two input frequencies: $nu_{IF} = nu_{LO}\nu_{RF}$.

4. **IF amplifier** The IF signal is then further amplified.

5. **Square Law Detector (SLD)** The IF signal is then brought to a detector stage. This stage converts the signal to a format that can be easily analyzed by our computers, producing a measurable output.

The spectral resolution of superheterodyne receiver

In sub-mm/radio spectroscopy, we call "channelthe smallest frequency element of a spectrum. The channel can be measured in units of frequency or velocity with respect to the center of the line, or even in units of wavelength (rare). In heterodyne receivers the minimum channel width corresponds also to the **spectral resolution** of the instrument (the minimum distance in the frequency of the two independent measurements). During the data analysis, the data may be **re-binned** in frequency channels that are bigger than the spectral resolution.

Noise of single-dish observations

By signal in astronomy, we refer to the translation of the information emitted by the source into something measurable by our telescopes, data that can then be processed by our computers to extract astrophysical information. The main goal of observational astronomy is to beat the noise and extract weak signals of astrophysical relevance from our observations. The noise in sub-mm/radio observations is described using the concept of equivalent temperature: given a power per unit bandwidth W_ν produced by *Johnson noise* in a resistor, we can define a **noise temperature** T_N as follows

$$W_\nu = kT_N \qquad (67)$$

Receiver noise and T_R

The power received by radio/(sub-mm) telescope is very small, and to detect it we must amplify it. With this large amplification, the internal noise of the system will also be amplified. This internal noise is called receiver noise and has the same statistical properties as the signal from the source on the sky: it is a continuum in nature and generated by random processes. It is convenient to measure it with the same units as the radiation, i.e. the temperature. So we specify the receiver temperature, TR. Just as for any thermal noise source, the noise power is:

$$W[\mathrm{W}] = kT_R \Delta\nu \qquad (68)$$

System temperature T_{sys}

The telescope collects the incoming power from the sky, specified by the antenna temperature T_A (or, even better, the antenna temperature corrected for atmospheric losses T_A^*, which combines linearly with the power generated by the receivers and their sum is called system temperature

$$T_{sys} = T_A^* + T_R + T_B^{CMB} + T_{at} + T_{wv} + T_{g'} \qquad (69)$$

where:
T_A^* Source antenna temperature corrected for atmospheric losses, represents the signal of astrophysical interest.
T_R Receiver noise corresponds to the electronic noise generated by the amplifiers in the receivers.
T_B^{CMB} Background signal of astrophysical origin mostly due to the Cosmic Microwave Background (CMB), originating from every direction in the sky.
T_{at} Radiation from the dry atmosphere.
T_{wv} Radiation from water vapor in the atmosphere.
T_g Radiation from the ground in the beam side lobes, depends also on the source elevation.

The radiometer equation: how the S/N depends on Δt and $\Delta\nu$

Introduction: detectors

A detector (or receiver) in (sub-)mm/radio/sub-mm astronomy) is a device that converts the incident energy transported by EM waves into data that can be measured and analyzed by astrophysicists. A general expression for the output of an instrument (e.g. a detector, or a detector + telescope system) at a given time is given by Lena

$$x = x_b + f\left[\int_{\Delta\nu} \Phi(\nu) d\nu \int_{\Delta\Omega} I_\nu(\Omega) P_n(\Omega) d\Omega\right] \qquad (70)$$

where $I_\nu(\Omega)$ is the specific intensity(or brightness, remember it depends on the direction i.e. the solid angle) reaching the system, $P_n(\Omega)$ is the angular response of the system (e.g. the normalized power pattern in the case of a (sub-)mm/radio telescope, also known as point spread function in single-aperture optical/IR telescopes), and $\Phi(\nu)$ is the spectral response of the system. x_b is the background noise delivered by the detector in the absence of an incident astronomical signal. Detectors are classified into the following categories:

1. **Photon detectors:** Incident photons release electrons during the interaction with the detector material. Charged-couple devices (CCDs) are the most used detectors of this type.

2. **Thermal detectors:** The energy of the incoming photons is converted into heat that modifies some of the properties of the detector. These types of detectors are intrinsically sensitive to a broad range of frequencies (broadband).

3. **Wave detectors:** Measure the electric field of the incoming EM wave (linear or coherent detectors). Generally used in radio astronomy. Contrary to photon and thermal detectors (which can only measure the signal's amplitude), wave detectors are coherent detectors, since they can measure the polarization and the wave's phase. Super-heterodyne receivers are wave detectors.

