NUR - Hand-in Exercise 2

Martina Cacciola (s4170814)

March 26, 2024

Abstract

In this document, I present the results for the Hand-in Exercise 2 for the course Numerical Recipes for Astrophysics.

1 Satellite galaxies around a massive central

The exercise is done in the script satellite.py. The necessary explanations of the methods used are in the comments of the code. For the purpose of this task, we implement a Random Number Generator (RNG) as follows:

```
class CombinedRNG:
      Pseudo-random number generator combining the XOR-shift and Multiply-With-Carry (MWC)
     We set the seed to the current time in milliseconds, as a source of variability
      def __init__(self , seed = None):
         # If no seed is provided, use the current time in milliseconds
         if seed is None:
             seed = int(time.time() * 1000)
         # Initialize the states for XOR-shift and MWC methods
         self.xor\_state = seed
         self.mwc\_state = seed
12
      def xor_shift(self):
14
         # XOR-shift method
         # The seed is converted to a 64-bit unsigned integer
         #By performing a shift of 21 bits to the left and XORing the result with the
17
      original state
         # Then, a shift of 35 bits to the right and again XOR with the previous result
         self.xor_state ^= (self.xor_state >> 35)
20
21
         # Finally, a shift of 4 bits to the left and XOR with the previous result
         22
         return self.xor_state
23
24
      def mwc(self):
25
         # Multiply-With-Carry (MWC) method
         # Set the multiplier
27
         a \, = \, 4294957665
         # Extract the upper and lower 32 bits of the state
29
         {\tt mwc\_upper} \, = \, (\, {\tt self.mwc\_state} \, \, \& \, \, 0 \\ {\tt xFFFFFFF000000000} \,) \, >> \, 32
30
         mwc_lower = self.mwc_state & 0x00000000FFFFFFFF
31
         # Calculate the next state by multiplying the lower 32 bits with the multiplier
32
         # and adding the upper 32 bits
33
         x_next = a * mwc_lower + mwc_upper
34
         # Update the state using AND operation with a 64-bit mask
35
36
         # Only the lower 64 bits are kept, the rest are set to zero
         37
         return x_next >> 32
38
39
      def combined_rng(self):
40
         # Combine the XOR-shift and MWC methods by XORing their outputs
41
```

satellite.py

For question (a), we do the following:

```
## 1a)
  # Define the parameters
  a = 2.4
  b = 0.25
  c = 1.6
  xmax = 5
  Nsat = 100
  def integrand(x):
      # Take into account singularity at x=0
      # Consider the volume element is 4*pi*x^2 (4*pi is taken outside)
12
      if x = 0:
13
          return 0
      return Nsat * (x/b)**(a-3) * exp(-(x/b)**c) * x**2
17
  # Implemement a numerical integration
  # Using the trapezoidal rule
18
  def trapezoidal_rule(f, a, b, n):
20
      Inputs:\\
21
      f: Function to integrate
22
      a, b: Integration limits
23
      n: Number of intervals
24
      It works by approximating the integral of f(x) between a and b
      by the sum of the areas of trapezoids formed by the function and the x-axis
26
27
      # Width of each trapezoid
28
      \ddot{h} = (b - a) / n
29
      # Initialize the result with the average of the function at the limits
30
      result = 0.5 * (f(a) + f(b))
31
32
       for i in range (1, n):
          # Add the value of the function at each interior point
33
          #Each of the function values corresponds to the height of a trapezoid
34
           result += f(a + i * h)
35
36
      # Multiply the sum by the width of the trapezoids
      # This gives the total area under the curve
37
      result *= h
38
      return result
39
  # Romberg integration
42
  def romberg(f, a, b, n):
44
      Inputs:
45
      f: Function to integrate
46
      a, b: Integration limits
47
      n: Number of rows in the Romberg table
48
49
      R = np.zeros((n, n))
50
      # At each iteration, calculates an approximation of the integral
51
      # using the trapezoidal rule with 2 ** i intervals
52
      # This forms the first column of the Romberg table
54
      for i in range(n):
          R[i, 0] = trapezoidal_rule(f, a, b, 2 ** i)
55
      # At each iteration, calculates an improved approximation of the integral
56
      # using Richardson extrapolation
```

```
# Takes a weighted average of the current approximation and the previous one
       # The weights are chosen to cancel out as much of the error as possible
       # This fills the rest of the Romberg table
60
       for j in range (1, n):
61
            for k in range(j, n):
62
                R[\,k\,,\ j\,]\ \stackrel{.}{=}\ \stackrel{.}{R[\,k\,,}\ \stackrel{.}{j}\ -\ 1]\ +\ (R[\,k\,,\ j\ -\ 1]\ -\ R[\,k\,-\ 1\,,\ j\ -\ 1]\,)\ /\ (4\ **\ j\ -\ 1)
63
       return R[-1, -1] # Return the last element of the last row of the Romberg table
64
65
66
  # Apply the method above to the integral
  n\,=\,10
68
  m = 6
69
70 R = romberg(integrand, 0, xmax, m)
  result = R * 4 * pi
  A = Nsat / result
  with open ("normalization.txt", "w") as file:
       file.write("The result of the numerical integration is: " + str(A))
```

