# 1 Lyman-$\alpha$ linear approximation

This is more or less a collection of my own notes on this, so there might be uncertainties or gaps in the derivation. I would appreciate any comments.

The absorption due to any process can be derived using the following differential equation for the amount of photons ($n_\gamma$) of a given (emitted) frequency ($\nu_e$) evolving through time:

$$\dot{n_\gamma}(\nu_e, t) = -\Lambda\left(\nu_e \frac{a(t_e)}{a(t)}, t\right) n_\gamma(\nu_e, t) \rightarrow n_\gamma(\nu_e, t) = e^{-\tau} n_\gamma(\nu_e, t_e), \tag{1}$$

where we've taken into account that while the photon is travelling through the Universe, the emitted frequncy changes with time and will thus change the properties of the absorption function $\Lambda$.x

The deficeit of photon number density (or equivalently of flux) is thus given by the optical depth ($\tau$):

$$\tau(t) = \int_{t_e}^{t} \Lambda\left(\nu_e \frac{a(t_e)}{a(t')}, t'\right) dt', \tag{2}$$

where the integration limit goes from the time of emission ($t_e$) to a given (later) time at which we are observing. Note however, that $\tau$ is also a function of emitted frequency $\nu_e$.

Apart from the regular absorption (scattering) process, the correct equation should take into account the contribtuion due to stimulated emission of the absorbing gas, which leads us to the full expression:

$$\tau(t) = \int_{t_e}^{t} \Lambda\left(\nu_e \frac{a(t_e)}{a(t')}, t'\right) \left[1 - \exp\left(-\frac{h_p \nu_e a(t_e)}{k_B T(t') a(t')}\right)\right] dt'. \tag{3}$$

However, for the Ly$\alpha$ the temperatures are very low and the exponential term in the brackets can be neglected, thus neglecting stimulated emission contribution.

The absorption function $\Lambda$ is proportional to the amount of (neutral hydrogen) gas that is absorbing $n_{HI}$ and the cross-section of the Ly$\alpha$ transition:

$$\Lambda(\nu, t) = c n_{HI}(t) \sigma_\alpha(\nu). \tag{4}$$

The cross-section for Ly$\alpha$ describes the absorption line profile

$$\sigma_\alpha(\nu) = a_\alpha \Phi(\nu) = \frac{e^2}{4\varepsilon_0 m_e c} f_\alpha \Phi_\alpha(\nu) = \frac{\alpha_{fs} h c}{2 m_e c} f_\alpha \Phi_\alpha(\nu), \tag{5}$$

where $e$ is the electron charge, $m_e$ the electron mas, $f_\alpha = 0.4164$ is the oscillator strength of the transition, $h$ is the planck constant and $\alpha_{fs} = 7.2973525698 \times 10^{-3} \approx 1/137$ is the fine structure constant. The above relation for $a_\alpha$ comes from the Einstein coefficients[1] And $\Phi$ is the (normalized) profile function, can be written using the Voigt line profile:

$$\Phi_\alpha(\nu) = \frac{1}{\sqrt{\pi}} \frac{V(A, B)}{b_\nu} \approx \frac{1}{b_\nu \sqrt{\pi}} \left[e^{-B^2} + \frac{A}{\sqrt{\pi}(A^2 + B^2)}\right]. \tag{6}$$

where $b_\nu$ describes the (thermal) broadening of the line profile, $B$ describes the frequency dependance and $A$ depends on the natural line width:

$$B = \frac{\nu - \nu_\alpha}{b_\nu}, \quad A = \frac{\Gamma_\alpha}{4\pi b_\nu}, \quad b_\nu = \frac{b \nu_\alpha}{c}, \quad b^2 = \frac{2 k_B T}{m_H} + b_{\text{turb}}^2 \tag{7}$$

,

[1] In the literature it is usually written in cgs units as $a_\alpha = \pi e^2/(m_e c) f_\alpha$.

where $\Gamma_\alpha = A_{21} = 6.265 \times 10^8$ s$^{-1}$, where $A_{21}$ is the Einstein's coefficient for spontaneous emission.

In simulations the full relation is used (even some higher order corrections perhaps?) However it turns out that the effect of the natural line width can be neglected with reasonably good accuracy, when dealing with the flux power spectrum as the statistical property. Even though for each absorption (line) feature, we pass from a regime where $B$ dominates over $A$, there is always a regime towards the wings of the line where $A$, being a constant, will dominate over the rapdily declining $B$. However, this is only true for a very small portion of the spectrum, thus creating an effect on very small scales in the power spectrum. Thus, for most analysis, it is reasonable to assume that the profile function $\Phi$ is given only by the Gaussian term, negelcting dependance on $A$.

