

# Reionized bubbles around primordial galaxies: exercises

Joris Witstok ([joris.witstok@nbi.ku.dk](mailto:joris.witstok@nbi.ku.dk))

<https://github.com/joriswitstok/reionised-bubbles-tutorial>

## 1 Background problems: recombination lines in astrophysical nebulae

Given atomic number  $Z$ , the energy levels  $E_n$  in the classical Bohr (1913) model of an atom are

$$E_n \approx \frac{-13.6Z^2}{n^2} \text{ eV}. \quad (1)$$

Electronic transitions between energy levels  $n \rightarrow m$  then offer a natural theoretical explanation for the empirical Rydberg formula,

$$\frac{1}{\lambda} = RZ^2 \left( \frac{1}{m^2} - \frac{1}{n^2} \right), \quad (2)$$

which relates spectral line wavelengths  $\lambda$  to a pattern based on a set of two integers,  $n$  and  $m$  ( $n > m$ ).

**Problem 1.1.** Given the photon energy  $E = hc/\lambda$ , work out  $R_H \equiv R$ , the Rydberg constant for hydrogen ( $Z = 1$ ).

**Problem 1.2.**

- (a) Using this atomic model of hydrogen, calculate the wavelengths of the principal spectral lines in the Lyman, Balmer, and Paschen series ( $\text{Ly}\alpha$ ,  $\text{H}\alpha$ , and  $\text{Pa}\alpha$ ), corresponding to the  $n + 1 \rightarrow n$  electronic transitions for  $n \in \{1, 2, 3\}$ .
- (b) Across what redshift range can we observe each of these lines with the low-resolution PRISM disperser of the JWST/NIRSpec<sup>1</sup> instrument? And with the medium-resolution G235M grating?
- (c) How do the wavelengths of these spectral lines compare to the equivalent He II (singly ionised helium) electronic transitions?

**Problem 1.3.** Keeping in mind that electrons usually carry some kinetic energy, what is the energy of a photon produced in the event where hydrogen recombination ( $H^+ + e^- \rightarrow H^0$ ) directly results in the ground state ( $n = 1$ )? Discuss why for case-B recombination we can make the so-called ‘on-the-spot approximation’, where only recombinations leaving hydrogen in one of the excited states ( $n > 1$ ) are taken into account.

**Problem 1.4.** When evolving on the main sequence, spectral O-type stars may have a size ten times that of the Sun, effective surface temperature of 40 000 K, and emit hydrogen-ionising photons at a rate of  $\dot{N}_{\text{ion}} = 10^{49} \text{ s}^{-1}$ . Consider a surrounding H II region consisting of pure hydrogen, with a number density of  $n_H = 300 \text{ cm}^{-3}$ .

- (a) If the H II region is 4 parsec in diameter, would it be ionisation bounded? Check in the range of electron temperatures  $10\,000 \text{ K} < T_e < 30\,000 \text{ K}$ , where you can assume the (case-B) recombination rate of the gas to follow  $\alpha_{\text{rec}} = 2.54 \times 10^{-13} (T_e/10^4)^{-0.8} \text{ cm}^3 \text{ s}^{-1}$ .
- (b) If roughly two out of three of recombination events lead to the emission of a  $\text{Ly}\alpha$  photon (in the ionisation-bounded case), what is the  $\text{Ly}\alpha$  luminosity of the cloud? How does this compare to the bolometric luminosity of the central star, if assumed to be a perfect blackbody?

---

<sup>1</sup>The wavelength coverage of the various NIRSpec instrument modes is documented here: <https://jwst-docs.stsci.edu/jwst-near-infrared-spectrograph/nirspec-instrumentation/nirspec-dispersers-and-filters>.

## 2 Background problems: measuring distances in an expanding universe

**Problem 2.1.** Assume the time evolution of an ionised bubble radius,  $R_{\text{ion}}(t)$ , follows the differential equation given by [Cen & Haiman \(2000\)](#):

$$\frac{dR_{\text{ion}}^3}{dt} = 3H(z)R_{\text{ion}}^3 + \frac{3f_{\text{esc, LyC}}\dot{N}_{\text{ion}}}{4\pi\bar{n}_{\text{H}}} - C_{\text{HII}}\bar{n}_{\text{H}}\alpha_{\text{B}}R_{\text{ion}}^3. \quad (3)$$