satellite.py

The normalization factor A we obtain is the following:

```
The result of the numerical integration is: 9.05216643171278
```

normalization.txt

. For question (b), we generate 3D satellite positions that statistically follow the satellite profile n(x). We sample them from the probability density function p(x), using the inverse transform sampling. We do this with the following code:

```
## 1b)
  # Define the number density of galaxies
  # This represents the number of galaxies per unit volume at a given radius x
  def n(x):
      return A * Nsat * (x/b)**(a-3) * exp(-(x/b)**c)
  \# Define the number of galaxies N(x) in a shell of radius x and thickness dx
  # We compute N(x) by multiplying the number density n(x) by the volume element 4*pi*x^2
  def N(x):
      return 4 * np. pi * x**2 * n(x)
  # Define the probability distribution function p(x) given that p(x)dx = N(x)dx / Nsat
13
  # It gives the probability of finding a galaxy at a given radius x
14
  def p_x(x):
      return N(x) / Nsat
16
17
18
  # Implement inverse transform sampling to sample from the distribution
19
  def inverse_transform_sampling(pdf, n_samples, x_min, x_max):
20
      x_values = np.linspace(x_min, x_max, 10000)
      cdf_values = np.zeros_like(x_values)
22
      # Calculate the cumulative distribution function using numerical integration (
      trapezoidal rule)
      # The cdf at a point x is the integral of the pdf from x_min to x
24
      for i in range(1, len(x_values)):
25
          cdf\_values[i] = cdf\_values[i-1] + trapezoidal\_rule(pdf, x\_values[i-1], x\_values[i-1])
26
      cdf_values /= cdf_values[-1] # Normalize to ensure CDF ranges from 0 to 1
27
28
      # Generate random numbers uniformly distributed between 0 and 1
      random\_numbers = rng.uniform(0, 1, num\_samples=n\_samples)
30
31
      # Apply the inverse CDF to the random numbers
32
      # For each random number, find the corresponding value of x
33
      # such that the CDF of x is equal to the random number
34
      # This is done by finding the first x value for which the CDF is greater than the
35
      random number
36
      # The corresponding x value is then stored as a sampled point
      sampled_points = np.zeros(n_samples)
```

```
for i in range(n_samples):
           for j in range(len(cdf_values)):
39
                if random_numbers[i] < cdf_values[j]:
40
                    sampled_points[i] = x_values[j]
41
                    break
42
43
       return sampled_points
44
45
  # Generate 10,000 sampled points
46
  n_samples = 10000
   x_{min}, x_{max} = 10**-4, 5
48
   h = (x_max - x_min) / n_samples
49
   sampled_points = inverse_transform_sampling(p_x, n_samples, x_min, x_max)
   # Calculate the bin edges
52
   bin\_edges = np.logspace(np.log10(x\_min), np.log10(x\_max), 21)
53
   # Calculate histogram and divide each bin by its width
   hist, _ = np.histogram(sampled_points, bins=bin_edges, density=False)
   bin_widths = np.diff(bin_edges)
57
   hist_density = hist / (bin_widths * n_samples / Nsat) # Correcting for the
       normalization offset
   \# Plot the analytical function N(x) and the histogram of the sampled points on a \log-\log
   x_values = np.logspace(-4, np.log10(x_max), 20)
   N_x-values = [N(x) \text{ for } x \text{ in } x-values]
62
63
   plt.figure(figsize = (10, 6))
65
   plt.loglog(x\_values, N\_x\_values, label='Analytical N(x)', color='r')
  plt.bar(bin_edges[:-1], hist_density, width=bin_widths, label='Sampled Points Histogram'
   plt.xlabel('Relative Radius x')
  plt.ylabel('Number of Galaxies N(x)')
69
  plt.ylim(10**-3, 10**5)
