CHEM2000 Lab Manual

Paolo Raiteri

March 9, 2026

Contents

1. Introduction

With advances in technology, researchers can collect vast amounts of data quickly, but the true value lies in making sense of this information. Effective data analysis allows scientists to uncover patterns, test hypotheses, and draw meaningful conclusions, driving scientific discovery and innovation. It enables the transformation of raw data into actionable tasks, guiding decision-making and advancing knowledge across disciplines. In this data-rich age, the ability to analyze and interpret data is essential for scientific progress and solving complex global challenges.

With this in mind, we have decided to move away from bench-top lab experiments and designed some virtual laboratory activities focused on data analysis and data presentation. These laboratories mimic/use real experimental data but removing the time to perform the real experiments, allows us to put a stronger focus on the post processing of relatively large data sets and data presentation.

Although the numerical labs in this unit can be solved using simple `Excel` spreadsheets or even by hand with some algebra and a pocket calculator, we would encourage you to use `Python` and Jupyter notebooks for this work, which can be used to develop important skills in data analysis, data processing, and computing. Some of you have already learned `Python` in a first year unit, and this laboratories will provide you with the opportunity to hone those skills and make Jupyter notebooks and `Python` day-to-day tools for all your calculations and data analysis. However, if you are not comfortable with programming and using `Python` **you can use `Excel` for these labs. The assessment is centred on the data analysis and data presentation and not on the code you may have written. No code needs to be included in the report.**

1.1 Why learning Python?

`Excel` is a widely used spreadsheet program that offers various functionalities for data management and analysis. However, it is important to note that `Excel` is primarily designed as a general-purpose tool and not specifically tailored for statistical analysis. While `Excel` does provide some statistical functions and tools, its capabilities in this area are limited and it lacks the robust capabilities and flexibility required for advanced analysis and data manipulation that dedicated statistical software or programming languages like `Python` have.

This is why we strongly encourage you to learn `Python` (or R), a powerful and versatile programming language widely used in scientific computing, data analysis, and machine learning. `Python` offers extensive libraries like `NumPy`, `lmfit`, and `scipy` that provide sophisticated statistical tools and data handling capabilities. We believe knowing `Python` would be crucial for future employment in chemistry, physics, and engineering due to its powerful applications in these fields. In fact, `Python` is used for molecular modeling, chemical informatics, simulation of complex systems, data processing, visualization, automation, control systems, and computational engineering. Its simplicity and readability make it an excellent choice for both beginners and experienced programmers. As industries increasingly rely on data and automation, acquiring proficiency in `Python` can significantly enhance your job prospects and career growth. By investing time in learning `Python`, you will gain valuable skills that are highly sought after in both academia and industry, opening up new possibilities for data analysis and visualization in your future work.

1.1.1 Using Python for this unit and beyond

There are many different ways to use `Python` for data analysis, which one is best for you largely depends on your computer skills, on the operating system of your computer, and how comfortable you are with manually installing software on it.

Arguably, `Jupyter Notebooks` provide the most user-friendly interface for beginners. They are an interactive computing environments that allow you to create and share documents containing live code, equations, visualizations, and explanatory text. They are particularly useful for data analysis, scientific computing, and programming education. In a `Jupyter Notebook`, you can write and execute code in various programming languages (such as `Python`, `R`, or `Julia`) alongside rich text explanations, making it easy to document your work step-by-step. This format enables you to combine computational output with narrative descriptions, creating a comprehensive and reproducible record of your analysis or experiment. `Jupyter Notebooks` are widely used in academic and professional settings for tasks like data cleaning, statistical modeling, machine learning, and creating reports with embedded visualizations. `Jupyter Notebooks` are available through various interfaces, for example

- Google Colab [here]
- Visual Studio Code [here]
- Command line installation [Linux and Mac OSX]

Using Google Colab (short for Colaboratory) is probably the easier way for you to get going, the only thing you need is a Google account. Google Colab is a free, cloud-based platform that allows you to write and execute `Python` code through your browser. It provides a Jupyter Notebook environment with free access to GPU resources, making it ideal for machine learning and data science projects. Colab integrates seamlessly with Google Drive for easy file storage and sharing. It comes pre-installed with many popular data science libraries, eliminating the need for local setup and configuration. This platform is especially useful for students and researchers who need computational power without access to high-performance hardware. We think this would be the ideal tool to use for this unit, if you choose to use `Python`.

Moreover, Google Colab provides AI-powered coding assistance that can be a valuable learning tool when used responsibly. These features include

- Code completion: Colab offers AI-powered code suggestions as you type. This can help with syntax and common coding patterns.
- Code explanation: You can select code and ask the AI to explain what it does. This is useful for understanding unfamiliar code or concepts.
- Natural language to code: You can describe what you want to do in plain language, and the AI will attempt to generate relevant code.
- Debugging assistance: The AI can help identify errors and suggest fixes.

While these tools can enhance your learning experience, it is crucial to use them ethically. Always strive to write code yourself first, using AI assistance only when you are stuck or need clarification. The goal is to understand the code you are working with, not to have AI complete assignments for you. Use explanations as a starting point for deeper learning, and always cite any AI-generated code in your work. Remember, developing your own coding skills and problem-solving abilities is the primary objective of this course.

Also remember that AI-generated code may not function as you expect it, or meet the specific requirements of your assignment. It is essential that you carefully review, test, and debug any AI-suggested code to ensure it works correctly and achieves the intended results. Ultimately, you are responsible for what you submit in your report.

Visual Studio code should also be pretty straightforward to install, even on windows. However, unlike Google Colab, VS Code runs locally on your computer and no internet connection is required. VS Code provides a nice work environment, although the installation of most `Python` packages needs to be done manually. If you want to use VS Code, we can provide some assistance to set up your computer.

1.2 The numerical laboratories

The numerical laboratories will run according the the schedule shown in Section 1.3, and most activities will be spread over two weeks. In the first of those two weeks, we will cover the background of the experiment, objectives and procedure, while the second week will be a free drop-in session for you to finish the report and get help from the lab tutors. The drop-in sessions can be very beneficial for students who you want to improve their programming skills

Depending on the available IT resources, the students will have to run virtual experiments using `Jupyter Notebooks` hosted on a virtual machine, or will access pre-generated data through a `github` repository. The data for the simulated experiments will then be available in CSV files, which can be readily imported into `Excel` or `Python` for post-processing. The assessment is based on a portfolio of reports; check the unit outline for more details.

We start with a refresher of key statistics concepts and a numerical laboratory to practice this concepts. We will then have virtual laboratories covering fundamental topics in physical-chemistry

- Thermodynamics
- Kinetics
- Interfacial chemistry
- Chemical equilibrium

All the labs are based on physical chemistry concepts that you have already seen in the lectures (*e.g.* calorimetry, equilibrium, kinetics). These numerical labs have been designed to mimic real experiments, to a certain extent, which means that the measurements dote include random noise, which is designed to mimic the experimental uncertainty of the instruments and user errors.

1.3 Calendar

Week 1 –	**Introduction and Basic statistics numerical lab**
Week 2 –	**Basic statistics numerical lab**
Week 3 –	**Bomb calorimetry numerical laboratory** **First lab report due**
Week 4 –	**Bomb calorimetry numerical laboratory**
Week 5 –	**Lecture + workshop (Chemical potential)** **Second lab report due**
Week 6 –	**No lab** **Mid-semester test**
Week 7 –	**Public holiday**
Week 8 –	**Tuition free week**
Week 9 –	**Tuition free week**
Week 10 –	**Crystal violet numerical laboratory**
Week 11 –	**Crystal violet numerical laboratory**
Week 12 –	**Surface adsorption numerical laboratory** **Third lab report due**
Week 13 –	**Surface adsorption numerical laboratory**
Week 14 –	**Study week** **Fourth lab report due**

Public holidays

- 2nd March – Labour Day
- 3rd April – Good Friday
- 6th April – Easter Monday
- 7th April – Easter Tuesday
- 27th April – Anzac Day

2. Numerical laboratories

The hands-on activities for this course take the form of numerical laboratories. In these sessions, you will work with computer-generated datasets that mimic real experimental conditions, complete with inherent noise and the need for unit conversions. This approach allows you to develop crucial skills in data interpretation and report writing without the need for physical laboratory equipment, and the associated time for data acquisition. By manipulating and analyzing these realistic datasets, you'll gain experience in identifying trends, handling experimental uncertainties, and drawing meaningful conclusions from complex information. Throughout the course, you'll be asked to produce a report detailing the data analysis and your findings, mirroring the professional standards expected in scientific communication.

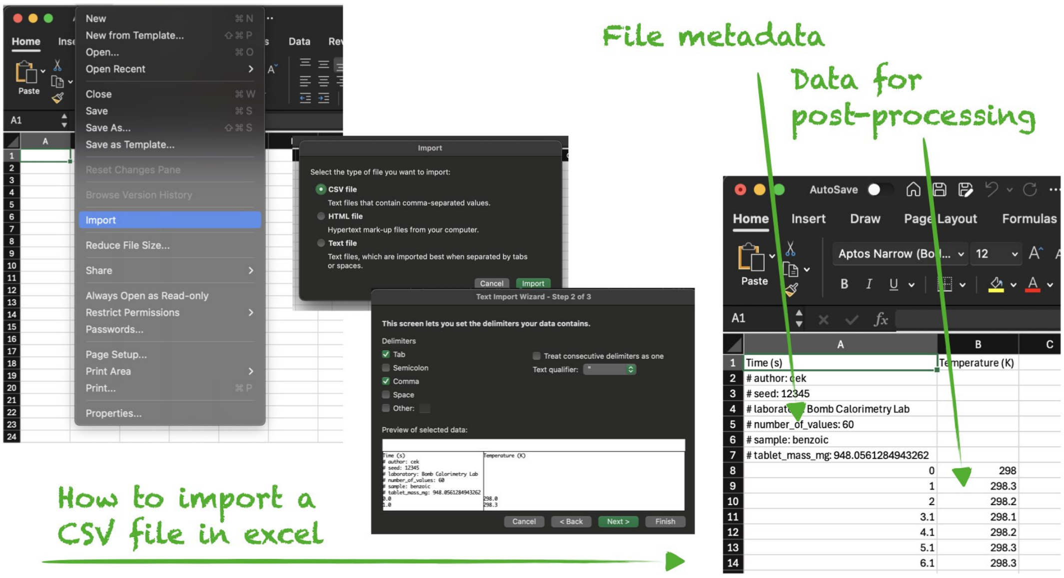
Depending on the computational resources available, you will either generate the data yourself through a suitable web interface, or you will be given a suite of data set, that you will then have to analyse. These data will be organised in folders, one per lab, and each file would correspond to an "experiment".

2.1 Structure of the data files

The files containing the data to be analysed will all be in CSV (Comma Separated Values) format, and they can be readily imported by Excel or read by Python. CVS files usually also have column names specified in the first line. The files produced in these labs also contain some metadata, with important details about the "expriment", *e.g.* the mass of the sample, the temperature of the experiment, etc. These metadata, are proceeded by the # symbol, which is usually interpreted as a commented by most codes.

```
Time (s),Temperature (K)
0.0,298.0
1.0,298.3
...
# author: cek
# seed: 12345
# laboratory: Bomb Calorimetry
# columns: ('Time (s)', 'Temperature (K)')
# number_of_values: 60
# sample: benzoic
# tablet_mass_mg: 948.0561284943262
```

If you are using excel, you can use the import function to load the data inside your spreadsheet (Figure 2.1) and the metadata, will appear in the cells between the column names and the data. On the other hand, if you are using python, you can import the data either by using `NumPy`. Here below there's a sample code that uses the `genfromtxt` method in `NumPy` to read the file. The code defines two functions `read_metadata` and `read_csv_file`, which you can use to extract the metadata from the file and store it into a dictionary and to read the values from a CSV file and store them into a `NumPy` array.