Charged-couple devices (CCD)

CCDs belong to the family of photon detectors. They were introduced in the 1970s, and have revolutionized this field, with their linear response, and the possibility to image multiple pixels simultaneously in a single frame. CCDs exploit the photoconductive effect, which is a manifestation of the photoelectric effect in semiconducting materials.

The photoconductive effect in semiconductors

Photoconducting effect (in solid material): This is a form of the photoelectric effect in solid material. In this case, the photon energy is absorbed by a bounded electron in the valence band, the energy is not sufficient to eject the electron, but it lifts the electron to the conduction band hence the charge is released internally and became a free carrier in the crystal, modifying its electrical conductivity.
Semiconductors are characterized by their band gap, which is the energy difference between the conduction band and the valence band, E_g. Conductors have a zero gap, meaning that electrons are always available to transfer charge. Insulators have very large gaps, implying zero conduction except under extreme circumstances. Semiconductors have intermediate gaps.

Metal-oxide-silicon (MOS) diodes: the pixels"of a CCD

A CCD is a 2-dimensional array of closely spaced metal-oxide-silicon (MOS) diodes, which are a special category of metal-insulator-semiconductor diodes. A MOS stores information as follows, when it is exposed to radiation, photons are absorbed in or near the depletion zone. Some CCDs are front-illuminated, while most CCDs employed in astronomy are typically made thinner and back-illuminated. In a back-illuminated MOS, the photons enter through the p-type Silicon.

The workings of a CCD: analogy with an array of buckets

Once the charges are stored in the depletion zone, they can be transferred by simply changing the voltage in a controlled sequence across neighboring pixels in the CCD. This is the main characteristic of a CCD: a detector that, when exposed to radiation, generates photo charges, which can be stored and moved across the semiconductor substrate in a controlled way. After the exposure, the electrons are shifted in a series of charge-coupled steps across the CD surface, then amplified, and finally read out of the CCD, and stored in computer memory. This is a destructive readout meaning that the signal can only be read once. Hence the steps are:
(1) conversion of individual photons to electric charges during the illumination
(2) storage of the charges over the desired exposure time
(3) transfer of the stored charges from pixel to pixel for the read-out procedure
(4) accurate read-out of the accumulated charge for each pixel element
Dark current Electrons in the conduction band of n-type Silicon will produce a signal even in the absence of incident radiation. This is the primary source of background noise in a CCD and is called dark current.

Sensitivity, saturation, and dynamic range

The sensitivity threshold of the detection system is the minimum signal that can be detected, i.e. the I_ν^{min} value for which signals $I_\nu < I_\nu^{min}$ are indistinguishable from the background noise.
The saturation level of the system is a maximum brightest I_ν^{sat} such as for $I_\nu > I_{nu}^{sat}$ the function $f(I_\nu)$ remains constant, hence the output signal from the detector stops growing with incident I_ν.
Affecting the sensitivity is also the domain of linearity of the detector, i.e. the range in signal I_ν over which the response function f is linear.
Another property that describes the sensitivity of a detector is its dynamic range, which is the ratio of maximum to minimum measurable signal, as measured in a single exposure.

Quantum efficiency

The performance of a detector is measured through its quantum efficiency(QE), defined as the number of detected photons divided by the number of incident photons

$$\eta = QE = \frac{N_{detected}}{N_{incident}} \qquad (71)$$

An alternative definition is the detective quantum efficiency (DQE) which extends the concept of QE to include the noise introduced by the detectors. The DQE is defined as the ratio between the squared output signal-to-noise ratio (S/N) to the squared input S/N

$$DQE = \frac{(S/N_{out})^2}{(S/N_{in})^2} \qquad (72)$$

Pixel size and sampling of the PSF

The pixel is the minimum measurable area of the surface of a detector. Pixels are not necessarily independent of one another. Actually, the flux measured by adjacent pixels cannot be independent, since it is important that the point spread function (PSF) of the instrument is properly sampled, at least ensuring Nyquist sampling. The Nyquist criterion implies that one pixel should span about half of the FWHM of the PSF, hence the FWHM of the PSF needs to be sampled by at least two pixels.

