Thermodynamic Simulations of Rigid Spherical 'Balls' of Gas in a Fixed Container

including of Inelastic Collisions in a Container, Ideal Gas Relationships, Conservation Laws, the Ideal Gas Laws, the Van der Waals Law, and the Maxwell-Boltzmann Distribution

Martin A. He Sunday 4th December, 2022

Abstract – 2D-thermodynamic simulations were performed for a fixed number of rigid *spherical ball* approximations of atoms, within a circular container of fixed area over a set number of collisions. This initially examined the collisions of a single ball within a circular container, before being extended to encompass multiple balls, and producing various plots of properties of the system and the ball over time. In particular, with relation to the laws governing an ideal gas, conservation of various properties, Van der Waals forces and the Maxwell-Boltzmann distribution. Our simulation was found to be in agreement with all theoretical distributions.

I Introduction

It has long been theorised that the erratic movement of various particles in mediums can be explained through particle collisions, mostly notably in 1827 where Robert Brown observed the movement of pollen grains suspended in water.

From this motion, he deduced that this must be the product of bombardment with water molecules. coined Brownian motion in his honour. These intermolecular interactions are outside the scope of this model, and the collisions which are simulated are purely rigid-body elastic collisions, without any attractive forces.

II Theory, Method & Data Collection

In all, the project was split between three main Python files: *ball.py*, which controls the movements of individual Ball objects, *simulation.py*, the primary file controlling the set-up of the simulation, as well as *plots.py* for the testing and generation of plots. Simulation contains the *Simulation* class, along with the *Plot* and *Event* classes, whilst Ball has the *Ball* and *Event* classes.

In order to optimise the data processing efficiency, the Python libraries Numpy, Matplotlib, Scipy, as well as other modules commonly found on Anaconda[1] were utilised, including: Pandas, Seaborn, Itertools, Heapdict, as well as built-in Python packages. These were chosen to extend the functionality of Python, in particular with the extensive use of Pandas *Dataframes* as opposed to say, Python *Arrays,* likewise with Numpy *Arrays* where appropriate[2]. This resulted in a fast, efficient runtime composed of a mere *16 functions* in the entire Simulation class. A progress bar was incorporated to aid the user experience.

A. *Ball.py* *(import ball as bl)*

More information can be found at martinhe.com/thermo-project.

This module is composed of 5 primary sections, including the: Initialisation, Information, Movement, Attributes and *Container* sections. The Initialisation section controls the configuration of the ball, including its main properties to be parsed, whilst the Information section gives the representation and string formats. Movement governs the functions which control how the ball travels in the *Container*, by calculating the time to the next collision for the ball *(time_to_collision)*, which makes use of the following equation shown below:

$$(\mathbf{r}_1 + \mathbf{v}_1\delta t - \mathbf{r}_2 - \mathbf{v}_2\delta t)^2 = (R_1 \pm R_2)^2 \tag{1}$$

Equation 1: The dynamic equation for two colliding idealised balls in an elastic collision, where ***r***, ***v***, and *R* are the position, velocity, and radius of the balls or container respectively. Evidently, there are many solutions from this equation.

Rearranging (1) to form a quadratic expression, we arrive at:

$$\delta t = \frac{-\vec{v}\cdot\vec{r} \pm \sqrt{(\vec{v}\cdot\vec{r})^2 - (\vec{v}\cdot\vec{v})\left(r\cdot\vec{r} - R^2\right)}}{\vec{v}\cdot\vec{v}} \tag{2}$$

Equation 2: Rearranged form of the dynamic equation, from (1).

We have that (2) evidentally gives four possibilities: no real solution of δt, only negative solutions for δt, a repeated root of 0, or a positive and negative solution. Next, we must further consider the possibility of a collision with a container, as well as with another ball. Evidentally, we can discount the complex and negative solutions to δt since they indicate the next collision is in the past, which reduces to an absurdity. We thus, only need to consider the latter two scenarios: for the repeated root, there is only a logical outcome where $\delta t = -b/2a > 0$. As well as this, we also have the positive and negative solution. In any case, when colliding with another ball, this results in (3):

$$\mathbf{v}_1' = \mathbf{v}_1 - \frac{2m_2}{m_1+m_2}\frac{\langle \mathbf{v}_1 - \mathbf{v}_2, \mathbf{x}_1 - \mathbf{x}_2\rangle}{\|\mathbf{x}_1 - \mathbf{x}_2\|^2}(\mathbf{x}_1 - \mathbf{x}_2)$$
$$\mathbf{v}_2' = \mathbf{v}_2 - \frac{2m_1}{m_1+m_2}\frac{\langle \mathbf{v}_2 - \mathbf{v}_1, \mathbf{x}_2 - \mathbf{x}_1\rangle}{\|\mathbf{x}_2 - \mathbf{x}_1\|^2}(\mathbf{x}_2 - \mathbf{x}_1) \tag{3}$$

Equation 3: Solution to the dynamic equation[3] applied to masses (m_1, m_2), positions (x_1, x_2) and velocities (v_1, v_2, v_1', v_2') of the balls, between two balls. The solutions between a ball and container can consequently be dervied.