Select the type of file you want to import:

● CSV file
Text files that contain comma-separated values.
○ HTML file
Hypertext mark-up files from your computer.
○ Text file
Text files, which are imported best when separated by tabs or spaces.

This screen lets you set the delimiters your data contains.

Delimiters
☑ Tab
☐ Semicolon
☑ Comma
☐ Space
☐ Other:

☐ Treat consecutive delimiters as one
Text qualifier: "

Preview of selected data:

```
Time (s)                           Temperature (K)
# author: cek
# seed: 12345
# laboratory: Bomb Calorimetry Lab
# number_of_values: 60
# sample: benzoic
# tablet_mass_mg: 948.0561284943262
0.0                                298.0
1.0                                298.3
```

	A	B	C
1	Time (s)	Temperature (K)	
2	# author: cek		
3	# seed: 12345		
4	# laborat[illegible] Bomb Calorimetry Lab		
5	# number_of_values: 60		
6	# sample: benzoic		
7	# tablet_mass_mg: 948.0561284943262		
8	0	298	
9	1	298.3	
10	2	298.2	
11	3.1	298.1	
12	4.1	298.2	
13	5.1	298.3	
14	6.1	298.3	

Figure 2.1: Example of how you can import .csv files in `Excel`.

We would suggest you to practice your coding skills and write a python function to read the metadata from the `.csv` file. However, you can use in your Jupyter notebook the one from the example below.

Python Code Example

```python
def read_data_from_file(filename):

    # Read file and separate comments from data
    comments = []
    data_lines = []

    with open(filename, "r") as f:
        for line in f:
            if line.startswith("#"):
                comments.append(line.replace("#","").strip())  # Store comment lines
            else:
                data_lines.append(line.strip())  # Store data lines

    # Extract header and data
    data = "\n".join(data_lines)  # Join remaining lines as CSV data

    # Convert CSV data to NumPy array
    from io import StringIO
    data_array = np.genfromtxt(StringIO(data), delimiter=",",skip_header=1)

    # Extract the metadata from the comment lines
    metadata = {}
    for l in comments:
        key, value = l.split('=')
        metadata[key] = value.strip()

    return data_array, metadata

data_array, metadata = read_data_from_file('lin_data.csv')
```

```
# Output results
print("Metadata:")
for k,v in metadata.items():
    print(f"{k} = {v}")
print("\nExtracted Data:")
print("X values:\n", data_array[:,0])
print("Y values:\n", data_array[:,0])
```

You could also use `Pandas` to read a CVS file. Alternatively, `Pandas` store the data as a `DataFrame`, which has to be be used in a different way.

Python Code Example

```
import numpy as np
from pandas import read_csv
def read_data_from_file(filename):
    """
    Read the metadata from a CSV file used in the CHEM2000 labs
    Parameters: str
        name of the file to read
    Return: dict
        dictionary with the metadata
    """
    comments = []
    with open(filename, 'r') as f:
        for line in f:
            if line.strip().startswith('#'):
                comments.append(line.replace('#','').strip())
    metadata = {}
    for l in comments:
        key, value = l.split('=')
        metadata[key] = value.strip()

    data_array = read_csv(filename,comment="#").to_numpy()
    return data_array, metadata

# Name of the file to read
data_array, metadata = read_metadata('lin_data.csv')

# Output results
print("Metadata:")
for k,v in metadata.items():
    print(f"{k} = {v}")
print("\nExtracted Data:")
print("X values:\n", data_array[:,0])
print("Y values:\n", data_array[:,1])
```

Both codes would produce the same output

Output

```
author = cek
seed = 12345
laboratory = Bomb Calorimetry
columns = ('Time (s)', 'Temperature (K)')
number_of_values = 60
sample = benzoic
tablet_mass_mg = 948.0561284943262
```

```
----------------------------------------
Values in the first column:
[ 0.   1.   2.   3.1  4.1  5.1  6.1  7.1  8.1  9.2 10.2 11.2 12.2 13.2
 14.2 15.3 16.3 17.3 18.3 19.3 20.3 21.4 22.4 23.4 24.4 25.4 26.4 27.5
 28.5 29.5 30.5 31.5 32.5 33.6 34.6 35.6 36.6 37.6 38.6 39.7 40.7 41.7
 42.7 43.7 44.7 45.8 46.8 47.8 48.8 49.8 50.8 51.9 52.9 53.9 54.9 55.9
 56.9 58.  59.  60. ]
Values in the second column:
[298.  298.3 298.2 298.1 298.2 298.3 298.3 298.3 298.1 298.3 298.4 298.5
 298.5 298.2 298.4 298.7 298.5 298.5 298.6 298.9 298.5 299.3 299.9 300.3
 300.5 300.7 301.  300.9 300.9 300.8 300.9 301.1 301.2 301.1 301.2 301.1
 301.3 301.  301.3 301.1 301.2 301.2 301.4 301.4 301.3 301.4 301.5 301.4
 301.4 301.4 301.5 301.6 301.3 301.3 301.5 301.5 301.4 301.6 301.5 301.5]
```

2.2 The lab report

The lab reports are typically due two weeks after the activity, see the Calendar (1.3) for more details.

In general, a well-written lab report should include several essential components. It begins with a clear title and abstract summarizing the experiment. The introduction provides background information and states the purpose or hypothesis. A detailed methods section outlines the experimental procedure. The results section presents data objectively, using tables anf figures. In the discussion, the results are interpreted and related to existing knowledge. The conclusion summarizes key findings and their significance and implications. Finally, a reference list cites all sources used. However, in this unit we will mainly consider two of the report essential components, **the result and discuss sections**, and unless it is explicitly requested in the laboratory activity, no introduction and method sections will be required.

Careful statistical analysis and accurate data reporting are the foundations of reliable scientific research. Proper statistical methods help researchers distinguish meaningful results from random chance, quantify uncertainty, and draw valid conclusions from their data. Moreover, accurate reporting of these analyses, including effect sizes, confidence intervals, and p-values when appropriate, allows other scientists to evaluate the strength and reliability of the findings. When researchers adhere to rigorous statistical practices and report their findings comprehensively, it builds confidence in individual studies and strengthens the overall scientific process. Conversely, flawed analyses or selective reporting can lead to misleading conclusions, wasting resources and potentially misdirecting future research efforts. Ultimately, the integrity of statistical analysis and data reporting is vital for maintaining trust in scientific discoveries and ensuring the steady advancement of knowledge across all fields of study. Therefore, we will strongly focus on **data analysis**, **data presentation** and some aspects of the **discussion**.

2.2.1 General considerations

The lab report must abide by the requirements to be uploaded into Turnitin for plagiarism checks, which include (but are not limited to)

- scanned images are NOT allowed
- photos of tables, figures or handwritten notes are NOT allowed
- inclusion of other printed documents is NOT allowed

If any of these items is detected a zero mark may be given to the report. Although it does not require more effort than including a figure made with excel or Python, the image of a graph made through screen capture can be included in the document. Tables should not be included as images.
In general, a lab report should

1. Discuss and show how every reported quantity has been calculated, either by reporting the equations used and/or using of figures. If the same procedure is used for multiple files, one or a few examples of the calculation procedure would suffice.
2. Report average quantities with their 95% confidence interval, while the individual values should be reported in a table or a figure. Quantities derived from fitting should also be reported with their 95% confidence interval.
3. the student-t test (or p-score) should be reported when comparing with literature values to show whether the results are consistent or not.
4. Each figure and table must have an explanatory caption (above or below).
5. Any questions listed for the laboratory should be answered.

Note that some of the items above may not apply to all reports.

2.2.2 Figures

A good figure effectively communicates complex information visually. It should have a clear, concise caption that describes its content. All axes must be properly labeled with units where applicable. The figure should be self-explanatory, with a legend explaining any symbols or color coding used. Data points or bars should be easily distinguishable, with error bars included where appropriate. The choice of graph type (*e.g.* line, bar, scatter plot) should match the data and the message being conveyed. Colors should be used judiciously and be accessible to colorblind individuals. The figure's resolution should be high enough for all elements to be clearly visible, and its overall design should guide the viewer's eye to the most important information. Figures should be numbered consecutively, and referenced in the report. In Figure 2.2 you can see an example of a good figure, although some aspects of it could be improved.

2.2.3 Tables

A well-designed table presents data in a clear, organized manner. It should be as simple as possible while still conveying all necessary information, avoiding excessive lines or shading that could distract from the data. Similarly to a figure, it should have a clear, concise caption that describes its content. Column headings should be specific and include units of measurement where relevant. Rows and columns should be logically ordered, often with independent variables in columns and dependent variables in rows. Data should be aligned consistently (*e.g.* decimal points in a column should line up). The precision of reported numbers should be uniform and appropriate for the uncertainty of the measurements. Footnotes can be used to provide additional information about specific data points or to explain abbreviations. In Figure 2.3 you can see an example of a good figure, although some aspects of it could be improved. Tables should be numbered consecutively (separately from figures), and referenced in the report.

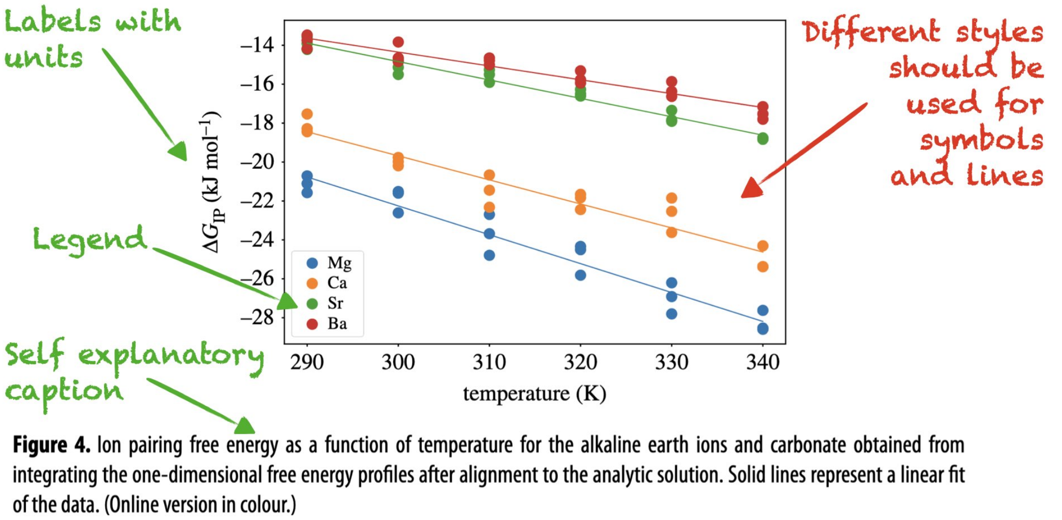


Figure 4. Ion pairing free energy as a function of temperature for the alkaline earth ions and carbonate obtained from integrating the one-dimensional free energy profiles after alignment to the analytic solution. Solid lines represent a linear fit of the data. (Online version in colour.)

Figure 2.2: Example of a figure for the report. Key important elements that should be included are marked in green, while red notes mark places where it can be improved.