- (a) If we only consider the cosmic expansion represented by the first term on the right-hand side, solve for  $R_{\text{ion}}(t)$ . Work out the  $e$ -folding time at redshift  $z = 7$  assuming a [Planck Collaboration et al. \(2020\)](#) cosmology, where the Hubble parameter is  $H(z = 7) \approx 856.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . Is this effect important at this redshift? (Hint: compare to the age of the Universe at this redshift.)
- (b) If we instead neglected the cosmic expansion and recombinations, how would  $R_{\text{ion}}(t)$  evolve in terms of the ionising photon production rate  $\dot{N}_{\text{ion}}$  and escape fraction  $f_{\text{esc, LyC}}$  of the central galaxy? (Hint: see [Mason & Gronke \(2020\)](#) or [Witstok et al. \(2024\)](#).)
- (c) Given the present-day ( $z = 0$ ) cosmic mean hydrogen number density of  $\bar{n}_{\text{H}} \approx 1.88 \times 10^{-7} \text{ cm}^{-3}$ , how large does an ionised bubble need to be for recombination events to balance the ionising output of a  $z = 7$  galaxy, if it produces ionising photons at a constant rate of  $\dot{N}_{\text{ion}} = 10^{54} \text{ s}^{-1}$ , of which  $f_{\text{esc, LyC}} = 10\%$  escape? You can assume a clumping factor  $C_{\text{HII}} = 3$  and the IGM within the bubble to have mean density and recombination rate as in Problem 1.4(b) with  $T_e = 20\,000 \text{ K}$ . What if the galaxy were at  $z = 12$ ?

## 3 Application to the JADES spectroscopic galaxy survey

### 3.1 Summary of the exercise

In this exercise, we will use the method outlined in [Witstok et al. \(2024\)](#) to estimate the size of an ionised bubble. This will involve calculating the Ly $\alpha$  transmission curve for the two-zone model from [Mason & Gronke \(2020\)](#), where a photon trajectory starts inside a spherical ionised bubble (with a radius  $R_{\text{ion}}$ ) centred on the source, before travelling through the IGM characterised by a ‘global’ neutral fraction,  $\bar{x}_{\text{H I}}$ . Then, given the intrinsic strength of Ly $\alpha$  predicted by converting the measured H $\beta$  line luminosity, it is possible to work out a lower limit on  $R_{\text{ion}}$ , the bubble radius.

### 3.2 Gather spectroscopic data and measure the redshift and line fluxes

We will be using NIRSpec spectra obtained as part of the JWST Advanced Deep Extragalactic Survey (JADES; [Eisenstein et al. 2023](#)) in the Great Observatories Origins Deep Survey (GOODS) extragalactic legacy fields located in the North (GOODS-N) and South (GOODS-S). Specifically, we will start by looking at source ID 1899 in a medium-depth tier of JADES covering the GOODS-N field, which was identified by [Witstok et al. \(2025\)](#) as one of the most distant known Ly $\alpha$  emitting galaxies (LAEs; see also [Tang et al. 2024](#); [Navarro-Carrera et al. 2024](#)).

NIRSpec spectroscopy in JADES is taken in the PRISM disperser operating at low spectral resolution across 1–5  $\mu\text{m}$ , as well as in the medium-resolution G140M, G235M, and G395M gratings (see Problem 1.2). Fully reduced and science-ready observations from JADES are publicly available on the Mikulski Archive for Space Telescopes (MAST). A dedicated MAST web page (<https://archive.stsci.edu/hlsp/jades>) describes how we can access the high-level science products (HLSP) via the PYTHON package ASTROQUERY.

- (a) Run the relevant code block in the notebook to download the PRISM/CLEAR, G140M, and G395M spectra of ID 1899 onto your local machine.
- (b) Open the .png preview of the PRISM spectrum. This shows the one-dimensional spectrum, which is extracted from the two-dimensional spectrum (how the dispersed light will show up on the detector). Discuss why the central trace with positive signal-to-noise ratio (SNR) is surrounded by two negative traces on either side?
- (c) Plot the one-dimensional spectra by reading in the downloaded \*x1d files in FITS (Flexible Image Transport System) format.

