1. a) The ~~An~~ astronomer Edwin Hubble made the amazing discovery that all the other galaxies he could see were moving away from the Milky Way. He also found that the further away other galaxies were, the quicker they travelled away. This means that if we work backwards, we figure that the universe must have started from a single point a very long time ago. In fact, it's still growing every day!

b) Hubble's discovery falsifies the "steady state" theory of the universe. He found that all observed galaxies were moving away from the Milky Way, which contradicts the notion that the universe had been and always would be mostly the same in any direction. Thus, by deduction, we can reject the "steady state" theory and replace it with Hubble's observation (which currently remains our best provisional theory).

c) Assuming the observable universe is spherical with a radius of 46.5 billion light-years:

$$\text{volume of observable universe, } V = \frac{4}{3}\pi r^3$$
$$= \frac{4}{3}\pi(46.5\times10^9)^3$$
$$\approx 4.2\times10^{32}\ \text{lyr}^3$$

An average-sized galaxy has 100 billion stars. There are about 125 billion galaxies in the observable universe. So:

$$\text{stars in the observable universe, } n = (100\times10^9)\times(125\times10^9)$$
$$= 1.25\times10^{22}$$

Assuming stars are evenly spaced:

$$\text{average number of stars in observable universe} = \frac{n}{V}$$
$$= \frac{1.25\times10^{22}}{4.2\times10^{32}}$$
$$* \approx 3\times10^{-11}\ \text{stars/lyr}^3$$

2. a) Exponential: $y = Ce^{kt}$

This would be a reasonable choice as the data seems to follow an exponential decay trend. C would be positive, as temperature is always positive, and k would be negative for exponential decay.

Power: $y = ax^{p}$

A power function would also be suitable as the data ~~decreases less rapidly~~ suggests that temperature decreases less rapidly as time gets larger. Thus, p would be negative to model this decrease. a would be positive as temperature is always positive.

b) Two data points from graph: (0.75, 2.5) and (2.62, 1).

$$\text{gradient, } m = \frac{\text{rise}}{\text{run}} = \frac{1-2.5}{2.62-0.75} = -0.802\ldots \approx -0.8 \quad \text{(unitless)}$$

~~To find y-intercept, let x = 0.~~

$y - y_1 = m(x - x_1)$: let $x_1 = 2.62$, $y_1 = 1$

$$\therefore y - 1 = -0.8(x - 2.62)$$
$$= -0.8x + 2.1\cancel{0}$$
$$y = -0.8x + 3.1 \quad \text{(unitless)}$$

$\therefore y = -0.8x + 3.1$, where y is the natural logarithm of the temperature and x is the natural logarithm of the time

c) let $y = \ln(T)$, $x = \ln(t)$

$$\therefore \ln(T) = -0.8\ln(t) + 3.1$$
$$T = e^{-0.8\ln(t) + 3.1}$$
$$= e^{3.1} \times t^{-0.8}$$

~~$= 22.197\ldots t^{-0.8}$~~

~~$\therefore T \approx 22t^{-0.8}$, where T is temperature (in Kelvins) and t is time (in gigayears)~~

$\therefore T = At^{-0.8}$, $A = e^{3.1}$

where T is temperature in Kelvin, and t is time in gigayears

$\therefore$ best fit function is a power function.

3. a) An estimate of the area under the curve with a triangle:

$$A = \frac{1}{2}bh$$

$$= \frac{1}{2} \times 22 \times 360$$

$$= 3960 \text{ MJy sr}^{-1}\text{cm}^{-1}$$

This ~~techniq~~ technique is reasonably accurate, as the ~~the~~ downward slope of the graph is well estimated by a triangle. Additionally, the area of the triangle is proportional to its height, which matches the fact that the AUC of the CMB intensity spectrum is proportional to the total power of the CMB.

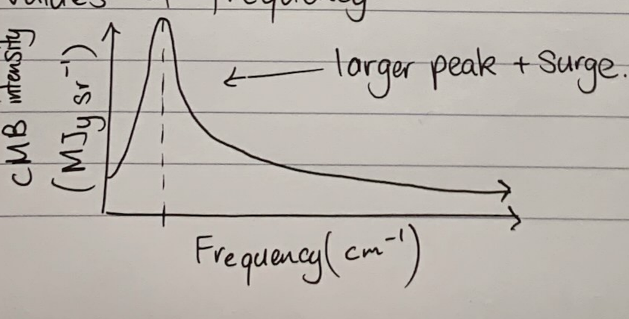
b) Using the trapezoidal method:

$$\cancel{AUC \approx \frac{201+328}{3.63-2.27}}$$

$$AUC \approx \frac{201+328}{2}(3.63-2.27) + \frac{328+381}{2}(4.99-3.63) + \frac{381+369}{2}(6.35-4.99)$$

$$= 1351.84 \text{ MJy sr}^{-1}\text{cm}^{-1}$$

c) The temperature of the universe was very high right after the big bang, but it has been decreasing ever since. As the frequency of the peak of the CMB spectrum is directly proportional to the temperature of the universe, the CMB spectrum would have had a much larger peak several billion years ago. Mathematically, this is equivalent to having a larger "surge" for the smaller values of frequency.