Also it is always assumed that the thermal broadening part dominates at all times over the turbulent one - this seems to agree well with the simulations and observations. However it might play some role on the small scale physics if present.

Since we are dealing with an absorption of cosmological nature, it is convenient to rewrite the integral for optical depth in redshift or comoving distance along the line of sight. This are the steps taken in most textbooks, and even assumed when dealing with simulations. By a simple change of coordinates (and not using any assumptions so far) we can rewrite the integral as a function of redshift:

$$\tau(t) = \int_{t_e}^{t} \Lambda(\nu(t'), t')dt' = \int_{t_e}^{t} n_{HI}(t')\sigma_\alpha(\nu(t'))cdt' = \tag{8}$$

$$= \int_{a_e}^{a} n_{HI}(a')\sigma_\alpha(\nu(a'))\frac{cda'}{a'H(a')} = \tag{9}$$

$$= \int_{z}^{z_e} n_{HI}(z')\sigma_\alpha(\nu(z'))\frac{cdz'}{H(z')(1+z')} \tag{10}$$

However, as stated earlier, optical depth $\tau$ is a function of time it took a photon to pass a region where the absorption occured, and also of the emitted frequency. In other words, optical depth tells us that for a given frequency $\nu_e$, the amount of absorption will be equal to a sum of absorption function since the beginning of the absorption to the end. Thus we could've taken the limits of the integration from point $A$ to point $B$, which would bracket the time in which the absorption occured.

In the context of Ly$\alpha$ forest the absorption is present from the word go, as we have already indicated by setting the lower integration limit to the time of emission $t_e$. But from the same reasoning, we can take the upper integration limit to today, thus setting $z = 0$. The optical depth is thus only a function of the emitted frequency.

Assuming that the line profile is a very peaked function around the absorbing frequency $\nu_\alpha$, we can extend the upper limit of the integral to infinity. This is a fair assumption, since for one in simulations and in observations, we are never dealing with the absorption close to the quasar when talking about the Ly$\alpha$ forest, because the continuum is hard to model and proximity effects take place. This assumption will also make our life easier later on, since an integral from 0 ot $\infty$ is much easier to deal with than the one with variable integration limit.

Since in observations we are not dealing with the emitted frequency but with the observed ones, we can make a change of the variables and get the following expression:

$$\tau(\nu_e, z(t)) = \int_{z(t)}^{z_e} n_{HI}(z')\sigma_\alpha\left(\nu_e\frac{(1+z')}{(1+z_e)}\right)\frac{cdz'}{H(z')(1+z')} \tag{11}$$

$$\tau(\nu_e) = \tau(\nu_e, 0) = \int_0^{z_e} n_{HI}(z')\sigma_\alpha\left(\nu_e\frac{(1+z')}{(1+z_e)}\right)\frac{cdz'}{H(z')(1+z')} \approx \tag{12}$$

$$\approx \int_0^{\infty} n_{HI}(z')\sigma_\alpha\left(\nu_e\frac{(1+z')}{(1+z_e)}\right)\frac{cdz'}{H(z')(1+z')} \tag{13}$$

$$\tau(\nu_o) = \int_0^{\infty} n_{HI}(z')\sigma_\alpha\left(\nu_o(1+z')\right)\frac{cdz'}{H(z')(1+z')}, \tag{14}$$

where we have assumed that the frequency of photons decays as $1/a$ (making the product of $a$ times the frequcny a constant through time):

$$\frac{\nu}{(1+z)} = \frac{\nu_e}{(1+z_e)} = \frac{\nu_o}{(1+0)} = \nu_o. \tag{15}$$

It is common to rewrite the integral over the redshift path to the one over the comoving coordinate length as follows:

$$\tau(\nu_o) = \int_0^{\infty} n_{HI}(z')\sigma_\alpha\left[\nu_o(1+z')\right]\frac{cdz'}{H(z')(1+z')} \approx \int_{-\infty}^{\infty} n_{HI}(z(x))\sigma_\alpha\left[\nu_o(1+z(x))\right]\frac{dx}{(1+z(x))} \tag{16}$$

where a flat and unpterubed Universe was assumed for the null geodesic ($cdt = adx$) for conversion between time and coordinate (or eqivalently redshift and coordinate). At this point I have explicitly written down that the redshift will be a function of $x$, indicating that the two are not unrelated, since we are following a light cone and thus a null geodesic trajectory. I will comment on how this is usually resolved in a moment.