The *Attributes* section simply sets and returns all the ball properties, whereas, the *Container* class simply inherits from the Ball class, whilst assigning a radius, mass, and count to it.

B. *Simulation.py* *(import simulation as sim)*

The Simulation module, the core of the project, contains the *Simulation*, *Event* and *Plot* classes. The *Event* object merely acts as a tuple record for collision events between any pair of balls, whilst the *Plot* class provides methods for plots and histograms. Seaborn was used alongside Matplotlib to add more flexibility to the plots which could be created, including 3D plots. The Simulation class is divided into *13 functions*, which each have many parameters to improve organisation and prevent clutter, with extensive documentation and judicious commenting. Prior to the Simulation class, the *gen_random_uniform* function is declared due to its universal application across classes. The first three functions are self-explanatory, and given by their docstrings, so we shall proceed to the *glossary* function which returns a dictionary of all quantities for plotting. This made use of a *heapdict*[4] due to the fast processing speed, as well as its additional methods of *popitem* to obtain data. The *property* function is used to define and return various properties, whilst *set_vel_ball* defines all the ball velocities. Where appropriate, the *enumerate* function was substituted for *np.ndenumerate* for its faster runtime, though in some cases this was not possible due to its tuple output resulting in *TypeErrors*.

The *next_collision* function is arguably the most important of the entire project, defining the logic to set up and perform it, by finding this *dt*, before moving the system to this point, and performing the collision. Its logic is convoluted, by essentially, it selects from the heapdict, before running through all possible collision pairs, ascertaining the minimum *dt* for collisions with other balls as well as a container, whilst checking for a situation where multiple collisions occur *à la fois*. Finally, the UTC time moves forward, with new plotting, and recording completed.

$$\mathbf{r}' = \mathbf{r} + \mathbf{v} * \mathrm{dt} \quad (2)$$

Equation 4: The simple logic for moving the colliding balls to a new position. [5]

The *randomiser* function provides all randomisation options for position and velocity. The *draw* function first initialises the patches, before proceeding to redraw all ball patches for every active tuple pair helpfully provided by the *itertools* library. The *record_property* function unsurpisingly records properties, including a complete dataset written into a *Pandas DataFrame*. Similarly *record_data_states* enables for a snapshot in time of data to be captured. Finally, the *run* method runs the above functions to perform the collisions and animation, recording data. A *progress* bar allows the user to keep track of the simulation. Finally, the *Event* and *Plot* classes enable for events to be recorded and plots to be generated respectively.

III Results & Analysis

N.B: the plots and full datasets can be obtained by visiting the following website: www.martinhe.com/thermo.

The obtained plots are attached in the *Appendix* section. Overall, we see that the plots are in accordance with the expected graphs for each task, and a further breakdown can be found below.

In **Figure 1**, we obtain an animation of the simulation. This demonstrates fluid Brownian motion for varying numbers of N. In **Figure 2**, we obtain the plot for the absolute distance distribution from the Origin, as well as the relative distances between every ball pairs and their corresponding histograms. Both are as expected (see Appendix and Figure 2 caption).

In **Figure 3**, have insight into conservation of quantities in the simulation. We can see that the pressure of the system resolves quickly into a steady state value after initial setup. For the momentum of collisions with the container, we see that momentum is similarly conserved, as is systemic kinetic energy. In **Figure 4**, we examine the various ideal gas relations, such as between pressure, temperature, volume, and N. We find linear fits with values corroborating theoretical explanations.

In **Figure 5**, we examine the Law itself through various isotherms of pressure against volume, for varying T and N. This results in decay curves as are expected according to the Theory. In **Figure 6**, we take a closer look at van der Waals' Law, discovering a upwards curve, for the Power Law over an effective area. Values for the results can be found in the Appendix Finally, in **Figure 7**, we look at the Maxwell-Boltzmann distribution which we find to be perfectly followed by the system for varying parameters. To demonstrate this, a 3D plot has been generated varying the number of collisions, and balls. Additional extension figures are shown in **Figure 8** onwards.

