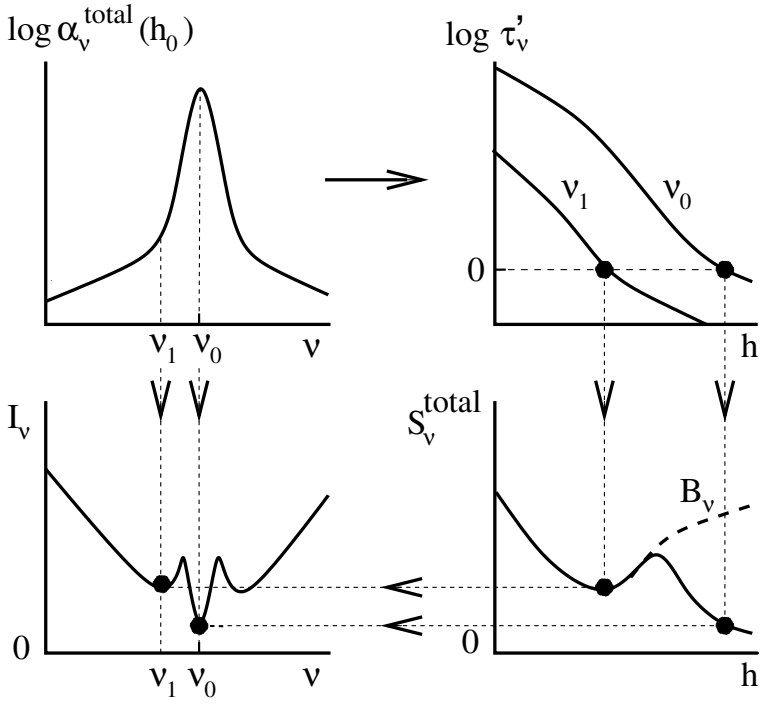
INTRODUCTION
TO
SOLAR SPECTRUM FORMATION

National Solar Observatory / Sacramento Peak
1993 Summer Course



Robert J. Rutten
Sterrekundig Instituut Utrecht
1993 Visiting Fulbright Scholar at NSO/SP

Contents

Bibliography. These notes follow the first chapter of the excellent textbook by Rybicki and Lightman (1979) and use the same notation (except for flux). Major other books of interest are the stellar-photospheres book by Gray (1992) (good on LTE line modeling and also on stellar observation techniques), Mihalas (1978) (the bible of NLTE stellar-atmosphere radiative transfreaks), and the introductory texts by Böhm-Vitense (1989) and Novotny (1973).

1 Introduction

THIS introduction to solar spectrum formation starts with microscopic radiation processes, macroscopic radiation quantities, the transport equation, and its formal solution for a given source function. The latter applies when LTE (Local Thermodynamic Equilibrium) holds, because the source function then simply equals the Planck function. In that case the emergent radiation can be attributed directly and simply to the local gas temperature where it originates. Much of the visible and infrared solar spectrum can be so explained.

However, most strong lines and ionization edges in the solar spectrum are much influenced by NLTE scattering. In the outer layers of the solar atmosphere the probability that the detected photon was scattered towards you becomes much larger than that it came to you directly after its creation from the thermal pool as LTE supposes. I treat NLTE spectrum formation by using idealized "two-level" atoms to illustrate the effect of such scattering. It makes strong solar-spectrum features form non-locally. And more complex, warranting explanation as given here.

Of course, all electromagnetic radiation is non-local; it conveys energy and information at the speed of light. In particular from distant objects to us, being our principal non-textile information carrier – with its spectral lines an informative encoder of physical circumstances within the distant object. I mean non-local here in the sense that the detected radiation is not characteristic for the location where it is seen coming from. Often, such non-locality arises from scattering. In a fog, one sees car headlight photons long before one sees the car. These photons are seen to come from the fog, but they were made in the car lights and they convey headlight radiation properties, not those of the fog. A spectrum of your nose in daylight, even in cloudy weather, contains thousands of lines made in the Sun, not by your nose, and represents a color temperature of 6000 K rather than 300 K. Most lines in the spectrum of the blue or cloudy sky are from the Sun, not from molecules in the air. And so, a spectral line from the Sun itself may also originate from another solar location than where it apparently comes from.

Conversion of photons may also happen, giving them the signature of a different spectral regime. The classical example is the bright Hα Balmer line from cold planetary nebulae. It portrays the presence of a hot, unseen central star that radiates mostly in the ultraviolet. The stellar Lyman-continuum photons ionize hydrogen atoms in the nebula, which subsequently recombine and emit Balmer photons in de-excitation cascades. These cold-nebula Balmer photons betray the existence of a hot central star.

Photon scattering and photon conversion occur in the Sun in many ways. Modern computers and solution methods make it possible to compute their effects in realistic detail. After the foundations were laid (mostly at Harvard and Boulder) in the 1960s and 1970s, numerical methods were developed to compute and analyze solar and stellar spectra in great detail. This text is a start-off tutorial for the underlying physics.

2 Solar radiation at the particle level

2.1 Solar radiation processes

APART from magnetodynamic complexities, the solar atmosphere represents a relatively simple medium: an ideal gas, mostly made up by hydrogen. No fluids or solids. A bit of helium, and just a sniff of all other elements.

Apart from magnetic complexities, the solar radiation that we receive is governed by only a few processes:

- bound-free (bf) ionization and recombination, contributing continua. Those of H^-, hydrogen atoms with an extra electron, make the gas in the solar photosphere much more opaque in the visible and infrared than the terrestrial air surrounding us. The extra electrons come from abundant "metals" with lower ionization energy than the 13.6 eV of hydrogen: Mg, Si, Fe, Al. Their bound-free ionization, in the upper photosphere mostly by ultraviolet radiation and therefore out of LTE, frees electrons to produce H^- at a density of 10^{-4} of the hydrogen density, set by the combined abundance of these electron-donor elements;
- free-free (ff) interactions. H^- ff (neutral hydrogen atoms encountering free electrons) dominates the solar infrared continuum beyond the H^- ionization limit at $\lambda = 1.6\,\mu$m. Hydrogen ff (free protons interacting with free electrons) dominates in the sub-mm and mm continua (in which the radiation escapes higher up in the atmosphere where hydrogen is ionized);
- Rayleigh scattering of photons off neutral hydrogen atoms. It contributes continuous opacity especially in the violet (as in our blue sky) and near-ultraviolet because its cross-section scales with λ^{-4};
- Thomson scattering of photons off free electrons, important only where hydrogen is ionized (and therefore in the atmospheres of hot stars that contain no H^-). Its small cross-section (6.7×10^{-25} cm^2 electron^{-1}) is constant with frequency;
- bound-bound (bb) interactions producing spectral lines. To the Sun these are most important as radiation-loss agents, especially Mg II h & k and Ca II H & K in the chromosphere, the hydrogen Lyman lines in the transition region from there to the corona, and numerous high-ion metal lines in the corona. To us they are important because they represent informative diagnostics of the physical conditions where they form.

2.2 Bound-bound process combinations

Thomson and Rayleigh scattering are always scatterings: photon in, the same photon out (apart from a Dopplershift) in a new direction. Isotropic redirection is a good approximation. Such scattering can put photons into the direc-

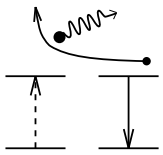
tion to the observer and so contribute to the local gas emissivity, but it can also redirect photons out of the beam of interest and so contribute extinction;

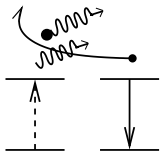
In contrast, ff interactions are always collisional, converting kinetic energy into a photon or vice-versa in an encounter with another particle. Assuming validity of the Maxwell distribution (usually the case in stellar atmospheres), this represents exchange between the local thermal pool and the radiation field.

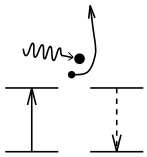
The bb and bf interactions can be of both type. They can go per photon (bb absorption and bf photoionization up; spontaneous plus induced bb emission and radiative bf recombination down), but they can also, instead, go per collision with another particle (collisional excitation and ionization up, collisional de-excitation and three-body recombination down).

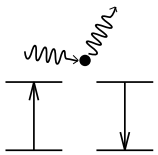
Let us discuss various bb pair combinations:

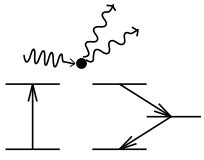
- collisional excitation up, collisional de-excitation down. No photon involved. Such pairs maintain thermal population ratios between the upper and lower levels;
- collisional excitation up, spontaneous radiative de-excitation (photon emission) down. This combination converts kinetic energy into radiation. With the Maxwell distribution this represents thermal photon creation;
- collisional excitation up, induced radiative de-excitation (photon emission) down. Similar photon creation, but the newly created photon comes along with the stimulating photon;
- radiative excitation up, collisional de-excitation down. Thermal photon destruction: the photon energy goes into the thermal pool;
- radiative excitation up, radiative de-excitation down (either spontaneous or induced). This is scattering, in principle comparable to Rayleigh and Thomson scattering but with an additional complexity: the possibility of a slight frequency change not due to Dopplershift but to losing coherency. In coherent scattering the re-emitted photon has the same frequency as the exciting one. In complete redistribution (CRD) the de-excitation is a fresh sample of the frequency distribution of the transition probability (the extinction "profile"). These are the extreme options; in partial redistribution (PRD) the actual frequency redistribution is evaluated. Most solar lines are well described by CRD, but the Lyman lines, Mg II h & k, Ca IIH & K and some others require PRD evaluation;
- radiative excitation up, radiative de-excitation down but into another level. This describes photon conversion from one wavelength domain to another. Similar multi-level complexity arises when the excited atom is subsequently ionized (radiatively or per collision), etc. To avoid such complexities, radiative-transfer articles and textbooks usually limit the discussion to "two-level" atoms in which whatever goes up has to come down in the same transition, without crosstalk to other wavelength domains.


creation


amplification


destruction


scattering


conversion

The same paired combinations as these bb ones can occur for bf transitions,

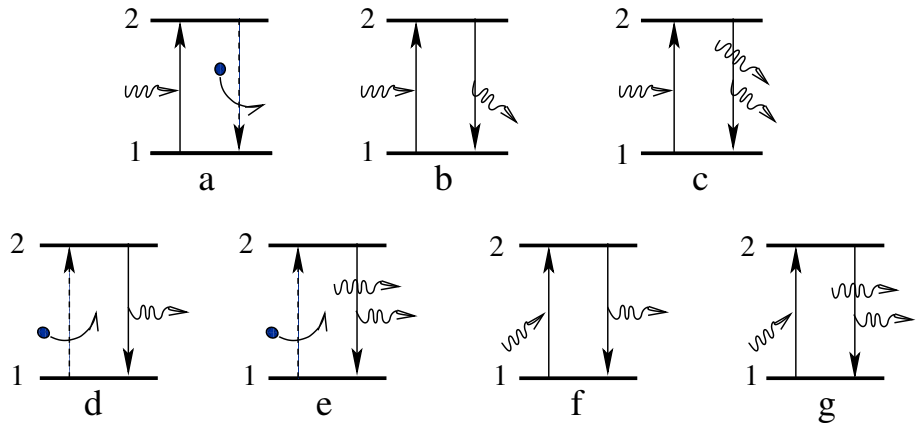


Figure 1: All up-down two-level-atom sequences involving photons in a given beam. The beam direction is to the right. The upper three pairs split the extinction of photons from the beam between collisional destruction and spontaneous and induced scattering. The lower four pairs split the emission of photons into the beam between spontaneous and induced thermal creation and scattering. The induced scattering pairs (c) and (g) both use a beam photon and one with arbitrary direction and so cancel against each other.

with as differences that the frequency range is much wider (an extended edge rather than a narrow line) and that in bound-free scattering the frequency redistribution over the edge profile is always complete (since the electron that is caught for recombination represents a fresh sample of the kinetic energy distribution).

2.3 Two-level-atom equilibria

For two-level atoms, each extinction (radiative excitation) is followed either by collisional de-excitation (photon destruction) or radiative de-excitation (photon scattering). This split depends on what happens after the initial extinction. Figure 1 details such splitting for the extinction of photons in a given beam and for the emission of photons into that beam. It shows the full set of up-down pairs involving photons:

(a) thermal extinction = radiative excitation by a beam photon followed by collisional de-excitation,

(b) spontaneous scattering extinction = radiative excitation by a beam photon followed by spontaneous de-excitation,

(c) induced scattering extinction = radiative excitation by a beam photon followed by induced de-excitation,

(d) spontaneous thermal emission = collisional excitation followed by spontaneous emission of a photon into the beam,

(e) induced thermal emission = collisional excitation followed by induced emission of a photon into the beam,

(f) spontaneous scattering emission = radiative excitation followed by spontaneous emission of a photon into the beam,

(g) induced scattering emission = radiative excitation followed by induced emission of a photon into the beam.

The type of equilibrium that may be attained in a two-level-atom gas depends on which of these pairs dominate:

- *LTE = large collision frequency, small photon losses*
  - up: mostly collisional = thermal creation (d + e);
  - down: mostly collisional = large destruction probability (a);
  - photon travel: short distances only (the photons acting as "honorary gas particles"), or only a negligible leak of escaping photons.

  This simple equilibrium holds very well in the solar interior and is a good approximation for continua and lines from the lower photosphere. The atomic level populations are given by the Boltzmann and Saha equations, the source function by the Planck function. Treated in the first half of this text;
- *NLTE = statistical equilibrium or time-dependent populations, large photon loss with much scattering*
  - photon travel: non-local radiation important;
  - much scattering, with coherent or complete or partial frequency redistribution;
  - multi-level processes: photon conversion, sensitivity transcription;
  - time dependence: ionization/recombination speed imbalances.

  Statistical equilibrium (meaning time-constant atomic level populations) may still apply in the upper photosphere, but in the chromosphere and transition region the populations are governed by time-dependent ionization and recombination, especially for helium and hydrogen. Source functions are often dominated by the radiation field, for multi-level interlocking also at other frequencies. Treated in the second half of this text;
- *coronal equilibrium = hot, tenuous*
  - up: only collisional = thermal creation (only d);
  - down: only spontaneous (only d);
  - photons either escape, or drown in the Sun, or scatter to other wavelengths in the H I, He I, or He II bf edges.

  This simple equilibrium describes coronal EUV radiation from highly-ionized metals very well. Each observed photon is thermally created, but LTE does not apply because the corona is optically thin. Note that the dependencies on temperature and density are the reverse of those in Boltzmann-Saha LTE equilibria: the degree of ionization is a function of temperature only, the degree of excitation depends on both the temperature and the electron density. Dielectronic recombination (in which part of the energy of the incoming electron is used to excite an already bound electron) is important. Not treated in this text.

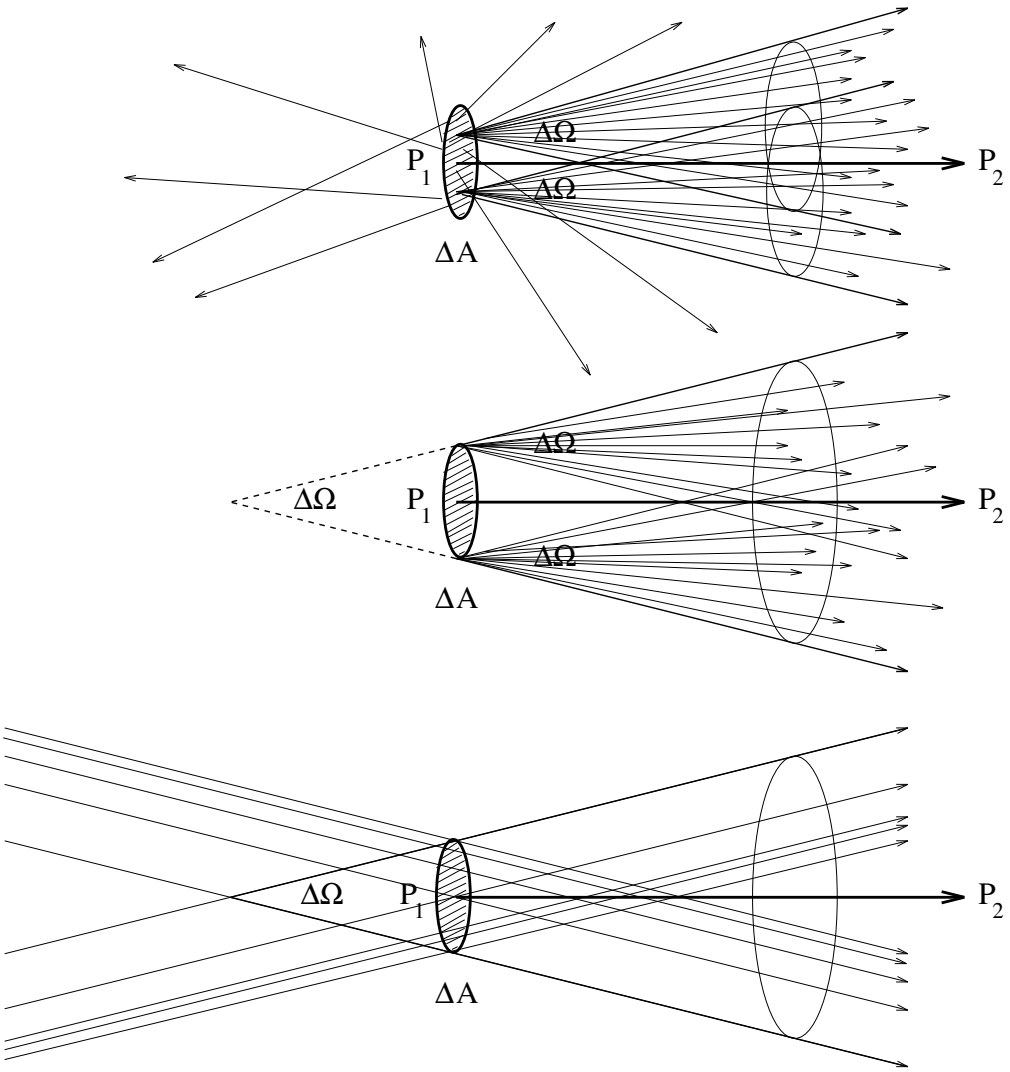


Figure 2: Cones ("pencils") of radiation. Photons are emitted by a circular surface with area ΔA around P_1 in all directions. The photons that leave a particular point of ΔA with directions within solid angle $\Delta\Omega$ around direction P_1P_2 constitute a cone of radiation emerging from that point (top). The cones from all such points on ΔA merge into a larger, truncated cone with opening angle $\Delta\Omega$ (middle). Likewise for beams of parallel rays from elsewhere that pass through A_1 with the same opening angle $\Delta\Omega$ (bottom). The angle is the same in the propagation direction towards the right and in the line-of-sight direction towards the left. The amount of energy in the cone is proportional to ΔA and $\Delta\Omega$, and also to the duration Δt and the frequency bandwidth $\Delta\nu$ of the measurement. After Figure 2-2 of Novotny (1973).

