

II Neutral Atomic Hydrogen (HI) Regions

This chapter discusses the physics of regions dominated by neutral (or only weakly ionized) atomic species. Since neutral atomic hydrogen is the dominant species, we generically refer to such gas as “Neutral Hydrogen” or “HI” Regions, but bear in mind that the gas also contains metals in neutral and weakly ionized forms that play an important role.

II-1 Interstellar UV & Visible Absorption Lines

The first observational evidence for an all-pervasive ISM came from observations of visible-wavelength absorption lines. These lines were the principal objects of general ISM studies before radio and space-borne observations became possible during the 1950s and later.

The strongest visible-wavelength absorption lines are:

$\text{NaI } 3^2\text{S} \rightarrow 3^2\text{P}^0_{1/2,3/2}$	$\lambda\lambda 5890, 5896\text{\AA}$	“D” lines of neutral sodium
$\text{CaII } 4^2\text{S} \rightarrow 4^2\text{P}^0_{1/2,3/2}$	$\lambda\lambda 3933, 3968\text{\AA}$	“H & K” lines of singly-ionized calcium

Both of these are resonance lines arising from the ground state in these ions. Other, weaker, visible lines of importance (all discovered in the 1930s and 40s) include TiII, CaI, KI, LiI, CH, NH, CN, CH^+ , and C_2 . Notice that there were about as many interstellar diatomic molecules known to early visible-wavelength studies as atomic species.

The first UV satellites (e.g., Copernicus) and later IUE and HST have observed strong UV absorption lines from the ISM. Because the typical excitation energies of ground-state resonance transitions are a few eV, most atomic species have resonance absorption lines (electric dipole transitions out of the ground state with $\Delta S=0$, $\Delta L=\pm 1$) in the near UV (longward of 1000\AA). These include:

MgII $\lambda 2800\text{\AA}$	HI Lyman series
(analog of CaII & NaI lines)	(primarily $\text{Ly}\alpha$)
OI, OII, OIII, OIV, OVI, OVII	H ₂ Lyman & Werner bands
CI, CII, CIII, CIV	

In addition, rare elements like Kr, Ga, Ge, As, Se, Sn, Te, Tl, Pb, Cu, Co, Mn, Zn, and Al are all seen in (weak) absorption lines. In general, the UV is the best place to study the gas-phase contents of the general ISM.

The **Diffuse Interstellar Bands** (DIBs) are the final and most mysterious of the UV/Visible absorption components of the ISM. Since their discovery by Merrill in 1938, about 200 DIBs have been identified in stellar spectra, with the strongest appearing at $\lambda 4430\text{\AA}$. They have not been identified conclusively with any atomic or molecular species (neutral or ionized). They are characterized by being extremely broad (by the standards of interstellar absorption lines). Some ideas are exotic molecular bands, transition from stuff on dust grain surfaces, exotica like ionized Fullerenes (3-D aromatic C molecules shaped like geodesic spheres), but none have produced consistent predictions of wavelengths.

Observations of Interstellar Absorption Lines

At high spectral resolutions ($R=\lambda/\Delta\lambda \geq 10^4$) interstellar absorption lines resolve into narrow absorption lines that are Doppler shifted relative to each other.

For example, towards the star ϵ Orionis, the NaI D lines resolve into 5 distinct radial velocity components with velocities of [+3, +11.3, +17.6, +24.7, +27.7] km s⁻¹ relative to the Local Standard of Rest (LSR). This observation by Adams in the 1940s was the basis of the “cloud picture” of the ISM. Some features of the clouds of absorption lines are that the strongest lines are associated with Galactic rotation, and that there is no dependence of line strength on Galactic longitude.

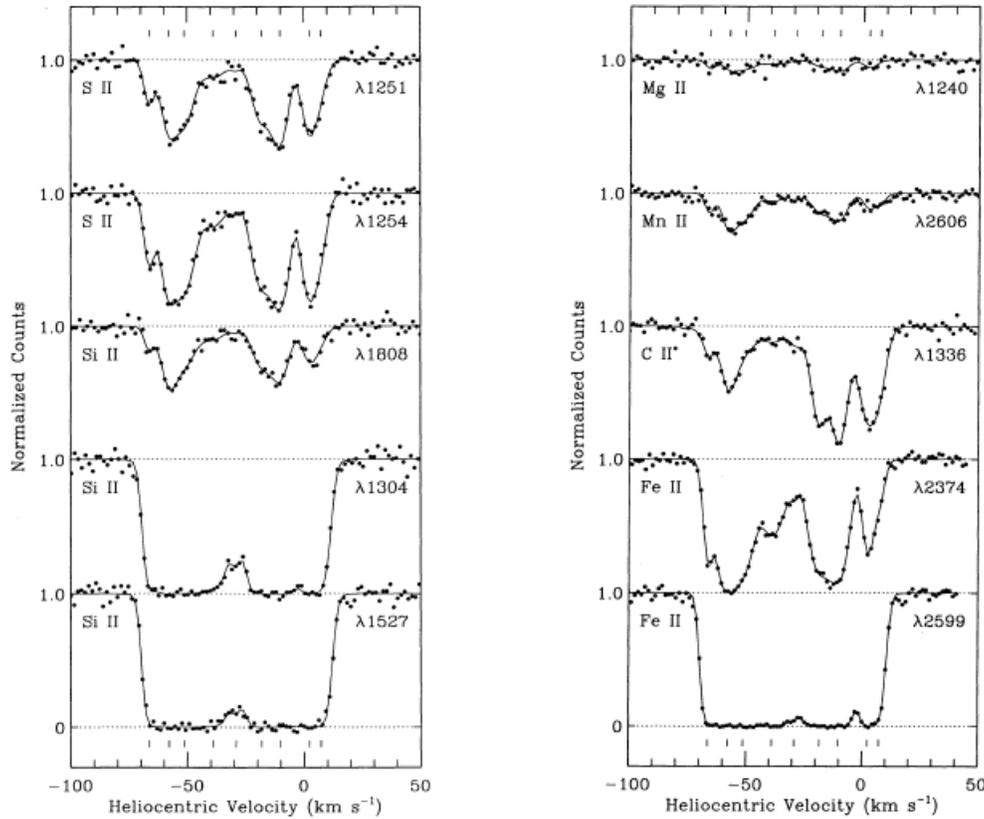
The distribution of cloud velocities with respect to the LSR seems to be empirically well-described by a simple exponential velocity distribution:

$$\psi(v) \propto e^{-|v|\sqrt{2}/\sigma}$$

where $\psi(v)$ is the number of clouds seen in the velocity range $(v, v+dv)$, and σ is the rms dispersion among cloud velocities

$$\sigma \equiv \langle v^2 \rangle^{1/2}$$

The observed dispersion is ~ 8 km s⁻¹, which is small compared to most populations of stars except for O and B stars.



Interstellar absorption lines towards the halo star HD93521 observed by Spitzer & Fitzpatrick [1993, ApJ, 409, 299] with HST and the Goddard High-Resolution Spectrograph (GHRS). These spectra show multiple velocity components and effects of line saturation in different species.

Radiative Transfer in Lines

The transfer equation is

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

The right-hand side consists of 2 parts: the **absorption term** $\kappa_\nu I_\nu$ and the **emission term** j_ν .

We can rewrite this in a useful form by defining the **Optical Depth**

$$d\tau_\nu = \kappa_\nu ds$$

and the **Source Function**

$$S_\nu \equiv j_\nu / \kappa_\nu$$

which give us:

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu$$

Integrating this equation with respect to optical depth gives the formal solution:

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

Depending on the circumstances, this is the form we will see used throughout this course.

In this section we are primarily interested in absorption and emission by atoms and molecules in the neutral phases of the ISM, so we need to introduce two useful quantities: the **line absorption coefficient**, κ_{lu} , and the **line emission coefficient**, j_{ul} .

Line Absorption Coefficient (κ_{lu})

The line absorption coefficient describes absorption of a photon by radiatively exciting an electron in an atom or molecule in a lower energy level (l) into an upper energy level (u):

$$\begin{aligned} \kappa_{lu} &= \int_{line} \kappa_\nu d\nu \\ &= n_l \int_{line} s_\nu d\nu = n_l s_{lu} \end{aligned}$$

Here we have introduced the **atomic absorption cross-section** $s_\nu = \kappa_\nu / n_l$. The line absorption coefficient has two components:

$$\kappa_{lu} = \frac{h\nu_{ul}}{c} (n_l B_{lu} - n_u B_{ul})$$

The first term on the right-hand side is the **rate of absorption** and the second term is the **rate of stimulated emission**. Since stimulated emission adds a photon back into the system, it enters as a “negative” absorption term – absorption because it involves interaction of the atom with the radiation field. This is to be distinguished from **spontaneous emission**, which is the electron spontaneously de-exciting independent of the incident radiation field.

The B 's are the **Einstein Coefficients** for absorption and stimulated emission, respectively. They are related via the **Einstein Relation** in thermal equilibrium

$$g_u B_{ul} = g_l B_{lu}$$

The Einstein B coefficients can in turn be written in terms of the **Einstein A Coefficient** for spontaneous emission (rate of radiative de-excitation) by

$$B_{ul} = \frac{c^3}{8\pi h\nu_{ul}^3} A_{ul}$$

where

$$A_{ul} = \frac{8\pi^2 e^2 \nu^2}{m_e c^3} f_{ul}$$

here f_{ul} is the **emission oscillator strength** of the transition which is related to the absorption oscillator strength via the statistical weights of the levels:

$$g_u f_{ul} = g_l f_{lu}$$

By convention there is no A_{lu} term.

This gives us an equation for the atomic absorption cross-section:

$$\begin{aligned} s_{lu} &= \int_{line} s_\nu d\nu = \int_{line} \frac{\kappa_\nu}{n_l} d\nu = \frac{h\nu_{ul}}{c} \left(B_{lu} - \frac{n_u}{n_l} B_{ul} \right) \\ &= \frac{h\nu_{ul}}{c} B_{lu} \left(1 - \frac{n_u}{n_l} \frac{g_l}{g_u} \right) \end{aligned}$$

Recall from Chapter I that the departure coefficients relate the true level populations (n) to the LTE level populations (n^*) via

$$n_l = b_l n_l^*$$

and the LTE level populations are related via the Boltzmann Equation:

$$\frac{n_u^*}{n_l^*} = \frac{g_u}{g_l} e^{-\Delta E_{ul}/kT}$$

Putting all the pieces together gives:

$$s_{lu} = s_{abs} \left(1 - \frac{b_u}{b_l} e^{-h\nu_{ul}/kT} \right)$$

Where we have defined the **integrated atomic absorption cross-section**, s_{abs} , to be:

$$s_{abs} \equiv \frac{h\nu_{ul}}{c} B_{lu} = \frac{\pi e^2}{m_e c} f_{lu}$$

Written this way, s_{lu} is just the integrated atomic absorption cross-section modified by a stimulated emission correction expressed in terms of the departure coefficients and an exponential in $h\nu/kT$, where T is the kinetic temperature of the system. The limiting behaviors are instructive:

Case 1: $h\nu_{ul} \gg kT$

In this limit the stimulated emission term vanishes and the line formation is dominated by pure absorption. Since in the ISM we expect most species to be in the ground state, very few species will be

in the excited state, contributing nothing to stimulated emission, so $s_{lu} \approx s_{abs}$. This is the situation for most ISM absorption lines at UV through near-IR wavelengths.