But first let's look at what assumptions go into deriving the finer points of neutral hydrogen dependence on the underlying matter perturbation.

The main assumption, which seems to hold, is the approximation of the photo-ionization equilibrium. In this scheme, the photoionizing flux of the (in this case UV - since we are dealing with Ly$\alpha$ line) photons, $\Gamma_{\gamma,HI}$, ionizies the neutral hydrogen, $n_{HI}$. This process is in equilibrium with the recombination (and no time delay between the two is assumed, since any such time delay due to recombination is assumed to be much smaller than the Hubble times on which we are observing). The recombination forms back neutral hydrogen, from the free protons (or ionized hydrogen atoms, if you prefer the terminology) with free electrons. The recimbination coefficient however is dependant on the temperature of the gas, which traces the underlying matter as well (i.e. where there is more gas it's hotter - we'll quantify this in a second). Always an assumptions is made that $n_p \approx n_e$, which is claimed to be due to thermal equilibrium :

$$\Gamma_{\gamma,HI}\, n_{HI}(x) = n_p(x)n_e(x)\alpha_\gamma(T) \approx n_p^2(x)\alpha_\gamma(T). \tag{17}$$

At this point no spatial or time dependence of the UV background present through the ionization coefficient $\Gamma_{\gamma,HI}$ is assumed, although it has been pointed out that it may play a role on large scales, especially at higher redshifts, closer to reionization (due to patchiness of the later).[2]

The temperature dependence of the recombination coefficient $\alpha_\gamma(T)$ is universally assumed to follow relation :

[2] see papers by Gontcho-a-Gonctho & Miralda-Escude, and Pontzen et al.

$$\alpha_\gamma(T) = \alpha_0 T_4^{-0.7}, \quad T_4 = \frac{T}{10^4\,\mathrm{K}}. \tag{18}$$

In the above the constant $\alpha_0$ is usually equal to $\alpha_0 \approx 4.3 \times 10^{-13}$ cm$^3$/s. Furthermore, it is assumed that all the protons come from reionization of hydrogen:

$$n_p(x) \approx n_H(x) = \frac{X\rho_b(x)}{m_H}. \tag{19}$$

where $X$ is the (mass?) fraction of hydrogen in the IGM (since it is believed IGM is more or less pure primoridal gas it is assumed $X$ to be equatl to $\sim 0.75$) . This assumption is based on the low temperature of the IGM gas we are probing (typical values $10^4$ K).

The density of baryons can be written as a background density contrast and some fluctuations on top of that :

$$\rho_b(x) = \bar{\rho}_b(z)\Delta_b(x) = \bar{\rho}_b(z)\left[1+\delta_b(x)\right], \tag{20}$$

$$\bar{\rho}(z) = \Omega_b(z)\rho_{\mathrm{crit}}(z) = \frac{3H(z)^2}{8\pi G}\frac{\Omega_{b,0}H_0^2(1+z)^3}{H(z)^2} \tag{21}$$

$$= \frac{3H_0^2}{8\pi G}\Omega_{b,0}(1+z)^3 \tag{22}$$

Combining the above results into one relation between the neutral hydrogen density and gas density and temperature we arrive at the following expression:

$$n_{HI}(x) = \frac{\alpha_0}{\Gamma_{\gamma,HI}}\frac{X^2}{m_H^2}\bar{\rho}_b^2(z)\left[1+\delta_b(x)\right]^2 T_4^{-0.7}(x). \tag{23}$$