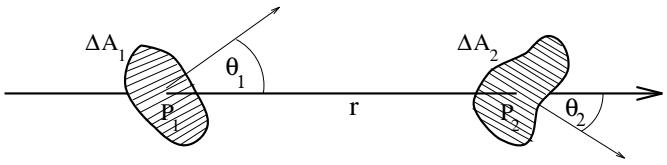


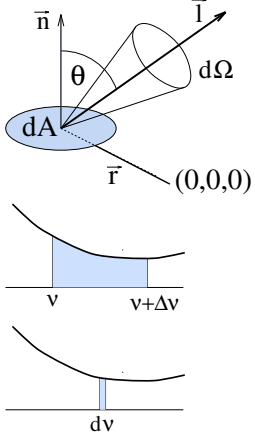
Figure 3: Demonstration that intensity does not vary with travel distance in vacuum. Consider all photons within frequency bandwidth $\Delta\nu$ that travel from left to right first through sampling area ΔA_1 at P_1 and then as well through sampling area ΔA_2 at P_2 during the measurement duration Δt (retard that over the photon travel time r/c from P_1 to P_2 if you want to be finicky). One such photon path is drawn. The first sample sees, according to (1), an amount of energy $\Delta E_1 \equiv I_1 \cos\theta_1 \Delta A_1 \Delta t \Delta\nu \Delta\Omega_1$. The second sees $\Delta E_2 \equiv I_2 \cos\theta_2 \Delta A_2 \Delta t \Delta\nu \Delta\Omega_2$. The two spread angles are $\Delta\Omega_1 = \cos\theta_2 \Delta A_2/r^2$ and $\Delta\Omega_2 = \cos\theta_1 \Delta A_1/r^2$, each seeing the other on its sky. Entering these makes the two expressions the same. Make all $\Delta's$ sufficiently small that you can believe these proportionalities. Since the same photons are measured at both locations, $\Delta E_1 = \Delta E_2$. Since the two expressions are the same, the two proportionality factors I_1 and I_2 must be identical.

3 Macroscopic radiation quantities

3.1 Intensity, mean intensity, flux

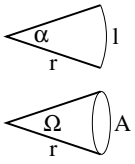
THE *intensity* I_ν is the coefficient of proportionality in:

$$\begin{aligned} \mathrm{d}E_\nu &\equiv I_\nu(\vec{r},\vec{l},t)\,(\vec{l}.\vec{n})\,\mathrm{d}A\,\mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega \\ &= I_\nu(x,y,z,\theta,\varphi,t)\,\cos\theta\,\mathrm{d}A\,\mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega, \end{aligned} \tag{1}$$



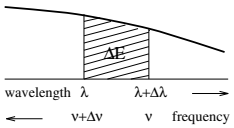
where $\mathrm{d}E_\nu$ is the amount of energy that is transported through area $\mathrm{d}A$, at the location $\vec{r}$ and with $\vec{n}$ the normal to $\mathrm{d}A$, during the time interval between t and $t + \mathrm{d}t$, in the frequency bandpass between ν and $\nu + \mathrm{d}\nu$ and in the solid angle $\mathrm{d}\Omega$ around direction $\vec{l}$. This definition expresses the heuristic insight that the measured amount of energy increases linearly with the detector area, measurement duration, spectral bandwidth, and amount of beam spreading, at least in the infinitesimal limits $\Delta \to \mathrm{d}$ where all these sampling dimensions are sufficiently small. Better names for I_ν are *surface brightness* and *specific intensity*. Dimension: [erg s^{-1} cm^{-2} Hz^{-1} ster^{-1}] or [W m^{-2} Hz^{-1} ster^{-1}]. This is the *monochromatic* intensity; the *total* intensity is $I \equiv \int_0^\infty I_\nu\,\mathrm{d}\nu$.

Intensity specifies the flow of energy along a beam of radiation both at departure and at arrival. It describes the radiation along a "ray" connecting the departure and arrival points – for example the solar origin and our detector. A single, infinitely thin ray doesn't contain energy, so we speak of a bundle or beam of rays. A "pencil of radiation" has angular spreading over a cone $\Delta\Omega$ around the ray direction. The rays travel towards us in the direction of propagation; their spreading is also measured when looking backwards along the line of sight. See Figure 2.

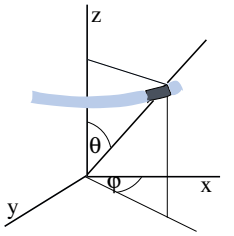


The measurement per steradian and per Hz are less intuitive than the measurement per cm^2 and per second, where the latter get rid of the actual detector size and measurement duration. A cone of rays spreads, but since intensity is measured per unit of spreading (a steradian being the spherical

analogue to the circular radian), the beam spreading does not affect its intensity, at least not in vacuum where there is no matter present to absorb or emit photons. The solar intensity is the same where it escapes at the solar surface and where it hits you on the beach, or in the focal plane of your telescope[1].See Figure 3. This property makes intensity the macroscopic quantity of choice to formulate radiative transfer with, *i.e.,* to describe processes by which matter and photons interact. Using intensity ensures that *only* such interactions affect the measure of radiation, not the distance over which it travels.



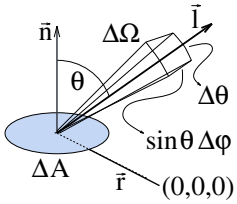
Conversion to wavelength for the measurement per unit of bandwidth: $I_\lambda = I_\nu\, c/\lambda^2$, with $\mathrm{d}\lambda$ and $\mathrm{d}\nu$ both defined positive for increase.



The *mean intensity* J_ν is defined as:

$$J_\nu(\vec{r},t) \equiv \frac{1}{4\pi}\int I_\nu\,\mathrm{d}\Omega = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} I_\nu \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi, \tag{2}$$

i.e., intensity averaged over all directions. Dimension: [erg cm^{-2} s^{-1} Hz^{-1} ster^{-1}], just as I_ν (averaging a quantity does not change its nature). This quantity is of interest to the atoms in the atmosphere as specifying the availability of photons with the proper energy for photo-excitation, regardless of where they come from. For *axial symmetry* with the z-as ($\theta \equiv 0$) along the axis of symmetry (only vertical stratification, "plane parallel layers"), with



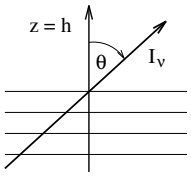
$$\mathrm{d}\Omega = 2\pi\sin\theta\,\mathrm{d}\theta = -2\pi\,\mathrm{d}\mu \tag{3}$$

and

$$\mu \equiv \cos\theta \tag{4}$$

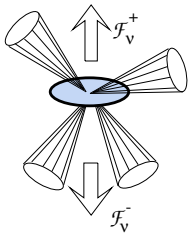
we find:

$$J_\nu(z) = \frac{1}{4\pi}\int_0^{\pi} I_\nu(z,\theta)\,2\pi\sin\theta\,\mathrm{d}\theta = \frac{1}{2}\int_{-1}^{+1} I_\nu(z,\mu)\,\mathrm{d}\mu. \tag{5}$$



The *monochromatic flux* $\mathcal{F}_\nu$ is:

$$\mathcal{F}_\nu(\vec{r},\vec{n},t) \equiv \int I_\nu\cos\theta\,\mathrm{d}\Omega = \int_0^{2\pi}\int_0^{\pi} I_\nu\cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi. \tag{6}$$



Dimension: [erg s^{-1} cm^{-2} Hz^{-1}] or [W m^{-2} Hz^{-1}]. This is the net flow of energy per sec through an area of one cm^2 which is placed at $\vec{r}$ perpendicular to direction $\vec{n}$, with:

$$\begin{aligned}
\mathcal{F}_\nu &= \int_0^{2\pi}\int_0^{\pi/2} I_\nu\cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi + \int_0^{2\pi}\int_{\pi/2}^{\pi} I_\nu\cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi \\
&= \int_0^{2\pi}\int_0^{\pi/2} I_\nu\cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi - \int_0^{2\pi}\int_0^{\pi/2} I_\nu\cos(\pi-\theta)\sin(\pi-\theta)\,\mathrm{d}(\pi-\theta)\,\mathrm{d}\varphi \\
&\equiv \mathcal{F}_\nu^+ - \mathcal{F}_\nu^-,
\end{aligned} \tag{7}$$

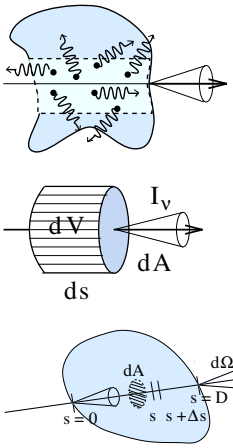
[1] Assuming loss-free optics that only affect the cone opening angle. But even with a perfect telescope you cannot observe the intensity of a distant star when its angular extent is below the resolving power. From unresolved stars we measure only irradiance, the flux at the telescope given by (6) but limiting the angular integration to the stellar image.

with the upward flux $\mathcal{F}_\nu^+$ and the downward flux $\mathcal{F}_\nu^-$ both positive. Isotropic radiation has $\mathcal{F}_\nu^+ = \mathcal{F}_\nu^- = \pi I_\nu$ en $\mathcal{F}_\nu = 0$. For axial symmetry:

$$\begin{aligned}\mathcal{F}_\nu(z) &= 2\pi \int_0^\pi I_\nu \cos\theta \sin\theta \, \mathrm{d}\theta \\ &= 2\pi \int_0^1 \mu I_\nu \, \mathrm{d}\mu - 2\pi \int_0^{-1} \mu I_\nu \, \mathrm{d}\mu \\ &= \mathcal{F}_\nu^+(z) - \mathcal{F}_\nu^-(z). \end{aligned} \tag{8}$$

The net flux is of interest to the star in maintaining equilibrium. When the whole energy flux is transported by radiation, radiative equilibrium holds when $\mathrm{d}\mathcal{F}/\mathrm{d}z = 0$, with $\mathcal{F} \equiv \int \mathcal{F}_\nu \, \mathrm{d}\nu$. Flux is often used by astronomers as the outward energy stream that leaves the stellar surface. And by observers as the energy stream from the (unresolved) object hitting their detector. These varied uses imply that one must carefully define where the measurement location is and what the positive direction is.

3.2 Emissivity, extinction, source function



The *monochromatic emissivity* j_ν is per cm^3 the constant of proportionality in:

$$\mathrm{d}E_\nu = j_\nu \, \mathrm{d}V \, \mathrm{d}t \, \mathrm{d}\nu \, \mathrm{d}\Omega, \tag{9}$$

with $\mathrm{d}E_\nu$ the energy which is added to the local radiation by local generation of photons, in a volume $\mathrm{d}V$, per bandwidth $\mathrm{d}\nu$, during a period $\mathrm{d}t$ and in the directions $\mathrm{d}\Omega$. Dimension j_ν: [erg cm^{-3} s^{-1} Hz^{-1} ster^{-1}]. This is another heuristic truth for sufficiently small ($\Delta \to \mathrm{d}$) sampling extents. We can also define j_ν as the amount of intensity that is added along a beam by local emission:

$$\mathrm{d}I_\nu(s) = j_\nu(s) \, \mathrm{d}s. \tag{10}$$

The *extinction coefficient* may (just as the emissivity) be defined per particle, per cm, or per gram.
Per particle:

$$\mathrm{d}I_\nu = -I_\nu \, \sigma_\nu \, n \, \mathrm{d}s \tag{11}$$

with σ_ν the monochromatic extinction coefficient = "cross-section" per particle ([cm^2]) and n the density of the absorbing particles ([cm^{-3}]).
Per unit of path length:

$$\mathrm{d}I_\nu = -I_\nu \, \alpha_\nu \, \mathrm{d}s \tag{12}$$

with α_ν the monochromatic linear extinction coefficient ([cm^{-1}]), identical to cross-section per unit of volume ([cm^2 cm^{-3}] = [cm^{-1}]).
Per unit of mass:

$$\mathrm{d}I_\nu = -I_\nu \, \kappa_\nu \, \rho \, \mathrm{d}s \tag{13}$$

with κ_ν the monochromatic mass extinction coefficient = cross-section per unit of mass ([cm^2 g^{-1}]) and ρ the mass density ([g cm^{-3}]).

Physicists are interested in single particles and use σ_ν. Astronomers are interested in gas opacities and prefer κ_ν. Here our interest is in the transport of radiation, so I follow Rybicki & Lightman and use α_ν.

The *source function* is:

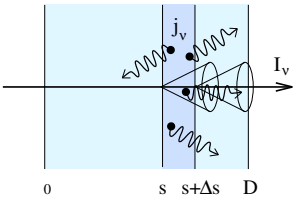
$$S_\nu \equiv j_\nu / \alpha_\nu. \tag{14}$$

Dimension: [erg cm^{-2} s^{-1} Hz^{-1} ster^{-1}], the same as intensity. As the name suggests it measures the energy of new photons in the given beam, normalized by the local extinction. At a given frequency multiple processes may contribute to the local emission and/or extinction. In that case:

$$S_\nu^{\text{total}} = \frac{\sum j_\nu}{\sum \alpha_\nu}, \tag{15}$$

as for the combination of continuum and line processes (Figure 7).

3.3 Radiation transport equation



The *equation for radiation transfer* is:

$$\mathrm{d}I_\nu(s) = I_\nu(s + \mathrm{d}s) - I_\nu(s) = j_\nu(s)\,\mathrm{d}s - \alpha_\nu(s)\,I_\nu(s)\,\mathrm{d}s \tag{16}$$

or

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = j_\nu - \alpha_\nu\,I_\nu \tag{17}$$

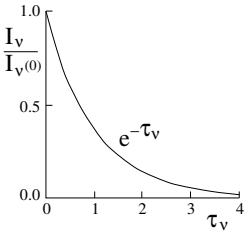
with s measured along the beam. This equation expresses the conservation of intensity along a beam in vacuum. It is only changed by matter–radiation interactions specified by j_ν and α_ν.