Case 2: $h\nu_{ul} \ll kT$

In this limit stimulated emission becomes very important. If we expand the exponential in the stimulated emission term to lowest non-trivial order in the exponent we get

$$s_{lu} = s_{abs} \left[1 - \frac{b_u}{b_l} \left(1 - \frac{h\nu_{ul}}{kT} \right) \right]$$

this can be re-arranged to give:

$$s_{lu} = s_{abs} \left(\frac{h\nu_{ul}}{kT} \right) \left[\frac{b_u}{b_l} - \frac{kT}{h\nu_{ul}} \left(\frac{b_u}{b_l} - 1 \right) \right]$$

This hardly seems much of an improvement until you consider two extreme limiting cases:

In Local Thermal Equilibrium (LTE), $b_l = b_u = 1$ (by definition), and we find

$$s_{lu} = s_{abs} \left(\frac{h\nu_{ul}}{kT} \right)$$

i.e., the effective pure absorption cross section is reduced by a factor of $h\nu/kT$ by stimulated emission. For example, for HI 21cm absorption at $T=100K$, $h\nu/kT \approx 6 \times 10^{-4}$.

In Extreme non-LTE:, $1 - b_l / b_u > h\nu_{ul} / kT$, so that

$$\left[\frac{b_u}{b_l} - \frac{kT}{h\nu_{ul}} \left(\frac{b_u}{b_l} - 1 \right) \right] < 0$$

Now the ‘‘absorption’’ term has become emissive! This condition occurs in a maser, when the level populations are driven so far out of TE that $n_u \gg n_l$.

Line Emission Coefficient (j_{ul})

The line emission coefficient describes radiative transitions from an upper excited level into lower levels:

$$j_{ul} = \int_{line} j_\nu d\nu$$

This is usually expressed in terms of the **Line Emissivity**:

$$4\pi j_{ul} = n_u h\nu_{ul} A_{ul}$$

This has units of $\text{erg s}^{-1} \text{cm}^{-3}$. The factor of 4π is the number of steradians in a sphere, removing the usual per-steradian unit from j_{ul} . The Einstein Coefficient A_{ul} is the radiative de-excitation rate for transitions from the upper to lower levels in units of s^{-1} .

This latter is the most common form in which we will express the line emissivity in a variety of situations in the ISM.

UV/Visible Absorption Line Formation

In Visible and UV interstellar absorption lines stimulated emission is unimportant because $h\nu_{ul} \gg kT$ for typical interstellar kinetic temperatures. These lines are therefore formed in the “pure absorption” limit, and the equation of radiative transfer has the simple solution:

$$I_\nu = I_{\nu,0} e^{-\tau_\nu}$$

Alternatively, we can express this in wavelength units; since UV and visible-light spectra are usually plotted in wavelengths (we’ll see frequency again in the radio regime):

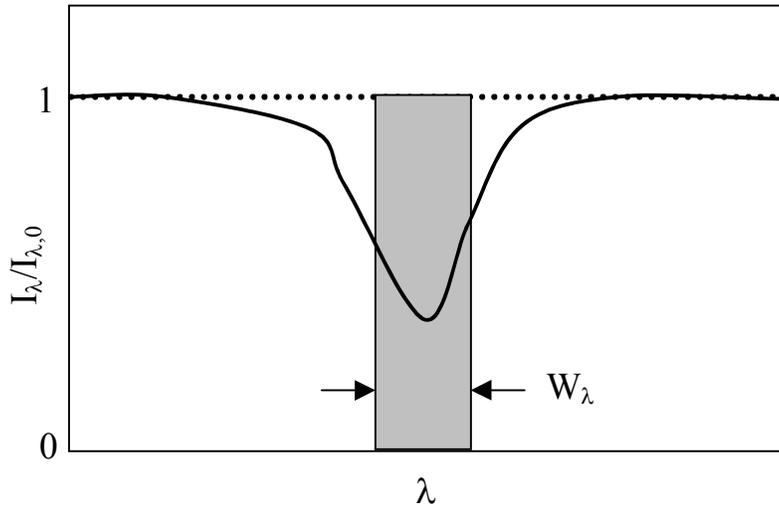
$$I_\lambda = I_{\lambda,0} e^{-\tau_\lambda}$$

Ideally, observation of an absorption-line profile can be turned into a measurement of the optical depth, τ_λ , for the line species. In practice, however, effects of finite instrumental resolution compared to the actual line width, limits on signal-to-noise, and so forth are such that we need to express the line strength in terms of an *integrated* observable, the **Equivalent Width**, W_λ , which is independent of spectral resolution:

$$W_\lambda \equiv \int \frac{I_{\lambda,0} - I_\lambda}{I_{\lambda,0}} d\lambda$$

Equivalent widths have units of Å or mÅ in the UV and visible bands. In words:

The Equivalent Width of an absorption line is the width that a line would have if it had a rectangular profile with zero intensity at the line center.



The “area” of the line here is defined as the integrated area of the absorption profile measured from the local continuum level, $I_{\lambda,0}$. Note the operative term here is “local”: equivalent widths are defined in terms of the unabsorbed continuum located immediately surrounding the absorption line of interest. In practice one does not normally measure the true “global” continuum shape of a spectrum, but instead estimates the local shape around the spectral features of interest. Note also that all information about the detailed line profile shape is lost in measuring an equivalent width (e.g., the line in the figure above is not symmetric).

In effect, this definition of the equivalent width “divides out” the spectrum of the background source. In practice, equivalent widths are measured by integrating the spectral line numerically after fitting a

local “pseudo continuum” using adjacent unabsorbed regions of the spectrum. If the spectrum is very complex (may stellar and/or interstellar absorption lines close to one another), defining this local continuum can be problematical. In general, uncertainty in how to measure the local continuum is the principal source of systematic error in measuring equivalent widths.

The equivalent width is an extremely useful quantity because what matters is the *fraction* of the light absorbed, not the total count of the photons absorbed, by the intervening material. By dividing out the spectrum of the background source, we have normalized the absorption-line profile. The equivalent width, then, measures the effective area of this normalized absorption-line, integrating over the detailed line profile shape. As a result, two stars with very different apparent brightnesses and intrinsic spectra, viewed along the same line-of-sight and path length through and the same interstellar cloud, will yield the same equivalent widths. In many cases, we will find that relative quantities are more useful to us than absolute measurements.

The Equivalent Width Curve of Growth

A traditional method of analyzing absorption line data is via the **Curve of Growth**.

Consider an atom at rest. The absorption coefficient, κ_λ , for a lower-to-upper transition will be:

$$\kappa_\lambda = \frac{\lambda_o^4}{8\pi^2 c} \frac{g_u}{g_l} A_{ul} \left[\frac{\gamma_u}{\gamma_u^2 + (\lambda - \lambda_o)^2} \right]$$

The term in []'s is the **Lorentzian** or “**Damping Profile**” that characterizes the natural broadening of the line due to quantum mechanical uncertainty. The damping width is characterized by the **radiation damping width**, γ_u , for the upper state is defined as the sum of all allowed downward radiative transitions out of that state:

$$\gamma_u \equiv \frac{\lambda_o^2}{4\pi c} \sum_{i < u} A_{ui}$$

The units of γ_u are typically given in Angstroms or microns as appropriate to the transition being considered. Each upper level has a different damping width.

An interstellar cloud is an ensemble of atoms all moving about with some combination of thermal and non-thermal motions (e.g., turbulence, bulk flows, etc.). The natural absorption-line profile is therefore Doppler broadened by the combination of all of these motions along the line of sight to the observer.

The distribution of line-of-sight velocities is $\psi(y)$, such that

$$\int \psi(y) dy \equiv 1$$

Here y is the dimensionless velocity parameter, $y=v/b$, the ratio of the line-of-sight velocity to the internal velocity dispersion, b . For a Maxwellian distribution of velocities in 3-space, the velocities project onto the line of sight with a Gaussian distribution:

$$\psi(y) = \frac{1}{\sqrt{\pi}} e^{-y^2}$$

For purely thermal motions with kinetic temperature T the **Doppler Velocity Dispersion** is:

$$b = \left[\frac{2kT}{m} \right]^{1/2}$$

Adding random turbulent velocities with a characteristic velocity dispersion of σ_{turb}^2 gives

$$b = \left[\frac{2kT}{m} + \sigma_{\text{turb}}^2 \right]^{1/2}$$

[NOTE: be careful not to confuse the dimensionless “b_j” departure coefficient with “b” the Doppler velocity dispersion. Sadly there are only 26 letters in the alphabet and so there has been considerable re-use – the key to survival in reading the ISM literature is to be careful of carrying forward definitions from one ISM phase to another and keeping aware of the context. The aggregation of a century of nomenclature doesn’t make it easy.]

The effect of the non-zero line-of-sight velocity of an individual atom is to Doppler shift the natural profile’s line center from λ_0 to $(1+v/c)\lambda_0$. The resulting optical depth, τ_λ , for the ensemble of atoms is the average of the individual κ_λ ’s over $\psi(y)$, hence:

$$\tau_\lambda = N_l \int_{-\infty}^{+\infty} \kappa_\lambda(y) \psi(y) dy$$

Here N_l is the **column density** of atoms with electrons populating the lower state out of which we observe an absorption line:

$$N_l = \int_0^L n_l(s) ds$$

This integral is taken along the line of sight (s) between the observer and the background source (e.g., a star) at distance L. Writing this out in full detail:

$$\tau_\lambda = N_l \int_{-\infty}^{+\infty} \frac{\lambda_0^4}{8\pi^2 c} \frac{g_u}{g_l} A_{ul} \frac{\gamma_u}{\gamma_u^2 + [\lambda - \lambda_0(1 + v/c)]^2} \psi(y) dy$$

This is rather complex, but we can simplify it by introducing four parameters:

$$\begin{aligned} y &= v/b \\ b_\lambda &= \frac{\lambda_0 b}{c} \\ u &= \frac{(\lambda - \lambda_0)}{b_\lambda} \\ a &= \gamma_u / b_\lambda \end{aligned}$$

These are the dimensionless velocity parameter, y , defined as before; the velocity dispersion, b , rewritten in units of wavelength, b_λ ; a dimensionless Doppler parameter, u , and the ratio of the natural (damping) width to the Doppler width, a . Substituting these definitions and assuming a Gaussian line-of-sight velocity distribution gives:

$$\tau_\lambda = \left[N_l \frac{\lambda_0^4}{8\pi^2 c} \frac{g_u}{g_l} \frac{A_{ul}}{b_\lambda} \right] \left[\frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{a^2 + (u - y)^2} dy \right]$$

The first term in []’s above is the optical depth at line center, τ_0 :

$$\tau_0 = N_l \frac{\lambda_0^4}{8\pi^2 c} \frac{g_u}{g_l} \frac{A_{ul}}{b_\lambda}$$

The second term in []s is the **Hjerting Function**, the convolution of the Lorentzian damping profile convolved with the Gaussian line-of-sight velocity distribution:

$$H(a, u) = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{e^{-y^2}}{a^2 + (u - y)^2} dy$$

The Hjerting function has no analytic solutions, and it is usually evaluated by numerical integration. We can, however, gain some useful insights by making a power-series expansion of $H(a, u)$ in a :

$$H(a, u) = H_0(u) + aH_1(u) + \dots + a^n H_n(u) + \dots$$

The first two terms are:

$$H_0(u) = e^{-u^2}$$

$$H_1(u) = \frac{1}{\sqrt{\pi}u^2}$$

$H_0(u)$ is a Gaussian profile describing the line core. $H_1(u)$ is the first “damping term” that describes the growth of the line wings (aka the “damping wings”) as the optical depth increases.