The above relation is used in the simulations as well. There have been attempts to test it validity, but are mainly focused on assuming that the gas density fluctuations probed ($\delta_b$) do not follow the underlying dark matter in a linear fashion. Usually a Jeans-like smoothing relation is assumed (on large scales) between the baryon density and the underlying dark matter density, e.g. (in k-space):

$$\delta_b(k) = \frac{\delta_{dm}(k)}{1+k^2/k_J^2}, \tag{24}$$

where $k_J$ is the Jeans scale. However, it has been pointed out (Bi & Davidson et al 1999?), that due to the fact we are probing slightly non-linear structure where the overdensity are around $\sim 1$, a more non-linear relation should be assumed (in real space this time):

$$1+\delta_b(x) = e^{\delta_L(x)-\sigma_L^2/2}, \tag{25}$$

where the $\delta_L$ stands for the linear (dark-matter) density fluctuations.The normalization is such that $\langle 1+\delta_b\rangle = 1$, as pointed out by a paper a few years ago[3] , one should always consider this kind of a normalization since it leads to a proper account of the normalizaton for the series expansion. The function that normalizes this log-normal transformation is the rms of the overdensity contrasts, $\sigma_L^2$. However in literature there is some dispute (or at least no consistency) on what exactly is this function. Most assume just the rms of the linear dark-matter field, while some including also a window function with the usual Jeans smoothin to this normalization factor.

[3]Frusciante & Sheth 2012

The last thing that is usually assumed in (semi-) analytical calculations is the temperature-density relation.This kind of relation is adopted after looking at the relevant distribution from simulations and it seems to be rather valid, in a tight regime of low densities and low temperature, with huge scatter in temperature towards the more dense part of the diagram. One could always choose a more complicated (non-deterministic) model than this, but it seems to work rather well, and the differences seem mostly relevant on small scales:

$$T(x) = T_0 \Delta_b^{\gamma-1}(x), \tag{26}$$

where $T_0$ is some mean gas temperature (simulations and observations suggest it to be around 14000 K, and $\gamma$ is the slope of the power law relation, and is estimated to be around 1.36. In principle it is assumed that $T_0$ and $\gamma$ follow a specific redshift evolution as well, i.e. that they are redshift dependant quantities[4].

Putting it all together we can derive the relation between the neutral hydrogen and the underlying baryon density as

$$n_{HI}(x) = \frac{\alpha_0}{\Gamma_{\gamma,HI}} \frac{X^2}{m_H^2} \bar{\rho}^2(z) \left( \frac{T_0}{10^4\,\mathrm{K}} \right)^{-0.7} [1 + \delta_b(x)]^{2-0.7(\gamma-1)} = n_{HI,0}(z) \Delta_b^p(x). \tag{27}$$

In a similar fashion one can describe the thermal broadening parameter as

$$b(x) = \sqrt{\frac{2k_b T(x)}{m_H}} = \sqrt{\frac{2k_b T_0(z)}{m_H}} \Delta_b^q(x) = b_0(z) \Delta_b^q(x), \tag{28}$$

where $q(z) = (\gamma(z) - 1)/2$. It might be important to include the redshift evolution of the two parameters of $T - \rho$ relation, but for the purpose of the integral it is probably negligable change (as in we can assume they're evaluated at some mean redshift for that integral).

To return to the expression for the optical depth and the integral, we had

$$\tau(\nu_o) = \int_0^\infty n_{HI}(z') \sigma_\alpha \left[ \nu_o (1 + z') \right] \frac{c dz'}{H(z')(1 + z')}. \tag{29}$$

What is usually assumed is that the line profile function $\Phi$ is peaked enough, to assumed that the path integral $dz'$ where the integral will return non-negligable result is small. In the light of this one can approximate the integration along the light cone as an integral at fixed (mean) redshift (time) over comoving coordinate long the line of sight[5]. This is an approximation, but a valid one (corrections are of the order of $v/c$, same size as higher-order relativistic corrections[6]).

$$z = \bar{z} + \Delta z = \bar{z} + \frac{H(\bar{z})}{c} (x - \bar{x}), \tag{30}$$

where in the last step we've used the null-geodesic relation of $(a(t)dx = cdt)$. The mean redshift $\bar{z}$ would correspond to the redshift of the simulation being analyzed or the redshift bin of the observational data. So in a sense this kind of approximation is used in analyzing the simulation outputs, as well as in data analysis. The comoving position $\bar{x}$ is the one that corresponds to the mean redshift $\bar{z}$.

But for the purpose of non-relativistic treatement the above assumption seem to work very well.

The above approximation is then used in the relation for the optical depth (to the first order) which gives us

[4] see Haardt-Madau papers and similar
[5] See for example the paper by Bi et al. 1997
[6] see paper Irsic, Di Dio & Viel 2015

$$\tau(\nu_o) = \int_{-\infty}^{\infty} n_{HI}(x,\bar{z})\sigma_\alpha(\nu)\,\frac{dx}{(1+\bar{z})}. \tag{31}$$

Note that the integral is extended from minus to plus infinity. The plus infinity was justified before, the minus can be justified in the similar sense - that the integral will be picking up only on a small portion of the line of sight, and that is not close to $z = 0$. The convolution integrals later on are easier to solve if we extend these integrals.