The *monochromatic optical path length* $\mathrm{d}\tau_\nu$ of a thin layer with thickness $\mathrm{d}s$ along the beam is

$$\mathrm{d}\tau_\nu(s) \equiv \alpha_\nu(s)\,\mathrm{d}s, \tag{18}$$

the *monochromatic optical thickness* of a layer with geometrical thickness D is:

$$\tau_\nu(D) = \int_0^D \alpha_\nu(s)\,\mathrm{d}s, \tag{19}$$



again measured along the beam. For pure extinction:

$$I_\nu(D) = I_\nu(0)\,\mathrm{e}^{-\tau_\nu(D)}. \tag{20}$$

The transition between little and much extinction lies at the $1/e$ value, *i.e.*, at optical thickness $\tau_\nu = 1$. A layer is called *optically thick* for $\tau_\nu(D) > 1$ and *optically thin* for $\tau_\nu(D) < 1$. The *mean photon free path in optical units* $<\tau_\nu(s)>$ is:

$$<\tau_\nu(s)> \equiv \frac{\int_0^\infty \tau_\nu(s)\,\mathrm{e}^{-\tau_\nu(s)}\,\mathrm{d}\tau_\nu(s)}{\int_0^\infty \mathrm{e}^{-\tau_\nu(s)}\,\mathrm{d}\tau_\nu(s)} = 1, \tag{21}$$

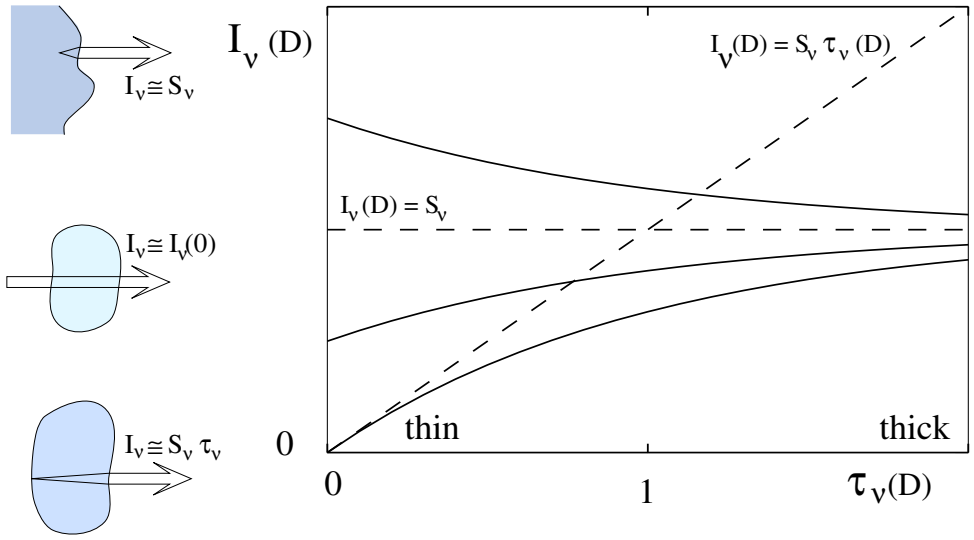


Figure 4: Emergent intensity $I_\nu(D)$ from a homogeneous object against its optical thickness $\tau_\nu(D)$, as given by (25). Optically thin objects without impinging radiation in the beam direction have $I_\nu = \tau_\nu(D)\, S_\nu$ (lowest curve with dashed limit). If a background intensity $I_\nu(0)$ illuminates the object from behind in the beam direction, there is enhancement of the intensity for $I_\nu(0) < S_\nu$ (middle curve), reduction for $I_\nu(0) > S_\nu$ (highest curve). $I_\nu(D)$ approaches S_ν for optical thickness $\tau_\nu(D) > 1$.

and the *mean photon free path in geometrical units* is, if the medium is homogeneous:

$$l_\nu = \frac{<\tau_\nu(s)>}{\alpha_\nu} = \frac{1}{\alpha_\nu} = \frac{1}{\kappa_\nu \rho}. \tag{22}$$

In an inhomogeneous medium this is a representative *local* mean free path.

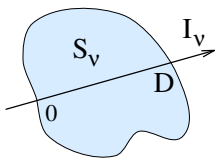
With τ_ν and S_ν the transport equation in differential form becomes:

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = S_\nu - I_\nu, \tag{23}$$

from which the transport equation in integral form follows:

$$I_\nu(\tau_\nu) = I_\nu(0)\, \mathrm{e}^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(t_\nu)\, \mathrm{e}^{-(\tau_\nu - t_\nu)}\, \mathrm{d}t_\nu. \tag{24}$$

For a homogeneous object with constant $S_\nu(s)$:



$$I_\nu(D) = I_\nu(0)\, \mathrm{e}^{-\tau_\nu(D)} + S_\nu \left(1 - \mathrm{e}^{-\tau_\nu(D)}\right). \tag{25}$$

If it is optically thick, then

$$I_\nu(D) \approx S_\nu \tag{26}$$

but if it is optically thin, then

$$I_\nu(D) \approx I_\nu(0) + [S_\nu - I_\nu(0)]\, \tau_\nu(D). \tag{27}$$

These results are illustrated in Figures 4 and 5. If the homogeneous object is optically thick, you observe its source function irrespective of the nature of the extinction coefficient. It won't show spectral lines.

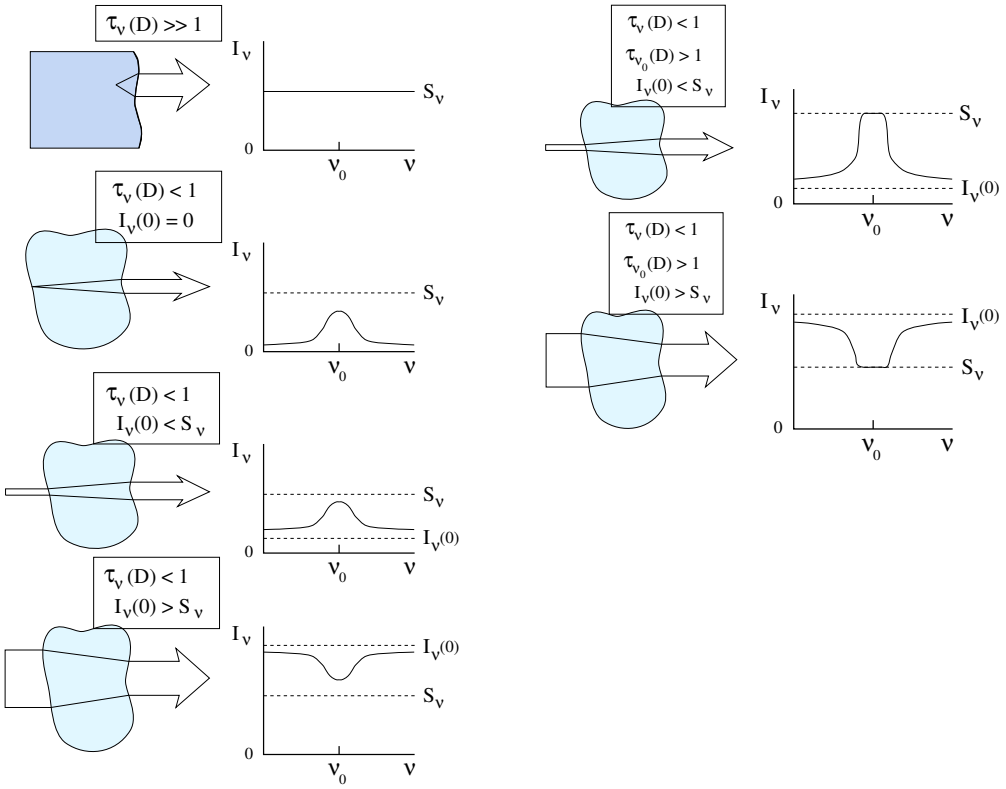
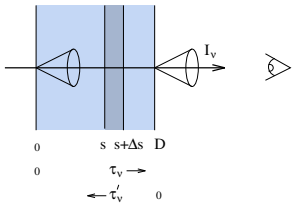


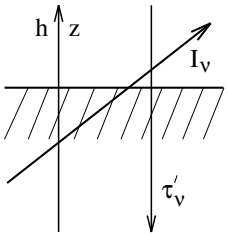
Figure 5: Spectral line formation in a homogeneous object with constant and equal continuum source function and line source function, as given by (25)–(27). An example would be an isothermal cloud obeying LTE. If the cloud is optically thick it shows no spectral lines since the same source function is sampled whether the thickness is even larger at the line frequency or less so in the continuum adjacent to the line. If the cloud is optically thin the additional line extinction enhances the positive or negative contribution by the cloud to the irradiation from behind. Right: when the cloud becomes optically thick at line center, the intensity saturates to the source function of the cloud.

The *monochromatic optical depth* is the optical pathlength along the line of sight, against the direction of photon propagation, from outside the object into it:

$$\mathrm{d}\tau'_\nu(s) \equiv -\alpha_\nu(s)\,\mathrm{d}s. \tag{28}$$



It is used instead of path length for objects that are hugely thick, such as stars, where it makes more sense to describe how deep one's view penetrates (*i.e.*, where the photons come out) then to count optical thickness from the backside.



In the case of axial symmetry (plane-parallel layers) the *radial optical depth* is used more commonly, also here. With the z-axis along the axis of symmetry the radial optical depth of a location with $z = z_0$ is:

$$\tau'_\nu(z_0) = \int_\infty^{z_0} \alpha_\nu\,\mathrm{d}z. \tag{29}$$

It represents the opaqueness increase along a line of sight perpendicular to the plane layers, with your eye in $z = \infty$. In axial symmetry and along a

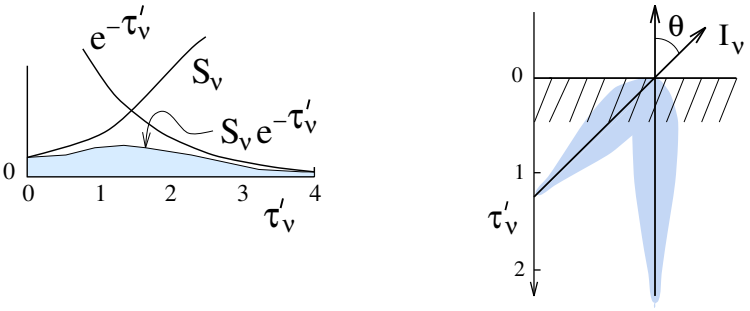


Figure 6: The Eddington-Barbier approximation. It does *not* mean that all observed photons escape at optical depth $\tau'_\nu = 1$, even though this is often said ("the photons come from $\tau = 1$"). The integrand $S_\nu \exp(-\tau'_\nu)$ extends over a broad domain in τ'_ν, from the surface with $\tau'_\nu = 0$ to $\tau'_\nu \approx 10$ where the factor $\exp(-\tau'_\nu)$ cuts it off. Photons escape from this entire slab; collectively they represent the value of the source function at $\tau'_\nu = 1$ in radial (vertical) viewing. In slanted viewing the characteristic path length $\Delta\tau = 1$ must be measured along the line of sight, yielding $\tau'_\nu = \mu$ as characteristic radial optical depth.

slanted line of sight with $\mu = \cos\theta$ the inwards directed ($\mu < 0$) intensity is, using $t'_\nu = \int_\infty^z \alpha_\nu(z')\,\mathrm{d}z'$ as integration variable:

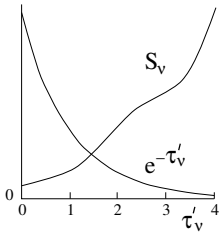
$$I_\nu^-(\tau'_\nu, \mu) = -\int_0^{\tau'_\nu} S_\nu(t'_\nu)\, \mathrm{e}^{-(t'_\nu - \tau'_\nu)/\mu}\, \mathrm{d}t'_\nu/\mu \tag{30}$$

and the outwards directed ($\mu > 0$) intensity is:

$$I_\nu^+(\tau'_\nu, \mu) = +\int_{\tau'_\nu}^{\infty} S_\nu(t'_\nu)\, \mathrm{e}^{-(t'_\nu - \tau'_\nu)/\mu}\, \mathrm{d}t'_\nu/\mu. \tag{31}$$

The *emergent intensity* is:

$$I_\nu^+(\tau'_\nu = 0, \mu) = \int_0^{\infty} S_\nu(t'_\nu)\, \mathrm{e}^{-t'_\nu/\mu}\, \mathrm{d}t'_\nu/\mu. \tag{32}$$



With

$$S_\nu(\tau'_\nu) = \sum_{n=0}^{\infty} a_n {\tau'_\nu}^n = a_0 + a_1\tau'_\nu + a_2{\tau'_\nu}^2 + ..., \tag{33}$$

and $\int_0^\infty x^n\, \mathrm{e}^{-x}\, \mathrm{d}x = n!$ one finds that:

$$I_\nu^+(\tau'_\nu = 0, \mu) = a_o + a_1\mu + 2a_2\mu^2 + \ldots + n!\, a_n\mu^n, \tag{34}$$

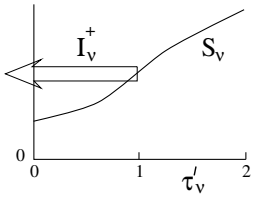
and the very important *Eddington-Barbier approximation:*

$$I_\nu^+(\tau'_\nu = 0, \mu) \;\approx\; S_\nu(\tau'_\nu = \mu), \tag{35}$$

which holds exactly if S_ν varies linearly with τ'_ν. The radiation which exits at the surface is then characterized by the conditions at one free path length below the surface. See Figure 6.

We conclude that a knowledge of the local source function $S_\nu(\tau'_\nu)$ is required in order to interpret the radiation from an object, both for the optically thick case with

$$I_\nu^+ \approx S_\nu(\tau'_\nu = 1) \tag{36}$$



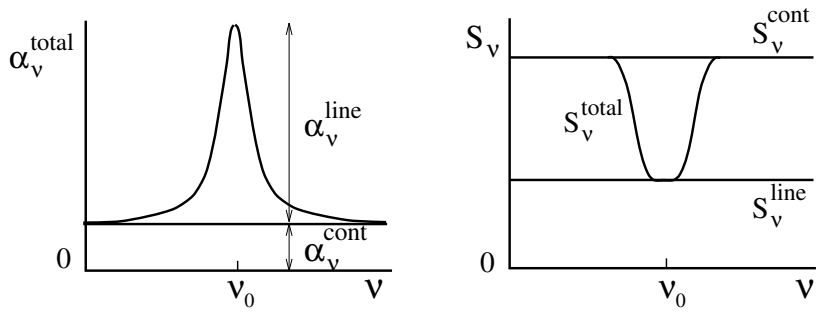
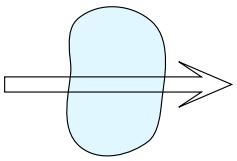


Figure 7: Combination of continuous and line extinction into the source function. The spectral line at left represents a narrow-band addition to the extinction at the line frequency. The total source source at right is the extinction-weighted combination given by (39).

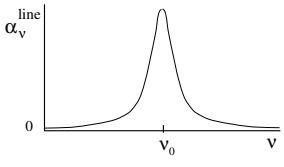
and for the optically thin case with

$$I_\nu(D) \approx I_\nu(0) + [S_\nu - I_\nu(0)]\ \tau_\nu(D). \tag{37}$$



In each case the τ scaling must be known as well. For the thick case, it sets the sampling depth at which the emergent radiation is characterized by the source function (Figure 6). In the thin case, the emergent emission or absorption scales linearly with the optical thickness of the object (Figure 5).

3.4 Spectral lines



A spectral line is always the result of a discrete bound-bound (bb) transition, *i.e.*, of an extra process above and beyond the continuous processes which operate at the relevant frequency. A "line" is not an infinitely sharp δ-function but has a narrow distribution in frequency. The line extinction coefficient itself has such shape, the "extinction profile" which is somewhat smeared out in frequency via various types of line broadening (Section 5.3). This statistical distribution is bell-shaped: a Gaussian core with broad wings. It affects the emergent line profile, but the latter is not equal to it (unless the object is optically thin and homogeneous as in Figure 5).

The existence of a bb transition probability at a given frequency has two consequences:

- the bb process provides an extra possibility for extinction, superimposed on the continuous extinction: at the frequency of a spectral line the extinction coefficient in the medium is larger than for the adjacent frequencies;
- the source function associated with the bb process may differ from that of the continuous processes at that frequency.

Taken together, these two possibilities can lead to a spectral line that is observable in the emergent intensity, either in emission or in absorption with respect to the continuous background.

The bb extinction peak contributes to the total optical thickness:

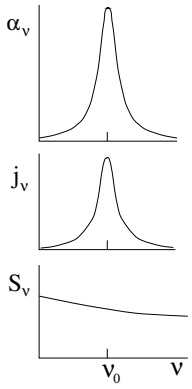
$$\tau_\nu = \int \alpha_\nu \, \mathrm{d}s = \int \alpha_\nu^{\mathrm{cont}} \, \mathrm{d}s + \int \alpha_\nu^{\mathrm{line}} \, \mathrm{d}s, \tag{38}$$

The source functions for the bb and the continuum processes are respectively:

$$S_\nu^l = \frac{j_\nu^{\mathrm{line}}}{\alpha_\nu^{\mathrm{line}}} \qquad \text{and} \qquad S_\nu^c = \frac{j_\nu^{\mathrm{cont}}}{\alpha_\nu^{\mathrm{cont}}},$$

and the total source function is the opacity-weighted combination of these process source functions (Figure 7):

$$S_\nu^{\mathrm{tot}} = \frac{\sum j_\nu}{\sum \alpha_\nu} = \frac{j_\nu^{\mathrm{cont}} + j_\nu^{\mathrm{line}}}{\alpha_\nu^{\mathrm{cont}} + \alpha_\nu^{\mathrm{line}}} = \frac{\alpha_\nu^{\mathrm{cont}} S_\nu^{\mathrm{cont}} + \alpha_\nu^{\mathrm{line}} S_\nu^{\mathrm{line}}}{\alpha_\nu^{\mathrm{cont}} + \alpha_\nu^{\mathrm{line}}}. \tag{39}$$



The existence of extra bb processes also increases the emissivity j_ν but we choose the combination extinction + source function in place of extinction + emission to describe line formation because this is a more orthogonal choice. Extinction measures the local processes that govern opacity; the source function measures quantities that govern the production of new photons into the beam. For example, when LTE holds with $S_\nu^{\mathrm{cont}} = S_\nu^{\mathrm{line}} = B_\nu(T)$ there is no bb signature whatsoever in the total source function.