Noise

In reality, at any given t, the signal x given by eq. 70 is just one realization of a random or stochastic process. The fluctuations of x around its real value are responsible for the so-called noise. These are due to the quantum and discrete nature of the EM radiation, for the physics of the detector process itself.

Photon noise: quantum and thermal limits

These are irreducible fundamental fluctuations due to the interactions of light with matter, whatever the nature of the detector is.

$$\langle \Delta W_\nu^2 \rangle = \langle W_\nu \rangle h\nu \left[1 + \frac{1}{(e^{h\nu/kT} - 1)}\right] \qquad (73)$$

There are two limiting cases for the fluctuations of power (photon noise) described by eq.73: the quantum and thermal noise limits. Generally:

1. The quantum limit holds in the visible and UV regions, or at higher frequencies (X-ray, γ-rays)

2. Both quantum and thermal noise limits are relevant in the IR/sub-mm regime

3. Thermal noise dominates in the millimeter and radio bands, even if the system is cooled down using cryogenics.

Quantum noise limit (shot noise)

For $h\nu \gg kT$ are independent and obey Poisson statistics, with fluctuations proportional to the square root of the number of photons:

$$h\nu \gg kT \rightarrow \langle \Delta W_\nu^2 \rangle \approx h\nu \langle W_\nu \rangle \qquad (74)$$

$$\langle \Delta N^2 \rangle = \langle N \rangle \qquad (75)$$

In this regime $h\nu \gg kT$, fluctuations are referred to as quantum noise, or shot noise, and correspond to the quantum limit. Quantum noise dominates at high frequencies and low temperatures, and in general, frequencies corresponding to the visible/optical band and higher.

Thermal noise

In the limit $h\nu \ll kT$ we are in the thermal noise regime.

$$h\nu \ll kT \rightarrow \langle \Delta W_\nu^2 \rangle kT \qquad (76)$$

The thermal noise generated by a resistor at temperature T (also known as Johnson noise) is given by the Nyquist theorem:

$$W_\nu = kT \qquad (77)$$

Thermal noise dominates at high temperatures across the entire EM spectrum, and at any temperature in the low-frequency regime of radio/mm observations.

Background noise

An observing system received simultaneously both signal from the astronomical source and a background (x_b in eq. 70). The background signal could have different origins and physical characteristics:

1. **Dark current** This is the photocurrent produced by a solid-state detector (such as a photomultiplier of the pixels of a CCD) when it is not exposed to any direct radiation source.

2. **Cosmic rays** Despite the name, cosmic rays are actually high-energy particles (e.g. protons, ions) emitted from astronomical sources (e.g. the Sun, supernovae explosions, AGNs) moving close to the speed of light. The impact of the photoconduction material of a CCD creates a large number of charges by photoionization, creating a background signal that will affect some of the pixels of the CCD.

3. **Thermal background emission** Radiation hits the receiver even when it is not exposed to any astrophysical source. Such radiation can be due to thermal emission from the Earth's atmosphere, or from telescope optics.

Noise equivalent power (NEP)

The sensitivity of a detector due to its background noise is often defined in terms of its Noise Equivalent Power (NEP). The NEP of a detector is the signal power that gives a signal-to-noise ratio of 1 for a 1 Hz bandwidth (i.e. a second of integration). In an ideal system (i.e. in absence of additional noise sources) the NEP is due to the statistical fluctuations of the background radiation.

$$NEP = \sqrt{\frac{2\Delta\nu W_b h\nu}{\eta}} \qquad (78)$$

In practice, the NEP of an ideal detector will have an additional multiplicative factor of $\sqrt{2}$ with respect to the expression given in eq. 78, given by the so-called **generation-recombination (G-R) noise**, which is caused by fluctuations in the rate of generation and recombination of the electrons in the semiconductor. Hence:

$$NEP[W] = 2\sqrt{\frac{\Delta\nu W_b h\nu}{\eta}} \qquad (79)$$

Though it is common use to give the NEP in units of $[W/\sqrt{Hz}]$, hence removing the dependency on $\sqrt{\Delta\nu}$

$$NEP[W/\Delta\nu] = 2\sqrt{\frac{W_b h\nu}{\eta}} \qquad (80)$$

Additional sources of detector noise

1. **Amplifier noise** Amplifiers in the detector and other devices used to convert the signal from the detector to a computer-readable form introduce an additional source of the noise.