Furthermore, the frequency $\nu$ dependence on redshift is also corrected for the peculiar velocities, so instead of just $\nu = \nu_o(1+z)$ we have

$$\nu = \nu_o(1+z)\left(1+\frac{v_{\rm pec}(x)}{c}\right) \approx \nu_o(1+\bar{z})\left(1+\frac{u(x)}{c}\right), \quad u = \frac{H(\bar{z})}{1+\bar{z}}(x-\bar{x}) + v_{\rm pec}(x), \tag{32}$$

with $u$ being the redshift space coordinate and where the peculiar velocities $v_{\rm pec} = v_b$ trace the underlying baryon field. In fact it is often assumed that even if the baryon density doesn't linearly trace the (linear) dark matter field, the velocities do.

Peculiar velocity of the observer (us) is neglected here, as it would only shift the redshift space coordinate $u$ by a constant amount (see Hui et al. 1997 for details).

Also note in the last relation we have approximated it to the first order, where we've dropped a term of the order of $\Delta z\, v_b$.

With the above in mind one can define a specific $u_o$ which corresponds to exactly the frequency of the Ly$\alpha$ absorption:

$$\nu_\alpha = \nu_o(1+\bar{z})\left(1+\frac{u_o}{c}\right), \tag{33}$$

where $u_o$ now corresponds to the observed redshift space coordinate, with the above relation linking it observed frequency $\nu_o$.

If we change the integration variable to the redshift space coordinate, we can rewrite the optical depth as

$$\tau(u_o) = A(\bar{z})\int_{-\infty}^{\infty} du\ (1+\delta_b(u))^p\ \frac{1}{b(u)\sqrt{\pi}} e^{-\frac{(u-u_o)^2}{b^2(u)}} \left|\frac{du}{dx}\right|^{-1} \tag{34}$$

$$\approx \tau_0(\bar{z})\int_{-\infty}^{\infty} du\ (1+\delta_b(u))^p\ \frac{1}{b(u)\sqrt{\pi}} e^{-\frac{(u-u_o)^2}{b^2(u)}} \left(1-\frac{\partial v_b}{\partial u}\right), \tag{35}$$

with $b = b_0(1+\delta_b(u))^q$. In the above we have assumed that when $z \to 0$ in the integral limits, so does $u \to -\infty$, which is not strictly speaking correct. It assumes, once again, that the exponential function is peaked around $u_o$, so that $u - u_o$ is small, otherwise we are not getting any signal. This means that the true lower limit can be replaced by a $-\infty$.

We have also crammed all the constatns that depend only on the mean redshift into $A$ and $\tau_0$ such that

$$\tau_0(z) = A(z)\frac{1+z}{H(z)} = \frac{1}{H(z)} n_{HI,0}(z)\lambda_\alpha a_\alpha \tag{36}$$

$$\approx 1.26468942 \times \left(\frac{\Gamma_{\gamma,HI}}{10^{-12}\,\mathrm{s}^{-1}}\right)^{-1}\left(\frac{X}{0.75}\right)^2\left(\frac{T_0}{10^4\,\mathrm{K}}\right)^{-0.7}\frac{(\Omega_{b,0}h^2)^2}{h}\left(\frac{H_0}{H(z)}\right)(1+z)^6, \tag{37}$$

where the lyman alpha wavelength is $\lambda_\alpha = 1215.6701$Å.

The constant at a typical redshift $z = 3$ has a value of

$$\tau_0(z = 3) = 0.665435, \tag{38}$$

for the cosmology of $\Omega_m = 0.274247$, $\Omega_b = 0.0458$, $\Omega_\Lambda = 0.725753$, $h = 0.7$ and for astrophysical parameters of $T_0 = 1.47088 \times 10^4$ K, $\Gamma_{\gamma,HI} = 10^{-12}$ s$^{-1}$, $X = 0.75$.

