

# HW2

William Strickland

October 11, 2019

10/10

## 1 Newman 6.11

**a**

Starting with Eqn. (6.81) from Newman, we can say:

$$x^* = x_i + \epsilon_i x^* = x_{i+1} + \frac{\epsilon_{i+1}}{(1+w)f'(x) - w} \quad (1)$$

Using:

$$x_{i+1} = (1+w)f(x_i) - wx_i \quad (2)$$

and

$$x^* = (1+w)f(x^*) - wx^* \quad (3)$$

We can solve for  $\epsilon_{i+1}$

$$x_i + \epsilon_i = x_{i+1} + \epsilon_{i+1} \quad (4)$$

$$x_i + \frac{\epsilon_{i+1}}{(1+w)f'(x_i) - w} = x_{i+1} + \epsilon_{i+1} \quad (5)$$

$$\frac{\epsilon_{i+1}}{(1+w)f'(x_i) - w} - \epsilon_{i+1} = x_{i+1} - x_i \quad (6)$$

$$\epsilon_{i+1} \left( \frac{1}{(1+w)f'(x_i) - w} - 1 \right) = x_{i+1} - x_i \quad (7)$$

Now we see that

$$\epsilon^* = \epsilon_{i+1} = \frac{x_{i+1} - x_i}{(1/((1+w)f'(x_i) - w) - 1)} \quad (8)$$



```

2.6843321302094658e-05
10
1.0907780881752593e-05
11
4.432542981991361e-06
12
1.8012590188298816e-06
13
7.319848429609976e-07
14
bash-4.2$ █

```

(a) Relaxation method, the first value is the difference of the old  $x$  and the new  $x$ , while the second value is the iteration number

```

0.796997075145081 1
1
0.7968323724281932 0.796997075145081
2
0.7968143476015649 0.7968323724281932
3
0.7968123729832619 0.7968143476015649
4
0.7968121566399141 0.7968123729832619
5
-

```

(b) Using the overrelaxation method, the number of iterations decreased from 14 to 5. The first two values are the old and new  $x$ s, while the second value is the iteration number

*clearer to show me their diff, as in other side*

**b**

We would like to find values of  $x$  which solve the equation

$$x = 1 - e^{-cx} \quad (9)$$

for  $c = 2$ . To do this, we use the method of relaxation. We want to also know how many iterations it takes to solve this. One can see in Figure 1a that using the relaxation method took 14 iterations to converge to the right solution. The relaxation method was used with a starting value  $x_0 = 1$ .

**c**

Now we would like to compare our results using the overrelaxation method. This method uses a parameter  $\omega$ , which helps us converge to the correct value of  $x$  faster, as seen in Figure 1b. This method was computed for a starting value  $x_0 = 1$ , and  $\omega = 0.5$ .

**d**

Using  $\omega < 1$  will help in cases where we overestimate the solution  $x$ . If we use  $\omega > 1$ , we will bounce around the solution for a long time before getting to the true solution. However, using  $\omega < 1$  will take an overestimate, and bring us closer to the original value.

## 2 Newman 6.13

**a**

Planck's radiation law is given by:

$$I(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \quad (10)$$

Differentiating this function with respect to the wavelength, we can find the wavelength that maximizes the intensity of emitted radiation.

$$\frac{dI}{d\lambda} = 5e^{-\frac{hc}{\lambda k_B T}} + \frac{hc}{\lambda k_B T} - 5 = 0 \quad (11)$$

Substituting

$$x = \frac{hc}{\lambda k_B T}$$

We can solve equation 11 for the wavelength.

$$\lambda = \frac{hc}{x k_B T} = \frac{b}{T} \quad (12)$$

**b**

Using the binary search method, we can solve this equation for lambda. My binary search started in the interval  $[x=1, x=7]$ , and I got a value for

$$b = 0.002900 \quad (13)$$

**c**

Plugging this value back into equation 12, and solving for T with a wavelength of 502 nm, we get

$$T = 5776K \quad (14)$$

This value agrees very well with the expected value for the temperature of the sun, which is 5778 K<sup>1</sup>

---

<sup>1</sup><https://web.archive.org/web/20100715200549/http://nssdc.gsfc.nasa.gov/planetary/factsheet/sunfact.html>