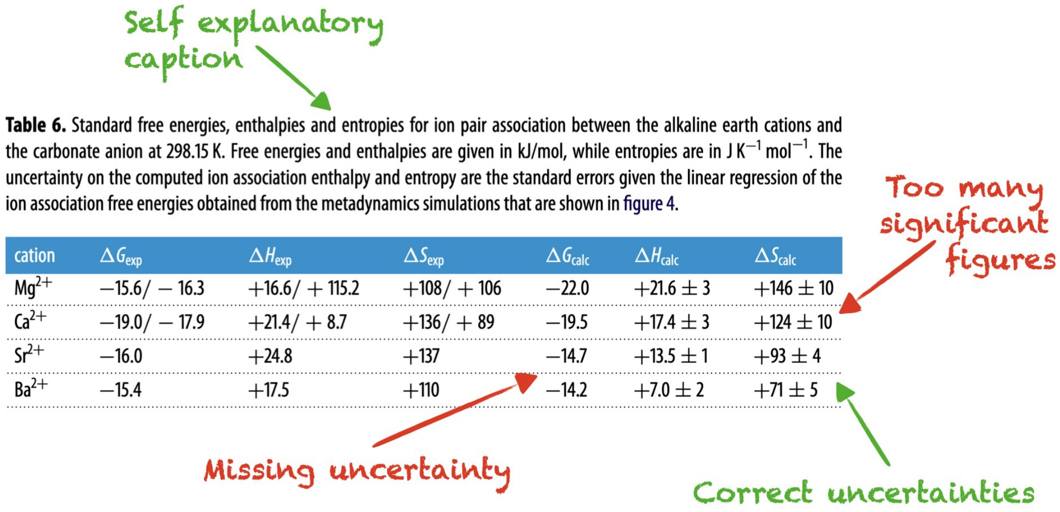


Table 6. Standard free energies, enthalpies and entropies for ion pair association between the alkaline earth cations and the carbonate anion at 298.15 K. Free energies and enthalpies are given in kJ/mol, while entropies are in J K^{-1} mol^{-1}. The uncertainty on the computed ion association enthalpy and entropy are the standard errors given the linear regression of the ion association free energies obtained from the metadynamics simulations that are shown in figure 4.

cation	ΔG_{exp}	ΔH_{exp}	ΔS_{exp}	ΔG_{calc}	ΔH_{calc}	ΔS_{calc}
Mg^{2+}	$-15.6/-16.3$	$+16.6/+115.2$	$+108/+106$	-22.0	$+21.6 \pm 3$	$+146 \pm 10$
Ca^{2+}	$-19.0/-17.9$	$+21.4/+8.7$	$+136/+89$	-19.5	$+17.4 \pm 3$	$+124 \pm 10$
Sr^{2+}	-16.0	$+24.8$	$+137$	-14.7	$+13.5 \pm 1$	$+93 \pm 4$
Ba^{2+}	-15.4	$+17.5$	$+110$	-14.2	$+7.0 \pm 2$	$+71 \pm 5$

Figure 2.3: Example of a table for the report. Key important elements that should be included are marked in green, while red notes mark places where it can be improved.

2.3 Marking rubric

Each lab report is marked out of 20. Feedback about your report is available on blackboard via Turnitin, or ask your lab demonstrator.

The lab reports are marked based on four main categories:

- Data Analysis (8 marks)
  This item considers if the data have been analysed properly and the correct answers have been obtained. Including the comparison with expected values, if relevant.

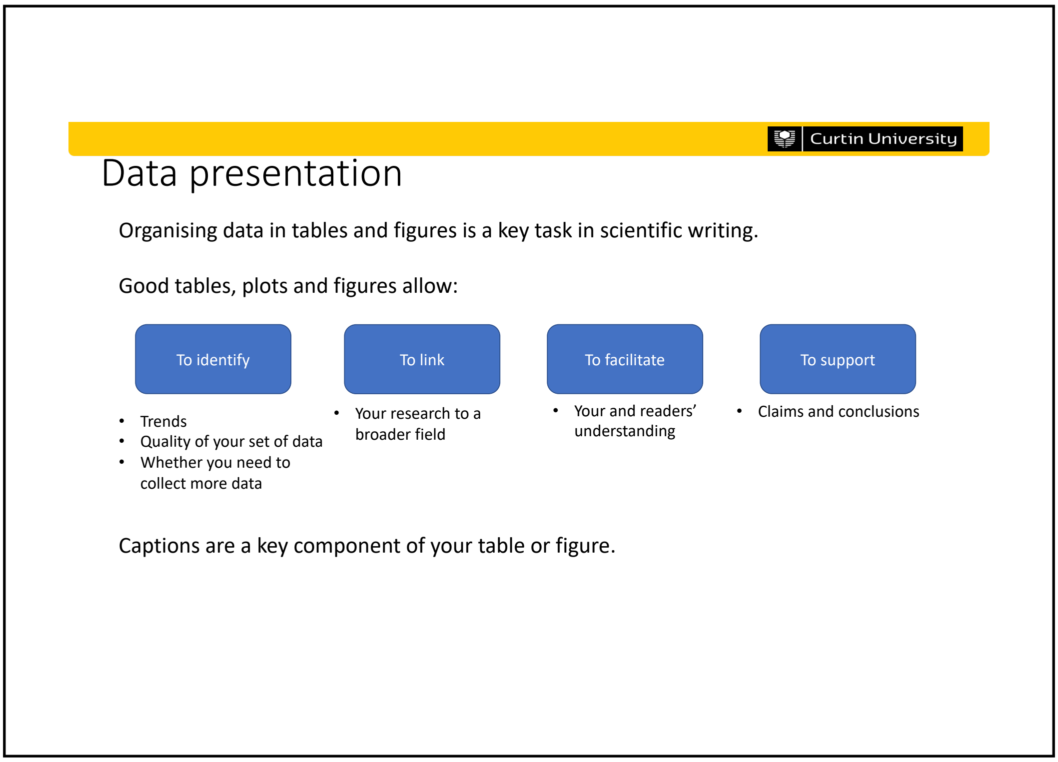
- Method (4 marks)
  This item considers the description of how the results have been obtained, including the presence of all the relevant equations. For the Chemical equilibrium laboratory, the method would also include a short description of the procedure used to numerically solve chemical equilibrium problems.
- Figures and Tables (4 marks)
  This item considers the quality of data presentation in tables and figures, including their captions.
- Uncertainty Analysis (4 marks)
  This item considers the presence of uncertainties on the computed values, error propagation and the correct use of significant figures.

Table 2.1: CHEM2000 Laboratory Assessment Rubric

Criteria	**Absent (0%)**	**Limited (25%)**	**Developing (50%)**	**Competent (75%)**	**Proficient (100%)**
Data Analysis (8 pts)	No evidence of the required skills, knowledge, or criteria is demonstrated. The work shows no attempt or understanding of the expectations.	Basic elements are missing or incorrect. Shows minimal understanding with significant gaps and errors that impede effectiveness.	Shows partial understanding and attempts to meet requirements. Core elements are present but need refinement and greater consistency.	Demonstrates solid understanding and meets most requirements effectively. Minor improvements could enhance quality but core expectations are fulfilled.	Consistently demonstrates thorough understanding and mastery. Work is comprehensive, well-executed, and meets or exceeds all requirements.
Method (4 pts)	No evidence of the required skills, knowledge, or criteria is demonstrated. The work shows no attempt or understanding of the expectations.	Basic elements are missing or incorrect. Shows minimal understanding with significant gaps and errors that impede effectiveness.	Shows partial understanding and attempts to meet requirements. Core elements are present but need refinement and greater consistency.	Demonstrates solid understanding and meets most requirements effectively. Minor improvements could enhance quality but core expectations are fulfilled.	Consistently demonstrates thorough understanding and mastery. Work is comprehensive, well-executed, and meets or exceeds all requirements.
Tables and Figures (4 pts)	No evidence of the required skills, knowledge, or criteria is demonstrated. The work shows no attempt or understanding of the expectations.	Basic elements are missing or incorrect. Shows minimal understanding with significant gaps and errors that impede effectiveness.	Shows partial understanding and attempts to meet requirements. Core elements are present but need refinement and greater consistency.	Demonstrates solid understanding and meets most requirements effectively. Minor improvements could enhance quality but core expectations are fulfilled.	Consistently demonstrates thorough understanding and mastery. Work is comprehensive, well-executed, and meets or exceeds all requirements.
Uncertainty (4 pts)	No evidence of the required skills, knowledge, or criteria is demonstrated. The work shows no attempt or understanding of the expectations.	Basic elements are missing or incorrect. Shows minimal understanding with significant gaps and errors that impede effectiveness.	Shows partial understanding and attempts to meet requirements. Core elements are present but need refinement and greater consistency.	Demonstrates solid understanding and meets most requirements effectively. Minor improvements could enhance quality but core expectations are fulfilled.	Consistently demonstrates thorough understanding and mastery. Work is comprehensive, well-executed, and meets or exceeds all requirements.

2.4 Report Writing Checklist

For more resources on scientific writing please check UniSkills



Curtin University

Data presentation

Organising data in tables and figures is a key task in scientific writing.

Good tables, plots and figures allow:

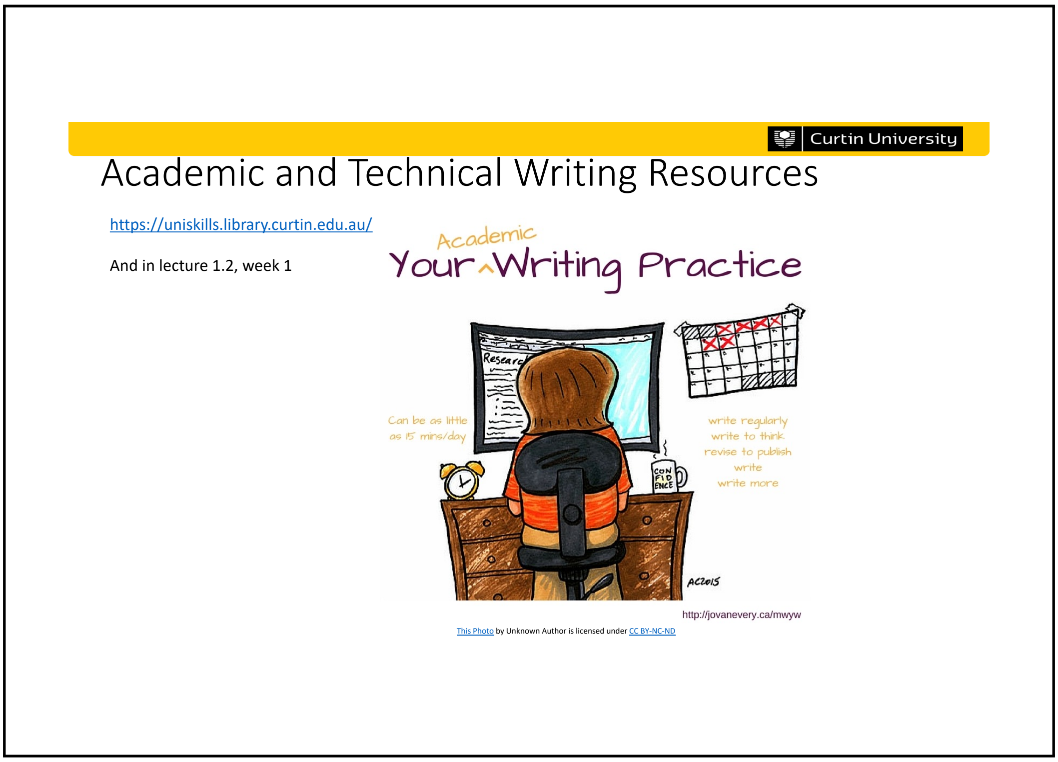
To identify
- Trends
- Quality of your set of data
- Whether you need to collect more data

To link
- Your research to a broader field

To facilitate
- Your and readers' understanding

To support
- Claims and conclusions

Captions are a key component of your table or figure.