IV Conclusion

From our extensive data collection, we have demonstrated our simulation to be in agreement with the theoretical values. As a result, the following simulation can be accurately examine the provided tests and parameters can be freely altered if one wishes.

Yet, many improvements could be made to this experiment if time permitting. 3D plots would be generated for all data specified above, to see how varying the system parameters affects the strength of these distributions, whilst in order to improve computational efficiency[6], almost everything would be converted to *Numpy Arrays* of defined size. This would vastly improve upon the current method appending to a *List*. Likewise, more repeats would be run of the simulation, and an average obtained to reduce random uncertainties. Additional plots could be generated for a situation describing the Brownian motion of pollen grains suspended in water grains: the origin of the Theory.

V Appendicies & Figures

A. *Appendices*

1 Anaconda Environment

The Anaconda environment contains all the relevant Python packages as listed above in this report. For more information on

how to go about installing these packages in other environments, please see the README file.

2 Numpy vs. Python Efficiency

It is oft said, and assumed that any corresponding Numpy feature will be faster than its Python counterpart, however over the duration of this project, I was surprised to find this to not always be the case, the most notable example being appending to Python lists as opposed to numpy arrays, where there can be up to a staggering 2000× difference in processing time.

3 Solutions to the Dynamic Equation in 2D

These solutions to the Newtonian collision between two rigid disc balls can be dervied, as found on the following website.

4 Maxwell-Boltzmann Distribution[7]

The Maxwell-Boltzmann distribution in 2 dimensions as given by its radial component in cylindrical polar co-ordinates takes the following form, where the initial factor is v^{d-1} where $d = 2$.

$$p(v) \propto v \exp\left(-\frac{\frac{1}{2}mv^2}{k_B T}\right)$$

Consequently, we are able to determine the form our graph takes.

B. *Figures*

Below, please find attached an assortment of *Figures* demonstrating the capabilities of the Simulation program.

#	*Description*	*Notes*
Figure 1	Simulation animation.	GIF is found in /packages folder.
Figure 2	Distribution of relative ball distances between *Balls*.	Gaussian-like relationship.
Figure 3	Distribution of absolute ball distances from the O.	Linear relationship.
Figure 4	Various ideal gas relations.	Contains 3 plots.
Figure 5	Ideal Gas Law, varying T and N.	Run for 4 scenarios.
Figure 6	Van der Waals' Law, obtaining a and b.	
Figure 7	Maxwell-Boltzmann distribution for different $N_{collisions}$ and N_{balls}.	
Figure 8	Various conservation laws.	Contains 3 plots.
Figure 9	3D Maxwell-Boltzmann distribution for different numbers of Balls.	Found at: martinhe.com/thermo

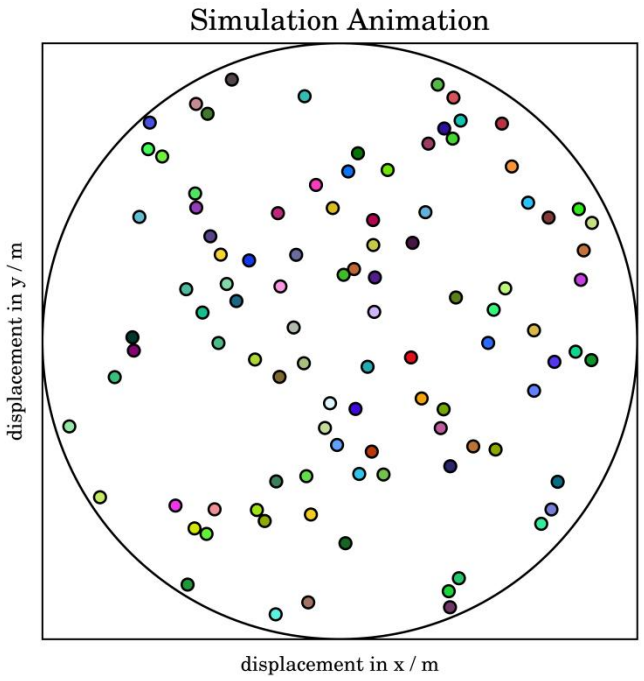


Figure 1: A still from the simulation animation *(plot_simulation_animation)* with N = 50 balls, collisions = 500, $r_{container}$ = 10m, which is not denoted for aesthetic reasons. Balls are randomly coloured, and well-distributed.