### 3 Fitting $\chi^2$ of the Schechter function

This question involved work with the Schechter function, which tells you the inverse of the density of some galaxy with a mass  $M_{gal}$ . This function looks the following:

$$n(M_{gal}) = \phi^* \left( \frac{M_{gal}}{M^*} \right)^{\alpha+1} e^{\left( -\frac{M_{gal}}{M^*} \right)} \ln(10) \quad (15)$$

Where  $\phi^*$  is the amplitude,  $M_{gal}$  is the characteristic mass scale, and  $\alpha$  is the low mass slope of the function. We are given measurements on the Schechter function for given values of  $M_{gal}$ , which are shown in figure 3, along with the errors of  $n(M_{gal})$ . Our job is to conduct a  $\chi^2$  test for the given schechter function and the measured values, and create a best fit of this data for free parameters  $\phi$ ,  $M_{gal}$ , and  $\alpha$ .

First we do this in two dimensions, where we want to minimize a function of two variables,  $x$ , and  $y$ , shown below:

$$f(x, y) = (x - 2)^2 + (y - 2)^2 \quad (16)$$

We know the minima of this function, so we would like to write a program that will find these minima. Using the gradient descent method in 2 dimensions, I was able to write a program that converged to the correct values of the minima of this function. The convergence is shown in figure ???. Next we would like to extrapolate this method to 3D, and minimize the function  $\chi^2$ . For the values of the parameters then minimize  $\chi^2$ , we can then plot the schechter function as a function of  $M_{gal}$ . This will give a best fit of our data. The gradient descent method relies heavily on the starting points of the three parameters however.

In Figure 4, I show what the schechter function looks like after the initial parameters are inputted. One can see that this gives a poor fit. We then conduct gradient descent on  $\chi^2$ , which is the function we'd like to minimize, and this tells us values of the the three fitting parameters that minimize  $\chi^2$ . The gradient was calculated using central difference algorithm, using a fixed step size  $h$  for each of the three parameters. I also worked with  $\log(M_{gal})$  and  $\log(\phi)$ . This made it easier to deal with very large numbers.

Going from Figure 4, to Figure 5, one can see that the gradient descent method worked and was able to output good fitting parameters, leading to a good fit.

We would like to examine how sensitive the method is for values away from the correct ones. In figure 7a, 7b, and 7c, I was able to change each fitting parameter a small amount away from their expected values and see how good the fit remains. By inspection, we can see that if you change one of the parameters a little bit (by about 10%, the fits remain pretty accurate, as no noticeable change can be recorded. I found, however, that there were values of the initial conditions that lead the gradient descent method to diverge. These limits were shown to be around  $10^{-6} < \phi < 10^{-1}$ ,  $10^8 < M_{gal} < 10^{14}$ , and  $-3 < \alpha < 2.5$

I will note that there is a better way of seeing the sensitivity of  $\chi^2$  than the way I did it. Since my method could only vary one parameter at a time, keeping the other two fixed, you do not see the affect of varying two parameters at the same time. A better way to see this would be to create a contour plot of  $\chi^2$  as a function of two parameters, leaving one fixed. This would tell you the minima in the function, as well as good locations to start your gradient descent method, as some initial conditions will lead the gradient descent method to diverge.

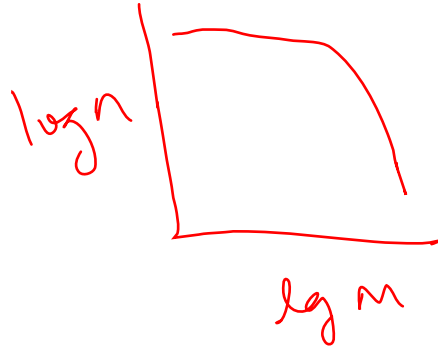
plotting this is  
better

```
[1.9998670772004215, 1.9998670772004215]
[1.9998936617603371, 1.9998936617603371]
[1.9999149294082696, 1.9999149294082696]
[1.9999319435266156, 1.9999319435266156]
[1.9999455548212925, 1.9999455548212925]
[1.999956443857034, 1.999956443857034]
[1.9999651550856272, 1.9999651550856272]
[1.9999721240685018, 1.9999721240685018]
[1.9999776992548015, 1.9999776992548015]
[1.9999821594038412, 1.9999821594038412]
[1.999985727523073, 1.999985727523073]
[1.9999885820184584, 1.9999885820184584]
[1.9999908656147667, 1.9999908656147667]
[1.9999926924918132, 1.9999926924918132]
[1.9999941539934505, 1.9999941539934505]
[1.9999953231947605, 1.9999953231947605]
[1.9999962585558084, 1.9999962585558084]
```