2. **Readout noise** Specific property of CCDs. The readout noise occurs in the measurement to f the voltage induced by the charges. Each charge transfer has its own fluctuations, characterized by standard deviation σ_R, given in RMS electron number. This type of noise does not depend on the exposure time, it is present in all images, the same amount regardless of exposure time.

3. **Flicker (or 1/f) noise** Occurs when the signal is modulated in time, either because of intrinsic variations or because of chopping techniques (e.g. source and background are alternately observed.

Astrophysical spectra and line parameters

The astrophysical spectrum shows the dependency of the specific intensity $I_\nu(\theta, \phi)$ of an astrophysical source (where θ, ϕ are the angular coordinates) on the frequency ν (or wavelength λ of the radiation. An instrument that can perform both imaging and spectroscopy at the same time is used: both the spectroscopic and the spatial information allows us to distinguish different sources in our data. Spectroscopy allows studying both the continuum of emission from sources and the emission and absorption lines. The continuum component is characterized by a specific intensity I_ν that varies slowly with ν, while the line components show rapid variations of I_ν over a narrow spectral range (i.e. over a small $\Delta\nu$ or $\Delta\lambda$).

Optical galaxy spectra

The optical and ultraviolet (UV) observing bands are populated by electronic transitions of atoms and ions, produced in the ionized ISM of galaxies. Some of the transitions are reported in square brackets: there are the so-called forbidden lines. These lines have a very low associated probability to be deexcited spontaneously. A characteristic feature of optical and UV spectra of galaxies with star-forming gas are the hydrogen recombination lines.
Spectrum of a star-forming galaxy Presence of several emission lines, tracing the emission from gas in the ISM. The stellar continuum is quite flat, and it shows a few absorption lines due to both stars and interstellar gas.
Spectrum of a passive (=with very little or no ongoing star formation) galaxy Emission lines due to the ISM are not visible because there is very little gas left, and the spectrum is dominated by the stellar continuum and stellar absorption lines.

Spectral solution

The ability of a spectrograph to separate the specific intensity I_λ or the flux density F_λ measured at neighboring wavelengths (or frequencies, for a spectrum showing I_ν vs ν) is not infinite, but it is given by the spectral resolving power of the instrument. The spectral resolving power is defined as the ratio:

$$R = \frac{\lambda}{\Delta\lambda_{instr}} \qquad (81)$$

where λ is the observed wavelength, and $\Delta\lambda_{instr}$ is the instrumental spectral line width or spectral resolution, i.e. the minimum separation between two spectral lines that can be resolved. Equivalent equation:

$$R = \frac{\nu}{\Delta\nu_{instr}} \qquad (82)$$

From the Doppler formula, we can rewrite eq. 82 becomes

$$R = \frac{\nu}{\Delta\nu_{instr}} \equiv \frac{c}{\Delta\nu_{instr}} \rightarrow \Delta\nu_{instr} = \frac{c}{R} \qquad (83)$$

where c is the speed of light ($c \sim 300,000 kms^{-1}$). Hence from eq. it becomes immediate to estimate that resolving power of R = 1000 corresponds to a spectral resolution of $\Delta\nu_{instr} \sim 300 kms^{-1}$.

Spectroscopic techniques

As for any astrophysical detection system, the information contained in the astrophysical spectra is strictly dependent on the angular portion of the celestial sphere that is probed by the spectrograph. The output of a spectrograph is generally described by the expression eq. 70

$$x = x_b + f\left[\int_{\Delta\nu} \Phi(\nu) d\nu \int_{\Delta\Omega} I_\nu(\Omega) P_n(\Omega) d\Omega\right] \qquad (84)$$

where $\Phi(\nu)$ describes the spectral response of the system, over the observed spectral bandwidth $\Delta\nu$. The output is always convolved also with the angular response $P_n(\Omega)$ of the system.

Dispersive spectroscopy

The basic elements of a dispersive spectrograph are summarised below:
Slit or aperture It is a mask with a narrow rectangular aperture that isolates the portion of the field of view whose radiation is sent to the spectrograph. The goals of a slit are: (I) to isolate the source(s) for spectroscopy, so as to avoid overlap of the spectra in the resulting image, and (ii) to provide a stable spectral resolution. A narrower slit gives higher resolution, but more light loss.
Collimator Makes a beam of light coming from a slit parallel that the grating is illuminated by a parallel beam of light (hence avoiding the effect of having variable angles of incidence as a function of positions on the grating, which can decrease the spatial resolution of the resulting spectrum).
Disperive element Disperses the radiation in wavelength.
Camera A goal of collecting the spectrally-dispersed beams from a dispersive element, which are still collimated, and make them converge to that the spectrum is imaged onto the detector.
Detector Images the dispersed light from the spectrograph, forming a 2D spectrum image.
Some dispersive elements are: prisms, diffraction gratings, and grisms.