   plt.legend()
  plt.title('Analytical N(x) vs Sampled Points Histogram')
   plt.grid(True)
   plt.savefig('my_solution_1b.png', dpi=600)
   ## 1c)
76
77
   # Define a selection method following these rules:
78
   # Select each galaxy with same probability
   # Not draw the same galaxy twice
   # Not reject any drawn galaxy
   def reservoir_selection(sampled_points, n_galaxies):
       selected_galaxies = []
83
       num_samples = len(sampled_points)
84
       for i in range(num_samples):
85
           # If we haven't selected n_galaxies yet, add the current galaxy
86
87
           if i < n_galaxies:
               selected_galaxies.append(sampled_points[i])
88
89
           else:
               # If we have already selected n_galaxies, decide whether to replace one of
       them
               # Generate a random index j between 0 and i
91
               j = rng.combined_rng() \% (i + 1)
92
               # If j is less than n_galaxies, replace the galaxy at index j with the
93
       current galaxy
                if j < n_galaxies:
94
                    selected_galaxies[j] = sampled_points[i]
9.5
       return selected_galaxies
97
   # Define a sorting method (quicksort)
  # Takes an array and two indices: low (starting index of the array) and high (last index
   def quick_sort (arr, low, high):
       # If the low index is less than the high index, there are elements in the array to
101
```

```
be sorted
        if low < high:
           # partition function called to partition the array around a pivot.
            # The pivot's final position in the sorted array is returned as pi
            pi = partition (arr, low, high)
106
            # quick_sort function recursively called for the parts of the array before pi
       and after pi
            quick_sort (arr, low, pi - 1)
108
            quick_sort(arr, pi + 1, high)
        return arr
   # This function partitions the array around a pivot
112
   def partition(arr, low, high):
113
       # Choose the pivot as the median of the first, middle, and last elements
115
       mid = low + (high - low) // 2
       pivot = arr[mid]
       # If the first element is greater than the middle element, swap them
118
        if arr [low] > arr [mid]:
119
120
            arr [low], arr [mid] = arr [mid], arr [low]
       # If the middle element is greater than the last element, swap them
        if arr[mid] > arr[high]:
123
            arr[mid], arr[high] = arr[high], arr[mid]
125
       # If the first element is greater than the middle element, swap them
126
        if arr[low] > arr[mid]:
127
            arr [low], arr [mid] = arr [mid], arr [low]
128
129
       # Now the middle element is the median of the first, middle, and last elements
130
       pivot = arr[mid]
131
132
       # Swap the pivot with the high element
133
       arr [mid], arr [high] = arr [high], arr [mid]
       # i is set to one index less than the low index
136
       i = (low - 1)
       # The array is iterated from the low to the high index
138
139
        for j in range(low, high):
           # If the element is less than or equal to the pivot, the element is swapped with
140
        the one at index i + 1
            if arr[j] <= pivot:</pre>
141
                i = i + 1
142
                arr\,[\,i\,]\;,\;\; arr\,[\,j\,]\;=\; arr\,[\,j\,]\;,\;\; arr\,[\,i\,]
143
       # After all elements have been checked, the pivot is swapped with the element at
145
       index i + 1
       # placing the pivot in its correct position in the sorted array
146
       arr[i + 1], arr[high] = arr[high], arr[i + 1]
```