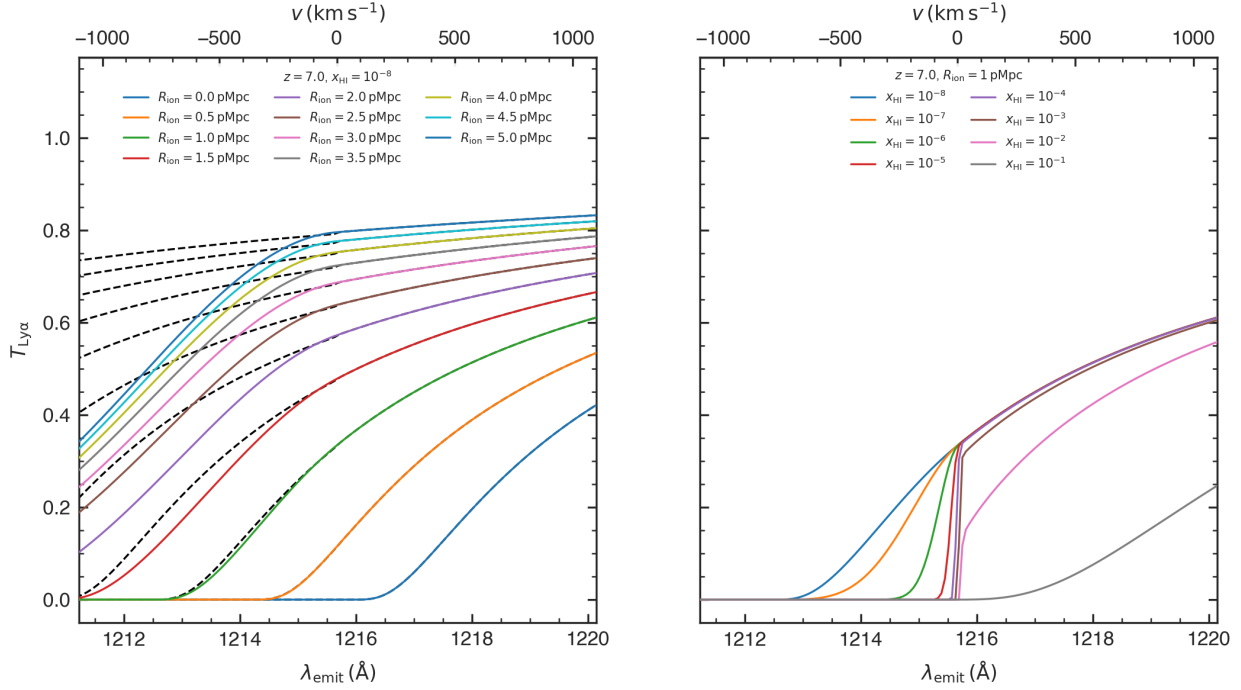


Figure 1: IGM transmission curves, reproduced as Figure 1 from [Mason & Gronke \(2020\)](#).

Several strong rest-frame optical emission lines are observed at the red of the spectrum covered by the G395M grating, from which we can measure a redshift,  $z = \lambda_{\text{obs}}/\lambda_{\text{emit}} - 1$ . At sufficiently high redshift (as is the case for ID 1899), the strong H $\alpha$  line shifts out of the NIRSpec wavelength range. However, there is still the H $\beta$  at a rest-frame wavelength of 4862.71 Å, as well as two strong emission lines of [O III] are located at 4960.295 Å and 5008.24 Å. The luminosity ratio between the [O III] lines is determined by atomic physics to be  $F_{5008}/F_{4960} \approx 2.98$  across a range of temperatures and densities applicable to H II regions (where the majority of the [O III] emission is produced).

- (d) What is the highest redshift H $\alpha$  can be seen with NIRSpec? And H $\beta$ ?
- (e) Given its spectral resolution of at least  $R \approx 1000$  in the observed G395M spectrum, what is the expected precision (its standard deviation  $\sigma_z$ ) of our redshift measurement? Assume the line spread function (LSF) is Gaussian, with full-width at half maximum (FWHM)  $\Delta\lambda$  relating to  $R$  via  $R \equiv \lambda/\Delta\lambda$ . (Hint: you can convert the FWHM of a Gaussian to the standard deviation  $\sigma$  by dividing by a factor  $2\sqrt{2 \ln(2)} \approx 2.35$ .)
- (f) Construct a model of Gaussian line profiles to measure the redshift from the H $\beta$  and [O III] lines observed at 4.4-4.9  $\mu\text{m}$  in the G395M spectrum. You can assume the spectral resolution to be constant value of  $R = 1500$  across this wavelength range, and an intrinsic line width ranging across  $0 < \sigma_{\text{int}} < 500 \text{ km s}^{-1}$ . Discuss: does the uncertainty on the redshift measurement match your estimate from above and if not, why?
- (g) With the measured systemic redshift, fit a Gaussian profile to the Ly $\alpha$  line in the G140M spectrum (NB: the line centre may be resonantly scattered away from the systemic redshift; see Appendix A.3 for the relevant conversions). Comparing the case-B luminosity ratio between Ly $\alpha$  and H $\beta$  ( $L_{\text{Ly}\alpha}/L_{\text{H}\beta} \approx 23.3$ ; [Saxena et al. 2023](#)) to the observed ratio, determine the escape fraction of Ly $\alpha$ ,  $f_{\text{esc, Ly}\alpha}$ .