x  
n

Figure 2: Convergence to the minima of the sample function in equation 16. The starting values were  $x = 1$  and  $y = 1$ .  $\gamma$  was  $10^{-2}$  and I used a fixed step size  $h = 10^{-1}$

$\log n$



$\log m$

this is the log

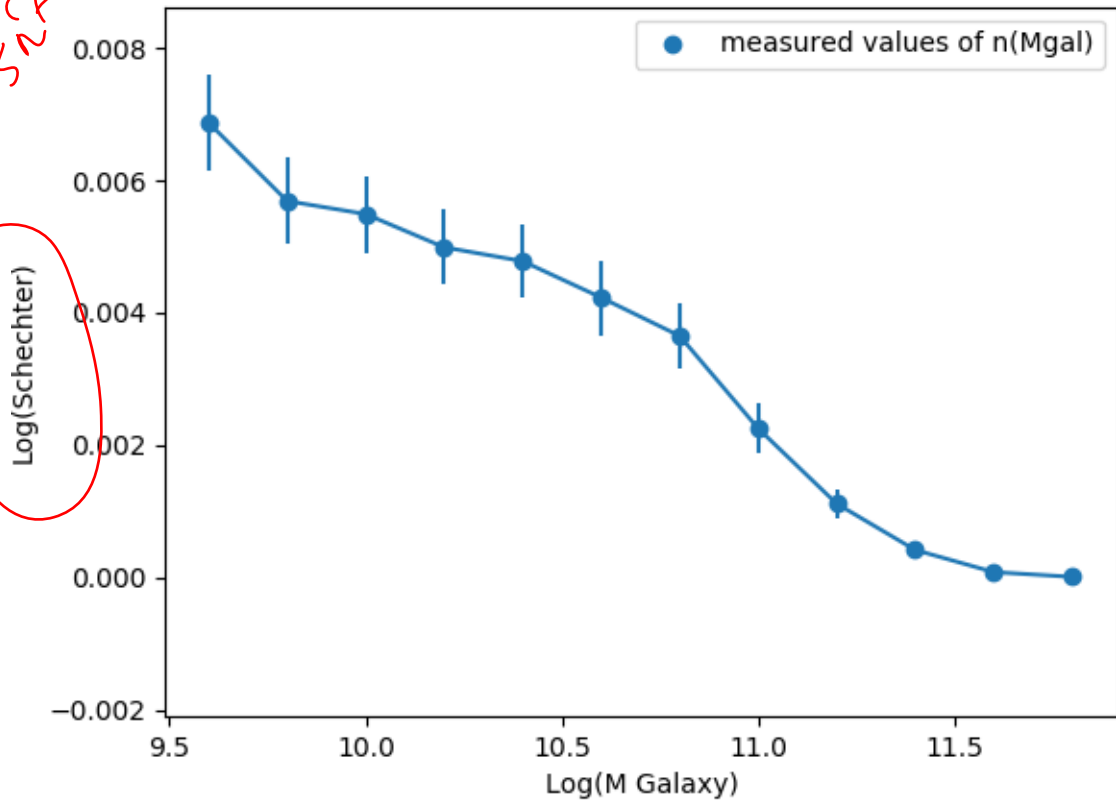


Figure 3: Schechter function measurements  $n(M_{gal})$  for different values of  $M_{gal}$ . Error bars are the errors of  $n(M_{gal})$