4 LTE solar spectrum formation

4.1 LTE solar line formation

In an optically thick inhomogeneous medium in which the source function is not everywhere the same, the influence of the extra bb processes on the extinction mixes with the influence of the extra bb processes on the source function. The extra line extinction provides for a shallower depth of formation of the emergent intensity at the line wavelength than for the adjacent continuum. The source function at the line wavelength is sampled at a shallower level; it will in general be different there from its value at the larger depth of formation of the emergent continuum intensity.

We now examine the case that the continuum source function S_ν^c and the line source function S_ν^l are everywhere equal to each other, but vary along with depth in the medium. We so focus on the effect of the extra bb extinction, and neglect the effect a line may have on the total source function. This description is applicable to LTE line formation, because then $S_\nu^l = S_\nu^c = B_\nu(T)$.

The effect of the extra line extinction is that the location of the optical depth $\tau_\nu' = 1$, which is representative of the escaping photons, lies farther outward at the line frequency than for the continuum on each side of the line in the

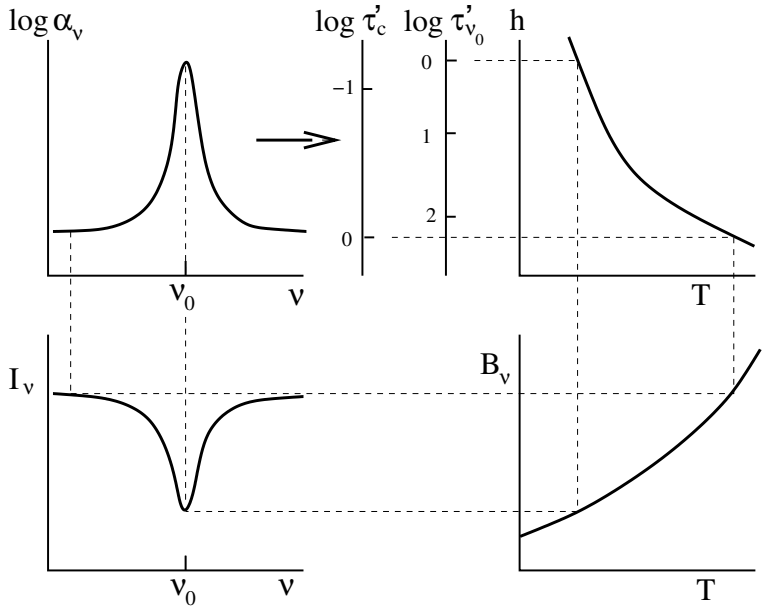


Figure 8: LTE line formation in the solar atmosphere. The depth-dependent extinction coefficient (first panel) determines the optical depth scales (top center). They differ per frequency, depending on the extinction profile. Generally extinction decreases roughly exponentially with height (h) with the diminishing density of the absorbing particles, so that $\log \tau'_\nu$ decreases approximately linearly with h at each frequency. The larger the extinction α_ν, the farther out the Eddington-Barbier representative height of formation $\log \tau'_\nu = 0$. The emergent intensity at each frequency (lower left) maps the temperature at this sampling height per frequency folded through the temperature dependence of the Planck function (bottom right; linear in the Rayleigh-Jeans domain and exponential in the Wien domain). When the temperature diminishes with height, absorption lines result. Emission lines are formed when the temperature increases outward.

spectrum. In other words: at the line frequency your sight penetrates less deep into the medium because extra bb extinction blocks your view.

From an optically thick medium in which $S^c_\nu = S^l_\nu$ one observes spectral lines only if the source function varies with depth, so that sampling different depths results in different intensities. In Figure 8 the source function $S_\nu(h) \approx B_\nu(T[h])$ decreases outward so that the resulting spectral lines are in absorption. Thus, in the case of optically thick LTE line formation, absorption lines result when and only when the temperature decreases outwards. The Planck function decreases outward with the temperature; the higher location that is characteristic for the emergent line-center radiation in Eddington-Barbier fashion samples a smaller Planck function than the continuum formation region does.

Thus, the observed shape of a spectral line from a stellar atmosphere does not portray directly the variation of the extinction coefficient with frequency, but reflects the variation of the source function with depth as sampled by the optical depth scaling which is set by the local extinction coefficient.

Only if a very weak line adds negligible opacity to the continuous opacity

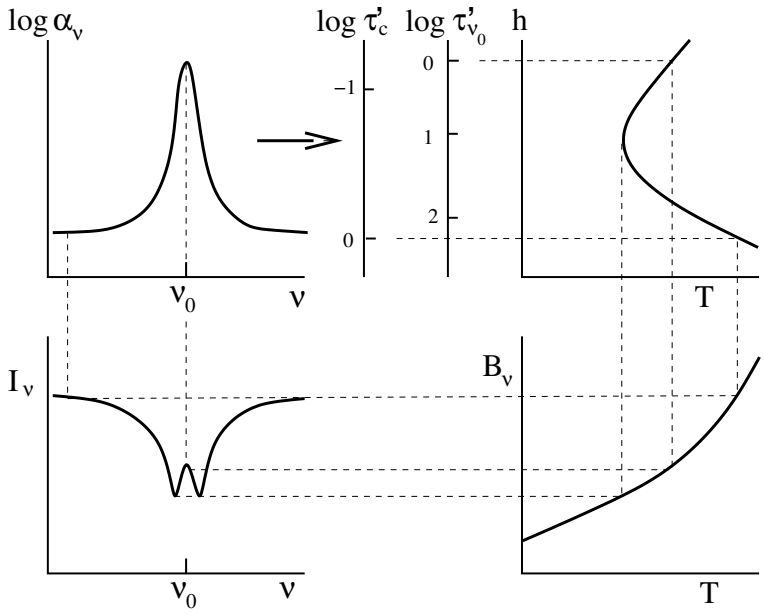


Figure 9: LTE line formation in an atmosphere with a temperature minimum. The core of a very strong line senses the outward temperature rise and maps that into a "self-reversal".

will its shape reflect the line extinction coefficient along the line of sight. For example, this is the case for weak telluric lines that the Earth's atmosphere adds into the observed solar spectrum.

The effect of the line extinction at the depth of the emerging photons can be quite large because the bb extinction coefficient in the center of the line profile often exceeds the continuous extinction coefficient by orders of magnitude. A "strong" spectral line provides sampling throughout a large height range. If a very strong line were formed in LTE in an atmosphere with a temperature minimum it would show a core reversal as in Fig. 9. In the solar case, lines that strong are not formed in LTE, and photospheric lines that display emission humps are all formed out of LTE. But the LTE description given here holds rather well for many weaker lines in the optical and infrared parts of the solar spectrum, in particular with respect to their source functions. For example, most optical Fe I lines have $S_\nu^{\rm line} \approx B_\nu$, but their extinction $\alpha_\nu^{\rm line}$ is smaller than LTE predicts via the Saha-Boltzmann equations, due to radiative overionization in the ultraviolet.

Local motions and magnetic fields in specific layers of the atmosphere cause changes in the local extinction profile. Motions cause Dopplershifts; magnetic fields cause Zeeman splitting. These changes of the extinction coefficient affect the mapping of the local source function into the emergent line profile.

For example, if the whole atmosphere moves upward the result is simply that the whole line shifts blueward. If a deep layer moves downward but a higher layer moves upward, the core of the emergent line is displaced to the blue while the wings shift to the red; the line becomes asymmetrical. If such motions are correlated with temperature changes, the emergent intensi-

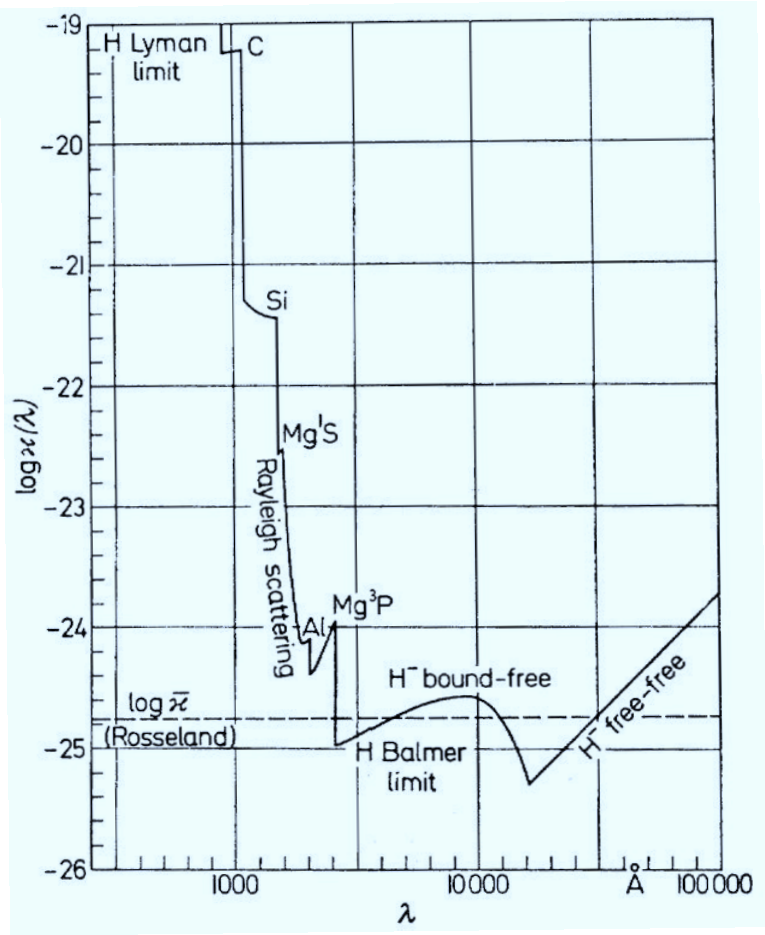


Figure 10: The continuous extinction coefficient in the photosphere of the Sun. From E. Böhm-Vitense.

ties change accordingly. Similar profile encoding follows from the presence of magnetic fields. Thus, in an optically thick atmosphere, the encoding of Doppler motions and magnetic fields into observable line profile information takes the form of changing the local extinction coefficient which affects the mapping of the local source function into emergent intensity.

4.2 LTE solar continuum formation

The discussion above holds not only for spectral lines (in which the extinction varies rapidly across the narrow-band profile), but also for the differences in emergent continua between different wavelength regions. For visible and infrared wavelengths LTE is a reasonably good assumption for the formation of the continuum, at least with respect to the source function. For example, the H^- ff and H I ff processes have LTE source functions since they are always collisional, but their opacity may be far out of LTE due to NLTE hydrogen ionization.

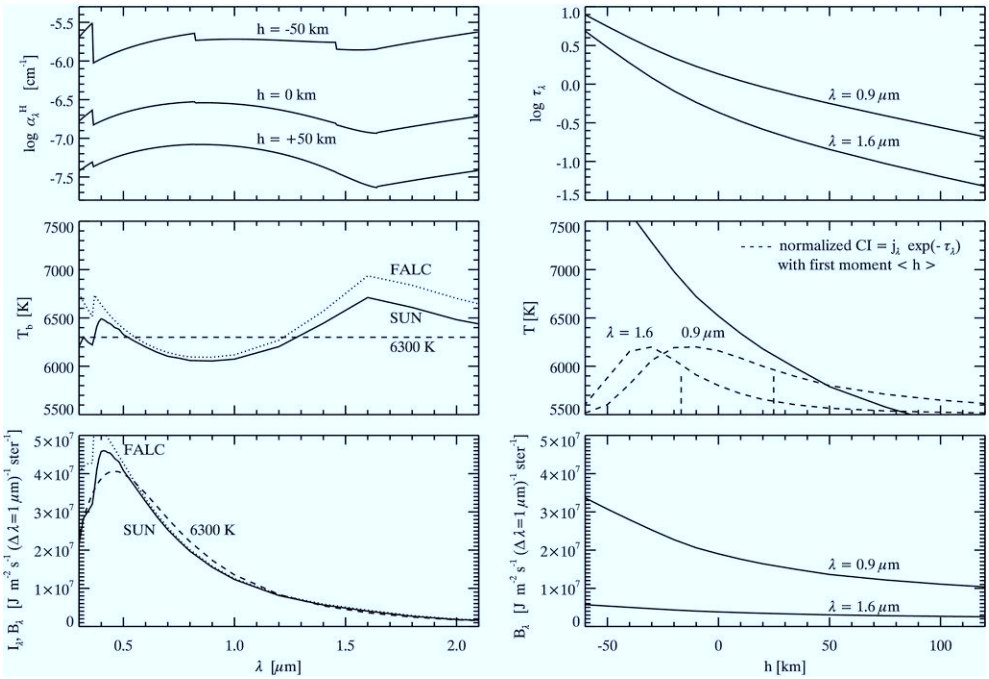


Figure 11: A six-panel computation of solar continuum formation in the visible and near-infrared. Assumptions: LTE, continuous extinction only from H I bf, H I ff, H^- bf, and H^- ff transitions, and the FALC up date by Fontenla *et al.* (1993) of the famous VAL3C standard model of Vernazza *et al.* (1981). *First panel:* total continuous hydrogen extinction at $h = -50,\ 0, +50$ km. H I contributes significantly only at the high temperature in the deepest layers; the top curve shows the Balmer, Paschen and Brackett edges. The steep decrease with height comes from the H^- dependence of the electron density, which decays much steeper than the neutral hydrogen density due to increasing hydrogen recombination. *Second panel:* two resulting $\tau_\lambda(h)$ curves, respectively for the maximum of the H^- bb extinction at $\lambda = 0.9$ μm and the opacity minimum at the H^- bf ionization threshold of $\lambda = 1.6$ μm. *Second row, righthand panel:* the FALC temperature stratification in the deep photosphere. The dashed curves show the normalized intensity contribution functions, *i.e.*, the integrands $\mathrm{d}I_\lambda/\mathrm{d}h = B_\lambda\, \mathrm{e}^{-\tau_\lambda}\, \alpha_\lambda = j_\lambda\, \mathrm{e}^{-\tau_\lambda}$ at the two wavelengths. The dashed lines mark the corresponding mean heights of formation $< h > \equiv \int_0^\infty h\, (\mathrm{d}I_\lambda/\mathrm{d}h)\ \mathrm{d}h\ /\ \int_0^\infty (\mathrm{d}I_\lambda/\mathrm{d}h)\ \mathrm{d}h\ = \int_0^\infty h\, B_\lambda\ \mathrm{e}^{-\tau_\lambda}\ \mathrm{d}\tau_\lambda\ /\ \int_0^\infty B_\lambda\ \mathrm{e}^{-\tau_\lambda}\ \mathrm{d}\tau_\lambda$. The Eddington-Barbier approximation suggests that these should coincide with the $\tau_\lambda = 1$ locations, but the latter lie deeper from lack of linearity. Note the wide extent of the integrands: the large width of stellar-atmosphere radiative-transfer kernels often make the notion "height of formation" naïve or inapplicable. *Second row, lefthand panel:* emergent spectra in the form of brightness temperatures T_b. The computed FALC values at $\lambda = 0.9$ and 1.6 μm agree closely with the FALC temperature at the corresponding mean formation height. The computation overestimates the observed solar spectrum (disk-center continuum between lines, taken from Allen, 1976), or the latter is too low. The computation fails for $\lambda < 5$ μm due to the neglect of metal ionization edges. *Bottom panels:* emergent spectra at left, Planck functions for the two wavelengths at right. For this wide spectral range the variation of the Planck function with wavelength dominates, so that the lefthand plot is not obviously a mapping of the extinction variation with wavelength. The slight differences with $B_\lambda(6300\,\mathrm{K})$, added as reference, appear much clearer in the T_b panel, especially at longer wavelengths where a given I_λ difference translates into larger temperature difference through the smaller Rayleigh-Jeans temperature sensitivity than in the Wien part.