4. a) The units can be found using dimension analysis:

$e^{\beta v/T}$ is unitless, so $\beta v/T$ must be unitless. v/T has units $cm^{-1}K^{-1}$, so β must have units $K\cdot cm$.

$I(v)$ has units $MJy\cdot sr^{-1}$, so αv^3 must have units $MJy\cdot sr^{-1}$ too, since the denominator is unitless. v^3 has units cm^{-3}, so α must have units $MJy\cdot sr^{-1}\cdot cm^3$.

b) let $I(v) = 320\ MJy\cdot sr^{-1}$.

$\Rightarrow I(v) - 320 = 0$

$$\therefore \frac{39.76v^3}{e^{1.44v/2.75}-1} - 320 = 0$$

let $v_0 = 6cm^{-1}$. Then by Newton's method,

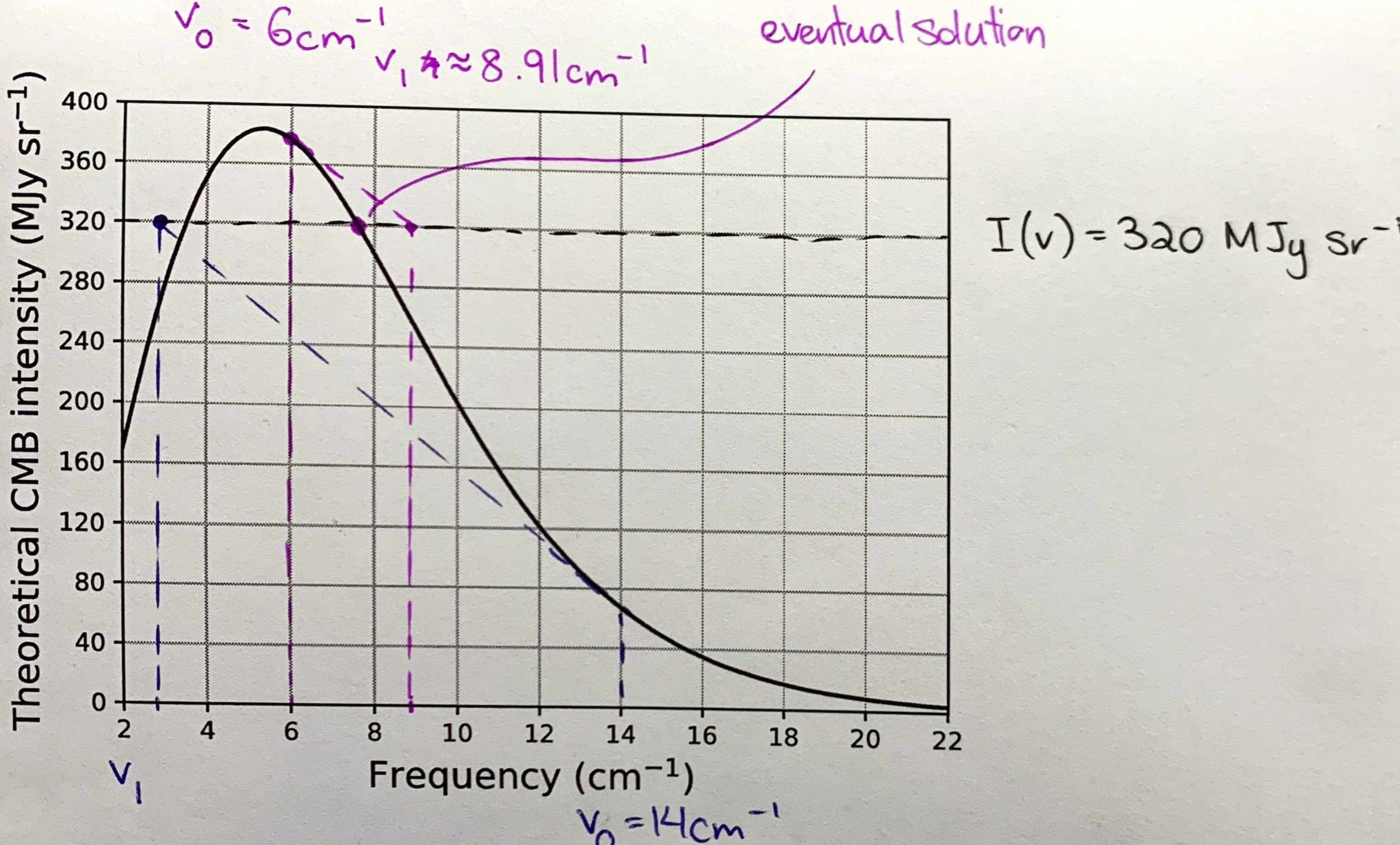
$$v_1 = v_0 - \frac{I(v_0)-320}{I'(v_0)}$$