Curtin University

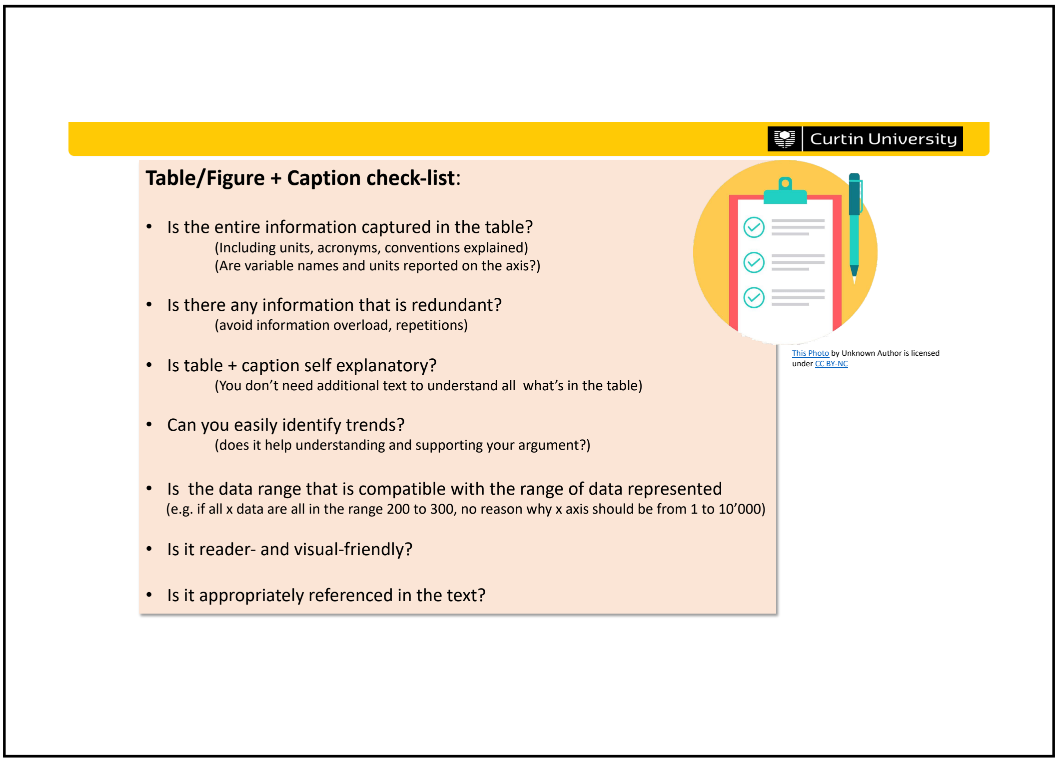
Academic and Technical Writing Resources

https://uniskills.library.curtin.edu.au/

And in lecture 1.2, week 1

http://jovanevery.ca/mwyw

This Photo by Unknown Author is licensed under CC BY-NC-ND

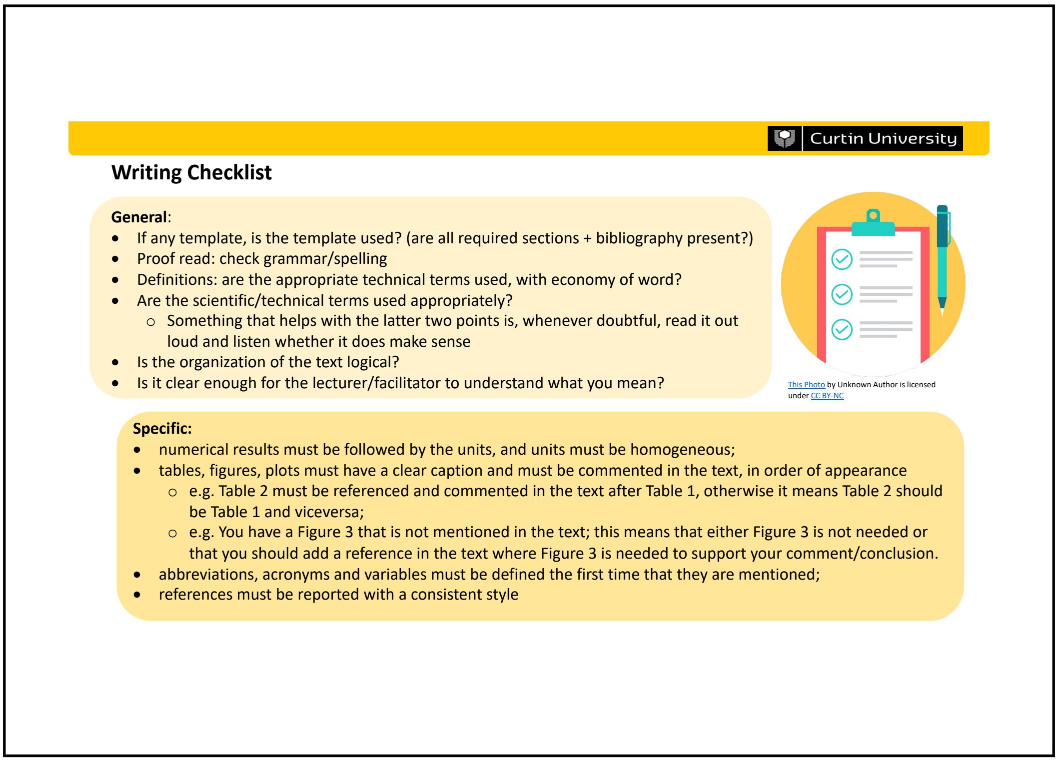


Curtin University

Table/Figure + Caption check-list:

- Is the entire information captured in the table?
  (Including units, acronyms, conventions explained)
  (Are variable names and units reported on the axis?)
- Is there any information that is redundant?
  (avoid information overload, repetitions)
- Is table + caption self explanatory?
  (You don't need additional text to understand all what's in the table)
- Can you easily identify trends?
  (does it help understanding and supporting your argument?)
- Is the data range that is compatible with the range of data represented
  (e.g. if all x data are all in the range 200 to 300, no reason why x axis should be from 1 to 10'000)
- Is it reader- and visual-friendly?
- Is it appropriately referenced in the text?

This Photo by Unknown Author is licensed under CC BY-NC



Curtin University

Writing Checklist

General:
- If any template, is the template used? (are all required sections + bibliography present?)
- Proof read: check grammar/spelling
- Definitions: are the appropriate technical terms used, with economy of word?
- Are the scientific/technical terms used appropriately?
  - Something that helps with the latter two points is, whenever doubtful, read it out loud and listen whether it does make sense
- Is the organization of the text logical?
- Is it clear enough for the lecturer/facilitator to understand what you mean?

This Photo by Unknown Author is licensed under CC BY-NC

Specific:
- numerical results must be followed by the units, and units must be homogeneous;
- tables, figures, plots must have a clear caption and must be commented in the text, in order of appearance
  - e.g. Table 2 must be referenced and commented in the text after Table 1, otherwise it means Table 2 should be Table 1 and viceversa;
  - e.g. You have a Figure 3 that is not mentioned in the text; this means that either Figure 3 is not needed or that you should add a reference in the text where Figure 3 is needed to support your comment/conclusion.
- abbreviations, acronyms and variables must be defined the first time that they are mentioned;
- references must be reported with a consistent style

3. Statistics laboratory

This numerical lab consists a few small tasks, which cover the key statistics topics that were introduced in the previous chapter. They are also preparatory for the following labs, where you would have to use the same concepts in more complicated situations. In particular, if you are using python, it would be beneficial to solve some of this exercises by creating specific functions that can the be reused (maybe with small modifications) in the following labs.

3.1 Pre-lab questions

1. We want to calculate the average weight of 10 screws using a scale that has a precision of 1 g.

    a. Do we get a more accurate answer if we weight them individually or all together?

    b. Besides the time-cost, would it be different if you had 1000 screws?

    c. Can you calculate the standard error on the two measurements?

    Justify your answers.

2. Given the following set of observations

$$X = \{25, 35, 31, 22, 27, 38, 18, 28, 36, 26\}$$

    a. Calculate the average ($\overline{X}$), standard deviation, standard error and 95% confidence interval.

    b. Is your estimate for $\overline{X}$ comparable with the its expected value of 30?

    c. What is the expectation value of $\ln(\overline{X})$ and its uncertainty?

3.2 Tasks

This lab covers six topics, which will be described in more details in the following sections.

1. Average and standard error
2. Propagation of uncertainty
3. Comparison of averages
4. Linear Fit
5. Non linear fit
6. Outlier detection

3.2.1 Average and standard error

This task focuses on key statistical concepts: averages, standard errors, confidence intervals, and comparing with expected/experimental quantities

Specific objectives

- Calculate averages and standard errors using the given data.
- Calculate the confidence interval of the data using the t-statistics.
- Compare the averages of the values in each column of the file with the expected experimental values reported in the file metadata (t-test and p-score).

3.2.2 Propagation of uncertainty

This task explores error propagation in scientific measurements, examining how uncertainties in individual variables combine to affect the overall uncertainty of a calculated result. You will have to use techniques to quantify and analyze these compounded errors, essential skills for accurate data analysis and interpretation in experimental sciences.

Specific objectives

- Compute the averages ($\overline{X}$ and $\overline{Y}$) and 95% confidence intervals of the measurements reported in the file.
- Calculate the 95% confidence intervals for the following derived quantities
    1. $f = 3\overline{X}$
    2. $f = \ln \overline{X}$
    3. $f = \overline{X} + 2\overline{Y}$
    4. $f = \overline{X}\,\overline{Y}$
    5. $f = \overline{X} \ln \overline{Y}$

3.2.3 Comparison of averages

This task focuses on comparing averages obtained from different sources. Imagine a scenario where two students decide to speed up their work by splitting the experiments between themselves and do the measurements simultaneously using two different instruments. Determine if their measurements are consistent. If not, this could be an indication that one of the instruments has a systematic error, or that one of the students have made a mistake in the post-processing of the data.

Specific objectives

- Generate the data, each column correspond to one student
- Compute averages, standard error
- Compute the p-score to verify if the data are consistent

3.2.4 Linear fit (Beer-Lambert law)

This task focuses on the concept of linear regression. Many properties in chemistry exhibit a linear dependency on certain variables, with the Beer-Lambert law being one of the most well-known examples. This fundamental principle describes the linear relationship between the absorption of light and the concentration of a solution. The Beer-Lambert law is widely used in spectroscopy and forms the basis for numerous analytical techniques in chemistry and related fields, and it can expressed as

$$A = \log\left(\frac{I_0}{I}\right) = \varepsilon c l \tag{3.1}$$

where A is the absorbance, I_0 is the intensity of the incident light, I is the intensity of the transmitted light, ε is the molar absorptivity (or molar extinction coefficient), c is the concentration of the absorbing species and l is the path length of the sample (usually in cm).

Specific objectives

- Do a liner fit of the provided data
- Extract the value of the ε and its uncertainty assuming the optical path is 1 cm
- Compare it with the experimental value shown in the file metadata

3.2.5 Non-linear fit (Murnaghan equation of state)

This task focuses on non linear regression and the importance of usig the correct equation for the fitting. Many properties in chemistry exhibit non-linear relationships with various parameters, adding complexity to their analysis and interpretation. Non-linear relationships are common in reaction kinetics, phase transitions, and many other areas of chemistry, often requiring more sophisticated mathematical models for accurate predictions and understanding. The Murnaghan equation of state is a widely used model in materials science and condensed matter physics to describe how materials behave under pressure, particularly in terms of their volume change. It relates the pressure (P) to the volume (V) of a material, taking into account its bulk modulus (K) and its pressure derivative (K'). The Murnaghan equation of state can be represented in pressure form

$$P(V) = \frac{K}{K'}\left[\left(\frac{V_0}{V}\right)^{K'} - 1\right] \tag{3.2}$$

or in volume form

$$V(P) = V_0\left(1 + \frac{K'P}{K}\right)^{-1/K'} \tag{3.3}$$

where P is the pressure, V is the volume, V_0 is the reference volume (usually the equilibrium volume), K is the bulk modulus at the reference volume and K' is the pressure derivative of the bulk modulus

The Murnaghan equation is often also employed to fit energy-volume ($E - V$) data obtained from first-principles computer simulations. For this purpose, the energy form of the equation is more commonly used:

$$E(V) = E_0 + \frac{KV}{K'}\left[\frac{(V_0/V)^{K'}}{K'-1} + 1\right] - \frac{KV_0}{K'-1} \tag{3.4}$$

where E_0 is the energy of the equilibrium state, *i.e.* the minimum energy of the data.