Non-dispersive spectroscopy

Fourier transform spectrometers (FTS) and Fabry-Perot interferometers (FPI) (or étalon) perform non-dispersive spectroscopy.

Deduction of radiative transfer transport equation

For a ray passing through matter, the variation of I_ν, through distance ds is

$$\frac{dI_\nu}{ds} = -\alpha_\nu I_\nu + j_\nu$$

where $ds = \frac{d\tau_\nu}{\alpha_\nu}$. From the definition of optical depth (eq. 8, $d\tau_\nu = \alpha_\nu ds$, one can rewrite the radiative transfer transport equation

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + \underbrace{\frac{j_\nu}{\alpha_\nu}}_{S_\nu}$$

The source function: $S_\nu \equiv j_\nu/\alpha_\nu$, has the same units as I_ν and provides a formal description for the additional contribution to the brightness due to the new emission from the material that the beam of light encounters along the path ds (or the optical path $d\tau_\nu$). Hence, it describes this new source"of emission.

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu \Big| \cdot e^{\tau_\nu}$$

$$\frac{dI_\nu}{d\tau_\nu} e^{\tau_\nu} = -\underbrace{I_\nu e^{\tau_\nu}}_{\mathrm{J}_\nu} + \underbrace{S\nu e^{\tau_\nu}}_{\mathrm{S}_\nu} \tag{*}$$

We introduce temporarily definitions, J_ν and S_ν (not to be confused with j_ν and S_ν, to simplify calculations.

$$\frac{d\mathrm{J}_\nu}{\tau_\nu} = \frac{dI_\nu}{d\tau_\nu} e^{\tau_\nu} + I_\nu e^{\tau_\nu} \tag{**}$$

Replace (*) into (**)

$$\frac{d\mathrm{J}_\nu}{\tau_\nu} = -\mathrm{J}_\nu + \mathrm{S}_\nu + \underbrace{I_\nu e^{\tau_\nu}}_{\mathrm{J}_\nu} = \mathrm{S}_\nu \Rightarrow \frac{d\mathrm{J}_\nu}{\tau_\nu} = \mathrm{S}_\nu$$

Solutions

$$\mathrm{J}_\nu(\tau_\nu) = \mathrm{J}_\nu(\tau_\nu = 0) + \int_0^{\tau_n u} \mathrm{S}_\nu(\tau_\nu^{'}) d\tau_\nu^{'}$$

where τ is a parameter and τ' is the integration parameter.
Insert the original expressions for J_ν and S_ν

$$I_\nu(\tau_\nu^{'}) e^{\tau_\nu} = I_\nu(0) e^0 + \int_0^{\tau_\nu} S\nu(\tau_\nu^{'}) e^{\tau_\nu^{'}} d\tau_\nu^{'} \Big| \cdot \frac{1}{e^{\tau_\nu}} = e^{-\tau_\nu}$$

$$I_\nu(\tau_\nu^{'}) = I_\nu(0) e^{-\tau_\nu} + e^{-\tau_\nu} \int_0^{\tau_\nu} S\nu(\tau_\nu^{'}) e^{\tau_\nu^{'}} d\tau_\nu^{'}$$

$$I_\nu(\tau_\nu^{'}) = I_\nu(0) e^{-\tau_\nu} + \int_0^{\tau_\nu} S\nu(\tau_\nu^{'}) e^{-(\tau_\nu - \tau_\nu^{'})} d\tau_\nu^{'}$$