Recall that the equivalent width, W_{λ} , of the line is defined as:

$$W_{\lambda} = \int \frac{I_{\lambda,0} - I_{\lambda}}{I_{\lambda,0}} d\lambda$$

For pure absorption $I_{\lambda} = I_{\lambda,0}e^{-\tau_{\lambda}}$ this becomes

$$W_{\lambda} = \int 1 - e^{-\tau_{\lambda}} d\lambda$$

Substituting in the optical depth in terms of the Hjerting Function, we get the useful general form:

$$W_{\lambda} = b_{\lambda} \int 1 - e^{-\tau_0 H(a, u)} du$$

where this integral is evaluated over the line-of-sight *velocities* rather than over wavelength.

What we *measure* is W_{λ} , but what we want to derive is the optical depth with wavelength, τ_{λ} , which in turn measures how much of the given species is producing the absorption we see along the line of sight. This conversion is described by the **Equivalent Width Curve of Growth**.

There are two limiting cases of interest that describe the properties of the Curve of Growth:

Case 1: The natural width γ_u is much smaller than the Doppler width, b_{λ} ($a = \gamma_u/b_{\lambda} \leq 10^{-3}$).

If the optical depth at line center, τ_0 , is relatively small ($< 10^3$), then only the first term in the Hjerting function is important:

$$H(a, u) \approx e^{-u^2}$$

In this case the line profile is the Doppler (Gaussian) core with no significant contribution from the damping wings. There are two regimes of interest:

a) **Optically Thin** ($\tau_0 \ll 1$):

In this case, we can expand $e^{-\tau_{\lambda}}$ to lowest non-trivial order in τ_{λ} :

$$W_{\lambda} = \int 1 - e^{-\tau_{\lambda}} d\lambda \approx \int \tau_{\lambda} d\lambda$$

$$= b_{\lambda} \int_{-\infty}^{+\infty} \tau_0 e^{-u^2} du$$

Evaluating the integral, and dividing by the Doppler width, b_λ , leads to:

$$\frac{W_\lambda}{b_\lambda} = \tau_0 \sqrt{\pi}$$

In this case the equivalent width (in dimensionless units of the Doppler width) grows linearly with optical depth τ_0 . We call this the **Linear Part of the Curve of Growth**.

b) **Optically Thick, but** $\tau_0 < 10^3$:

The lines are optically thick, but not so optically thick (in the limit $a < 10^{-3}$) that the damping wings become important. In this case, we have:

$$\frac{W_\lambda}{b_\lambda} = \int_{-\infty}^{+\infty} 1 - e^{-\tau_0 e^{-u^2}} du$$

This equation is not analytically integrable, but the limiting behavior is

$$1 - e^{-\tau_0 e^{-u^2}} \rightarrow \begin{cases} 1 & \text{for } u \text{ small (line core "saturates")} \\ 0 & \text{for } u \text{ large} \end{cases}$$

The turnover point occurs when

$$\begin{aligned} \tau_0 e^{-u_0^2} &\approx 1, \text{ or equivalently:} \\ \ln \tau_0 - u_0^2 &\approx 0 \\ \therefore u_0 &\approx \sqrt{\ln \tau_0} \end{aligned}$$

Hence

$$\frac{W_\lambda}{b_\lambda} = 2u_0 + \dots \approx 2\sqrt{\ln \tau_0}$$

In this regime the equivalent width of the line grows very slowly even with large changes in the optical depth. We call this the **Flat Part of the Curve of Growth**.

Case 2: Large optical depth τ_0 and large column density N_I

In this case, we can no longer ignore contributions from the damping terms in $H(a,u)$. For example, HI Lyman series lines, like Ly α , have $a \approx 10^{-3}$ but it is very abundant so that the column density is very high. In this case

$$\tau_\lambda = \tau_0 H(a,u) = \tau_0 \left[e^{-u^2} + \frac{a}{\sqrt{\pi}u^2} + \dots \right]$$

The damping wings become important when

$$\frac{a}{\sqrt{\pi}u^2} \approx e^{-u^2}$$

For Ly α , damping becomes important for $u^2 \geq 9.8$ by this criterion.

Because the line core has become “saturated”, all subsequent “growth” of the line equivalent width will be due to the increasing contributions from the unsaturated line wings far from line center, hence:

$$\tau_\lambda \approx \tau_0 \frac{a}{\sqrt{\pi}u^2}$$

The equivalent width is then:

$$\frac{W_\lambda}{b_\lambda} = \int_{-\infty}^{+\infty} (1 - e^{-\tau_\lambda(u)}) du \approx \int_{-\infty}^{+\infty} (1 - e^{-\tau_0 a / \sqrt{\pi}u^2}) du$$

Rewriting this by setting $x^2 = \tau_0 a / \sqrt{\pi}u^2$ gives:

$$\frac{W_\lambda}{b_\lambda} = \left(\frac{\tau_0 a}{\sqrt{\pi}} \right)^{1/2} \int_{-\infty}^{+\infty} (1 - e^{-x^2}) d\left(\frac{1}{x}\right)$$

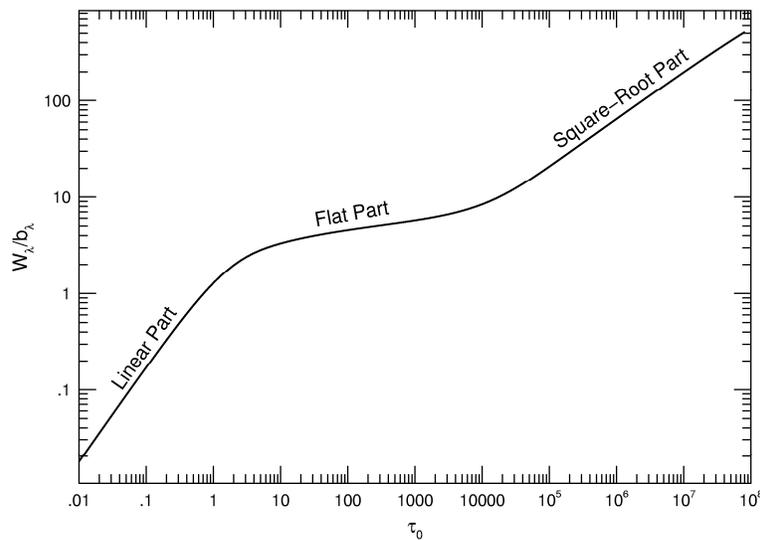
This is analytically integrable, hence

$$\frac{W_\lambda}{b_\lambda} = 2\sqrt{\pi} \left(\frac{\tau_0 a}{\sqrt{\pi}} \right)^{1/2}$$

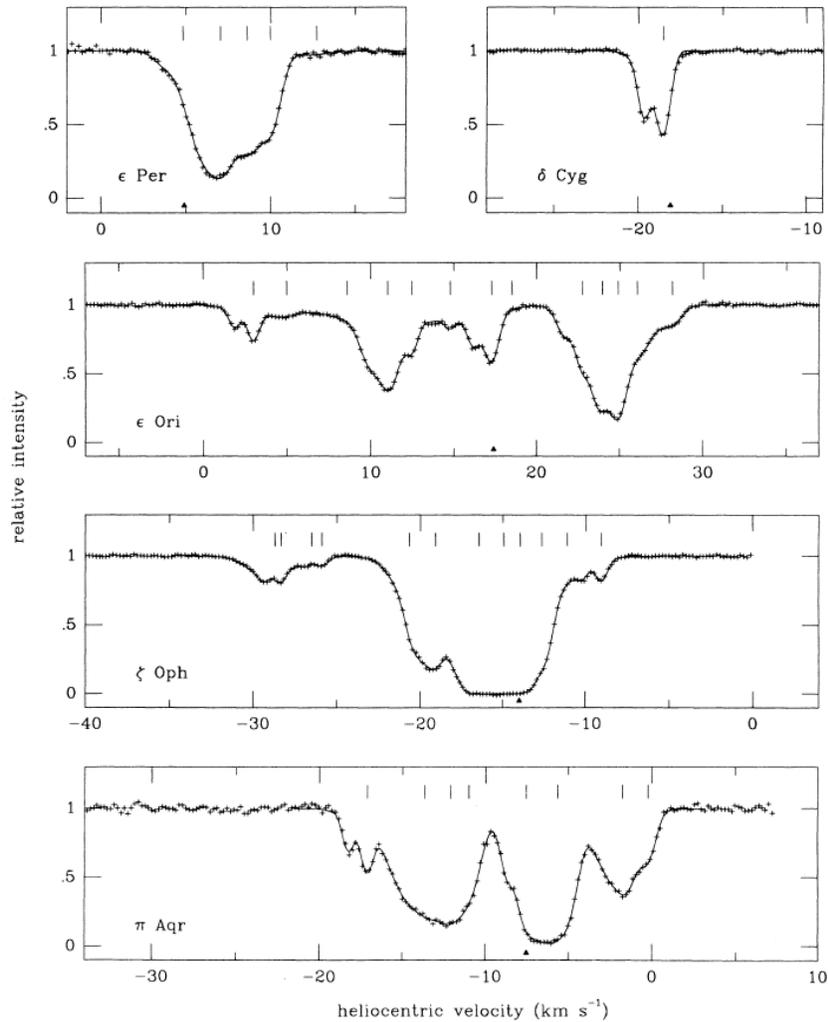
Here the equivalent width of the line grows like the square root of the optical depth and we say we are on the **Square-Root part of the Curve of Growth**.

These three regimes describe the behavior of the Equivalent Width Curve of Growth. At the transition regions, we of course need to evaluate the integrals numerically. A Curve of Growth computed numerically for interstellar HI Ly α is shown in the figure below, illustrating the three regions (linear, flat, and square root) derived above. In addition, we show spectra of the interstellar NaI D lines that exhibit the range of behaviors.

The most useful measurements occur at the extremes of the curve in the linear and square root parts of the curve. In the flat part, between the onset of saturation of the line core and the onset of significant growth of the damping wings, large changes in optical depth lead to very small changes in the measured equivalent width. Thus nearly saturated but undamped lines provide little useful data on the line-of-sight column densities.



Equivalent Width Curve of Growth for HI Ly α , showing the main regions derived in the text



Interstellar Na I-D absorption lines from Welty et al. [1994, ApJ, 436, 152]. These profiles show a mix of linear (δ Cyg), flat (ϵ Per & ϵ Ori), and square-root (ζ Oph) absorption lines.

