Sheffield Hallam University

Department of Engineering and Mathematics

Session:	2018/19
Module:	55-402182 Introduction to Condensed Matter Physics
Module Leader:	Simon Clark
Assignment number/title:	Assignment 2
Academic contact for guidance:	Tom Ostler T.Ostler@shu.ac.uk
Maximum word count or number of pages:	N/A
Percentage contribution to overall module mark:	10%
Deadline for submission:	3pm Thurs 28$^{\text{th}}$ March 2019
Method and Location for Submission:	Online Submission
Deadline for return of feedback:	Thurs 19$^{\text{th}}$ April 2019

Module learning outcomes to be assessed:

- Conduct individually investigations and analysis related to applications of condensed matter physics.

References/recommended reading:

See Library Reading Lists Online.

Please ensure that all sources of information used are referenced. For guidance see http://libguides.shu.ac.uk/referencing

This is an individual assignment.

All assessments are subject to SHU's collusion and plaigiarism regulations. Please refer to: https://students.shu.ac.uk/shuspacecontent/assessment/plagiarism

Please submit your work working out and Python program as separate files via Blackboard.

Any hand written work should be scanned and submitted electronically via Blackboard.

In this assignment, you will be performing statistical analysis of the sizes of polycrystalline grains of Cu(In,Ga)Se_2, which is a material that is a good light absorber and is used in thin film photovoltaic devices [1]. In Ref. [1], a method called electron backscatter diffraction was used to determine the grain size distribution of three samples. Atomic force microscope images of those devices are shown in Fig. 1.

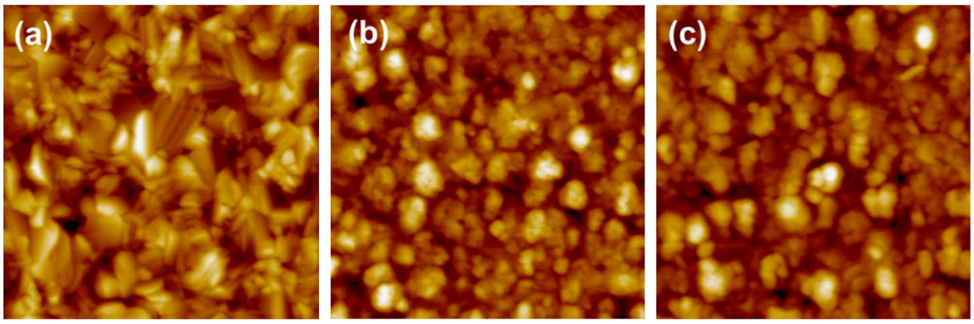

Figure 1: Atomic force microscope images of three think film Cu(In,Ga)Se_2 devices. The mean grain diameter of samples 1, 2 and 3 was $0.6\mu m$, $0.4\mu m$ and $0.4\mu m$, respectively.

The diameter of polycrystalline grains tend to follow a distribution known as a log-normal distribution. This is a type of skewed distribution as shown schematically below.

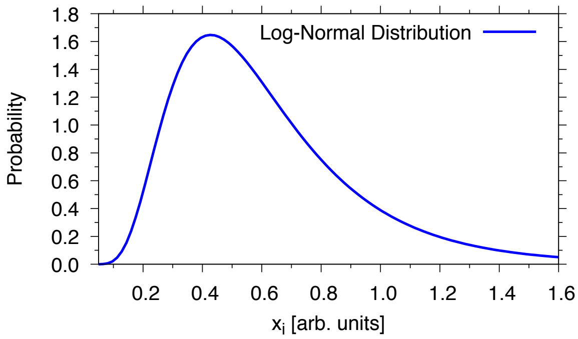

Figure 2: Schematic of a log-normal distribution. If the sizes of grains are given by x_i, then the probability of selecting one will follow a log-normal distribution.

On Blackboard, under the Assessment tab there is a file called *grain_data.txt*, containing the diameters (in micrometres) of a number of grains that you will need to analyse to complete this assignment. This assignment will build on what we have done in class, specifically:

- Loops
- Reading and writing files
- Arrays
- Use of numpy
- Plotting functions with matplotib

Question 1

In the first part of this assignment, you will be required to determine several statistical properties, the first of which is the mean of the diameters. The mean is given by:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

Where, x_i represents the size of each grain given by each line in the file and N is the number of entries in the file (the total number of data points).

a) Using equation 1 above, write a python code to determine the mean of the data. Your code must use a loop and write to a file, the number of data points in the file that you have read as well as the mean calculated by equation 1.

Note: you can check that your method determines the correct value using numpy. If you have an array of data (called dataarray here) you can calculate the mean by using an in-built numpy command:

```
numpy.mean(dataarray)
```

(15 marks)

b) Given your answer to part a), determine from the mean which device (1, 2 or 3) these data correspond to.

Note: If your answer to part a) is incorrect and you have not checked your answer using *numpy.mean*, you will not receive any error carried forward marks.

(5 marks)

c) Furthermore, use a loop to determine the smallest and largest value in the data set and hence the range and write them to a file. You must use a loop to determine these values.

Note: you can check that your method determines the correct values of the smallest and largest values using numpy. If you have an array of data (called dataarray here) you can calculate the maximum (largest) or minimum (smaller) by using an in-built numpy command:

```
numpy.amax(dataarray)
numpy.amin(dataarray)
```

(20 marks)

The standard deviation, σ, is given by the equation:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2}$$

d) Using equation 2 above, write a python code to determine the standard deviation of the data. Your code must use a loop and write to a file, the standard deviation calculated by equation 2.

Note: you can check that your method determines the correct value of the standard deviation using a numpy command. If you have an array of data (called dataarray here), you can calculate the standard deviation using the command:

```
numpy.std(dataarray)
```

(20 marks)

Question 2

a) Using matplotlib, create a fully labelled histogram plot of the data with 100 bins in pdf format. The displayed plot must be across a sensible range.

Note: to create a histogram plot (you will also need to add labels, legends etc) from an array of data (here called dataarray) use:

```
import matplotlib.pyplot as plt

# Some code here

plt.hist(dataarray,bins=100)
```

(35 marks)

b) What do you notice about the position of the peak in your graph and the mean calculated in question 1a)? Think about where the mean is in a normal distribution.

(5 marks)

References

[1] Rozeveld. S, Reinhardt. C, Bykov. E, and Wall. E. Measurement of Grain Boundary Properties in Cu(ln,Ga)Se_2 Thin Films. *Microscopy Today*, 26(3):32–39, 2018.