Now we will repeat the exercise for (much fainter) source ID 10013682 in the DEEP/HST tier of JADES, which was first identified as a remarkably strong LAE in the GOODS-S field by [Saxena et al. \(2023\)](#).

- (h) Given the lower SNR in the G395M spectrum, perform your redshift measurement on the lower-resolution PRISM data this time ( $R \approx 100$ ; how is the precision expected to change?). For the fit, you can assume  $R = 350$  at  $\lambda_{\text{obs}} > 4 \mu\text{m}$ .

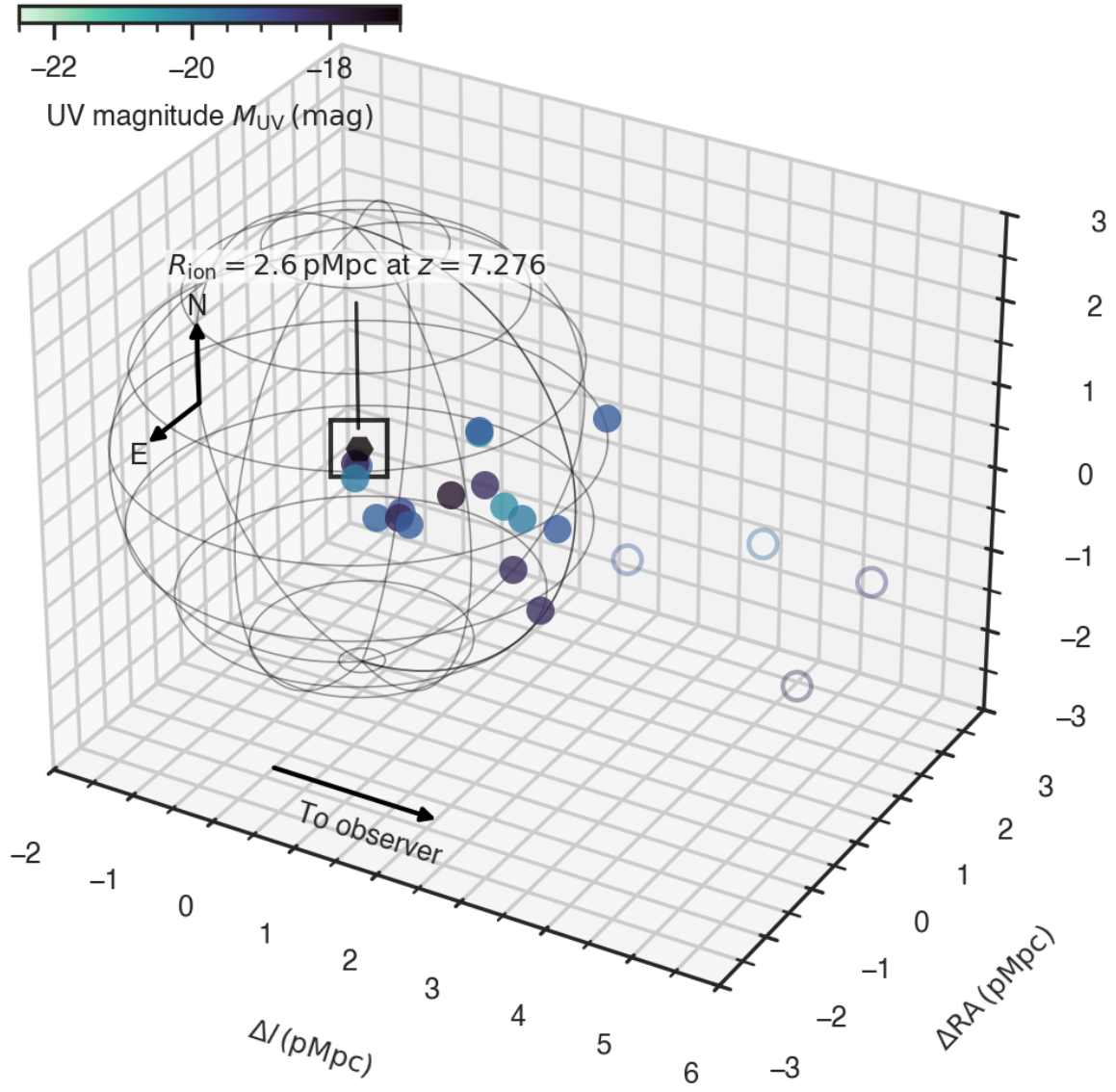


Figure 2: Neighbouring galaxies in the ionised bubble of ID 10013682 (partial reproduction of Figure 5 in [Witstok et al. 2024](#)).