Practical Considerations:

In practice, curve of growth methods are quite powerful, as they can relate the product $N\lambda f$ to a direct observable, W_λ , that is relatively insensitive to the choice of spectral resolution or fine details of the spectroscopic experiment. In principle, two different spectrometers working at very different resolutions and on different telescopes with different detectors should be able to measure the same equivalent widths to within the irreducible measurement uncertainties. However, because the equivalent width integrates over the detailed line profile shape, we do lose some information that might be useful.

Real interstellar absorption lines are often highly structured with a mixture of both saturated and unsaturated components because the line of sight to a particular star will often intersect a number of interstellar clouds with a wide range of column densities. While technically the equivalent width is insensitive to the spectral resolution (modulo effects of signal to noise which affects mainly the contrast of the line against the continuum), at lower spectral resolutions the saturated and unsaturated components will be blended, making interpretation of the composite line's equivalent width in terms of a single column density problematic at best.

In the case where lines are heavily saturated or show measurable damping wings (e.g., damped HI Lyman absorption systems), the equivalent width curve of growth method is unreliable. One cannot tell where the continuum should be placed, which leads to large systematic errors in measuring W_λ . In this case, various alternative methods have been used, for example the “Continuum Reconstruction Method” described by Bohlin et al. (1975, ApJ, 200, 402).

At very high spectral resolutions ($R > 10^4$), an alternative is to use the observed line profile and fit models to account for instrumental effects and the contributions from multiple absorption components with different column densities. The problem here is a lack of a complete database of UV spectra with sufficient resolution to employ these methods. One usually has to fall back on making assumptions about the intrinsic line profile shape.

The intermediate case occurs when the lines are fully resolved (i.e., when the line width is larger than 2–3 instrumental widths), but not necessarily resolved into fine velocity structure, a particularly powerful class of techniques has emerged to deal with these data. These techniques use direct integration of the observed **optical depth profiles**. Such methods make no a priori assumptions about either the detailed line shapes or the velocity distribution of the gas (unlike the case with the curve of growth, continuum reconstruction, and line profile modeling approaches).

A particularly good example of this type of analysis is the “Apparent Optical Depth Method” described by Savage & Sembach (1991 ApJ, 379, 245). This method does an excellent job of allowing discrimination of saturated structures in velocity space. When many different species are present in a spectrum, it provides a complete $N(v)$ profile by letting the different unsaturated parts fill in the gaps left by saturated species. This method provides a modern alternative to the traditional Curve of Growth analysis, and has been used in a number of recent absorption-line studies of the ISM.

Despite the practical caveats, the curve of growth still permits us to address a number of problems quantitatively in a way that illuminates what can be learned from IS absorption lines. The new methods lend us a greater degree of measurement precision, but no additional basic physical insight.

Applications of Interstellar Absorption Lines

1) Interstellar Ly α

The HI Lyman series lines are UV resonance lines that arise in radiative excitations out of the $n=1$ ground state of HI. First seen in the ISM in the early 1970s with the launch of the first UV satellites, they are seen along every line of sight, from early-type stars to late-type stars with chromospheric activity (you need some UV flux to observe them).

The optical depth of the HI Lyman series is very high, and on the damping (square-root) part of the Curve of Growth for most interstellar clouds, hence:

$$\frac{W_\lambda}{b_\lambda} = \left(\frac{\tau_0 a}{\sqrt{\pi}} \right)^{1/2} 2\sqrt{\pi}$$

The optical depth at line center, τ_0 , is given by

$$\tau_0 = N_l \frac{\lambda_0^4}{8\pi^{3/2} c} \frac{g_u}{g_l} \frac{A_{ul}}{b_\lambda}$$

From this result and the definition $a = \gamma_{ul}/b_\lambda$, the Doppler width, b_λ , cancels out, and we can write the equivalent width of the line as a function of the column density, N_l alone:

$$W_\lambda = \left(N_l \frac{\lambda_0^4}{2\pi c} \frac{g_u}{g_l} A_{ul} \gamma_u \right)^{1/2}$$

The γ_u 's are computed by summing over all the A's for the downward transitions out of the upper excited state of the particular absorption line of interest. For Ly α , which is the 2p \rightarrow 1s line, there is only 1 term in the computation of γ_{2p} (only one place for the electron to go, back to 1s). For Ly β (3p \rightarrow 1s) there are two terms (3p \rightarrow 1s and 3p \rightarrow 2s are possible radiative decay channels), and so forth. Putting in numbers for Ly α , we can solve for the column density along the line of sight as a function of the equivalent width:

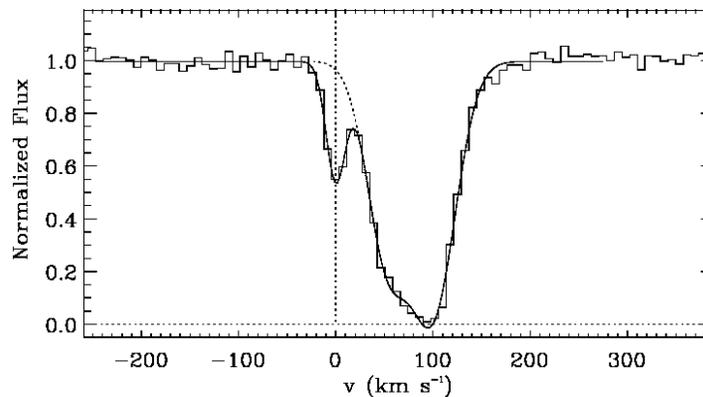
$$N_{Ly\alpha} = 1.867 \times 10^{18} W_\lambda^2 \text{ (cm}^{-2}\text{)}$$

for W_λ in units of Angstroms. Observations of stars at $d \approx 100$ pc give $W_\lambda(\text{Ly}\alpha) \approx 10 \text{ \AA}$, which implies a column density of Ly α absorbers of $N_{Ly\alpha} \approx 1.9 \times 10^{20} \text{ cm}^{-2}$. A line of sight of $100 \text{ pc} \approx 3 \times 10^{20} \text{ cm}$ implies a mean hydrogen column density of $\langle n_H \rangle \approx 0.6 \text{ cm}^{-3}$. For nearby stars, however, $\langle n_H \rangle \approx 0.1 \text{ cm}^{-3}$ or less is typical, indicating that we reside in a "local bubble" characterized by a lower average density.

We can also estimate the mean abundance of Deuterium in the ISM by measuring the Lyman-series absorptions from DI. The DI Lyman lines are shifted blueward of their respective HI counterparts by a small isotopic shift due to the neutron in the nucleus (the Rydberg constant is proportional to the reduced mass of the nucleus). The table below gives the isotopic shifts for the first 3 Lyman series lines of HI and DI:

	HI	DI	$\Delta\lambda$
Ly α	1215.67 \AA	1215.34 \AA	-0.33 \AA
Ly β	1025.72 \AA	1025.44 \AA	-0.28 \AA
Ly γ	972.54 \AA	972.27 \AA	-0.27 \AA

Since HI Ly α is so strongly saturated in the ISM, with very strong damping wings ($W_\lambda \approx 10 \text{ \AA}$), the DI line is lost in the saturated HI line core. You therefore need to measure the DI lines associated with higher-order HI Lyman lines, like Ly γ , Ly δ , etc., that have less strongly damped lines. For example, measurements of HI and DI absorption along the line of sight to β Centauri (where there is no detected H $_2$ absorption), one measures $D/H \approx (1.4 \pm 0.2) \times 10^{-5}$, smaller than 2×10^{-4} measured on Earth from the



Interstellar DI and HI Ly β absorption lines towards the white dwarf star WD0621-376 observed with FUSE. From Lehner et al. (2002, ApJS, 140, 81).

solar wind or ocean water. You can also measure HD molecular absorption bands (the analog of the H₂ Lyman bands), but it is hard to convert from HD/H₂ to D/H.

The Copernicus satellite (mid 1970s) was the first UV satellite that was sensitive to the higher-order Lyman series lines in the Far-UV (most UV satellites, except EUVE, were not sensitive below about 1100Å). Subsequent Far-UV studies have relied on either sounding rockets or short-duration space missions (e.g., ORFEUS-SPAS I and II, or IMAPS on the Shuttle). Lyman/FUSE (launched June 1999) is the first long-duration satellite mission since the Copernicus satellite 1970s to work in this Far-UV region (and is $\sim 10^5$ times more sensitive). So far, it has provided a number of good measurements of D/H in the local ISM from H_I and D_I Lyman series lines out to distances of 100pc from the Sun (and a few sight lines out to ~ 1 kpc with IMAPS). The data show a pretty constant $D/H \approx 1.2 - 1.7 \times 10^{-5}$, with the dispersion in values increasing with distance from the Sun.

2) ISM Gas-Phase Abundances

So many different species produce UV absorption lines in the ISM, from Hydrogen to rare metals, that we can get a pretty clear picture of the relative gas-phase abundances of the various elements from UV absorption-line studies. Among the scientific questions these permit us to address are chemical evolution of the ISM (in particular the mix of elements from different nuclear processes like α -process, neutron capture and proton capture) and the depletion of refractory elements onto dust grains.

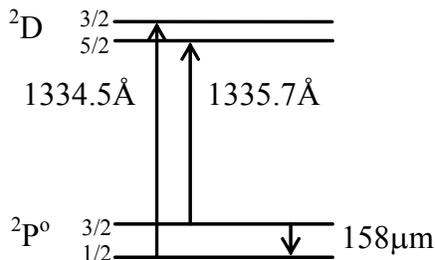
The current status of UV absorption-line abundances, especially as learned from the *Hubble Space Telescope* GHRS, has been reviewed by Savage & Sembach (1996, ARAA, 34, 279). A table of measured sight lines and the elements seen along them is reproduced on the following pages. The lines of sight to 2 stars, ζ Ophiuchus and ξ Persei, are particularly rich in interstellar absorption lines. These sight lines cross through regions of high column density, and many rare metal species (Cu, Zn, etc.) are very well studied there. Examples of spectra along the ζ Oph sight line are reproduced on the next page. Other sight lines have at least 5 or more different atomic species available for study.

A particular achievement of GHRS, which flew from 1990–1996, was the detection of unsaturated CII absorption (e.g., CII] $\lambda 2325$ Å absorption). In previous studies, only one line of sight, towards δ Scorpii, had unsaturated CII lines, the rest were all on the flat part of the curve of growth and produced no useful information (limits only). As a result, GHRS has permitted the first definitive interstellar gas-phase carbon abundances.

A particularly important result of the HST/GHRS data has been to greatly refine measurements of the depletion of gas-phase elements onto dust grains for sight lines that pass through both warm and cold phases of the ISM in both the disk and the halo of the galaxy. We will treat this subject in more detail later in the course when we discuss the properties of dust grains.