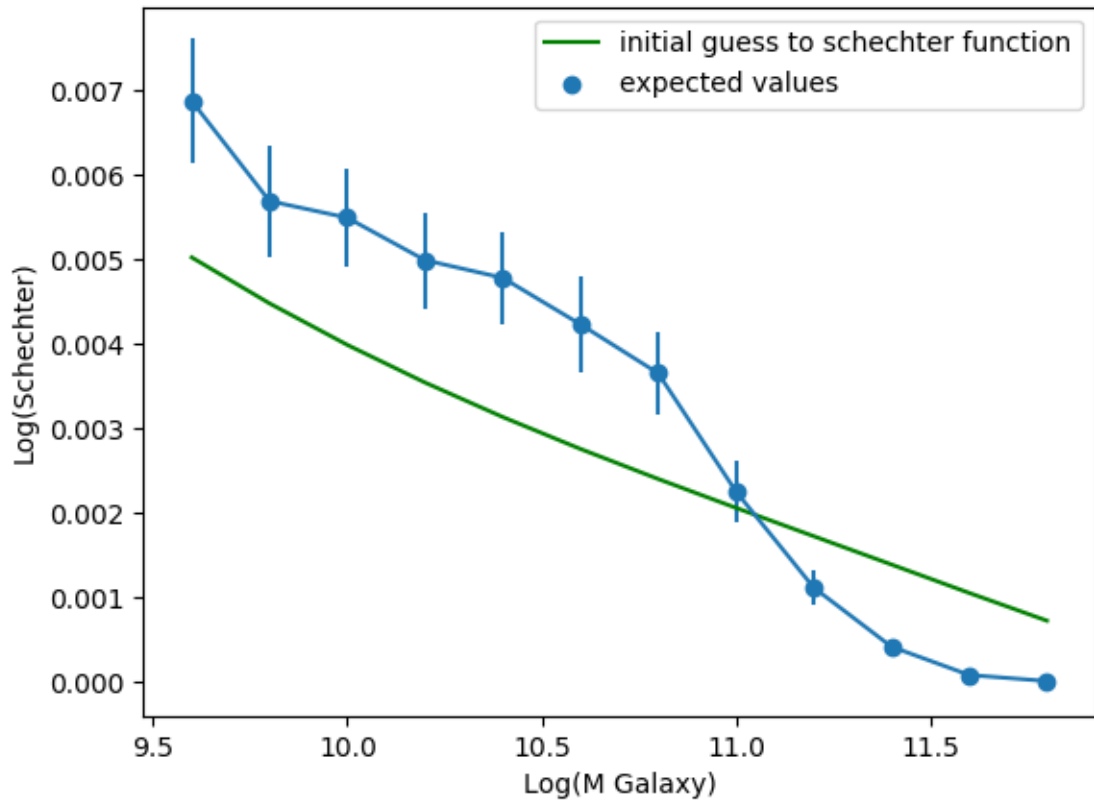


Figure 4: Schechter function evaluated for the initial guess parameters  $\phi = 10^{-3.23}$ ,  $M_{gal} = 10^{12}$ ,  $\alpha = -1.243$