- (i) With the measured systemic redshift, again measure the strength of the Ly $\alpha$  line in the G140M spectrum and estimate  $f_{\text{esc, Ly}\alpha}$ .

### 3.3 Validate against literature

From here on, we will focus on source ID 10013682 to study the ionised bubble it resides in. Compile the relevant measurements from [Witstok et al. \(2024\)](#), where ID 10013682 was analysed among a sample of high-redshift LAEs in JADES. We will need its sky coordinates, spectroscopic redshift, UV magnitude, the Ly $\alpha$  velocity offset  $\Delta v_{\text{Ly}\alpha}$  and escape fraction,  $f_{\text{esc, Ly}\alpha}$ , which are summarised in Table 1. How do your measurements from Section 3.2 compare to those reported in [Witstok et al. \(2024\)](#)?

### 3.4 Install prerequisites and calculate IGM transmission curves

Check whether you have installed the right modules by running the first code block. Download the PYTHON module LYMANA-ABSORPTION from [https://github.com/joriswitstok/lymana\\_absorption](https://github.com/joriswitstok/lymana_absorption). Verify that the code works by running the code block that reproduces Figure 1 from [Mason & Gronke \(2020\)](#), as shown in Fig. 2.

### 3.5 Calculate ionised bubble size

Finally, we will estimate the ionised bubble radius that ID 10013682 resides in. Following [Witstok et al. \(2024\)](#), we are looking for the value of  $R_{\text{ion}}$  at which the IGM transmission at the measured Ly $\alpha$  velocity offset is equal to the estimated Ly $\alpha$  escape fraction to find an effective lower limit on  $R_{\text{ion}}$ .

- (a) Predict the IGM transmission at the measured Ly $\alpha$  velocity offset  $\Delta v_{\text{Ly}\alpha}$  for a range of ionised bubble sizes (logarithmically spaced between 1 pkpc and 10 pMpc), and estimate (through interpolation) at which ionised bubble radius  $R_{\text{ion}}$  the IGM transmission becomes equal to the Ly $\alpha$  escape fraction.
- (b) Discuss why this effectively yields a lower limit on  $R_{\text{ion}}$ ? What are the caveats of estimating  $R_{\text{ion}}$  with this method?
- (c) Inexcusably, [Witstok et al. \(2024\)](#) do not report uncertainties on their inferred bubble radii. Just considering the stated uncertainty on  $\Delta v_{\text{Ly}\alpha}$ , what uncertainty range of  $R_{\text{ion}}$  do you find?
- (d) Given its absolute UV magnitude<sup>2</sup> of  $M_{\text{UV}} \approx -17$  mag and ionising photon efficiency directly measured to be  $\log \xi_{\text{ion}} = 25.66 \text{ Hz erg}^{-1}$  by [Saxena et al. \(2024\)](#), how long would it take ID 10013682 to create this ionised region itself if it had an escape fraction 100%? (Hint: use (i) Appendix A.2 to convert  $M_{\text{UV}}$  to a flux density  $F_{\nu, \text{UV}}$ , (ii) Appendix A.1 to obtain the luminosity density  $L_{\nu, \text{UV}}$ , (iii) the fact that  $\dot{N}_{\text{ion}} = \xi_{\text{ion}} L_{\nu, \text{UV}}$ , and (iv) your solution from Problem 2.1(b) to calculate the required age.)

### 3.6 Visualise the environment of the ionised bubble

The NIRCcam wide-field slitless spectroscopic mode allows an efficient way to obtain spectra of a large number of galaxies. This has been exploited in the FRESCO survey ([Oesch et al. 2023](#)) covering both the GOODS fields, allowing a ‘blind’ redshift search up to  $z \approx 8$ .