3) Thermal Balance

The CII $\lambda\lambda 1334.5, 1335.7$ Å doublet is produced by absorptions out of the C⁺ ground state, which is split into 2 fine structure levels, $^2P_{3/2}^o$ and $^2P_{1/2}^o$, into the same upper level. The strengths of these



absorption lines give the relative populations in each of these levels. This is of particular interest as transitions within this fine structure ground state are responsible for producing the [CII] λ 158 μ m emission line. Morton's 1975 observations of the CII λ 1334.5Å and 1335.7Å ratio showed a significant population of atoms in the $^2P_{3/2}^o$ state (the upper most of the fine structure level). This suggested that [CII] λ 158 μ m line could be a strong cooling line in HI regions. It was not until the 1980s that the Kuiper Airborne Observatory detected Far-IR [CII] emission from the ISM, as predicted by Morton.

The HI Hyperfine Structure Line

Most of our knowledge of the distribution of neutral atomic hydrogen in the ISM of the Milky Way and other galaxies come from observations of the strong 21-cm line. This line arises from transitions between the hyperfine structure levels in the ground state of Hydrogen, and is seen in both emission and absorption. This section reviews the physics of the HI hyperfine transition and of HI 21-cm line formation in both absorption and emission.

Hyperfine Splitting of HI: A little light quantum mechanics.

The $1s^2S_{1/2}$ ground state of HI is split by the interaction between the magnetic moments of the proton and electron spins.

The **electron magnetic moment** is:

$$\vec{\mu}_e = g_e \mu_B \vec{J}$$

where:

$$\begin{aligned} g_e &= \text{electron Lande factor} = -2.00232 \\ \mu_B &= \text{Bohr Magnetron} = \frac{e\hbar}{2m_e c} \\ \vec{J} &= \text{electron quantum number} = \vec{L} + \vec{S} \end{aligned}$$

The **proton magnetic moment** is:

$$\vec{\mu}_p = g_N \mu_N \vec{I}$$

where:

$$\begin{aligned} g_N &= \text{proton Lande factor} \cong 5.58 \text{ (experimentally)} \\ \mu_N &= \text{Nuclear Magnetron} = \frac{e\hbar}{2m_p c} \\ \vec{I} &= \text{proton quantum number, } |\vec{I}| = 1/2 \end{aligned}$$

A way to think of the Landé factors are that they express the difference between the classical and quantum descriptions of the magnetic moment of a sphere of uniform mass and charge density. The different Landé factors reflect the fact that the electron is a single particle with charge -1 , but the proton is a combination of 3 quarks with fractional charges that add up to $+1$. The quantization of the coupling of the μ 's results in the HI ground state being split into two states corresponding to parallel and anti-parallel spins of the electron and proton. The difference in energy, W , between the ground state and the hyperfine levels is given by:

$$W \propto \vec{\mu}_e \cdot \vec{\mu}_p$$

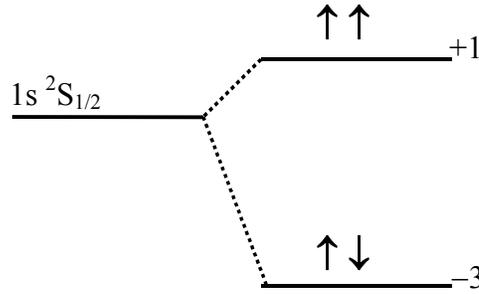
Following Condon & Shortly, for the $^2S_{1/2}$ terms, this is exactly

$$W \propto \frac{8}{3} \mu_e \mu_p \left[1, -\left(\frac{I+1}{I}\right) \right]$$

For HI, with $I=1/2$, this becomes

$$W \propto \frac{8}{3} \mu_e \mu_p [1, -3]$$

The result is that the hyperfine levels are split into two levels above and below the 1s ground level, as shown schematically below.



Schematic of the H α ground state hyperfine splitting, showing the electron and proton spins.

The hyperfine states are assigned a hyperfine quantum number, F:

$$\vec{F} = \vec{I} + \vec{J}$$

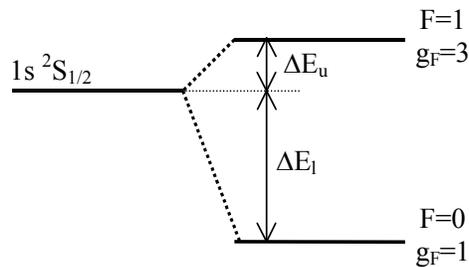
by analogy with L-S coupling for fine structure states. For the 1s $^2S_{1/2}$ state of H α is $F=(1/2\pm 1/2)=(1,0)$ for the upper and lower states, respectively. The hyperfine degeneracy, $g_F=(2F+1)$ is $g_F=(3,1)$ for upper and lower hyperfine states, respectively. The energy difference, ΔE , for a hyperfine state with quantum numbers (n,J,L,I,F) is given for any hydrogenic atom with nuclear charge Z and atomic number A by:

$$\Delta E = \frac{h\nu_0 Z^3}{n^3} \left[\frac{F(F+1) - I(I+1) - J(J+1)}{J(J+1)(2L+1)} \right]$$

$$\nu_0 = g_N \alpha^2 c R_A \left(\frac{m_e}{m_p} \right)$$

$$R_A = \frac{2\pi^2 e^4 m_e}{ch^3} \left(\frac{m_A}{m_A + m_e} \right)$$

These properties are illustrated schematically below:



Arrangement of the hyperfine energy levels in the H α ground state.

For the $F=1 \rightarrow 0$ transition in H α ($Z=1, A=1$), the total energy splitting ($\Delta E_u + \Delta E_l$) in frequency units has been measured in the lab using Hydrogen masers to be:

$$\nu = 1,420,405,751.786 \pm 0.010 \text{ Hz}$$

This is one of the few times you can quote a number in astrophysics to that kind of precision! In more convenient frequency and wavelength units, however:

$$\nu = 1420.4 \text{ MHz}$$

$$\lambda = 21.106 \text{ cm}$$

The $F=1 \rightarrow 0$ transition is a magnetic dipole transition, with transition probability

$$A_{10} = \frac{4\pi^2 e^2 h}{3m_e^2 c^2} \left(\frac{1}{\lambda^3} \right) \frac{1}{2F_1 + 1} \left| \langle \psi_0 | \vec{M} | \psi_1 \rangle \right|^2$$

$$\approx 2.85 \times 10^{-15} \text{ s}^{-1}$$

This relatively low transition probability gives a mean lifetime of the state of 1.1×10^7 years. The natural width of this transition is

$$\Delta \nu_{10} = \frac{A_{10}}{2\pi} \approx 4.5 \times 10^{-16} \text{ Hz}$$

Compared to the thermal Doppler width expected from an ensemble of HI atoms with kinetic temperature of 100K:

$$\Delta \nu_{th} = \nu_0 \sqrt{\frac{3kT}{mc^2}} \approx 4700 \text{ Hz}$$

For most astrophysical conditions, we expect the width of the HI 21 cm line (either in emission or absorption) to be dominated by the Doppler width due to thermal and/or turbulent motions.

Level Populations

We expect the level populations in the HI hyperfine levels to be far from LTE in most cases. As a starting point, the LTE populations are given by:

$$\frac{n_1^*}{n_0^*} = \frac{g_1}{g_0} e^{-h\nu/kT}$$

where “1” and “0” refer to the hyperfine quantum number, F , of the states, T is the kinetic temperature of the HI cloud, and ν is the frequency of the hyperfine transition. The non-LTE level populations are expressed in terms of an excitation temperature, T_S , which is traditionally called the “Spin Temperature” of the system:

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-h\nu/kT_S}$$

Putting in the values of the g 's, and using $h\nu/k = 68 \text{ mK}$, the transition energy in units of Kelvin:

$$\frac{n_1}{n_0} = 3e^{-0.068/T_S}$$

For most astrophysically interesting conditions, $T_S \gg 68 \text{ mK}$, and so we have to rely on subtle absorption effects in the line formation physics which, as we shall see, are very sensitive to small deviations in n_1/n_0 .

Optical Depth

In general, as we saw previously, the optical depth in a transition is given by

$$\tau_{\lambda} = N_0 \int \psi(y) \frac{\lambda_0^4}{8\pi^2 c} \frac{g_1}{g_0} A_{10} \frac{\gamma_{10}}{\gamma_{10}^2 + (\lambda - \lambda_0')^2} dy$$

where

$$\lambda_0' = \lambda_0 (1 - v/c)$$

Integrating over wavelength gives the integrated line optical depth commonly used in radio astronomy

$$\int_0^{\infty} \tau_{\lambda} d\lambda = N_0 \frac{\lambda_0^4}{8\pi c} \frac{g_1}{g_0} A_{10}$$

This equation relates total HI absorption to the column density of HI along the line of sight, N_0 .

Just to be sure of confusing everything, the traditional way of writing this in radio astronomy is to recast the distribution of velocities, $\psi(y)$, in frequency units:

$$\psi(y)dy = f(v)dv$$

and so we can write the optical depth at a particular wavelength, λ , as

$$\tau_{\lambda} = N_0 \frac{\lambda_0^4}{8\pi c} \frac{g_1}{g_0} A_{10} f(v) \frac{dv}{d\lambda}$$

since $dv/d\lambda = c/\lambda^2$ (eliminate the – sign by integrating backwards) we can write:

$$\tau_{\lambda} = N_0 \frac{3\lambda_0^2}{8\pi} A_{10} f(v)$$

or, finally purging the mixed wavelength/frequency units in favor of frequencies:

$$\tau_{\nu} = N_0 \frac{3}{8\pi} \frac{c^2}{\nu^2} A_{10} f(\nu)$$

where in all of the equations above, N_0 is the column density of HI in the F=0 (unexcited) hyperfine state. The g_F degeneracy values (3 and 1, respective for F=1,0) tell us that the mix of states in thermal equilibrium should be (modulo a T_S correction $\ll 1$):

$$N_0 = \frac{1}{4} N_{HI}$$

$$N_1 = \frac{3}{4} N_{HI}$$

So that:

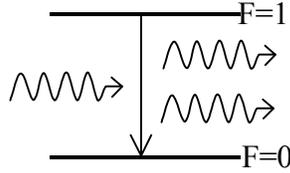
$$\tau_{\nu} = \frac{3}{32} \frac{c^2}{\pi \nu^2} N_{HI} A_{10} f(\nu)$$

This is the form of the HI 21cm optical depth most often used by radio astronomers.

Stimulated Emission

The calculation of the optical depth in the previous section was assuming the pure absorption case. However, while stimulated emission is insignificant at UV, optical, and near-IR wavelengths, it is very important at radio, mm, and (sometimes) far-IR wavelengths.

The effect of stimulated emission is for the radiation field to induce downward transitions from the upper excited states at a rate proportional to the local density of photons. This adds photons in the direction of the radiation field, making these photons coherent.