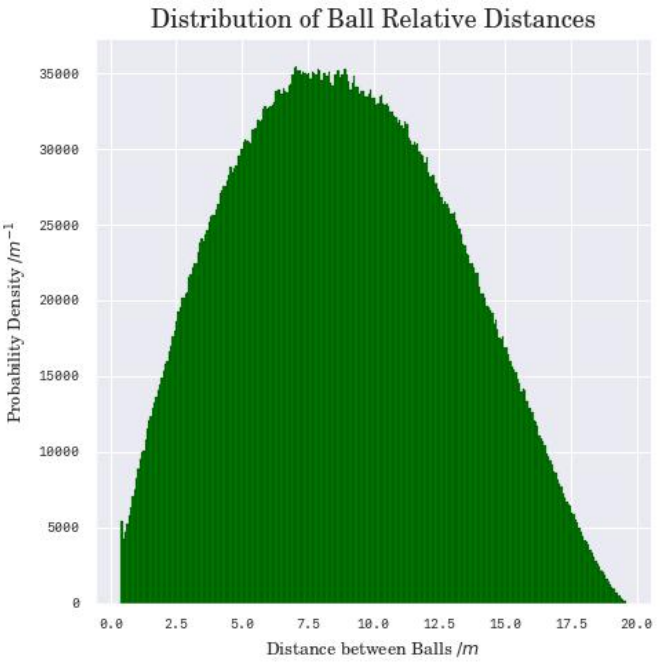


Figure 2: The plotted histogram distribution of relative ball distances between *Ball* objects, as well as the probability density in m^{-1}, with the parameters: N = 50 balls, $r_{container}$ = 10m, r_{balls} = 0.2m, $m_{balls} = 5\times10^{-26}$ kg, $range_{random\ speeds}$ = 500m.s^{-1}, collisions = 5000. This graph is as expected since this problem can be reconsidered in two manners: firstly, going out from 0 to r, we find that the number of possible points increases by r^2; secondly, we observe that after collision events with the Container, it traverses a path of similar *r* to before, resulting in a spike at *r=10*. Meanwhile, the probability decreases by a factor of r^{-1}, resulting in an overall linear trend. This shows a Gaussian-like trend.

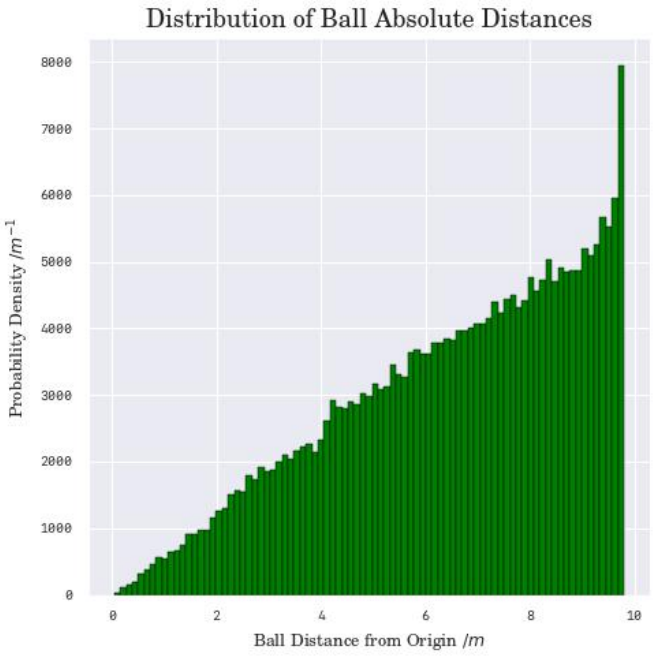
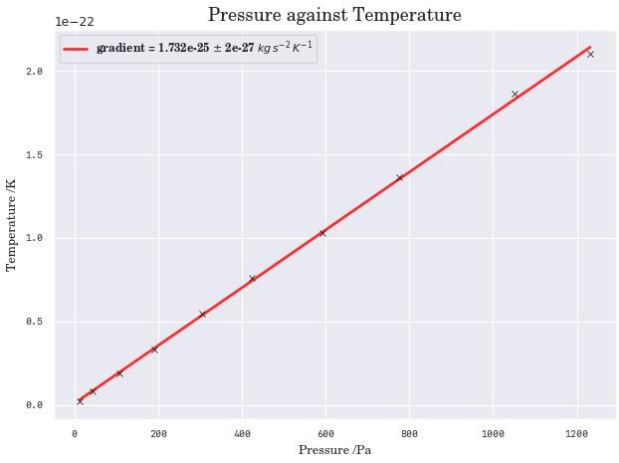
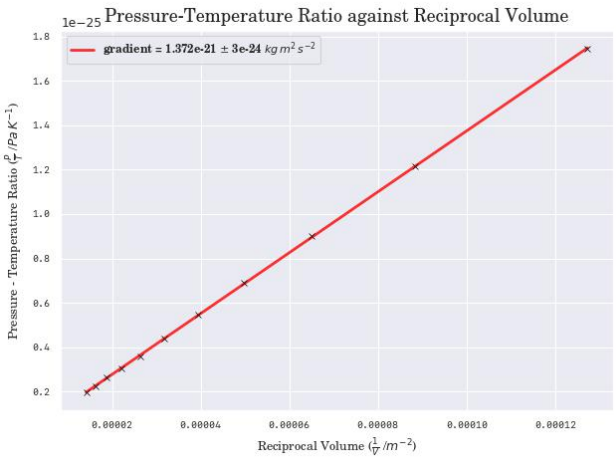