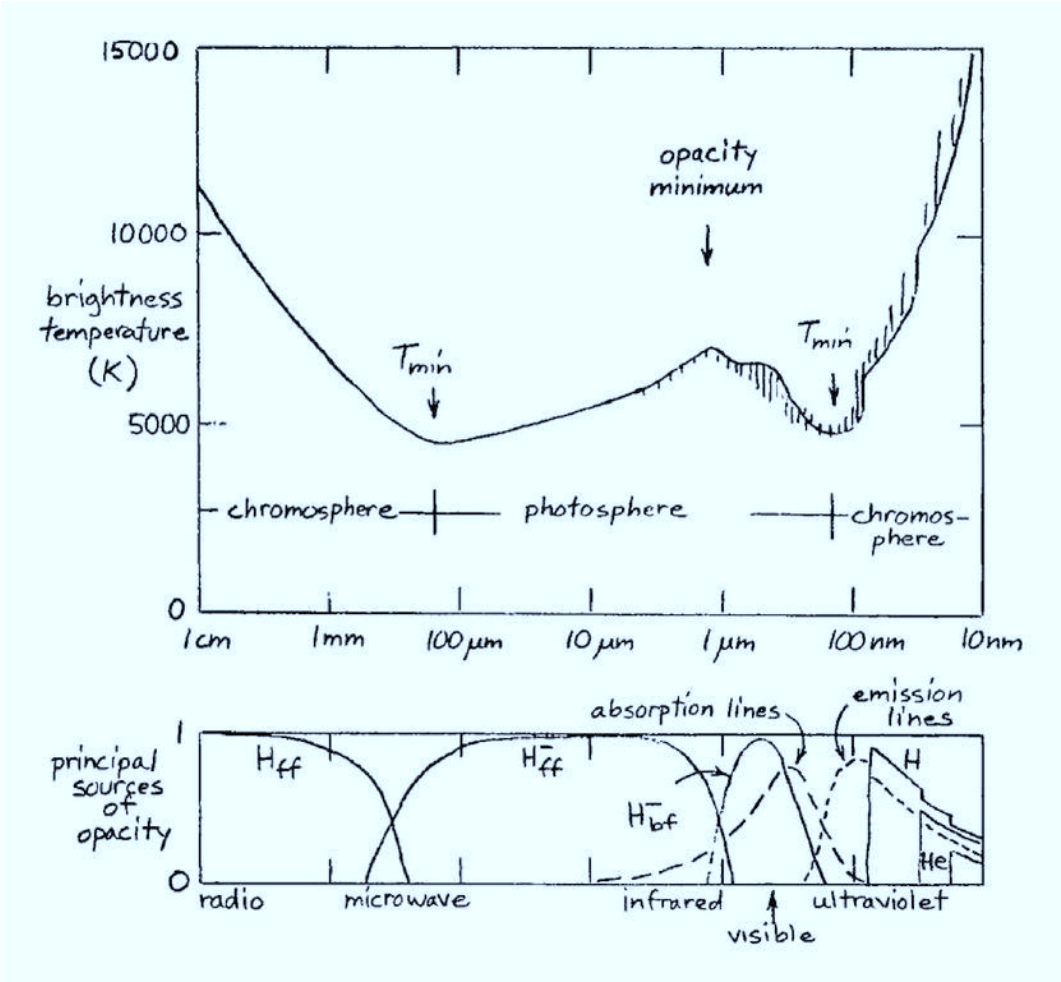


Figure 12: Brightness temperature of the VAL3C model for the Sun and the relative contribution of the principal sources of continuous extinction, always at the height where $\tau'_\nu = 1$, as a function of wavelength across the whole spectrum. From Avrett (1990).

Even in LTE the solar continuum intensities differ between different wavelength regions, first because different extinction samples different layers of the temperature stratification through the atmosphere; second, because the temperature sensitivity of the Planck function varies with wavelength.

Figure 10 shows the behavior of the continuous extinction coefficient with wavelength in the atmosphere. The curve changes somewhat in shape with z while the value at a given wavelength decreases roughly exponentially with z due to the density decay. In the ultraviolet at left, the extinction is dominated by bound-free edges of "metals" including Mg I, Al I, Si I and C I. These edges suffer from strong NLTE scattering. In the optical part of the spectrum, H^- bf provides most extinction, in near LTE because the energy jumps are relatively small. H^- ff takes over above $\lambda = 1.6$ μm in the infrared, with an LTE source function.

Figure 11 presents a computation of solar continuum formation in the visible

and near-infrared. The caption provides explanation. Note that the brightness temperature is defined as the formal temperature that reproduces the intensity when inserted into the Planck function, by $I_\nu \equiv B_\nu(T_b)$. In the case of an LTE total source function and validity of the Eddington-Barbier approximation, $T_b = T(\tau'_\nu = 1)$ for the disk-center intensity.

Figure 12 is an informative cartoon of the formation of the whole solar spectrum in terms of the brightness temperature emitted by the VAL3C standard atmosphere of Vernazza *et al.* (1981). The lower panel shows the fractional contributions to the opacity. The dashed curves represent the quasi-continuum contributed by the dense haze of spectral lines in the violet and ultraviolet. Optical depth $\tau'_\lambda = 1$ samples the temperature minimum of the VAL3C model at $\lambda = 160\,\mu$m and 160 nm. The deepest sampling is again at the H^- bf threshold at $1.6\,\mu$m. From there to $160\,\mu$m and 160 nm the sampling is increasingly higher in the photosphere, with the hump of the H^- bf opacity (Figure 10) causing a little dip in T_b.

5 NLTE formalisms

5.1 Einstein coefficients for bb processes

BETWEEN two energy levels u (upper) and l (lower) five processes are possible:
- spontaneous radiative de-excitation;
- radiative excitation;
- induced radiative de-excitation;
- collisional de-excitation;
- collisional excitation.

The same five processes occur in bf transitions. In ff transitions there is always collisional interaction, with kinetic energy loss or gain in the form of an emitted or absorbed photon. In the bb and bf processes the excitation or ionization energy operates as a temporal energy storage, acting as a memory in process sequences.

Spontaneous de-excitation:

$$A_{ul} \equiv \textit{transition probability for spontaneous de-excitation per second per particle in state } u. \tag{40}$$

The rate (number of de-excitations per second per cm^3) is given by $R_{ul} = n_u A_{ul}$ with n_u the density of the particles in state u (the *population*). The population diminishes as:

$$n_u(t) = n_u(0)\, e^{-A_{ul}t}.$$

For more de-excitation routes, the total damping constant is:

$$\Gamma_u \equiv \sum_l A_{ul}. \tag{41}$$

The mean lifetime of a particle in state u is Γ_u^{-1} s. With Heisenberg:

$$\Delta E = \hbar/\Delta t \approx \hbar\Gamma_u,$$

so that the energy spread per level is $\Delta\omega \approx \Gamma_u$. The corresponding emission profile function is the area-normalized *Lorentz profile:*

$$\psi(\nu-\nu_0) = \frac{\Gamma_u/4\pi^2}{(\nu-\nu_0)^2 + (\Gamma_u/4\pi)^2}. \tag{42}$$

Its wings fall off less steeply than a Gaussian, only quadratically: $\psi \sim 1/(\nu-\nu_0)^2$. This line broadening is called *radiation damping* or *natural line width*.

Radiative excitation:

$$B_{lu}\overline{J}^{\varphi}_{\nu_0} \equiv \textit{number of radiative excitations per second per particle in state } l. \tag{43}$$

The extinction profile function $\varphi(\nu-\nu_0)$ is employed in the frequency averaging of the angle-averaged exciting radiation field:

$$\overline{J}^{\varphi}_{\nu_0} \equiv \frac{\int_0^\infty J_\nu\,\varphi(\nu-\nu_0)\,\mathrm{d}\nu}{\int_0^\infty \varphi(\nu-\nu_0)\,\mathrm{d}\nu} = \int_0^\infty J_\nu\,\varphi(\nu-\nu_0)\,\mathrm{d}\nu. \tag{44}$$

Induced de-excitation:

$$B_{ul}\overline{J}^{\chi}_{\nu_0} \equiv \textit{number of induced de-excitations per second per particle in state } u, \tag{45}$$

analogous to B_{lu}, with

$$\overline{J}^{\chi}_{\nu_0} \equiv \int_0^\infty J_\nu\,\chi(\nu-\nu_0)\,\mathrm{d}\nu,$$

with $\chi(\nu-\nu_0)$ the area-normalized profile function for induced (stimulated) emission.

Collisional excitation and collisional de-excitation:

$$C_{ul} \equiv \text{number of collisional de-excitations per second per particle in state } u \tag{46}$$

$$C_{lu} \equiv \text{number of collisional excitations per second per particle in state } l. \tag{47}$$

The C's depend on the particle velocities and on the nature of the interaction. For example, for transitions of state i to state j via collisions with electrons:

$$n_i C_{ij} = n_i N_\mathrm{e} \int_{v_0}^\infty \sigma_{ij}(v)\, f(v) v\,\mathrm{d}v$$

with v_0 the threshold energy (minimum kinetic energy required is $(1/2)mv_0^2 = h\nu_0$), N_e the electron density, σ the collisional cross-section and $f(v)$ the velocity distribution (generally Maxwellian).

The *Einstein relations* are:

$$\frac{B_{lu}}{B_{ul}} = \frac{g_u}{g_l} \quad \text{and} \quad \frac{A_{ul}}{B_{ul}} = \frac{2h\nu^3}{c^2}. \tag{48}$$

These are two equations with three unknown radiation coefficients. They hold in any circumstance. Similarly for the collisional Einstein relation:

$$\frac{C_{ul}}{C_{lu}} = \frac{g_l}{g_u}\, \mathrm{e}^{E_{lu}/kT}. \tag{49}$$

The latter is easily understood. If thermodynamical equilibrium holds, the collision rate upwards must be equal to the collisional rate downward ("detailed balancing"). Since in TE the population ratio $n_u/n_l = (g_u/g_l) \exp -h\nu/kT$ (Boltzmann), the above relation must then hold. Einstein reasoned that if it holds somewhere, it should hold everywhere, also outside TE. He used the same argument to establish relations (48) in his derivation of the Planck function.

5.2 Line emissivity and extinction

First spontaneous de-excitation:

A_{ul} = number of spontaneous de-excitations per second per particle in the upper level,

$n_u\, A_{ul}$ = number of de-excitations per second per cm^3,

$h\nu_0\, n_u\, A_{ul}$ = energy radiated per second per cm^3,

$h\nu_0\, n_u\, A_{ul}\, \psi(\nu-\nu_0)$ = energy radiated per second per cm^3 per Hz,

$h\nu_0\, n_u\, A_{ul}\, \psi(\nu-\nu_0)/4\pi$ = energy radiated per second per cm^3 per Hz per steradian.

The line emissivity due to spontaneous de-excitation is therefore:

$$j_\nu^{\mathrm{spont}} = \frac{h\nu_0}{4\pi}\, n_u\, A_{ul}\, \psi(\nu-\nu_0). \tag{50}$$

Now radiative excitation. The total energy in a volume dV that is extinguished by radiative excitation during dt is:

$$\begin{aligned} \mathrm{d}E^{\mathrm{tot}} &= -h\nu_0\, n_l\, B_{lu} \overline{J}^{\varphi}_{\nu_0}\, \mathrm{d}V\, \mathrm{d}t \\ &= -h\nu_0\, n_l\, B_{lu}\, \mathrm{d}V\, \mathrm{d}t \int J_\nu\, \varphi(\nu-\nu_0)\, \mathrm{d}\nu \\ &= -\frac{h\nu_0}{4\pi}\, n_l\, B_{lu}\, \mathrm{d}V\, \mathrm{d}t \iint I_\nu\, \varphi(\nu-\nu_0)\, \mathrm{d}\Omega\, \mathrm{d}\nu, \end{aligned}$$

thus the energy dE^{bundle} that is extinguished during a time dt in a given bundle with intensity I_ν, opening angle dΩ and bandwidth dν in volume dV is:

$$\mathrm{d}E^{\mathrm{bundle}} = -\frac{h\nu_0}{4\pi}\, n_l\, B_{lu}\, I_\nu \varphi(\nu-\nu_0)\, \mathrm{d}V\, \mathrm{d}t\, \mathrm{d}\Omega\, \mathrm{d}\nu,$$

and from $\mathrm{d}V = \mathrm{d}A\ \mathrm{d}s$ and the definitions of intensity and extinction coefficient:

$$\alpha_\nu^{\text{excitation}} = \frac{h\nu_0}{4\pi}\, n_l\, B_{lu}\, \varphi(\nu{-}\nu_0).$$

Now stimulated emission. This is actually more similar to radiative excitation than to spontaneous de-excitation, since it similarly depends on the availability of photons measured by $\overline{J_\nu}$. Consequently, stimulated emission is treated as "negative extinction", *i.e.*, as a negative correction to the extinction when defining the extinction per cm or per gram:

$$\alpha_\nu^l = \frac{h\nu_0}{4\pi}\left[n_l B_{lu}\varphi(\nu{-}\nu_0) - n_u B_{ul}\chi(\nu{-}\nu_0)\right] \tag{51}$$

and so $j_\nu^l \equiv j_\nu^{\text{spont}} = (h\nu_0/4\pi)\ n_u\, A_{ul}\, \psi(\nu{-}\nu_0)$ without adding stimulated photons.

With $\alpha_\nu^l, B_{lu}, B_{ul}$ and A_{ul}, we have now four parameters that describe how readily a bb transition occurs: the bb transition probability. Usually one employs a fifth parameter: the *oscillator strength* f. The term stems from the classical description of a spectral line as a harmonic oscillator, in which the extinction coefficient per particle $\sigma(\nu)$ is:

$$\sigma(\nu) = \frac{\pi e^2}{m_\mathrm{e} c}\,\frac{\Gamma/4\pi^2}{(\nu{-}\nu_0)^2 + (\Gamma/4\pi)^2} = \frac{\pi e^2}{m_\mathrm{e} c}\,\varphi(\nu{-}\nu_0)$$

with

$$\sigma \equiv \int_0^\infty \sigma(\nu)\,\mathrm{d}\nu = \frac{\pi e^2}{m_\mathrm{e} c} = 0.02654\ \mathrm{cm}^2\ \mathrm{Hz}.$$

The oscillator strength f_{lu} is introduced as a correction factor to this classical value, without correction for stimulated emission (conventionally not applied to the extinction coefficient per particle):

$$\sigma^l = \int_0^\infty \frac{\alpha_\nu^l}{n_l}\,\mathrm{d}\nu = \frac{h\nu_0}{4\pi} B_{lu} \equiv \frac{\pi e^2}{m_\mathrm{e} c} f_{lu}.$$

For resonance lines such as Lyα the classical oscillator is a good approximation so that $f_{lu} \approx 1$, and so the oscillator strength has a reasonable numerical size. Other permitted transitions have $10^{-4} \le f_{lu} \le 10^{-1}$; forbidden transitions have $f_{lu} \le 10^{-6}$.

5.3 Line broadening

The profile functions ψ, φ and χ are not set by Heisenberg radiation damping as in (42) alone. Collisional damping and Doppler motions also contribute line broadening. Collisional disturbances are typically described either as nearby quasi-static Coulomb charges that affect the term layout of the atom, or as fast hitters disrupting the wave-train emission. Usually they add Lorentzian tail widening as in (42). Doppler shifts are added by line-of-sight velocity of the absorbing or radiating atom, due to the thermal motions and

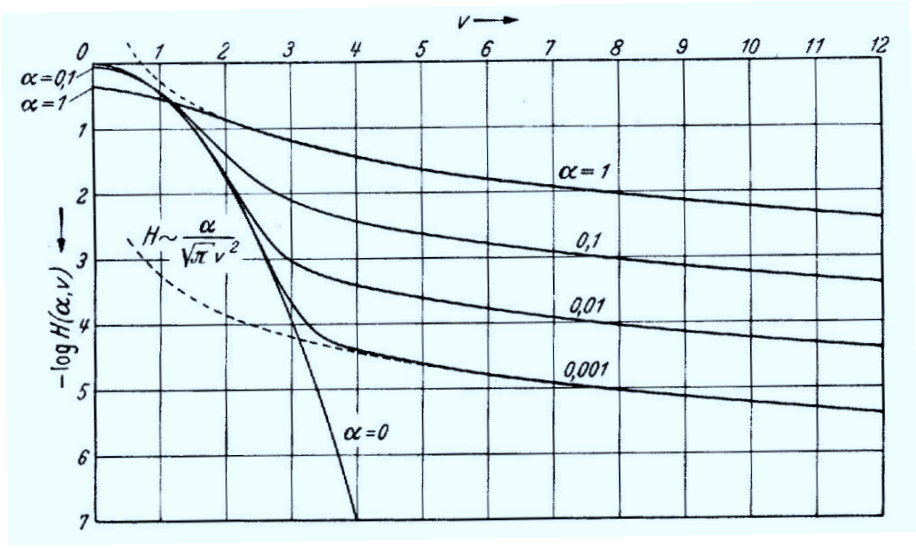


Figure 13: The Voigt function $H(a, v)$ for different values of the damping parameter a (written as α), on a logarithmic scale. From Unsöld (1955).

possibly to small-scale non-thermal motions if these are postulated to exist as so-called "microturbulence". Both have a Gaussian distribution, the thermal motions from the Maxwell distribution along the line of sight and the microturbulence through the definition of the Doppler width as

$$\Delta\nu_{\rm D} \equiv \frac{\nu_0}{c}\sqrt{\frac{2kT}{m} + \xi^2_{\rm micro}} \tag{52}$$

which expresses that convolution of two Gaussians produces a new Gaussian. It has full half-maximum width FWHM $= 2\sqrt{\ln(2)}\,\Delta\nu_{\rm D} = 1.665\,\Delta\nu_{\rm D}$.