satellite.py

We compare the distribution of the sampled points with N(x), the analytical function describing the number of galaxies in a shell of thickness dx at a given distance x. From Fig.1, we can see that the sampled points match the expected distribution. Nevertheless, we can see how the samples are not present in the first interval of x values. Inverse sampling works by transforming uniform random numbers into samples from the desired distribution, using the cumulative distribution function (CDF) of the PDF. The CDF maps values from the sample space to the interval [0, 1], and its inverse can be used to map values from [0, 1] back to the sample space. In the case of smaller values, a small range of uniform random numbers can map to a large range of small values in the sample space, leading to undersampling in these regions. This might explain why there is a discrepancy at smaller values in our histogram.

For question (c), we select 100 random galaxies from the ones sampled in point (b), using Reservoir method. It is a sampling algorithm that chooses a random sample, without replacement, of k items from a population of unknown size n in a single pass over the items. The 100 drawn galaxies are ordered from smallest to largest radius. We do this as follow:

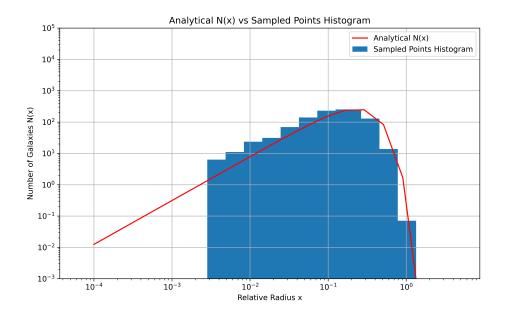


Figure 1: Plot in logarithmic scale showing N(x), function of the number of galaxies at given distance x, and the histogram of the 10000 sampled points. The sampled points match the analytical distribution, but there is a problem of undersampling in the region of smaller radii.

```
return (i + 1)
  # Select 100 galaxies
  n_{-}galaxies = 100
  selected_galaxies = reservoir_selection(sampled_points, n_galaxies)
  # Sort the galaxies using quicksort from smallest to higher radius
  sorted_galaxies = quick_sort(selected_galaxies,0,len(selected_galaxies)-1)
  # Plot the cumulative number of the chosen galaxies
  fig1c, ax = plt.subplots()
  ax.plot(sorted\_galaxies, np.arange(n\_galaxies))
  ax.set(xscale='log', xlabel='Relative radius
          ylabel='Cumulative number of galaxies'
          xlim = (x_min, x_max), ylim = (0, n_galaxies))
  plt.savefig('my_solution_1c.png', dpi=600)
  ## 1d)
  # Define Ridders' method for numerical differentiation
20
  def ridders_method(f, x, h, m, tol=1e-10):
      \# Compute first approx with central difference for high h D = [(f(x+h) - f(x-h)) \ / \ (2*h)]
22
23
      # Decrease h by a factor of 2 and calculate a new approximation. Repeat until you
25
      have m approximations
26
       for i in range(1, m):
          h\ /\!=\ 2
27
          D. append ((f(x + h) - f(x - h)) / (2 * h))
           for j in range(i):
29
               D[j] = (4**(j+1) * D[j+1] - D[j]) / (4**(j+1) - 1)
30
31
      # Terminate when the improvement over previous best approximation is smaller than
32
      the target error or if error grows
       best_approximation = None
33
       for i in range(1, len(D)):
34
           if abs(D[i] - D[i-1]) < tol:
```

```
best_approximation = D[i]
37
               break
38
           # Terminate early if the error grows and return the best approximation from
39
       before that point.
           if abs(D[i] - D[i-1]) > abs(D[i-1]):
40
               best\_approximation = D[i-1]
41
42
43
44
       if not best_approximation:
           print ("Tolerance not reached or error grew significantly. Consider increasing m.
```

satellite.py

The number of galaxies within a radius x are shown in Fig. 2. The radii selected are starting from higher values because of the previous undersampling. The increasing trend makes sense as the number of galaxies within a given radius should increase as the radius increases.

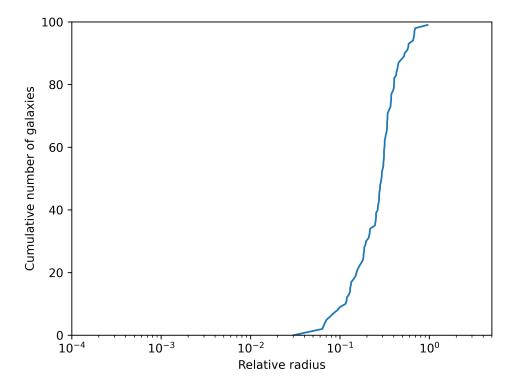


Figure 2: Plot showing the cumulative number of galaxies randomly selected with respect to the relative radius x. The number of galaxies progressively increases with larger distances.