For a homogenous medium where S_ν does not depend on τ_ν

$$I_\nu(\tau_\nu^{'}) = I_\nu(0) e^{-\tau_\nu} + S_\nu e^{-\tau_\nu} \underbrace{\int_0^{\tau_\nu} e^{\tau_\nu^{'}} d\tau_\nu^{'}}_{[e^{\tau_\nu^{'}}]_0^{\tau_\nu} = e^{\tau_\nu} - 1}$$

$$I_\nu(\tau_\nu^{'}) = I_\nu(0) e^{-\tau_\nu} + (S_\nu - e^{-\tau_\nu} S_\nu)$$

$\tau_\nu \gg 1 \Rightarrow_{absorption} I_\nu(\tau_\nu) \to S_\nu$
$\tau_\nu \ll 1 \Rightarrow_{emission} I_\nu(\tau_\nu) \to I_\nu(0)$

$$I_\nu(\tau_\nu) = S_\nu + e^{-\tau_\nu}(I_\nu(0) - S_\nu)$$

Planck function, $B_\nu(T) \to B_\lambda(T)$

Planck function as a function of wavelength $B_\lambda(T)$ can be derived from Planck's law as a function of frequency $B_\nu(T)$

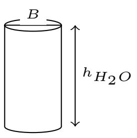
$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1}$$

From the expression for reduced brightness eq. 2, we obtain $B_\lambda(T) = \frac{\nu}{\lambda} B_\nu(T)$

$$B_\nu(T) = \frac{2hc^3}{c^2 \lambda^3} \frac{c}{\lambda} \frac{1}{\lambda} \frac{1}{e^{hc/\lambda kT} - 1}$$

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1}$$

Definition of Precipitable Water Vapor (PWV)



Imagine a column (cylinder) of liquid water with base B and height h_{H_2O}. The mass of this column of liquid water is:

$$m_{liquid} = \rho_l \cdot (B h_{H_2O}) \tag{1}$$

where ρ_l is the volume density of liquid water ($\rho_l = 1gcm^{-3}$). Now we want to calculate h_{H_2O} as a function of quantities that can be measured. Hence we need to find an expression for the mass of water vapor in the atmosphere that we can set equal to the equivalent mass of liquid in water. Consider a column of water vapor, with the same base B. The mass of water vapor can be calculated as

$$m_{vapor} = B \cdot m_{H_2O} \int_{h_0}^{\infty} n_{H_2O}(h) dh \tag{2}$$

where n_{H_2O} is number density of H_2O particles[particles/cm^3]. Volume density ρ_l is replaced with the product of m_{H_2O} [g/particle] and the number density $n_{H_2O}(h_0)$ [particles/cm^3]. Assume number density distribution follows an exponential law

$$n_{H_2O}(h) = n_{H_2O}(h_0) \cdot e^{-(h-h_0)/H} \tag{3}$$

The value of H ranges between 1 and 3 km.
Inserting eq. 3 into eq. 2 and solving the integral

$$m_{vapor} = B \cdot m_{H_2O} \int_{h_0}^{\infty} n_{H_2O}(h_0) \cdot e^{-(h-h_0)/H} dh$$

$$m_{vapor} = B \cdot m_{H_2O} \cdot n_{H_2O}(h_0) \cdot e^{h_0/H} \underbrace{\int_{h_0}^{\infty} e^{-h/H} dh}_{[-H \cdot e^{-h/H}]_{h_0}^{\infty}}$$

$$m_{vapor} = B \cdot m_{H_2O} \cdot n_{H_2O}(h_0) H \tag{4}$$

By imposing $m_{water} = m_{vapor}$, hence

$$h_{H_2O} = \frac{m_{H_2O} \cdot n_{H_2O}(h_0) \cdot H}{\rho_l} \equiv PWV \tag{5}$$

Deduction of the definition of noise-equivalent-power (NEP)

We derive an expression for the NEP in the case of a background-limited IR photodetector. Given a background W_b, the photon rate will be $W_b/(h\nu)$ [photons s^{-s}], and so the number of detected background photons in a time Δt is:

$$N_b = \eta \frac{W_b}{h\nu} \Delta t \tag{a}$$

where η is the quantum efficiency of the photoconductor. Following Poisson's statistics, the uncertainty (standard deviation) on N_b is $\sqrt{N_b}$, hence:

$$\sigma_{N_b} = \sqrt{N_b} = \sqrt{\eta \frac{W_b}{h\nu} \Delta t} \tag{b}$$

Recall the definition of NEP, which is the $1-\sigma$ standard deviation of the background power (in 1Hz bandwidth or per second). Hence we replace $W_b \to$ NEP, then we can replace $N_b \to \sigma_{N_b}$ in eq. a, and we get

$$\sigma_{N_b} = \eta \frac{\mathrm{NEP}}{h\nu} \Delta t \tag{c}$$

By equating eq. b and c we get

$$\eta \frac{\mathrm{NEP}}{h\nu} \Delta t = \sqrt{N_b} = \sqrt{\eta \frac{W_b}{h\nu} \Delta t} \to \mathrm{NEP} = \frac{W_b h\nu}{\eta \Delta t} \tag{d}$$