The convolution of these Lorentzian and Gaussian components into the extinction coefficient per particle is:

$$\begin{aligned}\sigma^l_\nu &= \left[\frac{\sqrt{\pi}e^2}{m_{\rm e}c}\frac{f}{\Delta\nu_{\rm D}}\,{\rm e}^{-(\Delta\nu/\Delta\nu_{\rm D})^2}\right] * \left[\frac{\gamma/4\pi^2}{(\nu'-\nu_0)^2 + (\gamma/4\pi)^2}\right] \\ &= \frac{\sqrt{\pi}e^2}{m_{\rm e}c}\frac{f}{\Delta\nu_{\rm D}}H(a,v)\end{aligned} \tag{53}$$

with

$$H(a,v) \equiv \frac{a}{\pi}\int_{-\infty}^{+\infty}\frac{{\rm e}^{-y^2}}{(v-y)^2 + a^2}\,{\rm dy} \tag{54}$$

$$y \equiv \frac{\xi}{\xi_0} = \frac{\xi}{c}\frac{\nu_0}{\Delta\nu_{\rm D}} = \frac{\xi}{c}\frac{\lambda_0}{\Delta\lambda_{\rm D}} \tag{55}$$

$$v \equiv \frac{\nu-\nu_0}{\Delta\nu_{\rm D}} = \frac{\lambda-\lambda_0}{\Delta\lambda_{\rm D}} \tag{56}$$

$$a \equiv \frac{\gamma}{4\pi\Delta\nu_{\rm D}} = \frac{\lambda^2}{4\pi c}\frac{\gamma}{\Delta\lambda_{\rm D}}. \tag{57}$$

$H(a, v)$ is the *Voigt function*. The corresponding area-normalized extinction profile is:

$$\varphi(\nu-\nu_0) = \frac{H(a, v)}{\sqrt{\pi}\,\Delta\nu_\mathrm{D}}. \tag{58}$$

A rough approximation for $a \ll 1$ is simply the sum of the Gaussian and Lorentzian wing components:

$$H(a, v) \approx \mathrm{e}^{-v^2} + \frac{a}{\sqrt{\pi}\,v^2}. \tag{59}$$

When partial frequency redistribution occurs the different contributions play different roles. Scattering may be coherent at line center as seen by the atom, but the observer sees the atom's Doppler shift. Their ensemble average is described as redistribution in the frame of the observer. The Doppler core so becomes an ensemble of photons with redistribution across it. The inner damping wings of strong lines are dominated by natural broadening without redistribution. At each wavelength the photon ensemble may be scattering coherently, again over a Doppler-width-wide segment in the observer's frame. Thus, inner line wings may have their own formation independent of the rest of the line. However, the outer wings form deeper in the atmosphere where the densities are higher so that collisional damping dominates. This affects a redistribution over the outer line wings. I do not detail the redistribution formalisms here; see *e.g.*, Jefferies (1968) and Mihalas (1978).

5.4 Line source function

The *line source function* S_ν^l is now given by:

$$S_\nu^l \equiv j_\nu^l/\alpha_\nu^l = \frac{n_u A_{ul}\psi(\nu-\nu_0)}{n_l B_{lu}\varphi(\nu-\nu_0) - n_u B_{ul}\chi(\nu-\nu_0)}.$$

With the Einstein relations, for an arbitrary radiative transition:

$$\begin{aligned} S_\nu &= \frac{\dfrac{A_{ul}}{B_{ul}}\dfrac{\psi}{\varphi}}{\dfrac{n_l}{n_u}\dfrac{B_{lu}}{B_{ul}} - \dfrac{\chi}{\varphi}} \\ &= \frac{2h\nu^3}{c^2}\frac{\psi/\varphi}{\dfrac{g_u n_l}{g_l n_u} - \dfrac{\chi}{\varphi}}. \end{aligned} \tag{60}$$

In complete redistribution (scattering without memory, the photon resulting from de-excitation is not correlated with the exciting photon) the three frequency distributions are equal: $\phi(\nu-\nu_0) = \psi(\nu-\nu_0) = \chi(\nu-\nu_0)$, and

$$S_\nu^l = \frac{n_u A_{ul}}{n_l B_{lu} - n_u B_{ul}} = \frac{2h\nu^3}{c^2}\frac{1}{\dfrac{g_u n_l}{g_l n_u} - 1}. \tag{61}$$

In (L)TE the Boltzmann law $n_u/n_l = (g_u/g_l)\,\mathrm{e}^{-h\nu/kT}$ makes S_ν^l equal to B_ν.

5.5 LTE, SE, NLTE

For LTE (Local Thermodynamic Equilibrium) the level populations are given by the Saha-Boltzmann equations from the local temperature T_e[2]; the populations are therefore coupled to the local Maxwell distribution. The source function (60) simplifies with Boltzmann to:

$$S_\nu = \frac{2h\nu^3}{c^2}\frac{\psi/\varphi}{\frac{g_u n_l}{g_l n_u} - \chi/\varphi} \tag{62}$$

$$= \frac{2h\nu^3}{c^2}\frac{1}{\left(\frac{g_u n_l}{g_l n_u}\right)^{\mathrm{TE}} - 1} \tag{63}$$

$$= \frac{2h\nu^3}{c^2}\frac{1}{\mathrm{e}^{h\nu/kT} - 1} = B_\nu(T). \tag{64}$$

The correction for induced emission is in LTE

$$1 - \frac{n_u B_{ul}\chi(\nu-\nu_0)}{n_l B_{lu}\varphi(\nu-\nu_0)} = 1 - \mathrm{e}^{-h\nu_0/kT}. \tag{65}$$

The LTE line extinction coefficient is

$$\alpha^l_\nu = \frac{\pi e^2}{m_e c} n_l f_{lu} \varphi(\nu-\nu_0)\left[1 - \mathrm{e}^{-h\nu_0/kT}\right] \tag{66}$$

with n_l from Saha-Boltzmann and with the dimensionless oscillator strength f_{lu} defined by

$$\sigma^l = \int_0^\infty \frac{\alpha^l_\nu}{n_l}\,\mathrm{d}\nu = \frac{h\nu_0}{4\pi}B_{lu} \equiv \frac{\pi e^2}{m_e c} f_{lu}. \tag{67}$$

The essence of LTE is that collisions maintain the energy partitioning of the matter states more locally than the partitioning of the photons, so that the matter does not, but the radiation does, depart from the local conditions as parametrized by the temperature:

$$S^l_\nu(\vec{r}) = B_\nu\left[T(\vec{r})\right] \qquad I_\nu(\vec{r}) \neq B_\nu\left[T(\vec{r})\right] \qquad J_\nu(\vec{r}) \neq B_\nu\left[T(\vec{r})\right] \qquad \mathcal{F}_\nu(\vec{r}) \neq 0. \tag{68}$$

Assuming statistical equilibrium (SE) implies time-independence of radiation fields and level populations. The latter then obey the *statistical equilibrium equations:*

$$\frac{\mathrm{d}n_i(\vec{r})}{\mathrm{d}t} = \sum_{j\neq i}^{N} n_j(\vec{r})P_{ji}(\vec{r}) - n_i(\vec{r})\sum_{j\neq i}^{N} P_{ij}(\vec{r}) = 0 \tag{69}$$

with N the total number of levels which influence the population of level n_i in some way or other, and P_{ij} given by:

$$P_{ij} = A_{ij} + B_{ij}\overline{J}_{\nu_0} + C_{ij}. \tag{70}$$

[2]The definition of LTE is that the material population distributions obey the TE laws at the local temperature. Bowers and Deeming (1984) write nonsense about it.

were the bar in $\overline{J}_{\nu_0}$ implies frequency averaging over the line. Analogous equations describe bf and ff transitions and other similar processes, for example with $\overline{J}_{\nu_0}$ averaged over a bf series limit continuum.

The population equations (69) are coupled with the transport equations

$$\frac{\mathrm{d}I_\nu(\vec{r})}{\mathrm{d}\tau_\nu(\vec{r})} = S_\nu(\vec{r}) - I_\nu(\vec{r}) \tag{71}$$

for all frequencies ν and all beams (directions) which influence any population of interest. The transition probabilities P_{ij} in the statistical equilibrium equations depend on $\overline{J}_\nu$, and so on I_ν in all directions, whereas the optical thicknesses τ_ν and the source functions S_ν in the transport equations depend on the populations. This coupling is non-linear.

NLTE or non-LTE simply means that LTE does not hold. Usually one then adopts SE, the Maxwell distribution and complete redistribution (CRD). The populations may differ from the local Saha-Boltzmann value.

For "partial" redistribution (NLTE-PRD) the scattering in a line is partially coherent, with Doppler redistribution over the line core but monofrequent scattering in the inner line wings as mentioned above. In that case the line source function is frequency dependent, and monochromatic solutions of the statistical equilibrium equations are required with a *redistribution function* which specifies the amount of cross-talk between different parts of the line.

NLTE population departure coefficients b_i are defined by:

$$b_l = n_l/n_l^{\mathrm{LTE}} \qquad b_u = n_u/n_u^{\mathrm{LTE}} \tag{72}$$

with n the actual population and n^{LTE} the Saha-Boltzmann population. With these the line source function becomes:

$$S_\nu^l = \frac{2h\nu^3}{c^2} \frac{\psi/\varphi}{\frac{b_l}{b_u}\,\mathrm{e}^{h\nu/kT} - \frac{\chi}{\varphi}} \tag{73}$$

and for complete redistribution:

$$S_\nu^l = \frac{2h\nu^3}{c^2} \frac{1}{\frac{b_l}{b_u}\,\mathrm{e}^{h\nu/kT} - 1}; \tag{74}$$

in the Wien approximation ($h\nu \gg /kT$):

$$S_\nu^l \approx \frac{b_u}{b_l} S_\nu^{\mathrm{LTE}} = \frac{b_u}{b_l} B_\nu. \tag{75}$$

The line extinction coefficient becomes:

$$\alpha_\nu^l = \frac{h\nu_0}{4\pi}\, b_l\, n_l^{\mathrm{LTE}} B_{lu}\, \varphi(\nu-\nu_0) \left[1 - \frac{b_u\, n_u^{\mathrm{LTE}}\, B_{ul}\, \chi}{b_l\, n_l^{\mathrm{LTE}}\, B_{lu}\, \varphi}\right] \tag{76}$$

$$= \frac{h\nu_0}{4\pi}\, b_l\, n_l^{\mathrm{LTE}} B_{lu}\, \varphi(\nu-\nu_0) \left[1 - \frac{b_u}{b_l}\frac{\chi}{\varphi}\,\mathrm{e}^{-h\nu/kT}\right] \tag{77}$$

$$= \frac{\pi e^2}{m_\mathrm{e} c}\, b_l\, n_l^{\mathrm{LTE}} f_{lu}\, \varphi(\nu-\nu_0) \left[1 - \frac{b_u}{b_l}\frac{\chi}{\varphi}\,\mathrm{e}^{-h\nu_0/kT}\right] \tag{78}$$

with $\chi/\varphi = 1$ for complete redistribution. In the Wien approximation the correction factor for induced emission is negligible:

$$\alpha_\nu^l \approx b_l \left(\alpha_\nu^l\right)^{\mathrm{LTE}} \tag{79}$$

with $\left(\alpha_\nu^l\right)^{\mathrm{LTE}}$ the Saha-Boltzmann value.

5.6 Scattering by two-level atoms

As we saw in Section 2.2, the five bb processes may be combined in sequential pairs:

- *photon scattering:* radiative excitation followed by spontaneous or induced radiative de-excitation;
- *photon creation:* collisional excitation followed by radiative de-excitation;
- *photon destruction:* radiative excitation followed by collisional de-excitation.

These hold also for bf and other processes such as synchrotron radiation. At repeated scatterings, photons step around in a random walk. Per extinction and re-emission, they just change direction without information exchange and so without adjustment to local conditions. In photon creation and photon destruction, however, photons are coupled to the local collision energy, *i.e.,* coupled to the thermal pool via the Maxwell distribution.

For a gas made up of two-level atoms the above pairs are all that can happen. Let us additionally assume for the moment that both levels are infinitesimally sharp so that the profile functions are just δ functions at the line frequency ν_0. The total transition probability for de-excitation per excited particle and per second is:

$$P_{ul}^{\mathrm{tot}} = A_{ul} + B_{ul}J_{\nu_0} + C_{ul}. \tag{80}$$

The partial extinction coefficients per particle for photon destruction (also called "true absorption") and photon scattering are given by the corresponding fractions:

$$\sigma_{\nu_0}^{\mathrm{a}} = \sigma_{\nu_0}^{l} \frac{C_{ul}}{P_{ul}} \tag{81}$$

$$\sigma_{\nu_0}^{\mathrm{s}} = \sigma_{\nu_0}^{l} \frac{A_{ul} + B_{ul}J_{\nu_0}}{P_{ul}} \tag{82}$$

with $\sigma_{\nu_0}^{l} = \sigma_{\nu_0}^{\mathrm{a}} + \sigma_{\nu_0}^{\mathrm{s}}$.

For the extinction α_{ν_0} per cm and κ_{ν_0} per gram the stimulated emission $B_{ul}J_{\nu_0}$ is inserted as negative extinction. The stimulated contribution to scattering cancels, as shown already by the (c) and (g) vignette equality in Figure 1. This cancellation of the (c) contribution in the extinction is desirable because otherwise summing the (a) + (b) + (c) pairs would overestimate the actual opacity.

With the subtraction only the destruction part of the correction affects $\alpha^{\rm a}_{\nu_0}$ and only the spontaneous emission survives in $\alpha^{\rm s}_{\nu_0}$:

$$\alpha^{\rm a}_{\nu_0} = \alpha^{l}_{\nu_0} \frac{C_{ul}\left[1 - {\rm e}^{-h\nu_0/kT}\right]}{P_{ul}} \tag{83}$$

$$\alpha^{\rm s}_{\nu_0} = \alpha^{l}_{\nu_0} \frac{A_{ul}}{P_{ul}} \tag{84}$$

with $\alpha^{l}_{\nu_0} = \alpha^{\rm a}_{\nu_0} + \alpha^{\rm s}_{\nu_0}$. The source function for collisional processes is the Planck function (assuming the Maxwell velocity distribution):

$$j^{\rm a}_{\nu_0} = \alpha^{\rm a}_{\nu_0} B_{\nu_0}. \tag{85}$$

For pure scattering in a two-level atom gas the total (scattered) radiation equals the amount of radiation that is extinguished through scattering (not absorbed but redirected out of a given beam):

$$\int j^{\rm s}_{\nu_0}\,{\rm d}\Omega = \int \alpha^{\rm s}_{\nu_0} I_\nu\,{\rm d}\Omega \tag{86}$$

so that, with $J_\nu = (1/4\pi)\int I_\nu\,{\rm d}\Omega$ and the isotropy of scattering:

$$j^{\rm s}_{\nu_0} = \alpha^{\rm s}_{\nu_0} J_{\nu_0}, \tag{87}$$

expressing that every photon that is scattered into the beam was already a photon before. Together, with extinction weighting for the two processes:

$$S^{l}_{\nu_0} \equiv \frac{\sum j_{\nu_0}}{\sum \alpha_{\nu_0}} = \frac{\alpha^{\rm a}_{\nu_0} B_{\nu_0} + \alpha^{\rm s}_{\nu_0} J_{\nu_0}}{\alpha^{\rm a}_{\nu_0} + \alpha^{\rm s}_{\nu_0}}. \tag{88}$$

The transport equation

$${\rm d}I_{\nu_0} = -\alpha^{\rm a}_{\nu_0} I_{\nu_0}\,{\rm d}s - \alpha^{\rm s}_{\nu_0} I_{\nu_0}\,{\rm d}s + \alpha^{\rm a}_{\nu_0} B_{\nu_0}\,{\rm d}s + \alpha^{\rm s}_{\nu_0} J_{\nu_0}\,{\rm d}s \tag{89}$$

becomes (or rather remains), using ${\rm d}\tau_{\nu_0} \equiv \alpha^{l}_{\nu_0}\,{\rm d}s = (\alpha^{\rm a}_{\nu_0} + \alpha^{\rm s}_{\nu_0})\,{\rm d}s$:

$$\frac{{\rm d}I_{\nu_0}}{{\rm d}\tau_{\nu_0}} = \frac{{\rm d}I_{\nu_0}}{(\alpha^{\rm a}_{\nu_0} + \alpha^{\rm s}_{\nu_0})\,{\rm d}s} = S^{l}_{\nu_0} - I_{\nu_0}. \tag{90}$$

The *photon destruction probability per extinction* is

$$\varepsilon_{\nu_0} \equiv \frac{\alpha^{\rm a}_{\nu_0}}{\alpha^{\rm a}_{\nu_0} + \alpha^{\rm s}_{\nu_0}}, \tag{91}$$

in Einstein coefficients

$$\varepsilon_{\nu_0} = \frac{C_{ul}\left[1 - {\rm e}^{-h\nu_0/kT}\right]}{C_{ul}\left[1 - {\rm e}^{-h\nu_0/kT}\right] + A_{ul}} = \frac{C_{ul}}{A_{ul} + B_{ul}B_{\nu_0} + C_{ul}}, \tag{92}$$

where in the second version the Planck function appears rather than the actual induced rate $B_{ul}J_{\nu_0}$ due to the cancellation of induced scatterings in $\alpha^{l}_{\nu_o}$.

The two-level line source function becomes

$$S^l_{\nu_0} = (1 - \varepsilon_{\nu_0}) J_{\nu_0} + \varepsilon_{\nu_0} B_{\nu_0}. \tag{93}$$

For real lines with wider profile functions than the δ function this equation describes the monochromatic source function for *coherent scattering* (no change in frequency between incident and outgoing photon):

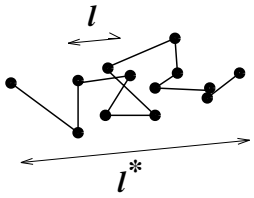
$$S^l_{\nu} = (1 - \varepsilon_{\nu}) J_{\nu} + \varepsilon_{\nu} B_{\nu}. \tag{94}$$

For example, it describes coherent Thomson or Rayleigh scattering in the presence of other processes that thermally produce photons at this frequency (say hydrogen Paschen bf transitions in the optical), and so it applies to the NLTE formation of hot-star continua.