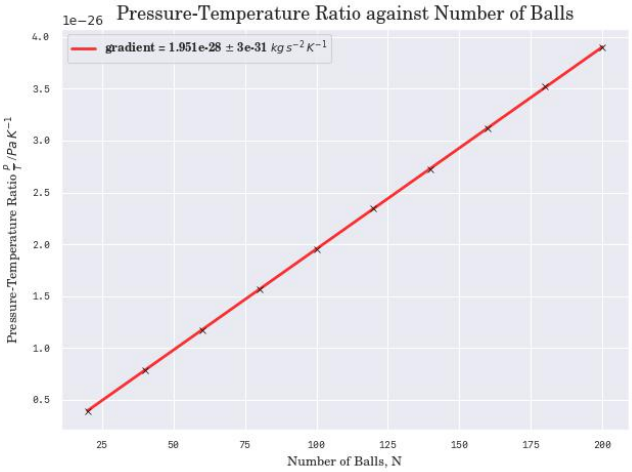
Figure 3: The plotted histogram distribution of absolute distances between the Origin and *Ball* objects, which can be seen by the significantly smaller y-axes ticks, as well as the x-axis range only covering 10m. This graph is also as expected, which indicates a Gaussian-like curve with a left skew, and peak around *8m* separation, between *0* and *20m*. This is as expected as for a large number of colliding balls, by the CLT, it should tend to Gaussian. The skewness can similarly be explained by the same collision logic given before, where it traverses similar distances right after it collides.



(4.1)



(4.2)



(4.3)

Figure 4.1: The variation of pressure against temperature.

Figure 4.2: The variation of the pressure-temperature ratio over the reciprocal of volume. This demonstrates a strong positive linear correlation.

Figure 4.3: The variation of pressure against temperature over the number of simulated *Ball* objects, N. This demonstrates a strong linear trend.

Figure 4: Various ideal gas relations, varying parameters as directed. Firstly, in *Figure 4.1*, we examine the graph of pressure *(y)* against temperature *(x)*, in terms of their linear relationship, as predicted by the Ideal Gas Law. As can be seen, the gradient is $(1.73 \pm 0.02) \times 10^{-25}$ kg.s^{-2}.K^{-1}, which is an excellent linear fit with such low uncertainty. We then proceeded in *Figure 4.2* to determine how varying the radius of the ball affects the pressure-temperature dependence, in the same linear fashion. Finally, in *Figure 4.3*, we can see that increasing the number of balls similarly increases the pressure-temperature ratio of each of the resulting graphs.