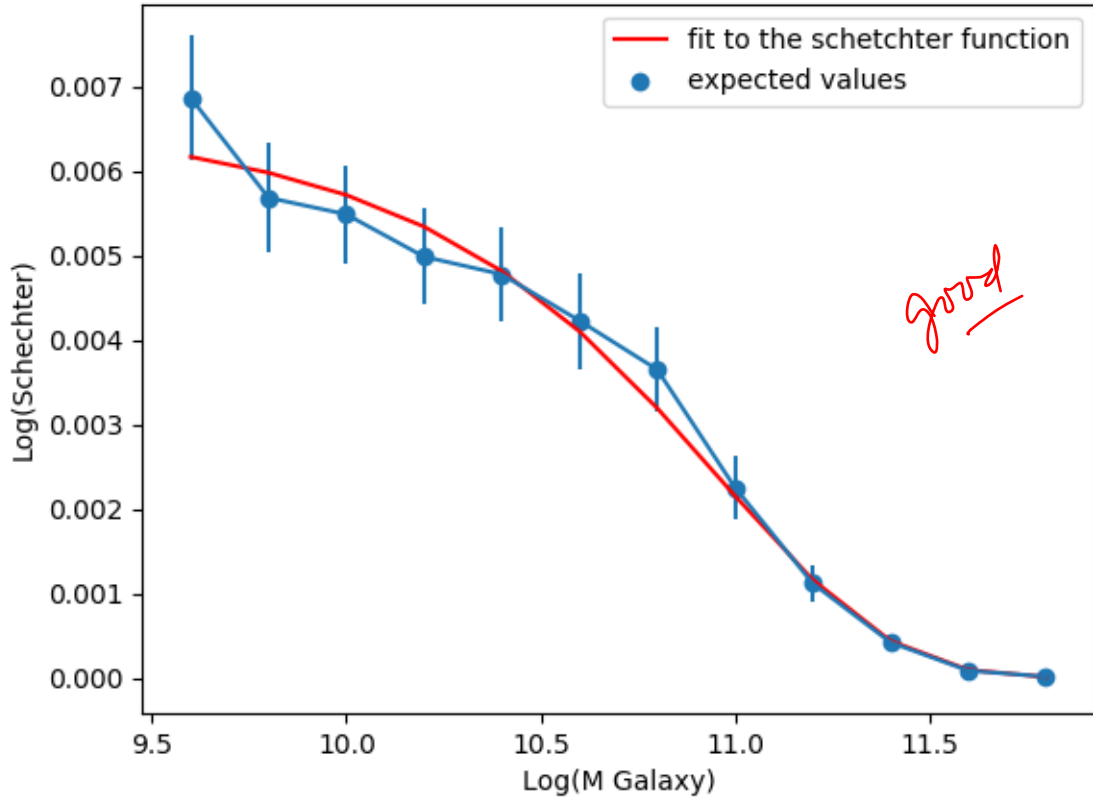


Figure 5: Fit of the Schechter function with three fitting parameters. My final values were  $\phi = 10^{-2.528}$ ,  $M_{gal} = 10^{10.95}$ ,  $\alpha = -0.943$  ✓ okay.

plot ↘

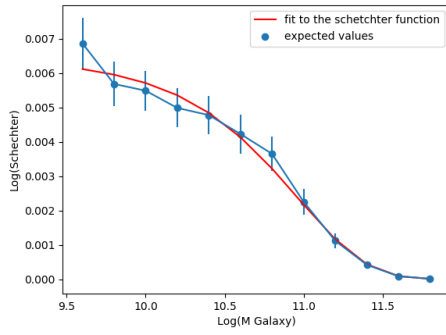
```

1.0025411669722928e-06 6.517683832640842e-07 6.517683832640842e-07
1.0023287404514747e-06 6.516294561720315e-07 6.516294561720315e-07
1.002116359227756e-06 6.514905592780451e-07 6.514905592780451e-07
1.0019040224129583e-06 6.51351692582125e-07 6.51351692582125e-07
1.00169173089526e-06 6.512128560842712e-07 6.512128560842712e-07
1.0014794842305719e-06 6.510740480081267e-07 6.510740480081267e-07
1.0012672819748047e-06 6.509352701300486e-07 6.509352701300486e-07
1.001055125016137e-06 6.507965206736799e-07 6.507965206736799e-07
1.0008430129104795e-06 6.506578031917343e-07 6.506578031917343e-07
1.0006309456578322e-06 6.505191141314981e-07 6.505191141314981e-07
1.000418923258195e-06 6.503804534929714e-07 6.503804534929714e-07
1.0002069452674789e-06 6.502418248288677e-07 6.502418248288677e-07
9.999950125738621e-07 6.501032245864735e-07 6.501032245864735e-07

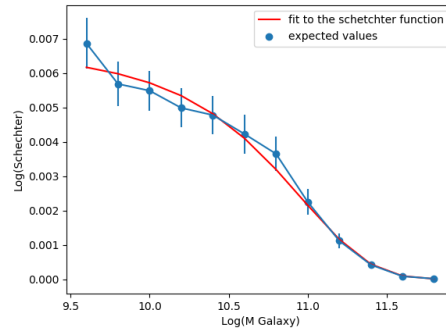
```

Figure 6: step sizes of my gradient descent method for  $\chi^2$ . these show a convergence to the value  $10^{-6}$  for values of the step size. We can see that a minimum  $\chi^2$  is found when the step size gets lower than this tolerance.