For question (d), we compute the derivative of the function n(x) using both the Ridders' method and the analytical procedure. The code is below:

```
# Now the middle element is the median of the first, middle, and last elements
pivot = arr[mid]

# Swap the pivot with the high element
arr[mid], arr[high] = arr[high], arr[mid]

# i is set to one index less than the low index
i = (low - 1)
# The array is iterated from the low to the high index
```

```
for j in range(low , high):
12
          # If the element is less than or equal to the pivot, the element is swapped with
       the one at index i + 1
          if arr[j] <= pivot:</pre>
13
              i = i + 1
14
               arr[i], arr[j] = arr[j], arr[i]
16
17
      # After all elements have been checked, the pivot is swapped with the element at
      index i + 1
      # placing the pivot in its correct position in the sorted array
      arr[i + 1], arr[high] = arr[high], arr[i + 1]
19
      # Return the final position of the pivot
20
      return (i + 1)
22
  # Select 100 galaxies
23
  n_galaxies = 100
24
  selected_galaxies = reservoir_selection(sampled_points, n_galaxies)
25
  # Sort the galaxies using quicksort from smallest to higher radius
27
  sorted_galaxies = quick_sort(selected_galaxies,0,len(selected_galaxies)-1)
  # Plot the cumulative number of the chosen galaxies
30
31
  fig1c, ax = plt.subplots()
32
  ax.plot(sorted_galaxies, np.arange(n_galaxies))
  ax.set(xscale='log', xlabel='Relative radius'
33
         ylabel='Cumulative number of galaxies'
         xlim = (x_min, x_max), ylim = (0, n_galaxies))
35
  plt.savefig('my_solution_1c.png', dpi=600)
36
37
  ## 1d)
38
39
  # Define Ridders' method for numerical differentiation
40
  def ridders\_method(f, x, h, m, tol=1e-10):
41
      # Compute first approx with central difference for high h
      D = [(f(x + h) - f(x - h)) / (2 * h)]
43
44
      # Decrease h by a factor of 2 and calculate a new approximation. Repeat until you
      have m approximations
```

satellite.py

The obtained results can be seen in:

derivative.txt

2 Heating and cooling in HII regions

The code relative to this exercise is in the script heating.py. The code relative to point (a) is the following:

```
## 2a)

# Define the equilibrium function

def equilibrium_1(T):
    term1 = psi * Tc * k
    term2 = (0.684 - 0.0416 * np.log(T / (1e4 * Z * Z))) * T * k

return term1 - term2

# Measure the time taken and find the root of the equilibrium function using the bisection method

def bisection_method(f, a, b, tol=1e-10, max_iter=100):
    ''''

Find the root of the function f(x) = 0 using the bisection method
    Takes the function f, the endpoints of the initial interval [a, b],
    and optional arguments tol and max_iter

''''
```

```
if f(a) * f(b) >= 0:
           print("Bisection method fails.")
           return None, None
18
19
      # Initialize values
20
      num_steps = 0
21
22
       start_time = time.time()
23
       while (b - a) / 2 > tol and num\_steps < max\_iter:
           c = (a + b) / 2
24
25
           if f(c) == 0:
               end_time = time.time()
26
               return c, num_steps + 1, end_time - start_time
28
           # Update the interval
29
           if f(c) * f(a) < 0:
30
               b = c
31
32
           else:
               a = c
33
           num_steps += 1
34
35
36
       end_time = time.time()
       return (a + b) / 2, num_steps, end_time - start_time
37
38
  # Define the constants
39

m k = 1.38e{-}16 \ \# \ erg/K
40
  aB = 2e-13 \# cm^3 / s
  Z = 0.015 \# metallicity
42
  Tc = 1e4 * Z**2 \# stellar temperature in K
43
  psi = 0.929
45
  # Measure the time taken and find the root of the equilibrium function using the
  a = 1
47
  b = 1e7
  T_eq, num_steps, time_taken = bisection_method(equilibrium_1, a, b)
49
50
51
52
  if T_eq is not None:
       output\_text = (f"The equilibrium temperature is {T_eq:.2f} K.\n"
53
54
                       f"The bisection method found the root in {num_steps} steps.\n"
                       f"The time taken was {time_taken:.10f} seconds.\n")
56
       output_text = "The bisection method did not converge."
57
  # Writing output to a file
  with open("2a.txt", "w") as file:
60
       file.write(output_text)
```

heating.py

The output of point (a) is in

```
The equilibrium temperature is 3.12 K.

The bisection method found the root in 56 steps.

The time taken was 0.0001060963 seconds.
```