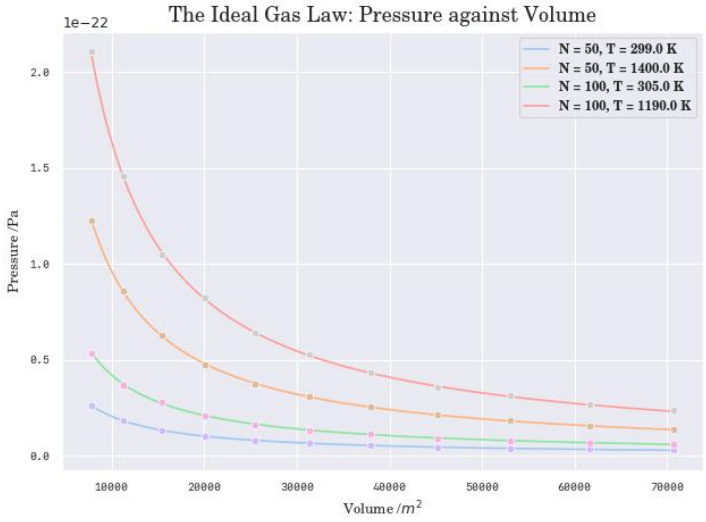


Figure 5: The Ideal Gas Law, varying T and N. This demonstrates that for greater number of balls N, there is a larger dependence of pressure on volume, as compared to lower values. Similarly, for higher temperatures T, there is a higher dependence as expected; however, evidently, N has a larger impact on this dependence than T as shown above.

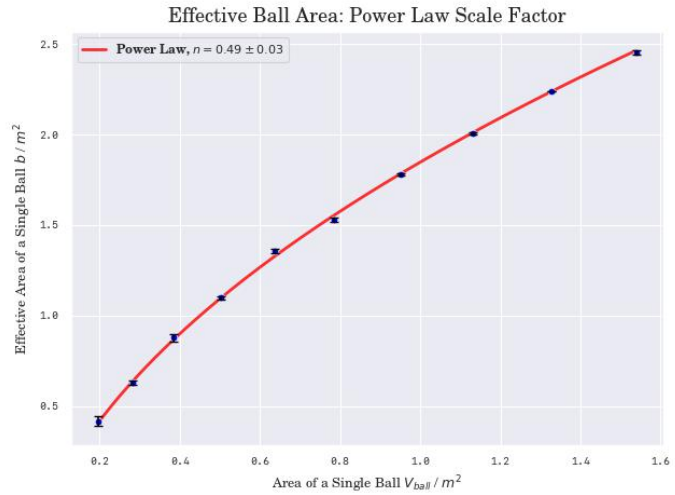


Figure 6: The Van der Waals graph, a power law. From the graph, we obtain the following power law of form $y = A.x^b + B$ with the parameters as given: $b = 0.49 \pm 0.03$, A = 2.61 $\pm$ 0.14, B = -0.77 $\pm$ 0.14. This is the case if we allow all parameters to be any values. By visual inspection, we find that the van der Waals fit to be very strong. However, if we instead consider, for fixed values of B, we instead result in the different values of the following: B = 10.31 $\pm$ 0.12. This graph instead passes through the Origin.

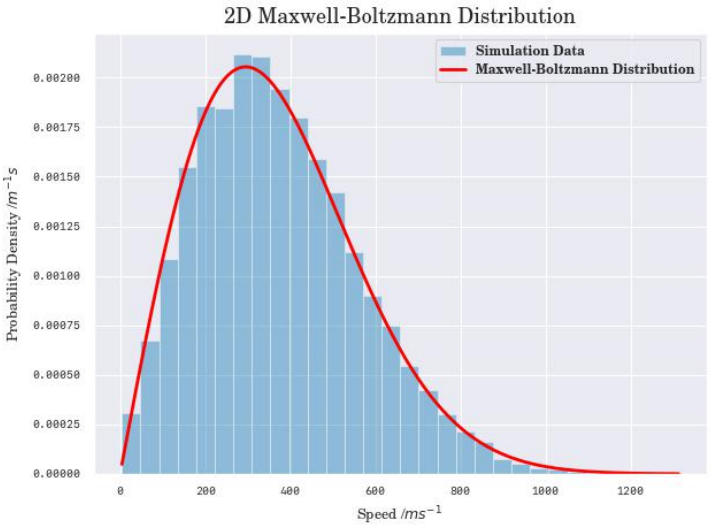
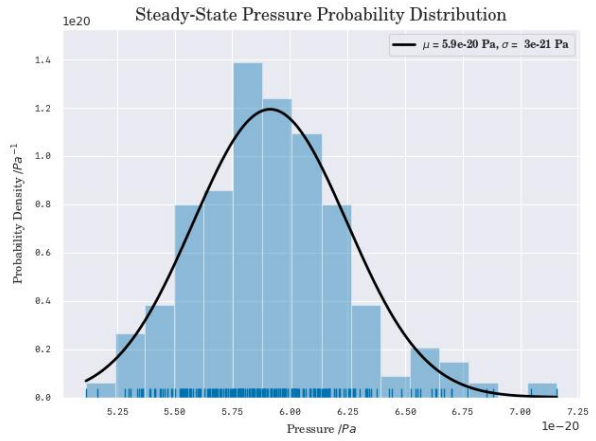
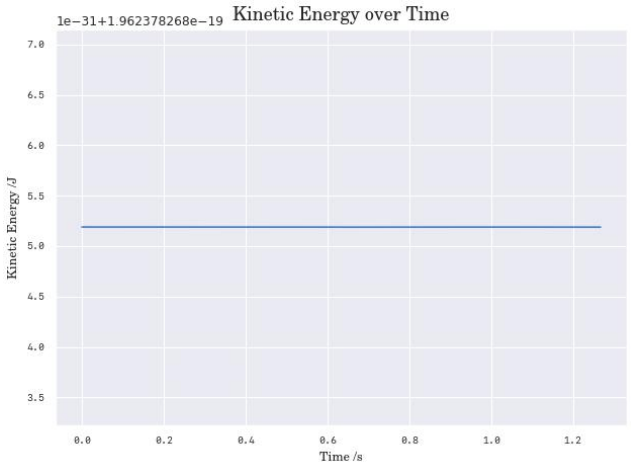