Schematic of Stimulated Emission

The intensity of the stimulated emission component, I_{SE} , is proportional to the intensity of the local radiation field:

$$I_{SE} \propto n_u B_{ul} I_\nu = n_1 B_{10} I_\nu$$

The absorption component is also proportional to the local radiation field:

$$\kappa'_\nu \propto n_l B_{lu} I_\nu = n_0 B_{01} I_\nu$$

The B's are related via the statistical weights:

$$g_u B_{ul} = g_l B_{lu}$$

So that, the net upward transitions (pure absorption out of F=0 less stimulated emission out of F=1 to F=0) is thus:

$$\begin{aligned} \kappa_\nu &\propto \left(n_0 B_{01} - n_1 \frac{g_0}{g_1} B_{01} \right) I_\nu \\ &\propto n_0 B_{01} I_\nu \left[1 - \frac{n_1}{n_0} \frac{g_0}{g_1} \right] \end{aligned}$$

Recall that the relative non-LTE level populations, n_1/n_0 , were written in terms of the spin temperature, T_S , as a Boltzmann-like equation

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-h\nu/kT_S}$$

hence

$$\kappa_\nu = \kappa'_\nu \left(1 - e^{-h\nu/kT_S} \right)$$

Recall now that $h\nu/k=68 \text{ mK} \ll T_S$, so expanding the exponential to lowest non-trivial order gives:

$$\kappa_\nu = \kappa'_\nu \left(\frac{h\nu}{kT_S} \right)$$

To account for stimulated emission, we must multiply the optical depth derived in the pure-absorption case (above) by a correction factor $h\nu/kT_S$, hence:

$$\begin{aligned}\tau_\nu &= \frac{3}{32} \frac{hc^2}{\pi\nu_0} \frac{N_{HI}}{kT_S} A_{10} f(\nu) \\ &= 7.95 \times 10^3 \int_{los} \frac{n_{HI} f(\nu)}{T_S} ds\end{aligned}$$

Here I've substituted the definition of the column density of HI, N_{HI} , in terms of the integral of the HI density, n_{HI} , along the line of sight

$$N_{HI} = \int_{los} n_{HI} ds$$

and put in the values of the physical constants, n_{HI} is in units of cm^{-3} , T_S is in K, and ds is in parsecs. Note that in general the HI density, spin temperature, and distribution of cloud velocities $f(\nu)$ are all functions of position, s , along the line of sight.

21-cm Line Formation

HI 21-cm Emission Lines

The equation of radiative transfer is

$$\begin{aligned}\frac{dI_\nu}{d\tau_\nu} &= -I_\nu + \frac{j_\nu}{\kappa_\nu}; \\ d\tau_\nu &= \kappa_\nu ds\end{aligned}$$

which has the solution:

$$\begin{aligned}I_\nu &= \int_0^\infty j_\nu e^{-\tau_\nu(s)} ds; \\ \tau_\nu(s) &= \int_0^s \kappa_\nu(s') ds'\end{aligned}$$

In radio astronomy, it is conventional to express the observed intensities in units of a *Brightness Temperature*, $T_B(\nu)$, because at $\lambda=21$ cm, we are in the Rayleigh-Jeans limit:

$$I_\nu = \frac{2k\nu^2}{c^2} T_B(\nu)$$

The radiation is not really blackbody, but the R-J limit allows us to express any I_ν in terms of an equivalent temperature for a blackbody that would give the same intensity. Brightness temperatures are actually what radio astronomers measure in their receivers at radio wavelengths, unlike the case in UV, visible, and IR wavelengths where (ideally) we count incident photons.

The emission coefficient expressed in temperature units (K/cm), J_ν , is given by

$$J_\nu = \frac{c^2}{2k\nu^2} j_\nu$$

Kirchoff's Laws relate κ_ν and J_ν through the spin temperature, T_S by:

$$J_\nu = \kappa_\nu T_S$$

Multiplying the transfer equation by a factor of $c^2/2k\nu^2$, and re-arranging gives:

$$\frac{dT_B}{d\tau_\nu} = T_S - T_B(\nu)$$

for the Rayleigh-Jeans limit. This equation has the solution:

$$\begin{aligned} T_B(\nu) &= \int_0^\infty T_S \kappa_\nu e^{-\tau_\nu(s)} ds \\ &= \int T_S e^{-\tau'_\nu} d\tau'_\nu \end{aligned}$$

where

$$\tau'_\nu(s) = \int_0^s \kappa_\nu(s') ds'$$

Thus τ_ν is the optical depth to infinity along the line of sight, while τ'_ν is the optical depth to distance s along the line of sight.

Assuming that T_S is constant along the line of sight (i.e., that we have an isolated isothermal cloud), the solution of the radiative transfer equation reduced to

$$T_B(\nu) = T_S (1 - e^{-\tau_\nu})$$

In general, however, both T_S and τ will be expected to vary with position, and solutions of the full transfer equation are required (they are not pretty).

Recall from our earlier discussion of UV and visible lines that in the limit $h\nu \ll kT$, we found

$$\kappa_\nu = n_l s_{lu} = n_l s_{abs} \left(\frac{h\nu}{kT_{exc}} \right)$$

where s_{abs} is the integrated atomic absorption cross-section for the transition. In this limit, the absorption is nearly completely compensated for by stimulated emission out of the excited state, so that the effective absorption coefficient depends on small variations of n_1/n_0 with spin (excitation) temperature, T_S . Numerically (from Kerr 1968 in Ch 10 of Middlehurst & Aller):

T_S (K)	n_1/n_0
10	2.9806
100	2.9981
1000	2.9998

These numbers show that as the spin temperature rises, the ratio of the hyperfine level populations approaches the ratio of the statistical weights (3.00). For the purposes of deriving column densities from HI line measurements, this means that assuming $n_1/n_0=3$ is a reasonable approximation.

Writing κ_ν in terms of the transition probability A_{10} :

$$\kappa_\nu = n_0 \frac{c^2}{8\pi\nu^2} A_{10} \frac{g_1}{g_0} \left(\frac{h\nu}{kT_S} \right) f(\nu)$$

Since $n_{\text{HI}}=4n_0$, the optical depth is:

$$\tau_\nu = \int_0^s \kappa_\nu ds' = \frac{N_{\text{HI}}}{4} \frac{c^2}{8\pi\nu^2} A_{10} \frac{g_1}{g_0} \left(\frac{h\nu}{kT_S} \right) f(\nu)$$

N_{HI} is the HI column density. Putting in the numbers for the HI 21-cm transition:

$$\tau_\nu = 2.59 \times 10^{-15} \frac{N_{\text{HI}}}{T_S} f(\nu)$$

Solving for N_{HI} :

$$N_{\text{HI}} = 3.88 \times 10^{14} \int_{-\infty}^{+\infty} T_S \tau_\nu g(\nu) d\nu \text{ cm}^{-2}$$

(the function $g(\nu)$ is the $1/f(\nu)$, but both are arbitrary placeholders for “line profile” in different units, so the notation can be similarly arbitrary). Since emission lines are often observed in frequency bins or “channels”, radio astronomers will often derive the column density in a particular frequency bin:

$$N_{\text{HI}}(\nu) = 3.88 \times 10^{17} T_S \tau_\nu \Delta\nu_{\text{kHz}} \text{ cm}^{-2}$$

Alternatively, one often sees the column density measured in a particular radial velocity bin:

$$N_{\text{HI}}(v) = 1.82 \times 10^{18} T_S \tau(v) \Delta v_{\text{kms}} \text{ cm}^{-2}$$

Here $\tau(v)$ is the optical depth in the radial velocity interval $[v, v+dv]$.

To measure the column density of HI along the line of sight, we need to measure T_S and τ_ν , but the only observable is $T_B(\nu)$. To see how optical depths, and hence column densities, are measured in practice, it is useful to examine two limiting cases:

Case 1: Optically Thin Limit, $\tau(v) \ll 1$

In this case,

$$T_B(v) = T_S \tau_\nu$$

so that

$$N_{\text{HI}}(v) = 1.82 \times 10^{18} T_B(v) \text{ in 1 km/sec bins}$$

or

$$N_{\text{HI}}(\text{tot}) = 1.82 \times 10^{18} \int_{\text{line}} T_B(v) dv$$

where the integral is often written as the “line intensity” expressed in units of “K km s⁻¹” (the somewhat odd unit you will find in many HI radio observation papers, especially older papers). Most extragalactic observations of the total HI content of galaxies determined from integrated 21-cm line measurements assume the optically thin case, and this is justified except for edge-on galaxies viewed in their mid-plane or large gas-rich merger remnants where lines of sight through the thickest parts of the galaxy are optically thick. It is also true of many, **but not all**, sightlines

through the disk of our own Galaxy – the exception being sight lines within a few degrees of the Galactic Center which are all optically thick.

Case 2: Optically Thick Limit, $\tau(\nu) \gg 1$

In this case

$$T_B = T_S$$

This happens because 21-cm photons emitted inside the cloud are absorbed within the cloud, and only those photons emitted from within 1 optical depth of the cloud surface escape. In this case the observed T_B is independent of column density, and depends on T_S . The column density must therefore become:

$$N_{HI} = 1.82 \times 10^{18} T_S \int_{-\infty}^{+\infty} \tau(\nu) d\nu$$

where now we must integrate the *optical depth* over velocity instead of integrating over the observed emission-line profile in brightness temperature.

Line Broadening

Because the natural width, $\Delta\nu_{10}$ is so small, Doppler broadening will dominate the line profile. This broadening takes two forms:

Thermal Broadening within a single cloud:

$$\Delta\nu_D^2 = 2 \left(\frac{kT}{m_H} \right)$$

in which case the optical depth profile will be:

$$\tau(\nu) = \tau_0 e^{-(\nu-\nu_0 / \Delta\nu_D)^2}$$

Bulk Motions of individual clouds:

$$\tau(\nu) = \sum_i \tau_{0,i} e^{-(\nu-\nu_0 / \Delta\nu_{D,i})^2}$$

Here the $\Delta\nu_{D,i}$ are the thermal Doppler widths of the individual clouds.

In practice, we often add an arbitrary turbulent velocity term in quadrature with the thermal term to specify the line width, much like what is done at UV and visible wavelengths with interstellar absorption lines.

The narrowest HI 21cm absorption lines seen in the Galaxy are roughly Gaussian in shape, but most emission lines are decidedly non-Gaussian, appearing as a superposition of many blended Gaussian components.

HI 21-cm Absorption Lines

Consider a background radio continuum source (e.g., a quasar or radio galaxy) with a brightness temperature T_{bg} . The solution of the equation of radiative transfer in the R-J limit is thus:

$$T_B(\nu) = T_{bg} e^{-\tau_\nu} + \int_0^{\tau_\nu} T_S e^{-\tau'_\nu} d\tau'_\nu$$

For a single, isothermal cloud, this reduces to

$$\begin{aligned} T_B(\nu) &= T_{bg} e^{-\tau_\nu} + T_S (1 - e^{-\tau_\nu}) \\ &= (absorption) + (emission) \end{aligned}$$

We have 2 unknowns to measure from these data, T_S and τ_ν . Three observational methods are used:

Method 1: Beam Switching

$T_B(\nu)$ is measured at two positions: on and off the source. The background source is also measured at adjacent line-free frequencies and interpolated to get T_{bg} at the line frequencies (“measuring the baseline”).