In this last expression, we can rewrite the observing time Δt as a function of the bandwidth $\Delta\nu$ by using the critical Nyquist sampling rate:

$$\Delta\nu = \frac{1}{2\Delta t} \to \Delta t = \frac{1}{2\Delta\nu} \tag{e}$$

Hence the last expression in eq. d becomes

$$NEP[W/\Delta\nu] = 2\sqrt{\frac{W_b h\nu}{\eta}}$$

The diffraction problem

The diffraction problem can be stated as follows: we need to derive the electric field of a radiaion beam after passing thorugh an aperture, at a distance z from such aperture. Let (x_1, y_1) be the coordinates of a point in the plane of aperture, and (x_0, y_0) of the corresponding diffraction point in the observation plane. Hence the problem is to derive $E(x_0, y_0)$ given $E(x_1, y_1)$.
The key is to apply the Huygen's principle: each point (x_1, y_1) in the aperture emits a spherical secondary wave, whose amplitude is determined by the incident primary wave. The net field is therefore given by the sum computed for all points (x_1, y_1) in the aperture of the spherical secondary waves propagating at a distance r_{01} multiplied by the indicent field at position (x_1, y_1)

$$E(x_0, y_0) = \sum_{\text{all points}(x_1, y_1)} (\text{secondary wave at } r_{01}) \times (\text{incident field at } (x_1, y_1))$$

The sum becomes an integral

$$E(x_0, y_0) = \int\int_{\text{Aperture } (x_1, y_1)} h(x_0 - x_1, y_0 - y_1) E(x_1, y_1) dx_1 dy_1 \tag{+}$$

where

$$h(x_0 - x_1, y_0 - y_1) = \frac{1}{i\lambda} \frac{e^{ikr_{01}}}{r_{01}}$$

with $k = \frac{2\pi}{\lambda}$ is the wave number, and $r_{01} = \sqrt{z^2 + (x_0 - x_1)^2 + (y_0 - y_1)^2}$. This is a complicated expression and for it to be useful we need to make some approximations:

1. **Paraxial approximation** Assume the aperture is much smaller than the distance z, i.e. $z \gg D$ (where D is the maximum distance between two points in the aperture, i.e. $D^2 = max(x_1^2 + y_1^2)$.
   We note that by factoring out z from the square root in the expression of r_{01} we have

   $$r_{01} = z\sqrt{1 + \left(\frac{x_0 - x_1}{z}\right)^2 + \left(\frac{y_0 - y_1}{z}\right)^2}$$

   in which we can use Taylor expansions $\sqrt{1+x} \approx 1 + \frac{1}{2} + \ldots (x \ll 1)$ to rewrite the term

   $$u = \left(\frac{x_0 - x_1}{z}\right)^2 + \left(\frac{y_0 - y_1}{z}\right)^2$$

   $$\begin{aligned} z\left[\sqrt{1+u}\right] \approx z\left[1 + \frac{1}{2}u = 1 + \frac{1}{2}\left(\frac{x_0 - x_1}{z}\right)^2 + \left(\frac{y_0 - y_1}{z}\right)^2\right] \\ = z\left[1 + \frac{(x_0 - x_1)^2}{2z^2} + \frac{(y_0 - y_1)^2}{2z^2}\right] \\ = z + \frac{(x_0 - x_1)^2}{2z} + \frac{(y_0 - y_1)^2}{2z}] \end{aligned}$$

   Now we wish to rewrite h in terms of the paraxial approximation applied to r_{01}

   $$h(x_0 - x_1, y_0 - y_1) = \frac{1}{i\lambda} \frac{e^{ikr_{01}}}{r_{01}} \approx \frac{1}{i\lambda} \frac{e^{ik\left[z + \frac{(x_0 - x_1)^2}{2z} + \frac{(y_0 - y_1)^2}{2z}\right]}}{z + \frac{(x_0 - x_1)^2}{2z} + \frac{(y_0 - y_1)^2}{2z}}$$