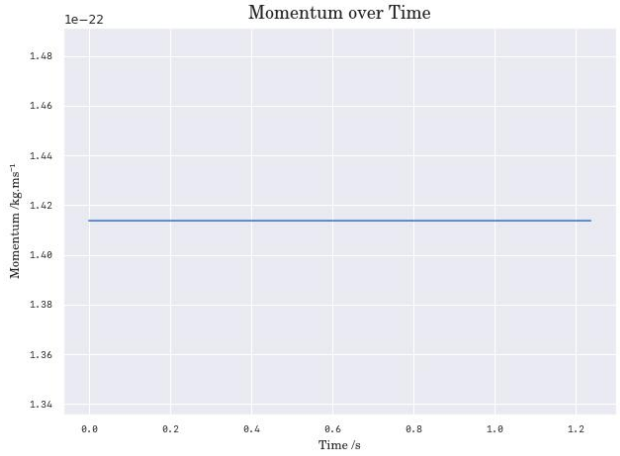
Figure 7: The Maxwell-Boltzmann distribution, over the *Simulation* with the distribution of speeds. This was generated with the same parameters, with 10000 collisions, and it shows that our fitting strongly corresponds to the actual theoretical distribution. We can infer the goodness of fit for our Maxwell Boltzmann by the temperature we obtain as compared to the temperature from the simulation, which are also very similar. Therefore the obtain the following fit, which has a scale factor of approximately 4300.



(8.1)



(8.2)



(8.3)

Figure 8.1: Steady-state systemic pressure distribution, for 50000 collisions. This works by taking an average every 50 collisions to obtain the pressure.

Figure 8.2: The time evolution of systemic kinetic energy.

Figure 8.3: The time evolution of momentum.

Figure 8: Various conservation laws. We examine the variation of momentum over time, the systemic kinetic energy over time, as well as the pressure over the container over time. These all exhibited behaviour which we would expect from such graphs.

Due to an acute lack of space, the 3D Maxwell-Boltzmann plots, along with the completed set of data, can be found at this site:

martinhe.com/thermo

VI References

[1] Anaconda (no date) *Anaconda package lists¶*, *Anaconda package lists - Anaconda documentation*. Available at: https://docs.anaconda.com/anaconda/packages/pkg-docs/ (Accessed: December 9, 2022).

[2] Ayar, M. (2022) *Python lists are sometimes much faster than numpy. here's proof.*, *Medium*. Towards Data Science. Available at: https://towardsdatascience.com/python-lists-are-sometimes-much-faster-than-numpy-heres-a-proof-4b3dad4653ad (Accessed: December 9, 2022).

[3] Russel, W.B., Saville, D.A. and Schowalter, W.R. (1991), p262, in *Colloidal dispersions*. Cambridge: Cambridge University Press.

[4] Heapdict (no date) *Heapdict, PyPI*. Available at: https://pypi.org/project/HeapDict/ (Accessed: December 9, 2022).

[5] Kingham, R. (no date) *Thermodynamics Snookered*. p3 Available at: https://bb.imperial.ac.uk/bbcswebdav/pid-2526807-dt-content-rid-12847107_1/courses/14925.202210/Computing/Scripts/Projects/.build/html/Snooker.html.

[6] *Python lists vs Numpy Arrays* (2022) *GeeksforGeeks*. Available at: https://www.geeksforgeeks.org/python-lists-vs-numpy-arrays/ (Accessed: December 9, 2022).