For *complete redistribution* (per radiative de-excitation a new sampling of the extinction profile) a similar derivation gives:

$$S^l_{\nu_0} = (1 - \varepsilon_{\nu_0}) \overline{J}^{\varphi}_{\nu_0} + \varepsilon_{\nu_0} B_{\nu_0} \tag{95}$$

with $\overline{J}^{\varphi}_{\nu_0}$ given by (44). This equation also applies to bound-free ionization edges with complete redistribution over the edge profile if one applies similar frequency averaging to ε_ν and B_ν.



The *total path length* which a photon travels in N scatterings through a homogeneous layer is:

$$l^*_\nu \approx \sqrt{N}\, l_\nu, \tag{96}$$

with

$$l_\nu = \frac{<\tau_\nu>}{\alpha_\nu} = \frac{1}{\alpha^{\rm a}_\nu + \alpha^{\rm s}_\nu} \tag{97}$$

the mean free path per scattering (see eq. 22). The destruction probability per scattering is ε_ν; a photon travels on the average $N = 1/\varepsilon_\nu$ scattering steps between creation and destruction. The characteristic travel length between creation and destruction, (or identity maintenance path, diffusion length, thermalisation length, effective mean free path) l^*_ν of a photon is therefore

$$l^*_\nu \approx l_\nu / \sqrt{\varepsilon_\nu}. \tag{98}$$

A homogeneous layer with thickness D and optical thickness $\tau_\nu = \alpha_\nu D$ has *effective optical thickness*:

$$\tau^*_\nu = D/l^*_\nu = \sqrt{\varepsilon_\nu}\, \tau_\nu \tag{99}$$

and is *effectively thin* if $\tau^*_\nu < 1$ and *effectively thick* if $\tau^*_\nu > 1$. Thus, an object may well be optically thick but effectively thin – like fog shining with headlight photons.

The effect of small ε on the emergent line profile is large. A hypothetical isothermal homogeneous atmosphere (constant B_ν and ε_ν) has for coherent scattering (see problem 1.10 of Rybicki and Lightman, 1979 or Section 6.1 of Mihalas, 1978):

$$S_\nu(\tau'_\nu = 0) = \sqrt{\varepsilon_\nu}\, B_\nu. \tag{100}$$

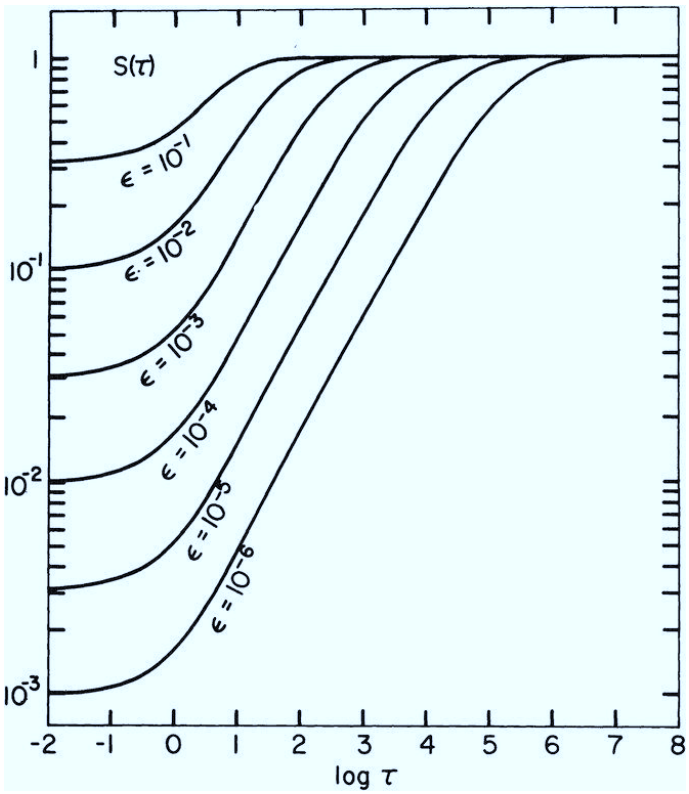


Figure 14: The ratio S/B for two-level scattering with complete redistribution in a homogeneous isothermal atmosphere, for different values of the collisional destruction probability ε (assumed constant with depth). For smaller ε the two-level source function $S = (1-\varepsilon)\overline{J}^{\varphi} + \varepsilon B$ follows the mean radiation field $\overline{J}^{\varphi}$ (not shown) more closely. $\overline{J}^{\varphi}$ and S uncouple from the Planck function $B = 1$ at larger depth for smaller ε; similarly to (96), this "thermalization depth" is given in τ units by $\Lambda \approx 1/\varepsilon$ for a Gaussian profile shape as assumed here (it would be even deeper, $\Lambda \approx 1/\varepsilon^2$, for a Lorentzian profile). This decoupling depth is far deeper than the $\log\tau = 0$ Eddington-Barbier depth. The surface value is given by (100) as $S(\tau = 0)/B = \sqrt{\varepsilon}$. The emergent intensity $I^+(\tau = 0) \approx S(\tau = 1)$ is nearly as small, far smaller than the Planck function. From the classic analysis by Avrett (1965).

The same result holds for scattering with complete redistribution as illustrated in Figure 14. Scattering lines with small ε_ν therefore fall far below the LTE value for the emergent line-center intensity. The physical reason is that the scattering reversedly propagates knowledge about the empty and cold universe outside the star to much deeper layers than $\tau'_\nu \approx 1$. The outside emptiness presents a NLTE photon sink, far from the local kinematic temperature. Radiation depart from a star implies NLTE loss of local conditioning. When scattering is important, that loss propagates inward over the effective mean free path ($l^* \approx l/\sqrt{\varepsilon}$ for coherent scattering, but $l^* \approx l/\varepsilon$ for complete redistribution over a Gaussian profile as in Figure 14). Below that depth there is yet no knowledge of this outside loss; there the radiation is still boxed in, thermally enclosed, and the intensity isotropic (in this isothermal case; in a real stellar atmosphere with inward increasing temperature there is slight linear anisotropy with $\overline{J_\nu}^{\varphi} = B_\nu$). At the surface, however, the radiation is peaked outward without backradiation, fully escaping out into the universe.

6 NLTE solar spectrum formation

6.1 Summary

IN summary: the emergent intensity is

$$I_\nu^+(\tau'_\nu = 0, \mu) = \int_0^\infty S_\nu(t'_\nu)\, \mathrm{e}^{-t'_\nu/\mu}\, \mathrm{d}t'_\nu/\mu, \tag{101}$$

the total source function is the weighted combination

$$S_\nu^{\mathrm{tot}} = \frac{j_\nu^{\mathrm{cont}} + j_\nu^{\mathrm{line}}}{\alpha_\nu^{\mathrm{cont}} + \alpha_\nu^{\mathrm{line}}} = \frac{S_\nu^c + \eta S_\nu^l}{1+\eta}, \tag{102}$$

the two-level line source function is for coherent scattering

$$S_{\nu_0}^l = (1-\varepsilon_{\nu_0})J_{\nu_0} + \varepsilon_{\nu_0} B_{\nu_0} \tag{103}$$

and for complete redistribution

$$S_{\nu_0}^l = (1-\varepsilon_{\nu_0})\overline{J}_{\nu_0}^\varphi + \varepsilon_{\nu_0} B_{\nu_0}, \tag{104}$$

with the frequency-averaged (or summed) mean intensity

$$\overline{J}_{\nu_0}^\varphi \equiv \frac{\int_0^\infty J_\nu\, \varphi(\nu-\nu_0)\, \mathrm{d}\nu}{\int_0^\infty \varphi(\nu-\nu_0)\, \mathrm{d}\nu} = \int_0^\infty J_\nu\, \varphi(\nu-\nu_0)\, \mathrm{d}\nu, \tag{105}$$

and, finally, the angle-averaged intensity is

$$J_\nu(z) = \frac{1}{4\pi}\int_0^\pi I_\nu(z,\theta)\, 2\pi \sin\theta\, \mathrm{d}\theta = \frac{1}{2}\int_{-1}^{+1} I_\nu(z,\mu)\, \mathrm{d}\mu. \tag{106}$$

So, to compute $I_\nu(0,\mu)$ in (101) one needs to know $I_\nu(z,\mu)$ in (106) at all depths, in all directions, and at all frequencies which contribute, in whatever manner, in setting S_ν and τ_ν.

In addition, actual atoms have more than two levels, so that the radiation field must be evaluated for any bb transition that is multi-level coupled to the line of interest. Similarly, the populations and number densities of the line-causing atoms (or ions or molecules) are also controlled by the pertinent amounts of ionization and recombination (dissociation and association). So these processes must also be included in detail, with similarly sophistication as above when they are out of LTE.

Such solving requires coupled solution of the population equations

$$\frac{\mathrm{d}n_i(\vec{r})}{\mathrm{d}t} = \sum_{j\neq i}^{N} n_j(\vec{r}) P_{ji}(\vec{r}) - n_i(\vec{r}) \sum_{j\neq i}^{N} P_{ij}(\vec{r}) \tag{107}$$

for all N relevant levels, with the rates for bb processes given by

$$P_{ij} = A_{ij} + B_{ij}\overline{J}_{\nu_0} + C_{ij} \tag{108}$$

and by similar expressions for bf and ff processes, simultaneously with the radiative transfer equations

$$\frac{\mathrm{d}I_\nu(\vec{r})}{\mathrm{d}\tau_\nu(\vec{r})} = S_\nu(\vec{r}) - I_\nu(\vec{r}) \tag{109}$$

for all frequencies ν and along all beams that take part in setting the relevant populations[3]. If statistical equilibrium cannot be assumed this whole problem gets time-dependent.

If LTE holds one simply puts $S_\nu^l = B_\nu$, forgets about such detailed solving, and does not read these notes beyond Section 4. LTE holds for transitions and locations with $\varepsilon_{\nu_0} \approx 1$. However, the stronger (and usually more interesting) lines tend to have $\varepsilon_{\nu_0} \ll 1$ for $z > z[\tau'^*_\nu \approx 1]$. Typically $\varepsilon_{\nu_0} \approx 10^{-4}$ near $\tau'_\nu = 1$. The same holds for the important bf edges of the electron-donor elements in the ultraviolet. They scatter as much as strong lines.

The above has here been formulated for the simplified case of axial symmetry, "1-dimensional" radiative transfer, in which the state parameters vary only along the z-axis. However, all of the physics above and the basic problems of radiative transfer are the same in more realistic 3-D situations. In 3D NLTE radiative transfer modeling one needs to solve for the radiation from any direction and at all frequencies at every spatial location. If the statistical equilibrium equations don't hold, one needs to solve the population equations and transport equations time-dependently (*i.e.*, solve for mass conservation). The research frontier is to do so within 3D time-dependent magnetohydrodynamic simulations of the solar atmosphere.

6.2 Example: VAL3C continua

Figure 15 copies two panels out of the extensive and informative collection of such figures for many more wavelengths in Figure 36 of Vernazza *et al.* (1981). The standard VAL3C model described in this monumental paper may not describe the real Sun, but it provides a perfect way to produce the time-averaged solar spectrum from a plane-parallel hydrostatic non-magnetic emulation of the Sun (that I call the computationally existing "star" VAL3C). This modeling obeys all the equations above, and so you should be able to explain every detail of the VAL3C spectrum formation, including these graphs.

In the lefthand figure the formation of the optical VAL3C continuum at $\lambda = 500\,\mathrm{nm}$ confirms the assumption of LTE in Figure 11: $S_\nu \approx B_\nu$ to large height. Notice the limb darkening (an atmosphere that produces absorption

[3]For example, in a recent analysis of NLTE departures in the solar infrared, Mats Carlsson and I constructed a model Mg atom with 66 levels (65 of Mg I, 1 of Mg II) and 315 Mg I lines, each sampled in 1–3 dozen frequencies, and solved the resulting equations for a mesh of some 70 z values (Carlsson *et al.*, 1992). Such setups used to be impossible; Mats' fast computer code (Carlsson, 1986) and the new workstations make them work. One iteration run took only 5 minutes on a DEC/5000 to reach a converged solution (but Mats ran it over 4000 times).

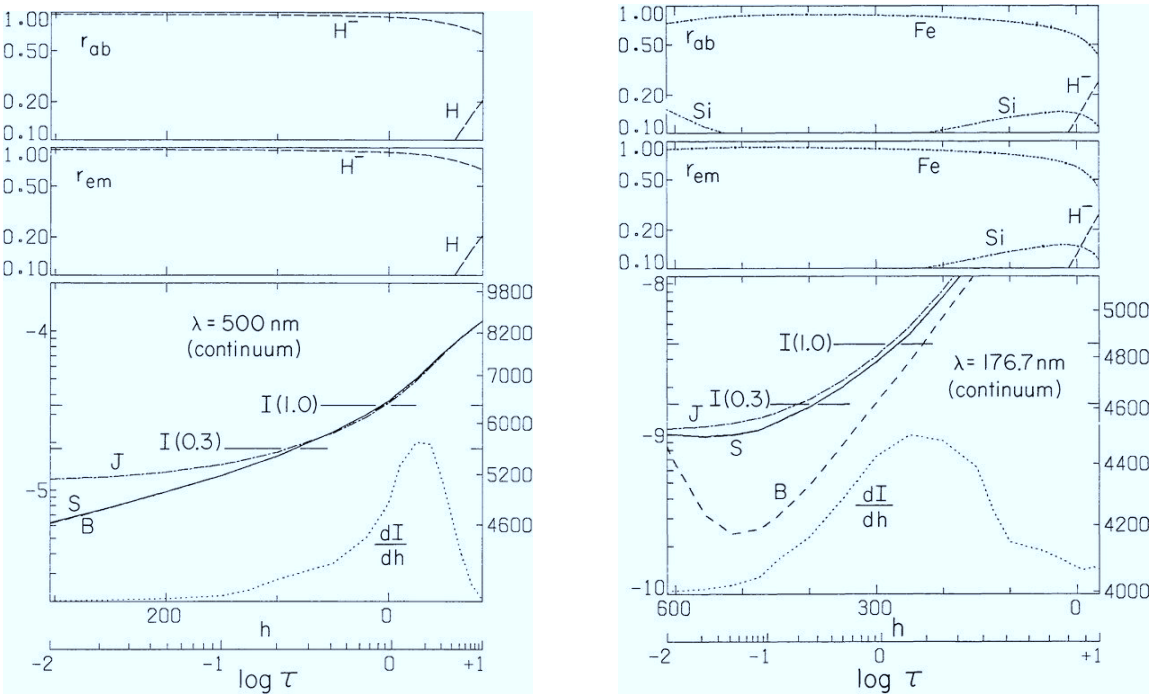


Figure 15: Continuum formation in the VAL3C atmosphere as function of height at $\lambda = 500\,\mathrm{nm}$ (left) and at $\lambda = 176,7\,\mathrm{nm}$ (right). The upper panels show the fractional contributions of the most important processes to the extinction (r_{ab}) and emissivity (r_{em}. At left H^- bf dointes, at right an Fe I ionization edge. The lower levels show J_ν, S_ν and B_ν in ergs cm^{-2} s^{-1} Hz^{-1} ster^{-1}. The dashes specify the emergent intensity at disk center ($\mu = 1$) and near the limb ($\mu = 0.3$). The dotted curve is the intensity contribution function $\mathrm{d}I/\mathrm{d}h = j_\nu \exp(-\tau_\nu)$. The $\log \tau_\nu$ scale is also given. At left LTE is an excellent approximation for the continuum source function but at right $S_\nu \approx J_\nu$ due to strong scattering in the Fe I edge.

lines from an outward declining temperature must also show limb darkening). The outward tail of the contribution function follows the density decay; the inward tail is set by the $\exp(-\tau_\nu)$ cutoff. The Eddington-Barbier approximation applies very well.

In the righthand figure the uncoupling of J_ν from B_ν (at the thermalization depth) occurs already outside the frame. At these short wavelengths the Planck function drops so steeply with temperature that it dips deep below J_ν on its way out towards the temperature minimum of the VAL3C atmosphere. This $J_\nu > B_\nu$ behavior is opposite to that in the isothermal atmosphere of Figure 14; in scattering, $J_\nu > B_\nu$ excesses arise for steep B_ν gradients. The scattering in the Fe I edge is so strong that S_ν follows J_ν (small ε_ν) and so the continuum is brighter than it would be in LTE, especially towards the limb. Since this is an important Fe I edge and the ionization follows the higher J_ν, Fe I is considerably overionized at this height[4]. Actual Fe I line opacities are smaller by about the J_ν/B_ν ratio than in LTE (equation (75)).

[4]But part of this overionization was undone in the FALC update by Fontenla *et al.* (1993) of the VAL3C model, since the outward temperature gradient became less steep. For detail see Rutten (1988).