Specific objectives

The objective of this task is to do a non linear fit of the provided data for quartz and extract the bulk modulus and its pressure derivative. With this exercise we want to demonstrate the importance of choosing the correct function to fit the data, which is based on the chemical/physical process that we are trying to model, and not by the shape of the curve. In fact, the data can also be well reproduced by a cubic equation, but in doing so we would not be able to extract the bulk modulus and its derivative from them. In particular you would have to

1. fit the $E - V$ data using a quadratic and a cubic function (show figures with data, fit and residuals)
2. fit the data using the Murnaghan equation (show figures with data, fit and residuals)
3. comment on the accuracy of the fits (R^2, χ^2 or residues)
4. report the values for K, K' and V_0.

Remember that for non linear fits, the initial guess for the fitting parameters can be critical to obtain a good fit.

3.2.6 Detection of outliers

Use Dixon's Q-test or Grubbs' test to identify outliers in the data.

Specific objectives

- Do a linear fit and check for outliers using the residuals

3.3 Lab report structure and content

3.3.1 General considerations

The lab report must abide by the requirements to be uploaded into Turnitin for plagiarism checks, which include (but are not limited to)

- scanned images are NOT allowed
- photos of tables, figures or handwritten notes are NOT allowed
- inclusion of other printed documents is NOT allowed

If any of these items is detected a zero mark may be given to the report. Although it does not require more effort than including a figure made with excel or Python, the image of a graph made through screen capture can be included in the document. Tables should not be included as images.
In general, a lab report should

1. Discuss and show how every reported quantity has been calculated, either by reporting the equations used and/or using of figures. If the same procedure is used for multiple files, one or a few examples of the calculation procedure would suffice.

2. Report average quantities with their 95% confidence interval, while the individual values should be reported in a table or a figure. Quantities derived from fitting should also be reported with their 95% confidence interval.
3. the student-t test (or p-score) should be reported when comparing with literature values to show whether the results are consistent or not.
4. Each figure and table must have an explanatory caption (above or below).
5. Any questions listed for the laboratory should be answered.

Note that some of the items above may not apply to all reports.

4. Bomb calorimetry - Thermodynamics

Calorimetry derives from the Latin word *calor* meaning heat, and the Greek *metry* meaning to measure. Thus, it is the science of measuring amounts of heat. All calorimetric techniques are based on the measurement of heat generated or consumed by a sample, depending on whether exothermic or endothermic processes are involved. A calorimeter can be a simple Styrofoam cup, because it is a container with insulated walls that prevents heat exchange with the environment, or more a complicated device such as a bomb calorimeter.

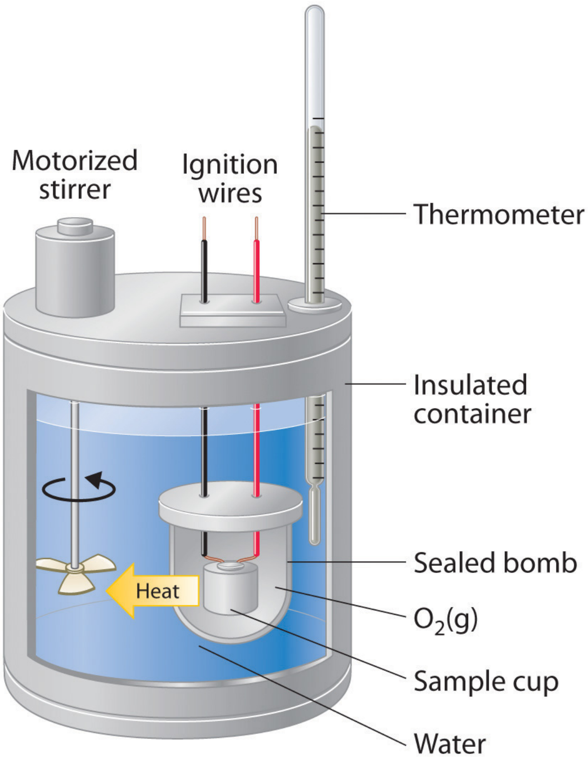


Figure 4.1: Schematic representation of a bomb calorimeter.

In the simplest version of this experiment, in order to measure heats of reactions, we enclose the reactants in a calorimeter, allow the reaction to proceed, and measure the temperature change (ΔT). Then the heat released or adsorbed in the reaction can be calculated. A calorimeter may be used under constant pressure (atmospheric), or constant volume. Regardless of the setup, the calorimeter adsorbs some of the released heat and its heat capacity must be measured to calibrate the instrument, which is usually done just before or after the actual measurement of the heats of reaction. This is done by transferring a known amount of heat into the calorimeter and measuring its temperature increase. Because the temperature differences are often very small, sensitive thermometers are preferred for these measurements.

A good explanatory video about this experiments made by Prof. Chris Ambidge at the University of Toronto can be found here.
(https://www.youtube.com/watch?v=bm3Tn6DXJrI)

4.1 Background

Conventionally, in thermodynamics we use two fundamental definitions: the system and the surroundings. The system is what is being measured, observed and investigated while the surroundings are everything else. An isolated system is referred to as adiabatic, which means that there is no exchange of matter or energy between the system and its surroundings. The energy considerations of a reaction include heat (denoted q), which is the energy absorbed by the system, and the work done on the system (denoted w). As heat is transferred into or out of a system this leads to a change in temperature, and the quantity of energy that causes a temperature change is determined with this equation

$$q = \sum c(T_f - T_i) \tag{4.1}$$

where q is the heat of the process, $\sum c$ is sum of the heat capacities of all the various components of the system, T_f is the final temperature and T_i is the initial temperature. Because it is difficult, if not impossible to separate the different contributions, in the following we will just use an "effective" heat capacity of the calorimeter, $C_V = \sum c$. Recalling the physical chemistry lectures, the heat capacity is the amount of heat required to raise the temperature of a substance (or of the entire calorimeter) by 1 K.

The enthalpy of any process is the heat absorbed by the system at constant pressure. The enthalpy of a reaction under isothermal conditions can be determined by conducting the experiment adiabatically, and then measuring the temperature change that occurs in the reaction vessel. The isothermal reaction at temperature T_i can be expressed as follows:

$$\mathrm{A(T_i) + B(T_i)} \xrightarrow{\Delta H_c} \mathrm{C(T_i) + D(T_i)} \tag{4.2}$$

where ΔH is the enthalpy of the isothermal reaction for each mole of reaction as written. In a calorimeter the system is insulated, and therefore no energy is lost to the surroundings. This is known as an adiabatic system and therefore in the calorimeter the energy of the reaction is used to heat the system with no loss. The adiabatic reaction can be expressed as follows:

$$\mathrm{A(T_i) + B(T_i) + S(T_i)} \xrightarrow{\Delta H_I} \mathrm{C(T_f) + D(T_f) + S(T_f)} \tag{4.3}$$

where ΔH_I is the enthalpy of the adiabatic reaction. A, B, C and D continue to be the reactants and products respectively and S is the parts of the system (thermometer, solvent, reaction vessel walls etc.) that maintain the same temperature as the reactants and products because of the insulated system. Enthalpy however is a state function, and therefore independent of the steps taken to get to the final stage. Thus, the enthalpy of the isothermal reaction will be the sum of the enthalpy of any steps that start and finish in the same state. Therefore in order to finish the adiabatic reaction under the same conditions as the isothermal reaction, the system merely needs to be cooled to the initial temperature.

$$\mathrm{A(T_f) + B(T_f) + S(T_f)} \xrightarrow{\Delta H_{II}} \mathrm{C(T_i) + D(T_i) + S(T_i)} \tag{4.4}$$

ΔH_I and ΔH_{II} are the enthalpies of the alternate steps taken to go from the initial stage to the final stage of the isothermal reaction, and therefore, as shown in the figure below, the enthalpy of the isothermal reaction is the sum of these two values.

$$\Delta H_c = \Delta H_I + \Delta H_{II} \tag{4.5}$$

An adiabatic reaction is a reaction contained in an isolated system, i.e. no energy (heat) is transferred out of the system. This means that for the first step in Figure 4.2, dq=0 and $\Delta H = 0$ adiabatic step is 0, and therefore ΔH_I is also 0. The cooling step indicated in Figure 4.2 does not need to be performed, but it can be simply calculated from the change in temperature of the water bath

$$\Delta H_{II} = C_V(T_f - T_i) \tag{4.6}$$

where C_V is the heat capacity of the whole bomb calorimeter.

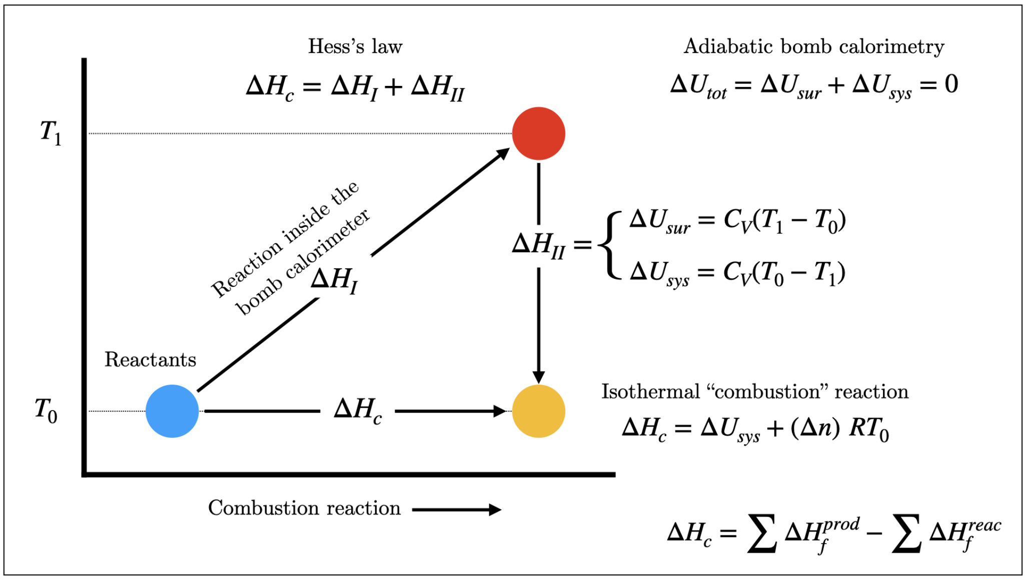


Figure 4.2: Schematic representation of the working principles of the calorimeter and key equations

4.2 Objectives

1. Calibration of the calorimeter
2. Calculation of enthalpy of combustion of sucrose
3. Calculation of enthalpy of combustion of naphthalene

4.3 Procedure

The experiments are performed using tablets of the reactant, which all have different masses. Recording and using the masses of the individual tablets is important because it allows we to connect the heat released by the combustion (extensive quantity) to the enthalpy of combustion (intensive/molar quantity), and to compute the change is the moles of gas present in the calorimeter.

The change in internal energy of the system is defined as:

$$\Delta U = dq - P\Delta V \tag{4.7}$$

However, because the experiment is performed a constant volume this becomes

$$\Delta U = dq_V \tag{4.8}$$

Therefore, by using the calorimeter to measure the change in temperature we can measure the change in internal energy of the system

$$\Delta U = dq_V = C_V(T_f - T_i) = C_V \Delta T \tag{4.9}$$

Where C_V is the heat capacity of the entire system at constant volume. Using the definition of enthalpy

$$H = U + PV \tag{4.10}$$

we can then write the combustion enthalpy change for a reaction as

$$\Delta H_c = \Delta U + \Delta(PV) = \Delta U + V\Delta P + P\Delta V \tag{4.11}$$

At constant volume, because $P\Delta V = 0$, this becomes

$$\Delta H_c = \Delta U + \Delta(nRT) \tag{4.12}$$

where we have assumed that the gas is ideal and $V\Delta P = \Delta(nRT)$. Where P is the pressure in kPa, V is the volume in litres, Δn is the change in the number of moles of gas particles during the reaction, R is the universal gas constant and T is the isothermal temperature. The enthalpy of the combustion reaction in the bomb calorimeter can then be calculated using change of internal energy ($\Delta U = C_v \Delta T$) and from the gas stoichiometry of the equation. The enthalpy of combustion can also be determined from the formation enthalpies of reactants and products

$$\Delta H_c = \sum \Delta H_f^{products} - \sum \Delta H_f^{reactants} \tag{4.13}$$

4.3.1 Calculation of the temperature change

Because each data set represents an experiment done with a different tablet, the data have to be analysed individually. The 95% confidence interval can then be determined by averaging the final quantity that is computed, that being either the calorimeter heat capacity of the molar enthalpy of combustion. During a calorimetry experiment, the water temperature is monitored as a function of time, and typical observations are reported in Figure 4.3, with the ignition of the sample occurring at the 20 seconds mark. In order to measure ΔT, we have to fit the linear portions of the data (before and after the ignition) and compute the temperature difference at the time of ignition. This corresponds to the difference between the values of the best fitting line at the time of ignition (Figure 4.3).