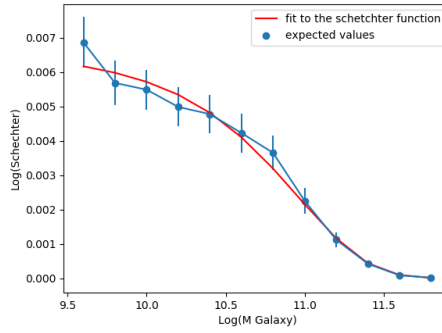




(a) Phi



(b) Alpha



(c) Mgal

changing all 3 at  
same time is  
more robust  
tests.  
(but deep)

Figure 7: Fits obtained by changing initial parameters, one at a time. For all the above plots I used a fixed value of  $\gamma$ , as well as fixed step sizes in the gradient of  $\chi^2$

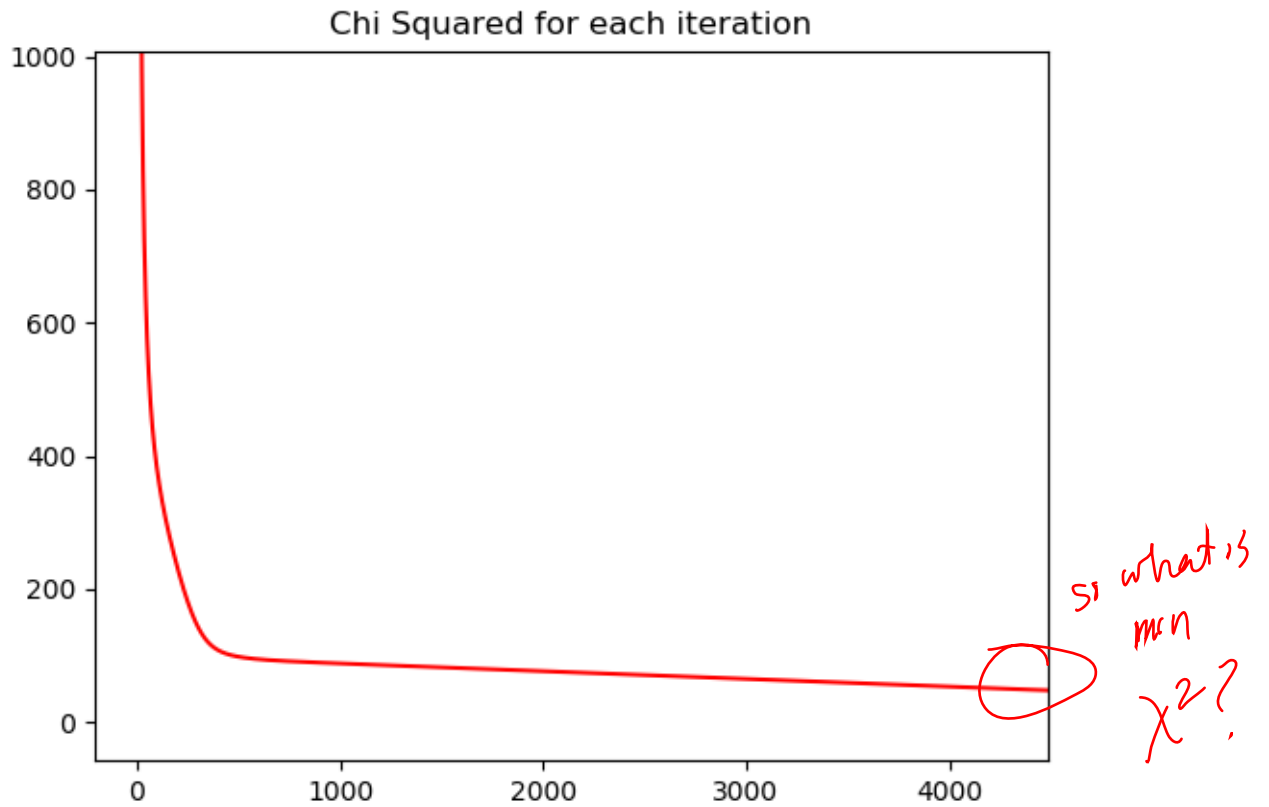


Figure 8:  $\chi^2$  evaluated for the initial guess parameters  $\phi = 10^{-3.23}$ ,  $M_{gal} = 10^{12}$ ,  $\alpha = -1.243$  at every step in gradient descent process.  $\gamma = 10^{-5}$  for this case, however, it should be stated that this method works for  $\gamma$  up to 0.1.<sup>2</sup>