We see that the optical depth is a sort of a convolution between the cosmology part from the redshift space distortions, cosmology and astrophysics part from the density of neutral hydrogen, and a Gaussian profile. It is not quite a convolution since the Gaussian profile has a density dependant width. What is then done in the linear approximation is assume that $\delta_b \ll 1$ (also means that $v_b \ll 1$ if we believe linear theory) and expand various functions:

$$\tau(u_o) = \tau_0(\bar{z}) \int_{-\infty}^{\infty} du \left[1 + p\delta_b(u) - \frac{\partial v_b}{\partial u} - q\left(1 - 2\frac{(u - u_o)^2}{b_0^2}\right)\delta_b(u)\right] W(u - u_o), \tag{39}$$

where the kernel function $W$ is now only explicitly dependant on $u$, so that

$$W(s) = \frac{1}{b_0\sqrt{\pi}} e^{-\frac{s^2}{b_0^2}}. \tag{40}$$

The last term in the optical depth equation can be further manipulated using per-partes integration, once we realize that the term is closely related to the second derivative of the kernel function $W$. The first two terms remain intact (in linear approximation) so we can write

$$\text{3rd term} = -\tau_0(\bar{z}) \int_{-\infty}^{\infty} du\, q\left(1 - 2\frac{(u - u_o)^2}{b_0^2}\right)\delta_b(u)\, W(u - u_o) = \tag{41}$$

$$= \tau_0(\bar{z})\frac{qb_0^2}{2} \int_{-\infty}^{\infty} \delta_b(u)\, \frac{\partial^2 W}{\partial u^2} du = \tag{42}$$

$$= \tau_0(\bar{z})\frac{qb_0^2}{2} \left[\frac{\partial W}{\partial u}\delta_b(u)\Big|_{-\infty}^{\infty} - W\frac{\partial \delta_b}{\partial u}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} W(u - u_o)\, \frac{\partial^2 \delta_b}{\partial u^2} du\right], \tag{43}$$

the last term here is what we need. The upper limit ($\infty$) of the per-partes limit tersm is obvious, as $W$ goes to zero faster than anything else - the only thing that is assumed here is that $\delta_b$ doesn't monotonously increase with $u$, which is a fair assumption. Same holds for lower limit.

Having dealt with the per-partes terms we are left with:

$$\tau(u_o) = \tau_0(\bar{z}) \int_{-\infty}^{\infty} du\, W(u - u_o) \left[1 + p\delta_b(u) - \frac{\partial v_b}{\partial u} + \frac{qb_0^2}{2}\frac{\partial^2 \delta_b}{\partial u^2}\right], \tag{44}$$

where $p = 2 - 0.7(\gamma - 1)$ and $q = (\gamma - 1)/2$. This is the result of the linear approximation found in the literature. What you can do then is define $\tau$ fluctuations and a power spectrum of the optical depth - this agree relatively well with the simulations of optical depth, on large scales the only difference is a constant bias, on small scales some $k$-dependence is a bit off.

Before going to Fourier space let's just review the above equation if we're working in 3D space. Everything stays the same, since the line-of-sight physics and integration doesn't change, except that quantities now depend on perpendicular direction $u_\perp = H(z)D_A(z)\Delta\theta$, where $\Delta\theta$ is the angular seperation on the sky. The optical depth can then be written as

$$\tau(\mathbf{u}) = \tau(u_\perp, u_o) = \tau_0(\bar{z}) \int_{-\infty}^{\infty} du_\parallel\, W(u_\parallel - u_o) \left[1 + p\delta_b(u_\perp, u_\parallel) - \frac{\partial v_b}{\partial u_\parallel} + \frac{qb_0^2}{2}\frac{\partial^2 \delta_b}{\partial u_\parallel^2}\right]. \tag{45}$$

Let's also define a optical depth fluctuations as

$$1 + \delta_\tau(\mathbf{u}) = \frac{\tau(\mathbf{u})}{\langle \tau(\mathbf{u}) \rangle}, \tag{46}$$

where we explicitly note that the mean optical depth may not necessarily be $\tau_0$. If the above relation were the full description of the optical depth then yes, the mean would be $\tau_0$. But in general the mean would also depend on the higher order moments of the underlying matter distribution, due to the non-linear relations between optical depth and the density field. If you evaluate the optical depth power spectrum from the numerical simulations you'd get a different value of the mean optical depth, which has to be taken into account. However to systematically use terms only up to linear order the mean of the $\tau$ field is $\tau_0$ (check yourself). Thus:

$$\delta_\tau = \int_{-\infty}^{\infty} du_\parallel \, W(u_\parallel - u_o) \left[ p\delta_b(u_\perp, u_\parallel) - \frac{\partial v_b}{\partial u_\parallel} + \frac{qb_0^2}{2} \frac{\partial^2 \delta_b}{\partial u_\parallel^2} \right]. \tag{47}$$