4.3.2 Calibration of the calorimeter

An adiabatic bomb calorimeter provides a constant volume environment for the measurement of the heat of combustion of compounds. An important first step in any calorimetry experiment is the calibration of the instrument, in order to achieve that, we perform a first set of experiments with benzoic acid. From the experimental data, we will determine T_i and T_f, and knowing the enthalpy of combustion of benzoic acid ($\Delta H_c = -3228$ kJ/mol) we have to determine the heat capacity of the calorimeter. Typically, a bomb calorimetry contains 2 L of DI water, so the expected heat capacity of the calorimeter should be of a bit larger than the heat capacity of 2 L of water.

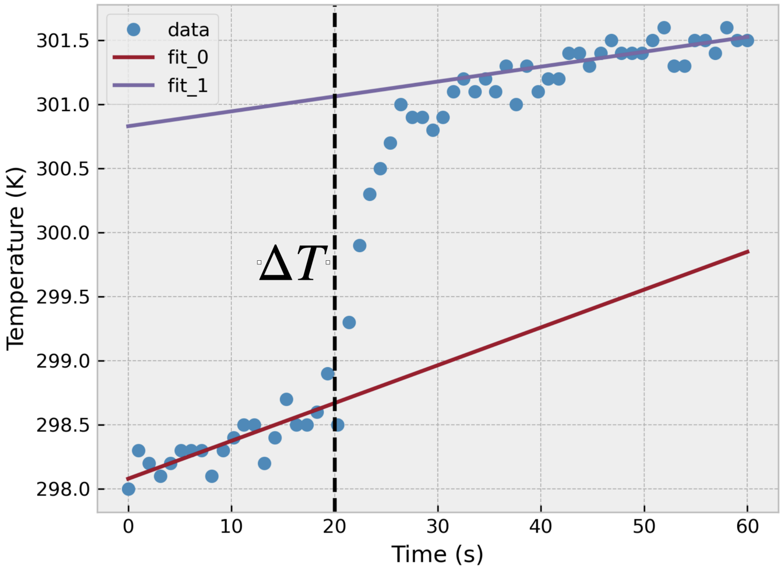


Figure 4.3: Example of the temperature data for the bomb calorimetry experiment, and the post processing

4.3.3 Calculation of enthalpy of combustion

Now that we have determined the heat capacity of the calorimeter, C_V, we can repeat the same experiment with different substances for which their enthalpy of combustion is unknown. In this case we will use data for succrose and naphthalene. Similarly to the previous experiments, we will determine T_i and T_f from the data and using the computed value of C_V we can determine the heat of combustion of the two species.

4.3.4 Notes

These experiments are normally carried out under 20 atm of excess oxygen in order to ensure complete and rapid combustion, which are not standard temperature nor pressure, and therefore the energy change measured is not the standard heat of reaction. Corrections would need to be made to convert to standard conditions, however the experimental uncertainty in this lab accounts for the errors introduced by the non-standard conditions.

4.4 Pre-lab questions

Answering these questions before starting your calculations will help you to verify that your procedure is correct by checking that your results are consistent with the expexted values.

1. Calculate is the heat capacity of 2 L of water?
2. Write the balanced combustion reactions for benzoic acid, sucrose and naphthalene.
3. Calculate the change in moles of gas present in the calorimeter upon combustion of 1 mole of benzoic acid, sucrose and naphthalene.
4. Obtain from literature the standard enthalpies of formation of O_2, H_2O, CO_2, sucrose and naphthalene.
5. Calculate the expected combustion enthalpy of sucrose and naphthalene

4.5 Lab report structure and content

4.5.1 General considerations

The lab report must abide by the requirements to be uploaded into Turnitin for plagiarism checks, which include (but are not limited to)

- scanned images are NOT allowed
- photos of tables, figures or handwritten notes are NOT allowed
- inclusion of other printed documents is NOT allowed

If any of these items is detected a zero mark may be given to the report. Although it does not require more effort than including a figure made with excel or Python, the image of a graph made through screen capture can be included in the document. Tables should not be included as images.
In general, a lab report should

1. Discuss and show how every reported quantity has been calculated, either by reporting the equations used and/or using of figures. If the same procedure is used for multiple files, one or a few examples of the calculation procedure would suffice.
2. Report average quantities with their 95% confidence interval, while the individual values should be reported in a table or a figure. Quantities derived from fitting should also be reported with their 95% confidence interval.
3. the student-t test (or p-score) should be reported when comparing with literature values to show whether the results are consistent or not.
4. Each figure and table must have an explanatory caption (above or below).
5. Any questions listed for the laboratory should be answered.

Note that some of the items above may not apply to all reports.

4.5.2 Specific items to be included in the report for this lab

At a minimum, the bomb calorimetry lab report should contain the following

1. one table per compound reporting the tablet mass and measured ΔT for each experiment
2. at least one figure demonstrating how the ΔT is calculated from the data
3. the heat capacity of the bomb calorimeter
4. the combustion and formation enthalpies (ΔH_c & ΔH_f) of sucrose and naphthalene
5. comparison with the experimental values for those quantities

5. Crystal violet - Kinetics

Crystal violet has many uses, which include its use as a dye, as a constituent in inks for printing, as an antibacterial and antifungal, to treat mouth ulcers and for the development of fingerprints in forensics. The dye has problems; the process requires the colour to be set in a highly basic washing soda and it has found that the dye looses it's colour over time during this process. The goal of this experiment is to investigate the decolourisation of the dye by studying the role of sodium hydroxide in the kinetics of crystal violet decolourisation. In this experiment a UV-Vis spectrophotometer is used to measure the absorbance as the reaction between crystal violet and hydroxide proceeds. The absorbance versus time data can then be used to determine the rate of the reaction with respect to both crystal violet and hydroxide ions.

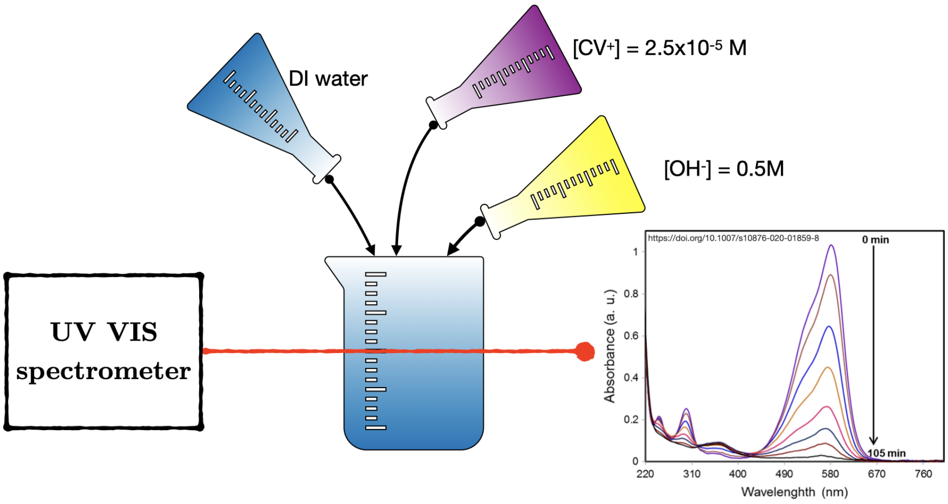


Figure 5.1: Caption

5.1 Background

5.1.1 Absorbance and concentration

Beer's law can be used to relate the absorbance, and the crystal violet concentration, [CV+] by the equation:

$$A = \varepsilon \; l \; [\mathrm{CV}^{+}] \tag{5.1}$$

where ε is the molar absorptivity of crystal violet at the wavelength of the maximum absorption peak, l is the length of the cuvette (l = 1cm) and [CV+] is the crystal violet concentration.

Since the absorbance, A, of the crystal violet is monitored as a function of time, it is not necessary to know the actual concentration of the crystal violet at any particular time. The absorbance, A, may therefore be used as a proxy of the crystal violet concentration in all calculations.

The integrated form equations for zero, first and second order can be rearranged and $[\mathrm{CV+}]_0/[\mathrm{CV+}]_t$ substituted for A_0/A_t. This will give theoretical expressions that relate the absorbance values to time and to the rate constant k_r.

The rate (or velocity) ν $[molL^{-1}s^{-1}]$ of a chemical reaction can be expressed in terms of loss of reactants or the formation of products. For the reaction:

$$n_aA + n_bB + \dots \rightarrow n_xX + n_yY + \dots \tag{5.2}$$

the rate can be written as

$$\nu = -\frac{1}{n_a}\frac{\mathrm{d[A]}}{\mathrm{d}t} = -\frac{1}{n_b}\frac{\mathrm{d[B]}}{\mathrm{d}t} = \frac{1}{n_x}\frac{\mathrm{d[X]}}{\mathrm{d}t} = \frac{1}{n_y}\frac{\mathrm{d[Y]}}{\mathrm{d}t} = \dots \tag{5.3}$$

where n_a, n_b, n_x and n_y, and [A], [B], [X] and [Y] are the stoichiometric coefficients and molar concentrations of the reactants and products. Note that the rate is always a positive quantity; therefore, the equation has a different sign to account for *loss* of reactants or *formation* of products.

5.1.2 Rate laws

The rate, ν, of a chemical reaction is frequently found to be proportional to the concentrations of the reacting species.

$$\nu = k_r[\mathrm{A}]^a[\mathrm{B}]^b[\mathrm{C}]^c \tag{5.4}$$

where k_r is the rate constant and a, b, and c denote the order of the reaction with respect to reactants A, B and C. The sum of the individual orders of the reactants is the overall order of the reaction. Most reactions are a complicated series of elementary reactions, and the measured rate is generally the rate of the slowest step. Therefore knowledge of the order of each reactant with respect to the overall reaction often allows the specific reaction mechanisms to be theorized. The rate of any reaction can vary with time, and therefore it is impossible to define a general rate of a reaction. Rather an instantaneous rate can be determined for any particular time. If one can determine the concentation of a reactant (or product) as a function of time, this instantaneous rate will be given by the tangent of the time-concentration curve.

The time *vs.* concentration curve corresponds to the *integrated* rate law, and it is often more useful than the instantaneous rate.

Reaction order	Rate law	Integrated rate law
zeroth	$\nu = k_r[\mathrm{A}]^0$	$[\mathrm{A}] = [\mathrm{A}]_0 - k_rt$
first	$\nu = k_r[\mathrm{A}]^1$	$\ln[\mathrm{A}] = \ln[\mathrm{A}]_0 - k_rt$
second	$\nu = k_r[\mathrm{A}]^2$	$\frac{1}{[\mathrm{A}]} = \frac{1}{[\mathrm{A}]_0} + k_rt$

Table 5.1: Summary of the most common rate laws for elementary reactions. The subscript 0 denoted the initial concentration and t is the time of when the concentration, [A], is measured.

For a **zeroth order** reaction the rate is constant and independent of the concentration of the reactant, [A]. Therefore a plot of the concentration of the reactant vs. time will be linear with a negative slope. The slope can be used to find the rate constant k_r, which will have units of (concentration/time).