[7] Kingham, R. (no date) *Thermodynamics Snookered*. p11 Available at: https://bb.imperial.ac.uk/bbcswebdav/pid-2526807-dt-content-rid-12847107_1/courses/14925.202210/Computing/Scripts/Projects/.build/html/Snooker.html.

VII Testing Logs

A compilation of testing woes, debugging, and many tears which were shed, as part of the painstakingly long debugging process.

Initial Aims

1. Change uses of lists [] to np.array where possible to improve efficiency. (I'm not very good at using np.append and np.zeros, so this will be quite the experience for me). If not possible to determine size, not change.

2. Remove uses of enumerate function, in favour of np.ndenumerate.

3. Check docstring formatting is consistent, as is the list of functions and properties is consistent (e.g. speed, KE, temperature, pressure, etc.).

4. Create a final minified version of the Python code when testing, to run more repeat in as efficient a manner as possible.

5. Use *np.linspace* instead of nested for-loops to increase efficiency.

6. Make sure singular/plural variable names is consistent between Python files. Similarly, only use the present form, and not present continuous forms of verbs. Each word should be separated by underscores, and variable names should be kept as short as possible.

7. Make sure spacing is appropriate between lines, and commenting is aligned. This includes indenting sub-headings, such as *Parameters* or *Variables*, and separating where possible with | . The function name and description should be kept in the initialisation line of the docstring.

8. Commenting should be judicious, in order to avoid cluttering the "flow" of the code and readibility. However, inline commenting is acceptable when spaced correctly. Commenting in every line is acceptable only for the initialisation and run functions.

9. Use as sparingly few functions as possible, for example have an entire function called *draw* or *record_property* with parameters instead to handle various cases of the function. This aids readability.

10. Implement an *Event* Class for each collision, to encode the vital information about ball pairs. Similarly, implement a *Plot* Class to contain all the methods for plotting graphs using Seaborn and Matplotlib.

Final Tests

Task 1 | Determine the time to next collision.

This task for the most part went off without a hitch, although whilst writing the code, I discovered a quicker method from Wikipedia.

Task 2 | Define *Ball* class with attributes for mass, radius, position.

This task went off smoothly, without a hitch. The various variable names, as well as the corresponding functions to designate these variables were created.

Task 3 | Define *Simulation* class with *next_collision* function.

On the other hand, this task proved to be more difficult than the prior tasks. This was due to the fact, that the *next_collision* function, which I initially split into three separate functions, was encountering constant errors, which I fixed by switching over from a *list* to a *heapdict*, as well as using the function *popitem* which resolved the problems.

Task 4 | Create *Simulation* object with radius 10, ball 1, mass 1, testing the *time_to_collision*, *next_collision* and *collide* methods.

This task was thankfully easier on account of the debugging in the previous section. However, this didn't go on without a hitch since I had to figure what was the best way to fix errors which were cropping up in the *collide* function. The mass and radius of the ball were later altered in order to accommodate the size of the *Container*.

Task 5 | Define the *run* method by running an animation.

In order to aid the presentation of the *Simulation*, a progress bar function was created to track the progress of the running of the simulation. This was accomplished by the use of an *fstring*, along with metrics such as the current collision number, the time elapsed and the expected time before the collisions were completed.

Task 6 | Check conservation of energy, and pressure on container.

Conservation of energy was initially not met in the simulation. In order to ensure this was met, floating point errors were accounted for, by adding in exception statements, among other changes. Similarly, pressure was fixed on account of these changes.

Task 7 | Modify *Simulation* class to handle multiple *Ball* objects, and to prevent overlapping *Balls*.

Luckily, I had taken into account the possibility of adding multiple balls, consequently, this was an easy feat to accomplish.

For the remaining tasks, please see the Comments in the Appendix.

Task 8 | Modify the *Simulation* class either in intialisation or via. a new method to provide a systematic way of initialising a system with multiple *Balls*, by avoiding overlapping *Balls*.

Task 9 | Plot histograms to absolute and relative distances.

Task 10 | Investigate KE and temperature relation. Also consider the velocity and pressure.

Task 11 | Check all conservation laws, the pressure on the container, and its dependence on T, as well as how V and N affect this.

Task 12 | Compare results to the Ideal Gas Law, and vary ball size.

Task 13 | Compare the velocity distribution to the Maxwell-Boltzmann distribution, measuring the variance, and plotting histograms.

Task 14 | Use Van der Waals' law to identify a and b.