Defining the Fourier transform as

$$\tau(\mathbf{u}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} \tau(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{u}} \tag{48}$$

Noting that the integral along the line of sight is now nothing but a convolution we can write the Fourier transform of the linearized optical depth fluctuations as:

$$\delta_\tau(\mathbf{k}) = \left[ p\delta_b(\mathbf{k}) - ik_\parallel v_b(\mathbf{k}) - \frac{qb_0^2}{2} k_\parallel^2 \delta_b(\mathbf{k}) \right] W_\mathbf{k}. \tag{49}$$

If the peculiar velocities are small, one can use linear relation (coming from linear evolution equations[7]). The linear relation says

$$v_b(\mathbf{k}) = i\frac{k_\parallel}{k^2} \mathcal{H} f \delta_b(\mathbf{k}) \tag{50}$$

What people used to do is also expand $F = \exp(-\tau)$ relation to first order, from where you get that the fluctuations in flux are just fluctuations in $\tau$. Introducing the fluctuations in the flux $1 + \delta_F = F/\langle F \rangle$, we can write

$$1 + \delta_F = \frac{e^{-\tau_0(1+\delta_\tau)}}{\langle e^{-\tau_0(1+\delta_\tau)} \rangle} \approx \frac{e^{-\tau_0} \left(1 - \tau_0\delta_\tau + \frac{1}{2}\tau_0^2\delta_\tau^2 + \dots\right)}{\langle e^{-\tau_0} \left(1 - \tau_0\delta_\tau + \frac{1}{2}\tau_0^2\delta_\tau^2 + \dots\right) \rangle} \approx 1 - \tau_0\delta_\tau + \frac{1}{2}\tau_0^2 \left(\delta_\tau^2 - \langle \delta_\tau^2 \rangle\right) + \dots, \tag{51}$$

thus in linear approximation the relation between (observed) flux fluctuations and optical depth fluctuation is simple and linear

$$\delta_F \approx -\tau_0\delta_\tau. \tag{52}$$

The negative sign means that for a positive optical depth fluctuation (which means a positive overdensity), a flux fluctuation is negative. This is simply a consequence of the denser the region, the more absorption in the flux.

The flux power spectrum is then (as usual)

$$P_F(k) = \langle \delta_F \delta_F^* \rangle = \tau_0^2 \langle \delta_\tau \delta_\tau^* \rangle. \tag{53}$$

[7] See Large Scale Structure Lectures by Emiliano Sefusatti

What we have computed above holds for 3D!. While usually one observed flux statistics averaged over all QSO lines of sight, and thus in 1D only. This is easily achieved by integrating flux over $k_\perp$ in the Fourier space.

$$P_F^{1D}(k_\parallel) = \frac{1}{2\pi}\int P_F^{3D}(\mathbf{k})d^2\mathbf{k}_\perp = \frac{1}{2\pi}\int_{k_\parallel}^{\infty} P_F^{3D}(k_\parallel, k)\, k\, dk = \frac{1}{2\pi}\int_{k_\parallel}^{\infty} K(k_\parallel, k)P_b(k)\, k\, dk, \tag{54}$$

with the integration kernel $K(k_\parallel, k)$ being

$$K(k_\parallel, k) = \tau_0^2 e^{-\frac{1}{2}k_\parallel^2 b_0^2}\left[p + \frac{k_\parallel^2}{k^2}\mathcal{H}f - \frac{qb_0^2}{2}k_\parallel^2\right]^2 \tag{55}$$

The above relation is a very nice example in describing the workings behind the Ly$\alpha$ absorption and relation to cosmology. However its usefulness and validity is very limited, and never used in practice, since numerical simulations work so much better. The reason why the relation doesn't work very well is because of the highly-nonlinear relation between flux and optical depth and the Taylor expansion done is not very justified for the larger optical depth fluctuations. However, the above relation is a motivation for empirical fits to simulations which give similar result (see Arinyo-i-Prats et al.) which gives a relation of a sort

$$P_F(k, \mu = k_\parallel/k) = P_L(k)\left(b_F + fb_\eta\mu^2\right)^2 D(k, \mu) \tag{56}$$