For a **first order** reaction the rate is directly proportional to the concentration of the reactant, [A]. Therefore a plot of the natural log of the concentration of the reactant vs.

time will be linear with a negative slope. The slope can be used to find the rate constant k_r, which will have units of (time)$^{-1}$.

For a **second order** reaction the rate is directly proportional to the square of the concentration of the reactant, [A]. Therefore a plot of the reciprocal of the concentration of the reactant vs. time will be linear with a positive slope. The slope can be used to find the rate constant k_r, which will have units of (concentration*time)$^{-1}$. Second order reactions can arise when a reaction is first order in two separate reactants. The method of isolation described next, applies to this scenario.

5.1.3 Isolation method

When an elementary reaction involves more than one reactant,

$$n_a A + n_b B + n_c C \rightarrow n_x X + n_y Y + \dots \tag{5.5}$$

the most general form of the instantaneous rate equation includes the concentration of all the reactants.

$$\nu = k_r [\mathrm{A}]^a [\mathrm{B}]^b [\mathrm{C}]^c \tag{5.6}$$

It is then much more difficult to analyse the rate equation in a way similar to that described above for simple zero, first and second order processes. The analysis can be greatly simplified by **isolating** one reactant. In the isolation method, all concentrations the rate equation are kept (as much as possible to) constant, except for one. This is usually achieved by having all but one reactant present in large excess, so the concentration hardly changes during the reaction. In this case, the limiting reagent (for example, A) is said to be isolated, and its concentration is measured as a function of time to determine the kinetics of the reaction. Then we can rewrite the rate equation as

$$\nu = k_r^{\dagger} [\mathrm{A}]^a \tag{5.7}$$

where $k_r^{\dagger}$ is a new "pseudo rate constant" incorporating the values of $[\mathrm{B}]^b$ and $[\mathrm{C}]^c$. Note that the value of $k_r^{\dagger}$ depends on the chosen "constant" values of [B] and [C].

$$k_r^{\dagger} = k_r [\mathrm{B}]^b [\mathrm{C}]^c \tag{5.8}$$

Now that A is **isolated**, we can apply the methods of analysis for a single reactant to any complicated reaction and determine the order of that reaction with respect to A and $k_r^{\dagger}$. This is done by plotting the relationships for zero, first and second orders, and choosing which plot fits best (*i.e.* which is the most linear).

Then, in order to find b, c and k_r we can look at a logarithmic form of the pseudo-rate constant

$$\log k_r^{\dagger} = b \log\,[\mathrm{B}] + c \log\,[\mathrm{C}] + \log k_r \tag{5.9}$$

This means that if $k_r^{\dagger}$ is found for different values of [B] (while [C] is kept constant and the conditions are still met for isolation of A), then b can be found from a plot of $\log k_r^{\dagger}$ versus $\log\,[\mathrm{B}]$. A near-integer value of slope can be rounded to the integer, as the order of reaction is expected to be an integer. Similarly, if $k_r^{\dagger}$ is found for different values of [C] (while [B] is kept constant and the conditions are still met for isolation of A), then c can be found from a plot of $\log k_r^{\dagger}$ versus $\log\,[\mathrm{C}]$. A near-integer value of slope can be rounded to the integer, as the order of reaction is expected to be an integer. Finally, now that b, and c have been found, k_r can be found using the intercepts of the log / log plots.

5.1.4 Activation energy

By using experiments run at different temperatures, we can determine the activation energy for the reaction. In fact, as discovered by Arrhenius, chemical reactions proceed faster at higher temperatures. This is reflected by the rate constant increasing with temperature according to an equation that is named after Arrhenius

$$k_r = A\exp(-E_a/RT) \tag{5.10}$$

where E_a is the activation energy and A is the frequency factor. The Arrhenius equation is commonly use in its linear form

$$\ln k_r = -\frac{E_a}{R}\frac{1}{T} + \ln A \tag{5.11}$$

Hence, when plotting the natural logarithm of the rate constant versus $1/T$ the slope of the line is $-E_a/R$ and the intercept is the natural logarithm of the frequency factor.

5.2 Objectives

1. Determine the reaction order with respect to CV
2. Determine the reaction order with respect to hydroxide
3. Determine the rate constant for the overall reaction

5.3 Procedure

Crystal violet, CV, and hydroxide react according to the following reaction:

$$\mathrm{CV}^+ + \mathrm{OH}^- \rightarrow \mathrm{CVOH} \tag{5.12}$$

Hence, the reaction rate can therefore be expressed in the form:

$$\nu = k_r[\mathrm{CV}^+]^a[\mathrm{OH}^-]^b \tag{5.13}$$

The isolation method is used by isolating the contribution of the crystal violet and then by evaluating the order of the reaction with respect to hydroxide separately. The rate of the reaction can therefore be written as:

$$\nu = k_r^\dagger[\mathrm{CV}^+]^a \tag{5.14}$$

where

$$k_r^\dagger = k_r[\mathrm{OH}^-]^b \tag{5.15}$$

or

$$\ln k_r^\dagger = \ln k_r + b\ln\,[\mathrm{OH}^-] \tag{5.16}$$

which shows that $\ln k_r^\dagger$ changes linearly with $\ln\,[\mathrm{OH}^-]$. Hence, by computing the pseudo rate constant at various concentrations of hydroxide we can determine the true rate constant (k_r) and the order with respect to hydroxide by doing a simple linear fit.

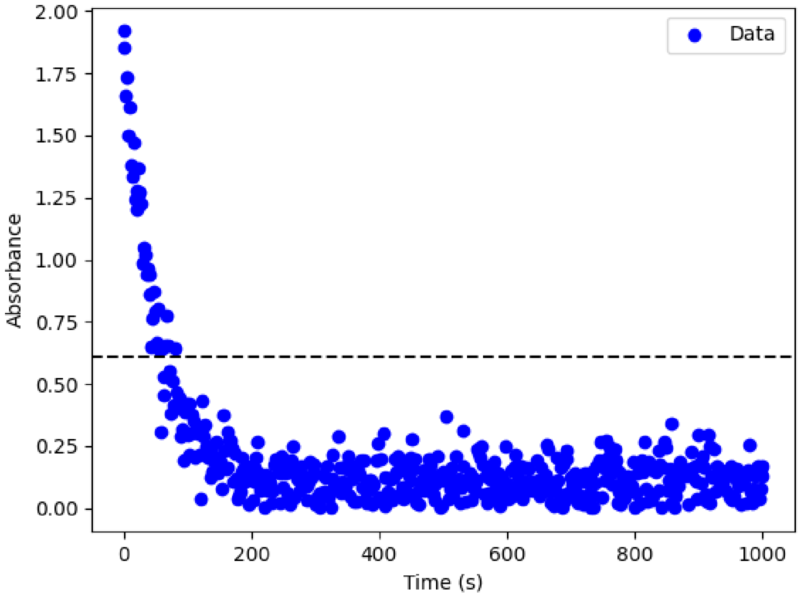


Figure 5.2: Sample data obtained from the CV virtual experiment. The dashed horizontal line, is an example of a threshold that can be applied to distinguish the data from the instrument noise

5.3.1 Calibration of the spectrometer

The data generated in this virtual laboratory represent the absorbance of the solution measured as a function of time, and they contain some noise. The absorbance can be used as a proxy for the concentration in all equations, since the two quantities are proportional to each other (Beer-Lambert law). Also note that, analogously to real experiments, if the concentration of the analyte drops below the instrument detection limit the data show only the instrument noise 5.2.

Therefore the first step in this laboratory is to determine the *limit of detection* (LOD) and the *limit of quantification* (LOQ) of the instrument. This can be done by running a series of "blank" experiments (with no CV solution) and extract some basic statistics about the instrument noise.

First prepare and analyse a few replicates (typically 2-3) of a blank solution containing no analyte (in this case no CV). Calculate both the mean (μ_{blank}) and standard deviation (σ_{blank}) of these blank measurements, as background noise typically has a non-zero average. The LOD is calculated as

$$\text{LOD} = \mu_{\text{blank}} + 3.3 \times \sigma_{\text{blank}} \tag{5.17}$$

where σ_{blank} is the standard deviation of the blank measurements. Similarly, the LOQ is calculated as

$$\text{LOQ} = \mu_{\text{blank}} + 10 \times \sigma_{\text{blank}} \tag{5.18}$$

Alternatively, the average background noise (μ_{blank}) can be subtracted from the signal obtained from each experiment. The horizontal dashed line in figure 5.2 corresponds to the limit of quantification (LOQ) of the experiment. $3.3 \times \sigma_{\text{blank}}$ and $10 \times \sigma_{\text{blank}}$ are conventional values, so it is important that you explain how you computed LOD and LOQ, and maintain it constant for all analysis.

5.3.2 Reaction order with respect to the isolated species

Take one data set (file) and determine whether the reaction is zeroth, first or second order with respect to the isolated species. This can be achieved by performing a linear fit of A, $\ln A$ and 1/A *vs* time and identify which one is more linear. Note that we have to fit only the part of the data set where the concentration is above the detection limit of the instrument and the signal is larger than the noise.

5.3.3 Reaction order with respect to the NON-isolated species

Compute the pseudo-rate constant for at least 5 experiments from solutions with different concentrations of the non isolated species, but at the same temperature. In order to do that we perform a linear fit of $\ln k_r^\dagger$ *vs* $\ln[B]$, where B is the concentration of the non isolated species and determine the order with respect to the NON-isolated species.

$$\ln k_r^\dagger = \ln k_r + b \ln [\mathrm{B}] \tag{5.19}$$

5.3.4 Rate constant

From procedure above we can also determine the true rate constant for the overall reaction. In fact, the intercept of the best fitting line of those data is $\ln k_r$.

5.4 Pre-lab questions

In the real experiment the samples will be prepared by mixing comparable amounts of liquid taken from a 2.5×10^{-5} M stock solution of CV and a 0.5 M stock solution of NaOH.

1. If we run two batches of experiments where we alternatively keep the amount of one of the two solutions constant, which species is isolated in the two sets?

2. Fill the table below by choosing appropriate volumes of the CV stock solution (Vx), of the OH stock solution (Vy) and of DI water (Vz) to produce 5 solution with that have constant concentration of Crystal Violet but different concentrations of hydroxide.

Vx (mL)	Vy (mL)	Vz (mL)	[CV+] (μM)	[OH$^-$] (M)
20	20	10	10	0.2
20			10	
30			10	
40			10	
10			10	

3. Fill the table below by choosing appropriate volumes of the CV stock solution (Vx), of the OH stock solution (Vy) and of DI water (Vz) to produce 5 solution with that have constant concentration of hydroxide but different concentrations of Crystal Violet.

Vx (mL)	Vy (mL)	Vz (mL)	[CV+] (μM)	[OH$^-$] (M)
20	20	10	10	0.2
	20			0.2
	30			0.2
	40			0.2
	10			0.2

5.5 Lab report structure and content

5.5.1 General considerations

The lab report must abide by the requirements to be uploaded into Turnitin for plagiarism checks, which include (but are not limited to)

- scanned images are NOT allowed
- photos of tables, figures or handwritten notes are NOT allowed
- inclusion of other printed documents is NOT allowed

If any of these items is detected a zero mark may be given to the report. Although it does not require more effort than including a figure made with excel or Python, the image of a graph made through screen capture can be included in the document. Tables should not be included as images.
In general, a lab report should

1. Discuss and show how every reported quantity has been calculated, either by reporting the equations used and/or using of figures. If the same procedure is used for multiple files, one or a few examples of the calculation procedure would suffice.
2. Report average quantities with their 95% confidence interval, while the individual values should be reported in a table or a figure. Quantities derived from fitting should also be reported with their 95% confidence interval.
3. the student-t test (or p-score) should be reported when comparing with literature values to show whether the results are consistent or not.
4. Each figure and table must have an explanatory caption (above or below).
5. Any questions listed for the laboratory should be answered.