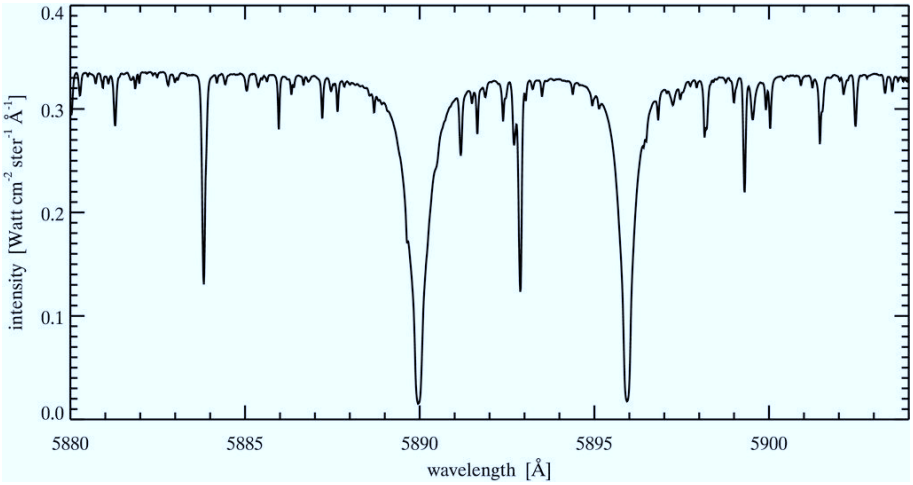


Figure 16: The Na I D lines in the solar disk-center intensity spectrum. These are the resonance lines (the strongest lines arising from the ground level) of Na I; the name "D" is due to Fraunhofer who alphabetized the more striking features in the solar spectrum. They correspond to the two transitions between the ground level and the first two excited levels of the neutral sodium atom, and are the same spectral lines which appear in the yellow sodium lamps along highways. These are the lines which gave Fraunhofer the idea that dark lines in the solar spectrum and bright lines in flame spectra might have something to do with each other, followed by the realization by Kirchhoff and Bunsen (while strolling the evening after they did sodium flame spectroscopy from Heidelberg on a fire in Mannheim) that solar line strength might be an indicator of solar sodium abundance, *i.e.*, that spectrometry works over distance. That was the birth of astrophysical spectroscopy. The solar sodium lines are in absorption: the solar brightness is lower in the lines than in the adjacent continuum.

The second figure is representative of VAL3C (and solar) continuum formation across the ultraviolet. The strong edges of the electron-donor elements that supply the ultraviolet opacity are well characterized as two-level scatterers, with the upper level a bound-free continuum rather than a bound one.

6.3 Example: the Na I D lines

Figure 16 shows the two yellow Na I D lines in the solar spectrum. Detailed numerical solution of the statistical-equilibrium equations for the excitation and ionization of sodium atoms in the solar atmosphere shows that the two-level approximation is valid for these resonance lines (Bruls, 1992, Bruls *et al.*, 1992). Therefore, the result in Figure 17 is readily understood. As the density drops, the collisional probability C_{ul} falls sharply with the height h; since A_{ul} is large for these resonance lines and does not depend on height, the photon destruction probability $\varepsilon \approx C_{ul}/A_{ul}$ also falls sharply with h; in deep layers $\varepsilon \approx 1$. The line source function follows $\overline{J}_{\nu_0}$ in the higher layers and approaches the Planck function only in the deep photosphere where $\varepsilon \approx 1$.

The line wings are formed in sufficiently deep layers that LTE holds for them (large density = frequent collisions). The line extinction α^l in the wings decreases monotonically with wavelength separation from line cen-

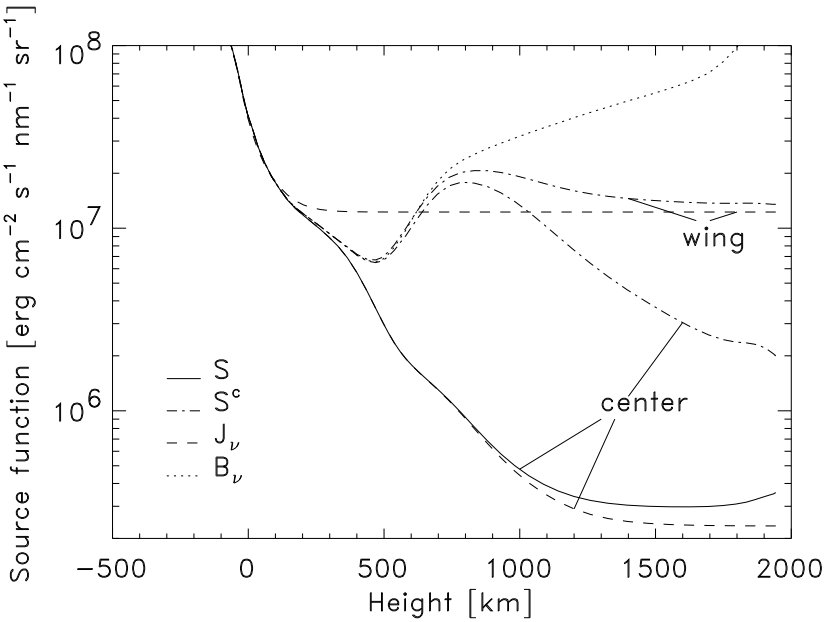


Figure 17: Formation of the Na I D_2 line in the solar spectrum. The source function follows the Planck function in the deepest layers; in higher layers the source function follows the angle-averaged intensity J_ν with $J_{\nu_0} < B_\nu$ due to photon losses. For the wings of the lines LTE holds to larger height until Thomson scattering becomes important. The line core, which has $\tau_\nu = 1$ near $h = 600$ km, is very much deeper than it would be in LTE. From Uitenbroek and Bruls (1992).

ter $\Delta\lambda = \lambda - \lambda_0$; the larger $\Delta\lambda$, the deeper the emergent intensity originates. Moreover, for sufficiently large $\Delta\lambda$, $\alpha^l \ll \alpha^c$: the continuum source function dominates the total source function and LTE formation is ensured because the H^-_{bf} extinction process is an LTE one.

The line cores have $\alpha^l \gg \alpha^c$. The cores are formed much higher, around $h = 600$ km; for them, the emergent intensity is determined in the regime where $S_\nu \approx \overline{J}_{\nu_0}$. This is much lower than the Planck function at that height. The source function (≈ line source function) is not influenced by the ambient temperature there; the existence of the temperature minimum does not affect the line source function or the emergent line profile. The line cores are entirely determined by the strong resonant scattering of photons formed deeper, and so their intensities drop much lower than they would under LTE. The drop is due to large photon losses in the line whose effect on the source function is noticeable until well below the $\tau' \approx 1$ formation depth, down to the $\tau'^\star \approx 1$ effective formation depth.

6.4 Example: the Ca II K line

Figure 19 sketches an extension of Figure 8 for the formation of the strong Ca II K line in the solar spectrum (Figure 18).

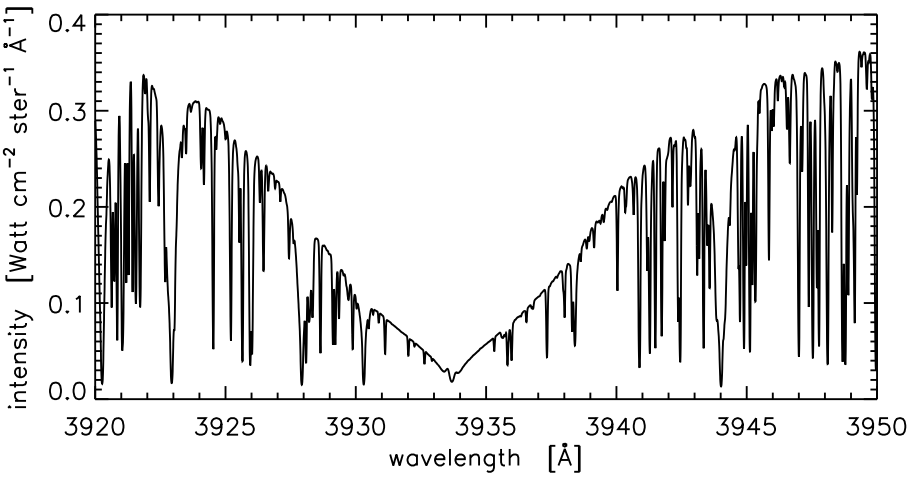


Figure 18: The Ca II K line in the solar spectrum. The broad deep absorption dip in this part of the solar spectrum was called "K" by Fraunhofer. It is a resonance line of the Ca^+ ion, and is the strongest line in the visually observable part of the solar spectrum. Superimposed on the broad line wings are many weaker spectral lines ("blends"); most arise from neutral "metals" such as Fe I. In its core the K line shows two minuscule peaks. These are extensively studied both for the Sun and for other cool stars, because they turn out to be excellent magnetometers. Their height is a good indicator for the number of magnetic fluxtubes that stick out of a star. In addition, their width present a very sharp luminosity indicator (the "Wilson-Bappu effect"). Why the peaks provide such good gauges is not understood.

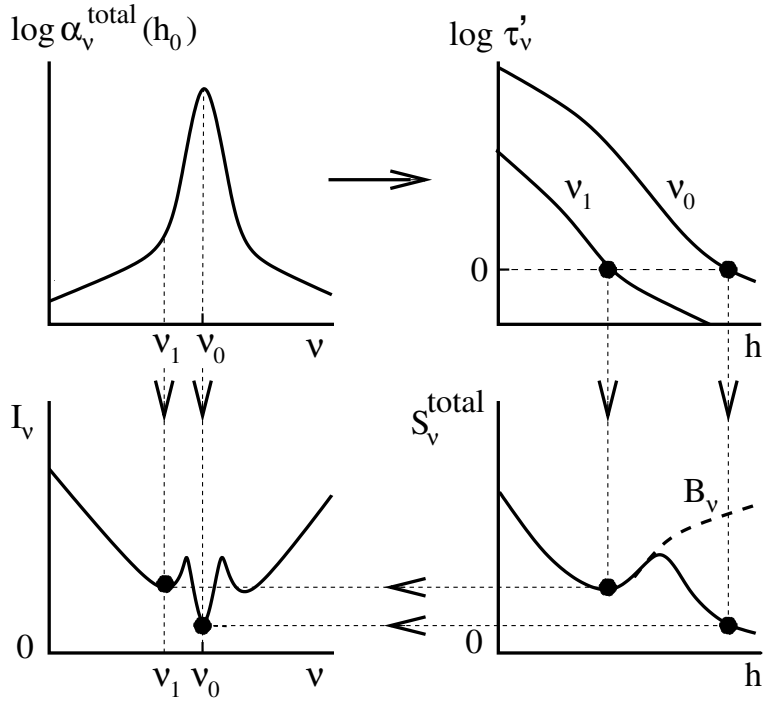


Figure 19: Cartoon for the formation of the Ca II K resonance line in the solar spectrum.

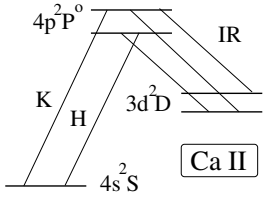
The extinction coefficient varies strongly with the wavelength because the bb processes offer an extra possibility for absorption and scattering. The size of the bb peak varies much with height, being dependent on the level popu-

lations which are sensitive to the density, the temperature and (in NLTE) the radiation field. The shape of the bb peak varies also with height, being dependent on the density (collisional damping) and on the temperature (collisional damping and Doppler broadening).

The extinction coefficient determines where the representative height of formation h, with $\log \tau'_\nu = 0$, lies. Each frequency has its own optical depth scale $\tau'_\nu(h)$, roughly exponential in h near $\log \tau_\nu(h) = 0$. The intensity $I_\nu(0, 1)$ of the emergent radiation is given in Eddington-Barbier approximation by the value of the monochromatic source function S_ν at the representative $\log \tau'_\nu = 0$ height.

The monochromatic (total) source function S_ν (bottom center) is the convolution of the continuum source function S^c and the line source function S^l (below right). (Because of this convolution the total source function is always frequency-dependent, even if the line source function S^l does not vary across the line profile which is the case for complete redistribution, assumed here.) Both source functions are determined by the temperature variation $T(h)$ and the amount of coupling to it, given by the destruction probabilities ε^c and ε^l which are small if scattering is dominant in the extinction. The line source function dominates in the Ca II K core because of the large amount of bb extinction in the layers where the radiation escapes; scattering is important there. In the deep photosphere where the outer line wings originate, the total source function is to a good approximation equal to the Planck function, because $\varepsilon^c_\nu \approx 1$.

The Ca II K line is just strong enough that S^l is sensitive to the temperature rise at the base of the chromosphere, before photon scattering losses win and decouple S^l from B_ν. The result is a small increase in S^l that is evident as two small emission peaks in the observed line profile (bottom left).



Photon conversion is likely from the Ca II H & K lines to the three "infrared" Ca II lines and vice versa, because they share common upper levels. The two lower levels of the infrared lines are metastable because there are no permitted radiative transitions from them to the ground level. Such photon conversion is not really important for the H & K lines because they have larger transition probabilities: the branching ratio from the common upper levels favors the resonance lines. Such conversion is more important for the three infrared lines because their extinction is smaller (Boltzmann): where they become optically thin, their line source function faithfully follows the Planck function because coupling occurs via conversion to the still optically thick H & K lines.

Actually, this diagram is old hat. It illustrates the formation of Ca II K according to Jefferies and Thomas (1960). With this analysis they started NLTE interpretation of spectral lines from the Sun and other stars. However, their description is by no means complete or final. For one thing, there is partial frequency redistribution so that the line source function itself is frequency dependent: different parts of the line each have their own source function,

which each in its own fashion differs from the Planck function (Uitenbroek, 1989). Furthermore, the small emission peaks at the K_2 wavelengths actually originate exclusively in regions on the Sun with an enhanced concentration of magnetic field. They turn out to be an excellent gauge of the amount of magnetic activity that cool stars display, correlating well with the amount of chromospheric and coronal emission. This sensitivity still needs to be explained. Furthermore, away from strong magnetic fields (in the internetwork areas between the strands of the magnetic network) the two peaks are asymmetric: the violet one is higher than the red one. I suspect that this is a signature of upward propagating and interfering acoustic shocks (Rutten and Uitenbroek, 1991), but this is not yet proven.

Thus, actual Ca II K line formation is much more complex than the one-dimensional explanation sketched here.

References

Allen, C. W.: 1976, *Astrophysical Quantities*, Athlone Press, Univ. London

Avrett, E. H.: 1965, in E. H. Avrett, O. J. Gingerich, and C. A. Whitney (Eds.), *The formation of spectral lines*, Procs. Second Harvard-Smithsonian Conf. on Stellar Atmospheres, Smithsonian Astrophys. Obs. Special Report 174, Cambridge, Mass., p. 101

Avrett, E. H.: 1990, in J.-O. Stenflo (Ed.), *Solar Photosphere: Structure, Convection and Magnetic Fields*, IAU Symposium 138 (Kiev), Kluwer, Dordrecht, p. 3

Böhm-Vitense, E.: 1989, *Introduction to stellar astrophysics. II. Stellar Atmospheres*, Cambridge Univ. Press, Cambridge UK

Bowers, R. L. and Deeming, T.: 1984, *Astrophysics I. Stars*, Jones and Bartlett, Boston

Bruls, J. H. M. J.: 1992, *Formation of diagnostic lines in the solar spectrum*, PhD thesis, Utrecht University

Bruls, J. H. M. J., Rutten, R. J., and Shchukina, N. G.: 1992, *Astron. Astrophys.* **265**, 237

Carlsson, M.: 1986, *A Computer Program for Solving Multi-Level Non-LTE Radiative Transfer Problems in Moving or Static Atmospheres*, Report No. 33, Uppsala Astronomical Observatory

Carlsson, M., Rutten, R. J., and Shchukina, N. G.: 1992, in M. S. Giampapa and J. A. Bookbinder (Eds.), *Cool Stars, Stellar Systems, and the Sun*, Proc. Seventh Cambridge Workshop, Astron. Soc. Pacific Conf. Series, Vol. 26, p. 518

Fontenla, J. M., Avrett, E. H., and Loeser, R.: 1993, *Astrophys. J.* **406**, 319

Gray, D. F.: 1992, *The Observation and Analysis of Stellar Photospheres*, Cambridge Univ. Press, U.K. (second edition)

Jefferies, J. T.: 1968, *Spectral Line Formation*, Blaisdell, Waltham, Mass.

Jefferies, J. T. and Thomas, R. N.: 1960, *Astrophys. J.* **131**, 695

Mihalas, D.: 1978, *Stellar Atmospheres*, W. H. Freeman and Co., San Francisco (second edition)

Novotny, E.: 1973, *Introduction to stellar atmospheres and interiors*, Oxford Univ. Press, New York

Rutten, R. J.: 1988, in R. Viotti, A. Vittone, and M. Friedjung (Eds.), *Physics of Formation of FeII Lines Outside LTE*, IAU Coll. 94, Reidel, Dordrecht, p. 185

Rutten, R. J. and Uitenbroek, H.: 1991, *Solar Phys.* **134**, 15

Rybicki, G. B. and Lightman, A. P.: 1979, *Radiative Processes in Astrophysics*, John Wiley & Sons, Inc., New York

Uitenbroek, H.: 1989, *Astron. Astrophys.* **213**, 360

Uitenbroek, H. and Bruls, J. H. M. J.: 1992, *Astron. Astrophys.* **265**, 268

Unsöld, A.: 1955, *Physik der Sternatmosphären*, Springer Verlag, Berlin (second edition)

Vernazza, J. E., Avrett, E. H., and Loeser, R.: 1981, *Astrophys. J. Suppl. Ser.* **45**, 635