2a.txt

For point (b), we have adopt again the bisection method, which guarantees convergence to a root of the function within a given interval, provided that the function is continuous and changes sign within that interval. It has a linear rate to convergence, which makes it slower with respect to other approaches, e.g. the secant method. We tried also this approach but the time taken for this specific problem was approximately the same and it experienced divergence from the expected root. The code follows the same structure has the one implemented in point (a), now taking into account the density variable $n_{\rm e}$. The code is reported below:

```
## 2b)

# Define the constants
```

```
_{4}|_{\rm k} = 1.38e-16 \# {\rm erg/K}
    aB = 2e-13 \# cm^3 / s
    Z = 0.015 \; \# \; \text{metallicity}
     Tc = 1e4 * Z**2 # stellar temperature in K
     psi = 0.929
     A = 5e-10 \# erg
    \texttt{epsilon\_CR} \, = \, 1 \hspace{-.05cm} \cdot \hspace{-.05cm} -15 \, \# \, \hspace{.05cm} \text{s} \hspace{.05cm} \hat{} -1
     # Define the equilibrium function
12
     def equilibrium_2(T, n_e):
             T4 = T / 1e4
14
             term1 = 0.54 * T4**0.37 * aB * n_e * n_e * k * T
             term2 = A * n_e * epsilon_CR
16
             term3 = 8.9e-26 * n_e * T4
17
              return term1 - term2 - term3
18
19
     # Bisection method
20
     21
             # Check if the function values at the interval endpoints have the same sign
22
              if f(a, n_e) * f(b, n_e) >= 0:
23
24
                      print("Bisection method fails.")
                      return None, None, None
25
26
              num\_steps = 0
27
             start_time = time.time()
28
29
             # Iterate until convergence or maximum iterations
              while (b - a) / 2 > tol  and num_steps < max_iter:
30
                     # Midpoint of the interval
31
                      c = (a + b) / 2
32
                       # Check if the midpoint is already the root
33
                      if f(c, n_e) = 0:
34
                               end_time = time.time()
35
                               return c, num_steps + 1, end_time - start_time
36
37
                     # Oheck if the root lies between a and c
38
39
                      if f(c, n_e) * f(a, n_e) < 0:
                                # Update the upper bound
40
                               b = c
41
42
                      else:
43
                             # Update the lower bound
                              a = c
44
45
                      num\_steps += 1
46
             end_time = time.time()
47
             return (a + b) / 2, num_steps, end_time - start_time
49
     # Densities to consider
50
     densities = [1e-4, 1, 1e4]
51
52
     with open('2b.txt', 'w') as f:
53
              for n_e in densities:
54
                     # Initial interval for bisection method
55
56
                      a_bisection = 1
                      b_bisection = 1e15
                      T\_eq\_bisection\;,\;\;num\_steps\_bisection\;,\;\;time\_taken\_bisection\;=\;bisection\_method\_ne(
58
             equilibrium_2, a_bisection, b_bisection, n_e)
60
                      # Write the result to the file
                      if T_eq_bisection is not None:
61
                               \label{eq:formula} \mbox{f.write(f"For n_e = \{n_e\} cm^-3, the equilibrium temperature is \{n_e\} cm^-3, the eq
62
             T_{eq}-bisection:.2 f K. n
                               f.write(f"The bisection method found the root in {num_steps_bisection} steps
63
              .\n")
                               f.write(f"The time taken was {time_taken_bisection:.6f} seconds.\n\n")
                      else:
65
                               n")
```

heating.py

The output of (b) is in

```
For n_e = 0.0001 cm^-3, the equilibrium temperature is 160561340829973.25 K.

The bisection method found the root in 54 steps.

The time taken was 0.000035 seconds.

For n_e = 1 cm^-3, the equilibrium temperature is 34243.13 K.

The bisection method found the root in 83 steps.

The time taken was 0.000058 seconds.

For n_e = 10000.0 cm^-3, the equilibrium temperature is 29.12 K.

The bisection method found the root in 83 steps.

The time taken was 0.000051 seconds.
```

2b.txt

These results represent the equilibrium temperature of the ionized gas in HII regions. This quantity is determined by the balance between heating and cooling processes, which involve various physical phenomena such as free-free emission (Λ_{FF}), heating by cosmic rays (Γ_{CR}), and MHD waves (Γ_{MHD}).

- For low-density gas ($n_e = 10^{-4} \text{cm}^{-3}$), the equilibrium temperature is extremely high (1.60 × $10^{14} K$). This is because, at low densities, the cooling processes (like free-free emission which depends on the square of the density) are inefficient, and the gas remains hot.
- For intermediate-density gas $(n_e = 1cm^{-3})$, the equilibrium temperature drops to 34243.13 K. At these densities, the cooling processes become more efficient, and the gas can cool down to lower temperatures.
- For high-density gas $(n_e = 10^4 cm^{-3})$, the equilibrium temperature is quite low (29.12 K). At high densities, the cooling processes are very efficient, and the gas can cool down to very low temperatures.