- (a) Using the catalogue provided by [Helton et al. \(2024\)](#), plot all nearby galaxies in the GOODS-S field that have been spectroscopically confirmed close to ID 10013682 (within approximately 5 pMpc).
- (b) Assuming the relations from [Mason & Gronke \(2020\)](#) to link the absolute UV magnitude  $M_{\text{UV}}$  and UV slope  $\beta_{\text{UV}}$  of a galaxy to its ionising photon production rate,

$$\dot{N}_{\text{ion}} = 1.65 \times 10^{54} 10^{\frac{M_{\text{UV}} + 20}{-2.5}} \left( \frac{912.0}{1500.0} \right)^{\beta_{\text{UV}} + 2} \text{ s}^{-1},$$

by how much do galaxies contained within the bubble increase the available ionising-photon budget from Section 3.5(c)? Does this alleviate the tension with the inferred size required to explain the Ly $\alpha$  emission?

<sup>2</sup>See [https://en.wikipedia.org/wiki/Absolute\\_magnitude](https://en.wikipedia.org/wiki/Absolute_magnitude).

## References

- Bohr N., 1913, [Philosophical Magazine](#), 26, 1
- Cen R., Haiman Z., 2000, [ApJ](#), 542, L75
- Eisenstein D. J., et al., 2023, p. [arXiv:2306.02465](#) ([arXiv:2306.02465](#))
- Helton J. M., et al., 2024, [ApJ](#), 974, 41
- Mason C. A., Gronke M., 2020, [MNRAS](#), 499, 1395
- Navarro-Carrera R., Caputi K. I., Iani E., Rinaldi P., Kokorev V., Kerutt J., 2024, p. [arXiv:2407.14201](#) ([arXiv:2407.14201](#))
- Oesch P. A., et al., 2023, [MNRAS](#), 525, 2864
- Oke J. B., Gunn J. E., 1983, [ApJ](#), 266, 713
- Planck Collaboration et al., 2020, [A&A](#), 641, A6
- Saxena A., et al., 2023, [A&A](#), 678, A68
- Saxena A., et al., 2024, [A&A](#), 684, A84
- Tang M., Stark D. P., Topping M. W., Mason C., Ellis R. S., 2024, [ApJ](#), 975, 208
- Witstok J., et al., 2024, [A&A](#), 682, A40
- Witstok J., et al., 2025, [MNRAS](#), 536, 27

## A Lookup

### A.1 Converting between flux and flux densities $F_\nu$ and $F_\lambda$

$$\begin{aligned} F &= \frac{L}{4\pi d^2} \\ F_\nu &= \frac{F}{\Delta\nu}, \quad F_\lambda = \frac{F}{\Delta\lambda} \\ \lambda &= \frac{c}{\nu} \\ \frac{d\lambda}{d\nu} &= -\frac{c}{\nu^2} \\ d\nu F_\nu &= d\lambda F_\lambda \\ F_\nu &= \left| \frac{d\lambda}{d\nu} \right| F_\lambda = \frac{c}{\nu^2} F_\lambda = \frac{\lambda^2}{c} F_\lambda \end{aligned}$$

### A.2 Magnitude system

With AB magnitude  $m_{\text{AB}}$  as defined by (Oke & Gunn 1983):

$$\begin{aligned} m_1 - m_2 &= -2.5 \log_{10} (F_1 / F_2) \\ m_{\text{AB}} &= -2.5 \log_{10} \left[ F_\nu \text{ (erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}) \right] - 48.60 \\ &= 31.4 - 2.5 \log_{10} [F_\nu \text{ (nJy)}] \end{aligned}$$

### A.3 Convenient conversions in frequency/wavelength/velocity space

$$\begin{aligned}
 \lambda &= \frac{\lambda_{\text{line}}}{1 - \frac{\Delta v}{c}}, \quad \nu = \nu_{\text{line}} \left( 1 - \frac{\Delta v}{c} \right) \\
 \Delta \lambda &= \lambda - \lambda_{\text{line}} = \frac{\lambda_{\text{line}}}{\frac{c}{\Delta v} - 1}, \quad \Delta \nu = \nu - \nu_{\text{line}} = -\frac{\Delta v}{c} \nu_{\text{line}} \\
 \Delta v &= \left( 1 - \frac{\lambda_{\text{line}}}{\lambda} \right) c = \frac{c}{1 + \frac{\lambda_{\text{line}}}{\Delta \lambda}} = -\frac{\Delta \nu}{\nu_{\text{line}}} c \\
 &= \frac{c}{1 + R} = \frac{zc}{1 + z} \approx zc \quad (z \ll 1)
 \end{aligned}$$