In the on-source beam, we measure

$$\begin{aligned}\Delta T_B(\nu)_{ON} &= [T_{bg} e^{-\tau_\nu} + T_S (1 - e^{-\tau_\nu})] - T_{bg} \\ &= T_B - T_{bg}\end{aligned}$$

In the off-source beam, we measure

$$\Delta T_B(\nu)_{OFF} = T_S (1 - e^{-\tau_\nu})$$

We can then solve these two equations for T_S and τ_ν :

$$\begin{aligned}\tau_\nu &= -\ln \left[1 - \frac{\Delta T_B(\nu)_{OFF} - \Delta T_B(\nu)_{ON}}{T_{bg}} \right] \\ T_S &= T_{bg} \left[\frac{\Delta T_B(\nu)_{OFF}}{\Delta T_B(\nu)_{OFF} - \Delta T_B(\nu)_{ON}} \right]\end{aligned}$$

The success of this beam-switching technique depends on the absorption region varying smoothly over very small scales, so that T_S and τ_ν are not substantially different on the on- and off-source beams.

Method 2: Observe Variable Sources

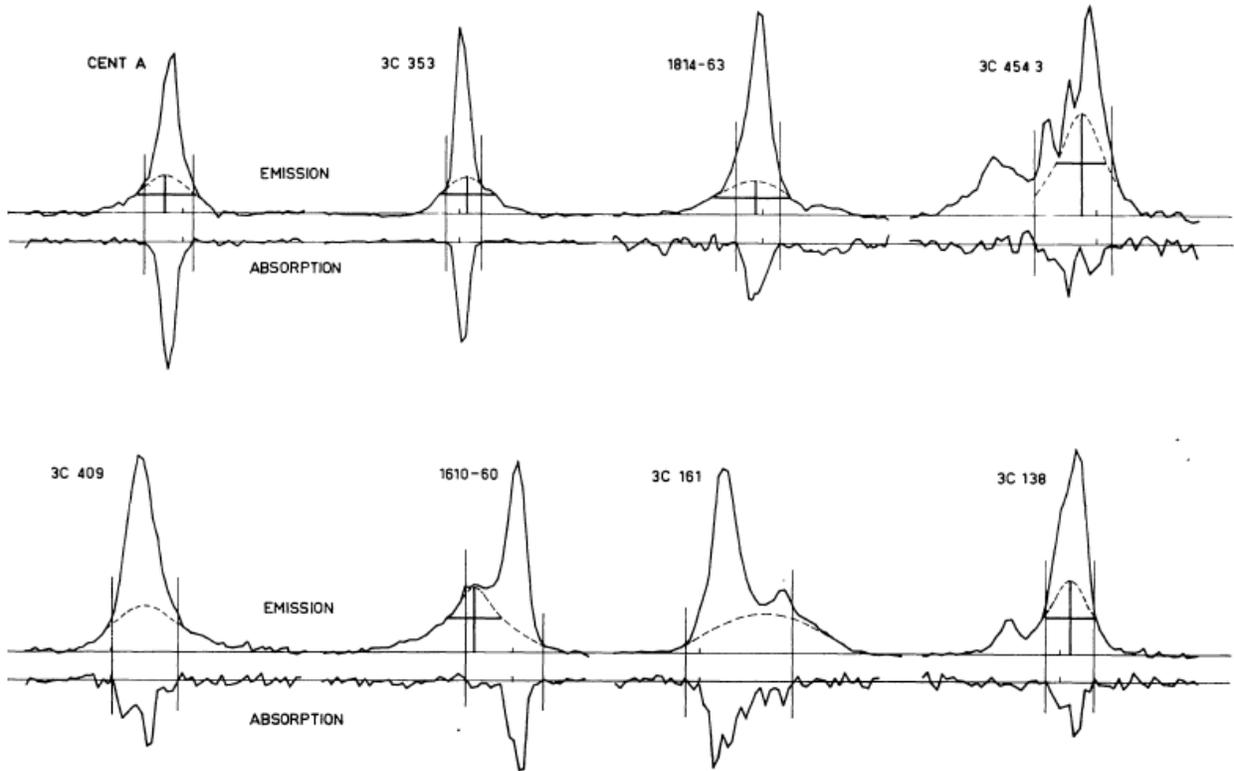
Observe a background source that is time variable, for example, a pulsar that turns off, a variable radio quasar or BL Lac object, etc. Pulsars tend to be very weak at 1420MHz, so this has only been done for ~40 of the 1000-odd known radio pulsars.

This provides you with a measurement of the T_S and τ_ν at the *same* physical location, obviating the assumption that the absorbing cloud varies smoothly between beam positions as in the beam-switching technique.

Method 3: Interferometry

Interferometric techniques (e.g., with the VLA) average out the diffuse emission, giving a direct measurement of $T_{bg} e^{-\tau_\nu}$. Another way of stating this is that interferometry acts as a *spatial filter* that resolves out most of the structure in the diffuse emission from the absorption-line cloud.

Interferometers are only sensitive to structures on scales of B/λ (the spatial frequency), where B is the baseline of the interferometer. The “imaging” capabilities of the antenna array are not used (e.g., one often observes by “phasing” the VLA antennas, reducing the array to a 3-element interferometer to boost sensitivity). Most Galactic HI clouds do not show significant structure on $<1'$ scales, whereas most [useful] background radio sources have typical angular scale sizes of an arcsecond (or less). Interferometry is a very powerful technique for probing HI absorption along lines of sight to radio quasars.

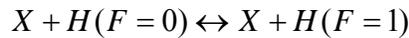


HI 21cm emission and absorption profiles towards 8 extragalactic radio sources from Radhakrishnan et al. (1972, ApJS, 24, 15). The dashed lines are fits to an optically thin emission-line component without a corresponding absorption component. The vertical lines delineate the velocity limits of the optically-thick absorption components. These are beam switching observations with the Parkes Telescope, so the vertical axis is $\Delta T_B(v)$. The narrow absorption lines correspond to discrete CNM clouds, and the broad optically-thin emission is diffuse WNM (intercloud) gas.

Excitation of 21-cm Emission

How do you excite HI into the excited hyperfine structure level?

The competing processes for determining the population of the upper excited hyperfine structure level are Collisional Excitation and Collisional De-excitation.



Because the lifetime of the excited state is so long ($t_0 \approx 1.1 \times 10^7$ years), radiative de-excitations are relatively rare. Collisions with other HI atoms and electrons dominate.

The cross-sections for collisional excitation/de-excitation are:

$$Q_{01}(v_0): \text{excitation, with threshold } h\nu / k = 0.068 \text{ K}$$

$$Q_{10}(v_1): \text{de-excitation (no threshold)}$$

Energy conservation demands

$$\frac{1}{2} m v_0^2 = \frac{1}{2} m v_1^2 + h\nu,$$

$$\text{with threshold: } v_0 = \sqrt{\frac{2h\nu}{m}}$$

Note also that electrons have a much larger speed than H atoms by a factor of $(m_p/m_e)^{1/2} \approx 45$.

The Principle of Detailed Balance

In LTE, any microscopic process is precisely balanced by its inverse, thus collisional excitation with rate

$$n_0^* n_x f(v_0) v_0 Q_{01}(v_0) dv_0$$

is precisely balanced by collisional de-excitations with rate

$$n_1^* n_x f(v_1) v_1 Q_{10}(v_1) dv_1$$

Energy conservation demands that

$$v_0 dv_0 = v_1 dv_1$$

so that the collisional equilibrium condition (no radiative processes) becomes:

$$n_0^* f(v_0) Q_{01}(v_0) = n_1^* f(v_1) Q_{10}(v_1)$$

For a Maxwellian velocity distribution

$$f(v) \propto v^2 e^{-mv^2/2kT}$$

which means that the equilibrium condition becomes

$$n_0^* v_0^2 e^{-mv_0^2/2kT} Q_{01}(v_0) = n_1^* v_1^2 e^{-mv_1^2/2kT} Q_{10}(v_1)$$

In LTE, the Boltzmann equation tells us

$$\frac{n_1^*}{n_0^*} = \frac{g_1}{g_0} e^{-h\nu/kT}$$

but now

$$\begin{aligned} h\nu &= \frac{1}{2} m v_0^2 - \frac{1}{2} m v_1^2 \\ \Rightarrow e^{-h\nu/kT} &= e^{-mv_0^2/2kT} \times e^{+mv_1^2/2kT} \end{aligned}$$

Thus the exponentials cancel, leaving us with the *Milne Relation*:

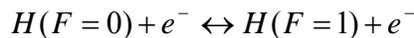
$$g_0 v_0^2 Q_{01}(v_0) = g_1 v_1^2 Q_{10}(v_1)$$

Since $h\nu/k=68\text{mK}$ is small for typical kinetic temperatures T , $v_0 \approx v_1$, and we have:

$$Q_{10}(v) \approx \frac{1}{3} Q_{01}(v)$$

Electron Collisions

Electron collisions dominate the excitation balance in H I if the electron density, $n_e > 0.03 n_H$.



There are two modes:

1. **Spin-Flip Collisions** due to the magnetic field of the passing electron flipping the spin of the bound electron. This has a very small cross-section and is thus rare.

2. **Electron-Exchange Collisions** in which the incoming electron is captured and the bound electron is kicked out. If the spins are opposite, the process looks schematically like this:



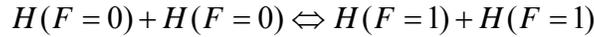
The cross-section for this process is $\langle Q_{10} \rangle = 4.7 \times 10^{-16} \text{ cm}^2$ at $T=100 \text{ K}$.

H-Atom Collisions

If the electron density is very small ($n_e \ll n_H$), Hydrogen-Hydrogen spin-change reactions will dominate the collisional excitation/de-excitation of the hyperfine states:



or



Mixed collisions result in no net change in excitation. The process was first described by Purcell & Field (1956, ApJ, 124, 542), with subsequent re-calculation by Allison & Dalgarno (1969, ApJ, 158, 423). The cross-section for this process is $\langle Q_{10} \rangle \approx 6 \times 10^{-16} \text{ cm}^2$ at 100K, and is a complicated function of the kinetic temperature (e.g., Allison & Dalgarno 1969, figure 2).

Collisionally-Dominated Excitation

The collisional de-excitation rate is

$$n_1 n_x \langle v_1 Q_{10}(v_1) \rangle_{\text{Maxwellian}} = n_1 n_x q_{10}(T)$$

Here we have introduced a new quantity, $q_{10}(T)$, the **Collision De-excitation Coefficient**, which has units of $\text{cm}^3 \text{ s}^{-1}$. The collisional excitation rate can be written similarly in terms of **Collisional Excitation Coefficient** $q_{01}(T)$:

$$n_0 n_x q_{01}(T)$$

The q 's are related by the integral of the Milne Relation:

$$\begin{aligned} q_{01}(T) &= q_{10}(T) \times \frac{g_1}{g_0} e^{-h\nu/kT} \\ &= 3e^{-0.068/T} q_{10}(T) \end{aligned}$$

where T is the *kinetic temperature* of the gas.

In the **collisional limit** (radiative processes are unimportant) detailed balance only includes the collisional rates:

$$n_0 n_x q_{01}(T) = n_1 n_x q_{10}(T)$$

The Milne Relation above gives the ratio of the q 's as a function of the Kinetic Temperature, and The ratio of the level populations is given in non-LTE by

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{-h\nu/kT_s}$$

where T_S is the **Spin Temperature**¹. Substituting both into the equation of detailed balance, we obtain the result

$$T_S = T_{kin}$$

When collisions dominate the excitation, the Spin Temperature is driven to the Kinetic Temperature of the gas. In such circumstances we say that the level populations have “thermalized”.