As z is very large, i.e. $z \gg$ we can neglect some terms

$$h(x_0 - x_1, y_0 - y_1) \approx \frac{1}{i\lambda z} e^{ik\left[z + \frac{(x_0-x_1)^2}{2z} + \frac{(y_0-y_1)^2}{2z}\right]}$$

Insert back in to eq. (+)

$$E(x_0, y_0) = \int\int_{\text{Aperture } (x_1,y_1)} \frac{1}{i\lambda z} e^{ik\left[z + \frac{(x_0-x_1)^2}{2z} + \frac{(y_0-y_1)^2}{2z}\right]} E(x_1, y_1) dx_1 dy_1$$

From here we assume $E(x_1, y_1) = constant$ meaning there is a constant incoming planar wave. We will also further simplify the exponent in the expression above

$$= z + \frac{(x_0 - x_1)^2}{2z} + \frac{(y_0 - y_1)^2}{2z} = z + \frac{(x_0^2 + y_0^2)}{2z} + \frac{(x_1^2 + y_1^2)}{2z} + \frac{(-2x_0 x_1 - 2y_0 y_1)}{2z}$$

$$E(x_0, y_0) = E_{in} e^{\left[\frac{(x_0-x_1)^2}{2z}\right]} \int\int_{\text{Aperture } (x_1,y_1)} e^{ik\left[\frac{(x_1^2+y_1^2)}{2z} + \frac{(-2x_0x_1-2y_0y_1)}{2z}\right]} dx_1 dy_1$$

From this we obtain the **Fresnel diffraction equation**

$$\boxed{E(x_0, y_0) \propto \int\int e^{ik\left[\frac{(x_1^2+y_1^2)}{2z} + \frac{(-2x_0x_1-2y_0y_1)}{2z}\right]} Aperture(x_1, y_1) dx_1 dy_1}$$

2. **Far field approximation** $\frac{kD}{2z} \ll 1$ This new approximation requires $\frac{D}{z}$ to be smaller than before

$$\frac{kD^2}{2z} \ll 1 \rightarrow \frac{D}{z} = \frac{2}{kD} = \frac{D}{z} = \frac{2z}{\frac{2\pi}{\lambda} D} = \frac{\lambda}{\pi D}$$

Rewriting the first term in the exponent in the Fresnel diffraction equation

$$\frac{(x_1^2 + y_1^2)}{2z} = \frac{D^2}{2z} \approx e^0 = 1$$

where we obtain the **Fraunhofer diffraction equation**

$$\boxed{E(x_0, y_0) \propto \int\int e^{-ik\frac{(-x_0x_1-y_0y_1)}{z}} Aperture(x_1, y_1) dx_1 dy_1}$$

The Fraunhofer diffraction equation shows that the observed electric field in the far field approximation is the Fourier transform of the field at the Aperture (actually two Fourier Transforms, in two variables). This is more evident by writing the equation in one dimension

$$E(x_0) \propto \int\int e^{-ik\frac{(-x_0x_1}{z}} Aperture(x_1) dx_1$$

In the case of a circular aperture with diameter $D = 2b$, we have

$$FT(Aperture_{circular}(x_1, y_1)) \propto \frac{J_1(k\rho b/z)}{k\rho b/z}$$

where $\rho = \sqrt{x_1^2 + y_1^2}$ is the radial coordinate in the (x_1, y_1) plane, and J_1 is the Bessel function of the first kind. The intensity of the diffracted wave is

$$I = I_0 \left[2\frac{J_1(k\rho b/z)}{k\rho b/z}\right]^2$$

The first minimum of the $(2J_1(x)/x)^2$ function is a t $x = 3.8317$ and the intensity pattern is called **Airy disk**. The distance from the center to the first zero (ρ_{spot}) is:

$$\rho_{spot} = 3.83\frac{z}{kb} = 3.83\frac{z\lambda}{\pi\lambda} = 1.22\frac{z\lambda}{D}$$

If we apply this formula to a telescope with aperture D and focal length f (hence $z \equiv f$) we have

$$\boxed{\rho_{spot} = 1.22\frac{f}{\lambda D} \rightarrow \Delta\theta = \frac{\rho_{spot}}{f} = 1.22\frac{\lambda}{D}}$$

where $\Delta\theta$ is the annuglar spread, λ is the wavleength of the incident radiation, D is the aperture diamenter, and θ is given in radians. The Airy pattern is therefore the Fraunhofer diffraction pattern from a ciruclar aperture.