$$= 6 - \frac{\frac{39.76(6)^3}{e^{1.44\times 6/2.725}-1} - 320}{-19.36}$$

$= 8.907...$

$\approx 8.91\ cm^{-1}$

c) If an initial estimate of $v_0 = 14cm^{-1}$ was used, the eventual solution would be on the "other side" of the surge, as there are two solutions for $I(v) = 320\ MJy\cdot sr^{-1}$



$v_0 = 6\,cm^{-1}$
$v_1 \approx 8.91\,cm^{-1}$
eventual solution
$I(v) = 320$ MJy sr^{-1}
Theoretical CMB intensity (MJy sr^{-1})
400
360
320
280
240
200
160
120
80
40
0
2
4
6
8
10
12
14
16
18
20
22
Frequency (cm^{-1})
v_1
$v_0 = 14\,cm^{-1}$

5. a)

	True Collision? Yes	No
Claimed collision	74 A	34 B
Claimed not a collision	56 C	12786 D

$N = A + B + C + D = 74 + 34 + 56 + 12786 = 12950$

$$\text{accuracy} = \frac{A+D}{N} = \frac{74+12786}{12950} = 0.993\ldots$$

$$\text{Sensitivity} = \frac{A}{A+C} = \frac{74}{74+56} = 0.569\ldots$$

$$\text{specificity} = \frac{D}{B+D} = \frac{12786}{34+12786} = 0.997\ldots$$

b) ~~Prevalence = A+C = 74~~

$$\text{Prevalence} = \frac{\text{no. of true collisions}}{\text{total}} = \frac{74+56}{12950} = 0.010\ldots$$

c) As the experts want to minimise time spent going through "test negative" images, they would favour a higher specificity - that is, the proportion of correct identifications for images that aren't actually collisions. This way, they could rely on those classifications without needing to check them.

Whilst Delta has the highest specificity of 99.9%, their sensitivity is only 67%. Beta has a slightly lower specificity of 99.3% but a significantly higher sensitivity of 92.1%. I think Beta has the best perfomance considering the goals of the project - ~~its~~ their specificity is only lower than Delta's by 0.6% (≈78 pictures out of 12950), so there is a very small difference, ~~This is justified by its Beta's~~ but this is justified by a higher sensitivity.

6. a) The code is using Euler's method with a step size of 0.5 gigayears to calculate an estimate for the density of a nebular over time. This is done with the differential equation $z' = Ae^{-z/C} - Bz$ with appropriate values for A, B, C. The results show the user inputting an initial value $z_0 = 0.1 \times 3\times10^{-17}\ kg \cdot m^{-3}$ and an end time of 10 gigayears. The program then loops through to estimate z for each time step, ending at $t = 10$ gigayears (20 steps) with an estimated $z(10) = 1.069\ldots \times 3\times10^{-17}\ kg \cdot m^{-3}$.

b) 1. Adding comments throughout the code to explain what each section is doing. For example, the function could have a comment describing how to use it. The while loop could indicate its use of Euler's method. Also, units should be ~~comm~~ added in comments to give the variables more meaning and clarity.

2. Adding/printing more information to the output to make the program more user-friendly. A brief summary of the program's purpose, adding units to the input prompt, and mentioning that each line represents a new step and estimate would all greatly benefit user experience.

c) The sceptic is saying that A, B, C are technically wrong as they are not precise enough (i.e. not the "true" values), and as such we cannot use this model, since it is incorrect.

My response would be that we cannot realistically obtain the "true" values for A, B, C. Though this makes the model technically wrong, it is still useful because we do not require a perfect level of precision. The model is "good enough" that we can obtain reasonably reliable results, and more importantly, that we can make use of these results.