Note that some of the items above may not apply to all reports.

5.5.2 Specific items to be included in the report for this lab

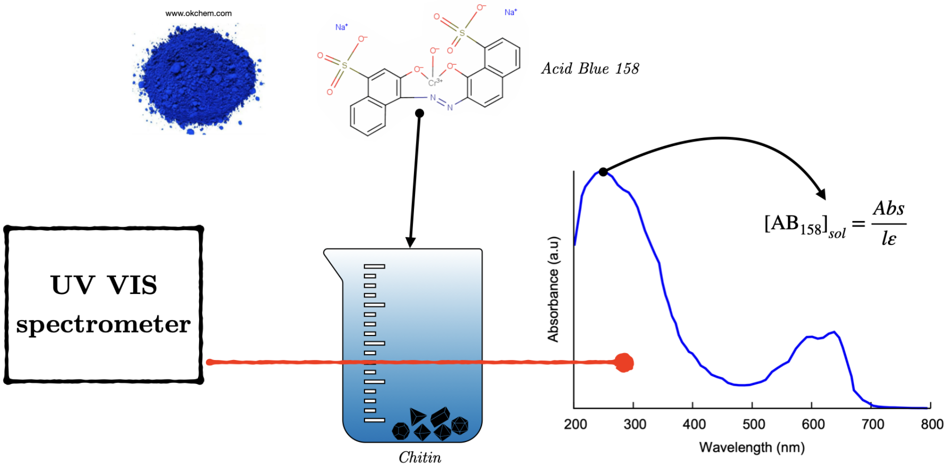
At a minimum, the crystal violet lab report should cover the following point

1. Show the full rate law for the reaction between Crystal Violet and hydroxide
    - what is the order of the reaction with respect to [CV$^+$] ?
    - what is the order of the reaction with respect to [OH$^-$] ?
    - what is the rate constant at 326 K

2. Report the activation energy for the reaction
3. Comparison between results for the rate law and activation energy for the reaction between crystal violet and hydroxide with the experimental values (search and reference literature value).
4. Comment on the differences between the plots of the absorbance *vs* time obtained at constant [CV] and constant [OH].
5. Does it matter if we use the rate constant or the pseudo-rate constant for the calculations of the activation energy? Justify your answer.

6. Surface adsorption

In the virtual laboratory below, we will be looking at the adsorption of the dye Acid Blue 158 on chitin in water. The simulated experiments mimic different conditions and will be used to determine the enthalpy of adsorption of the dye on the substrate. The molar mass of Acid Blue 158 is 584.91 g/mol.



https://doi.org/10.1080/02772248.2016.1178264

Figure 6.1: Caption

6.1 Background

The Langmuir isotherm is one of the simplest models that can be used to describe the adsorption of molecules on surfaces, either in the gas phase or in solutions. It is based on 5 key assumptions:

1. The surface is flat
2. The adsorbate is immobile on the surface
3. All adsorption sites are equivalent
4. There are no interactions between adsorbate molecules on adjacent sites
5. One one molecule can adsorb in a site (monolayer coverage)

The fundamental equation of the Langmuir adsorption isotherm can be derived using either thermodynamic or kinetic arguments. Here we will follow the thermodynamic route. The surface adsorption process can be regarded as an equilibrium problem, where the adsorbate molecules, A, on the surface are in dynamic equilibrium with those in solution

$$A + S \leftrightharpoons SA \tag{6.1}$$

where A are the free molecules in solution (or in the gas phase), S are the available adsorption sites and SA are the filled adsorption sites. The equilibrium constant for this chemical reaction is

$$K = \frac{[SA]}{[A][S]} \tag{6.2}$$

Although, the concentrations of free/occupied adsorption sites are somewhat ill-defined quantities, it is easy to see how their "concentration" would be related to the surface coverage. If we define the coverage, θ as the fraction of occupied surface sites,

$$[SA] \propto \theta \tag{6.3}$$

$$[S] \propto (1 - \theta) \tag{6.4}$$

$$[A] = c_{sol} \tag{6.5}$$

where we have introduced a slight change of notation by calling the c_{sol} the equilibrium concentration of the adsorbate in solution. We can then rewrite the equilibrium constant as

$$K_L = \frac{\theta}{(1-\theta)c_{sol}} \tag{6.6}$$

where K_L is the Langmuir constant, which contains all the unknown proportionality constant that connect the coverage with the "concentrations" that are in the definition of the equilibrium constant. This equation can then be rewritten to obtain the famous Langmuir isotherm equation

$$\theta = \frac{K_L c_{sol}}{1 + K_L c_{sol}} \tag{6.7}$$

where θ is the fraction of adsorption sites that are occupied, K_L is the Langmuir equilibrium constant, and c_{sol} is the equilibrium concentration of the adsorbate in solution. Because it is not possible to directly measure the fraction of occupied surface sites, a more practical version of that equation is

$$c_{surf} = \frac{Q K_L c_{sol}}{1 + K_L c_{sol}} \tag{6.8}$$

where c_{surf} is the concentration of the adsorbate that is on the surface, *i.e.* not in solution, and the new parameter Q corresponds to the *monolayer* coverage, *i.e.* the maximum concentration of molecules that can adsorb on the substrate. The linear form of the above equation, which uses the inverse of the concentrations, is more convenient for the fitting;

$$\frac{1}{c_{surf}} = \frac{1}{QK_L}\frac{1}{c_{sol}} + \frac{1}{Q} \tag{6.9}$$

The name *isotherm* stems from the fact that the experiments are performed at constant temperature and in principle both the Langmuir equilibrium constant, K_L, and the *monolayer* coverage, Q, can have a temperature dependence. Similarly to normal chemical reactions, by performing a series of experiments at different conditions it is possible to determine the enthalpy and entropy of the adsorption process using the van't Hoff equation.

$$K_L = e^{-\Delta G/RT} \tag{6.10}$$

$$\ln K_L = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \tag{6.11}$$

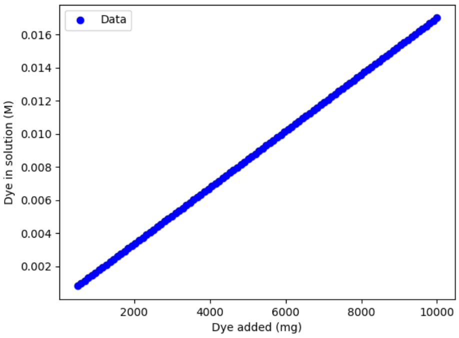


Figure 6.2: Example of the data obtained from this virtual lab.

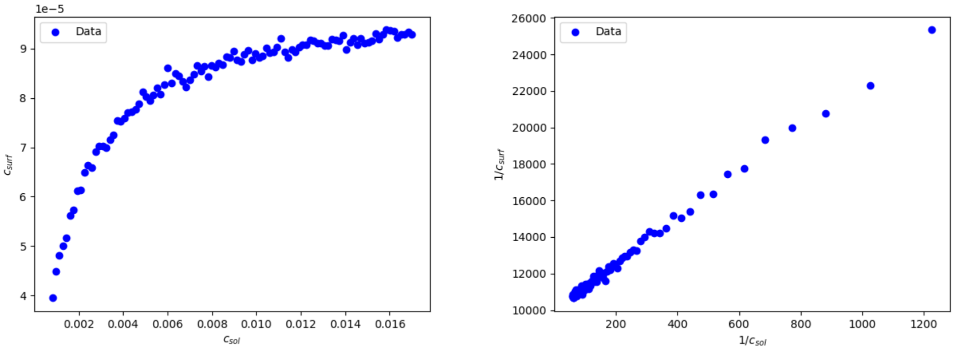


Figure 6.3: Example of the data needed for the post processing of this virtual lab.

6.2 Objectives

1. Calculation of the Langmuir constant (K_L) and the monolayer coverage (Q) at different temperatures
   a. Compare the fitted values obtained from fitting both forms of the Langmuir isotherm Eq. 6.8 and Eq. 6.9.
2. Calculation of the adsorption enthalpy
3. Comparison with the provided experimental value

6.3 Procedure

The experimental procedure consists of dosing increasing amounts of dye into the solutions containing the chitin powder. After each dose of the dye the system is let to equilibrate and the amount of dye left in solution is measured through a spectrophotometer. The data produced during the experiment consist of the amount of dye added to the solution and the measured dye concentration in solution.

6.3.1 Calculation of K_L and Q

After having manipulated the provided data to compute c_{surf} and c_{sol}, you can determine K_L and Q by fitting the appropriate quantities using Eq. 6.8 and Eq. 6.9. If you are using

Excel, you can get the uncertainties on the fitting parameters only from the linear fit. Hence, you would have to use the error propagation rules to obtain the uncertainties for K_L and Q from the uncertainties on the slope and y-axis intercept.

6.3.2 Adsorption enthalpy

Do a weighted fit of the K_L values calculated for the different temperatures using the Arrhenius equation. If you are using Excel, you need to use the linearised Arrhenius equation to be able to get the uncertainties for the fitted parameters.

6.3.3 Comparison with experiment

Do the Student's t test to compare your result with the experimental value of the adsorption enthalpy, $\Delta H = -19.5$ kJ/mol.

6.4 Pre-lab questions

1. Write the equation to compute c_{surf} from the experiment data (dye added in mg and c_{sol}).
2. Looking at the equations of the Langmuir isotherm (Eq. 6.8 and Eq. 6.9):
    (a) what are the dependent and independent variables?
    (b) what are the relationships between the slope and y-axis intercept and the ?
3. What are the equations to propagate the uncertainty on the slope and y-axis intercept to the uncertainties on K_L and Q?

6.5 Lab report structure and content

6.5.1 General considerations

The lab report must abide by the requirements to be uploaded into Turnitin for plagiarism checks, which include (but are not limited to)

- scanned images are NOT allowed
- photos of tables, figures or handwritten notes are NOT allowed
- inclusion of other printed documents is NOT allowed

If any of these items is detected a zero mark may be given to the report. Although it does not require more effort than including a figure made with excel or Python, the image of a graph made through screen capture can be included in the document. Tables should not be included as images.
In general, a lab report should

1. Discuss and show how every reported quantity has been calculated, either by reporting the equations used and/or using of figures. If the same procedure is used for multiple files, one or a few examples of the calculation procedure would suffice.

2. Report average quantities with their 95% confidence interval, while the individual values should be reported in a table or a figure. Quantities derived from fitting should also be reported with their 95% confidence interval.
3. the student-t test (or p-score) should be reported when comparing with literature values to show whether the results are consistent or not.
4. Each figure and table must have an explanatory caption (above or below).
5. Any questions listed for the laboratory should be answered.

Note that some of the items above may not apply to all reports.

6.5.2 Specific items to be included in the report for this lab

At a minimum, this lab report should contain the following

1. Figures showing the data with the fit function and the residuals, at least for one temperature.
2. Comparison of the fitted values obtained by the linear and non-linear fit.
3. A table with the compute values of K_L and Q at the different temperatures
4. Figure for the fit of the K_L values with the Arrhenius equation
5. Comparison with the provided experimental value

7. Appendix

7.1 Student's t distribution

Table 7.1: List of tabulated values to compute the confidence interval using the Student's t test distribution for sample sizes of $\nu = n - 1$ degrees of freedom and different confidence intervals. "One-sided" and "Two-sided" refer to the cases where only one side of the distribution is excluded or both sides are excluded. The "Two-sided" case is the most commonly used.

	Confidence Interval										
"One-sided"	75%	80%	85%	90%	95%	97.5%	99%	99.5%	99.75%	99.9%	99.95%
"Two-sided"	50%	60%	70%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%
ν =1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
$\nu = \infty$	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291
	Confidence Interval										
"One-sided"	75%	80%	85%	90%	95%	97.5%	99%	99.5%	99.75%	99.9%	99.95%
"Two-sided"	50%	60%	70%	80%	90%	95%	98%	99%	99.5%	99.8%	99.9%