Radiatively-Dominated Excitation

The opposite limit occurs at low densities when collisions are unimportant relative to radiative processes in the excitation equilibrium. In this case, the equation of detailed balance becomes

$$n_1 A_{10} + 4\pi J_\nu n_1 B_{10} = 4\pi J_\nu n_0 B_{01}$$

where $4\pi J_\nu$ is the radiation field density. With the HI 21-cm line we are in the Rayleigh-Jeans limit for most radiation fields, and so we can express the radiation-field density in terms of T_R , the **Color Temperature of the radiation field**:

$$J_\nu = \frac{2h\nu}{c^3} kT_R$$

From the Einstein Relations, the stimulated emission coefficient, B_{10} can be expressed in terms of the radiative transition probability, A_{10} :

$$4\pi J_\nu B_{10} = A_{10} \left(\frac{kT_R}{h\nu} \right) = A_{10} \left(\frac{T_R}{T_*} \right)$$

where

$$T_* = h\nu / k \approx 68 \text{ mK}$$

Using the Einstein relation $g_0 B_{01} = g_1 B_{10}$ the equation of detailed balance in the radiation-dominated case becomes:

$$n_1 A_{10} + n_1 A_{10} (T_R / T_*) = n_0 \frac{g_1}{g_0} A_{10} (T_R / T_*)$$

This can be solved for the ratio of the level populations:

$$\frac{n_1}{n_0} = 3 \frac{(T_R / T_*)}{(1 + T_R / T_*)}$$

Recalling that the non-LTE (“true”) level populations are given by:

$$\frac{n_1}{n_0} = 3e^{-0.068/T_S} = 3e^{-T_*/T_S}$$

and given that we are in the limit that $T_R \gg T_*$ the relative populations reduce to

$$\frac{n_1}{n_0} \rightarrow 3 \left(1 - \frac{T_*}{T_R} \right)$$

¹ In keeping with the convention we established in Chapter I this should be called the “Excitation Temperature”, but it is traditional to refer to it as “Spin Temperature” in radio astronomy. This is a recurrent theme we will encounter in ISM studies: each specialty has its own traditional nomenclature and notation that makes crossing between disciplines a challenge.

However, $T_S \gg T_*$, so we also have from the Boltzmann equation

$$\frac{n_1}{n_0} = 3e^{-T_*/T_S} \rightarrow 3\left(1 - \frac{T_*}{T_S} + \dots\right)$$

Thus $T_S \rightarrow T_R$, and in the radiation-dominated limit the spin temperature is driven to the color temperature of the ambient radiation field by the effect of stimulated emission. Under typical astrophysical conditions, T_R is the temperature of the cosmic microwave background (~ 2.725 K).

General Excitation

Reality lies between the two limits just discussed, and the excitation and de-excitation of the HI hyperfine levels are a mixture of collisional and radiative processes. If we assume for simplicity that only atomic collisions are important ($n_e \leq 0.03 n_H$), the condition of detailed balance becomes:

$$n_1(n_H q_{10} + A_{10} + 4\pi J_\nu B_{10}) = n_0(n_H q_{01} + 4\pi J_\nu B_{01})$$

Following a procedure similar to the one used in the limiting cases discussed above, this equation can be solved for T_S in the general case:

$$\frac{1}{T_S} = \frac{\left[\frac{1}{T_{kin}} + \frac{A_{10}}{n_H q_{10}} \left(\frac{T_R}{T_*} \right) \frac{1}{T_R} \right]}{\left[1 + \frac{A_{10}}{n_H q_{10}} \left(\frac{T_R}{T_*} \right) \right]}$$

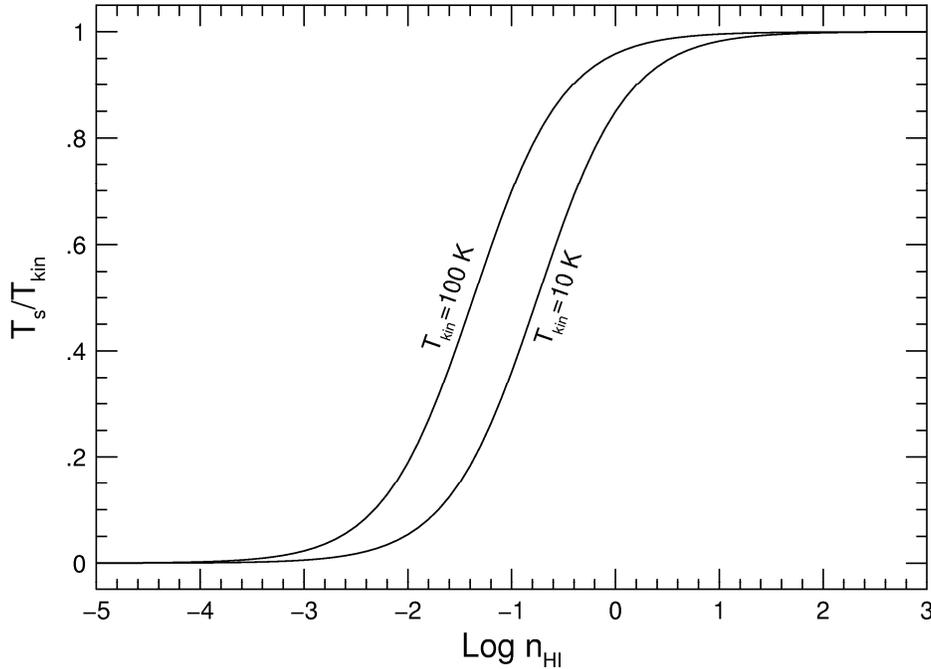
This form is somewhat more complicated than is usually found in most textbooks, but it has the virtue of emphasizing that the spin temperature is in effect a harmonic mean of the kinetic and radiation temperatures, with the radiation temperature weighted by a factor that includes the influence of stimulated emission and the relative importance of collisional vs. radiative de-excitation.

The relative importance of collisional and radiative de-excitation can be parameterized in terms of a **Critical Density**, n_{crit} , at which the rates are roughly the same

$$n_{crit} = \frac{A_{10}}{q_{10}(T)}$$

When $n_H > n_{crit}$, collisional de-excitation dominates and $T_S \rightarrow T_{kin}$, whereas in the low-density limit radiative processes dominate and $T_S \rightarrow T_R$ as before. Because of the temperature dependence on q_{10} , the critical density depends on temperature and does not have a simple analytic or even empirical power-law form (despite what you may find in some references). See for example the tabulation of Allison & Dalgarno. In general, the critical density is of order $10^{-(4-5)} \text{ cm}^{-3}$ for typical ISM conditions ($T = \text{few } 100\text{K}$).

The figure below shows a plot of T_S/T_{kin} for $T_{kin} = 10$ and 100K for the collisional rates tabulated by Allison & Dalgarno (1969). Note that T_S does not exceed the background radiation temperature until well above the critical density. For most ISM conditions, this won't be an issue as n_H is nearly always much larger than n_{crit} , but this is not true at the very low densities of the intergalactic medium.



Plot of T_S/T_{kin} for $T_{kin}=10$ and $100K$.

Lyman- α Pumping

In HI, absorption of Ly α photons is almost immediately followed by re-emission (resonant scattering). This process, however, does not necessarily return the electron to the same hyperfine structure level as before absorption of the Ly α photon.

Because most HI clouds are very optically thick to Ly α photons, a single Ly α photon is resonantly scattered many times before exiting the cloud. Field (1959, ApJ, 129, 551) showed that if the tiny energy difference of the hyperfine levels (6×10^{-6} eV compared to 10 eV for Ly α transition) is accounted for, a very slight slope in the center of the Ly α line profile develops that will drive $T_S \rightarrow T_{kin}$.

Watson & Deguchi (1985, ApJL, 281, L5) have shown that in low-density *extragalactic* HI clouds this process can occur, and can lead to small changes in the observed T_S that increases it compared to the expectations for the low-density limit.

In general, in the low-density limit ($n_H < n_{crit}$) if no other radiation sources are present:

$$T_S \rightarrow T_R = T_{CMB} \approx 2.725 \text{ K}$$

As other sources of T_R come into play, $n_H \rightarrow n_{crit}$, both collisional processes and Ly α pumping will start to drive $T_S \rightarrow T_{kin}$. In the Galaxy,

$$\begin{aligned} n_e &\approx 0.03 \text{ cm}^{-3} \\ n_H &\approx 0.1 - 1 \text{ cm}^{-3} \\ T_{kin} &\approx 100 \text{ K} \end{aligned}$$

Thus $T_S \approx T_{kin}$ is essentially always the case in HI regions in our Galaxy and other galaxies.

The CNM and WNM

HI 21cm studies of the Galaxy reveal that 21cm emission is virtually ubiquitous along every line of sight, but 21cm absorption is not. The typical $N(\text{HI}) > 5 \times 10^{19} \text{ cm}^{-2}$, and the general run of emission lines are broader than the absorption lines (see the spectra from Radhakrishnan et al. above). This is taken as evidence for two physically distinct neutral thermal phases coexisting in the ISM:

Cold Neutral Medium (CNM), composed of cold ($T < 100\text{K}$), dense ($n = 20\text{--}60 \text{ cm}^{-3}$), clouds and filaments with a small filling factor ($\sim 1\text{--}4\%$).

Warm Neutral Medium (WNM), composed of diffuse, relatively low-density ($\sim 0.3 \text{ cm}^{-3}$) gas with temperatures of $\sim 5000\text{K}$ with a much larger filling factor ($\sim 30\%$).

Observations suggest that the CNM has a typical column density of $5 \times 10^{19} \text{ cm}^{-2}$, with a median temperature of $T_S = 58\text{K}$, much colder than the canonical 100K . A column-density weighted average is $\sim 70\text{K}$, but some regions can be as cold as 15K and as warm as 250K . Detailed line profiles show that the CNM is highly turbulent and supersonic, with a typical Mach numbers of ~ 3 with considerable variation, and its morphology may thus be more sheet-like. The exponential scale height of the CNM is about 100pc . While the CNM occupies only a few percent of the volume of the ISM, it probably contains $\sim 30\%$ of the total mass of HI in the Galaxy.

The WNM fills about 30% of the ISM, with a higher typical column density of $\sim 10^{20} \text{ cm}^{-2}$. Along sight lines with no detectable absorption, $T_S = 5000\text{K}$, but using local UV absorption lines to distinguish between thermal and turbulent line broadening suggests temperatures could be as high as 7000K . Two vertical scale-height components are seen, one is roughly Gaussian with a scale height of $\sim 250\text{pc}$, and the other is exponential with a scale height of $\sim 500\text{pc}$. A thermal stability analysis analogous to the FGH criteria discussed in Chapter I suggests that about 50% of the WNM is thermally unstable ($T = 500\text{--}5000\text{K}$). Overall, the WNM contains about 40% of the total mass of HI in the Galaxy.