Contents

Chapter 1

Introduction

The main goal of this project is to improve EDS analysis. There are multiple ways to do this. One way is to make the analysis more transparent, which would make it easier to understand and use. A second option is to improve the input parameters of the analysis by control checking the instrument with a known sample. A third way is to make the quantitative analysis more accurate, which would improve EDS analysis. In this project, the main focus have been trying to improve the transparency of the analysis. Thus, the problem statement was formulated as:

Main problem statement. **How does different spectroscopy data processing influence the quantitative Cliff-Lorimer analysis in HyperSpy?**

Most of the time used in this project was spent on trying to understand the different methods of data processing in EDS analysis. Towards the end the different methods were applied on the same data set using the Cliff-Lorimer quantification in HyperSpy to see how the different methods influence the quantitative analysis. Solving the main problem statement with an open-source Jupyter Notebook would increase the transparency of the analysis and allow users to both adjust their analysis and understand better the different steps in the analysis. The main problem statement was broken down into five sub-problems. The sub-problems and a short description follow below.

Sub-problem 1. How accurate it the out-of-the-box quantification in AZtec and HyperSpy?

Sub-problem 2. What are done with the data at the different steps in the analysis when using HyperSpy?

Sub-problem 3. How are the peaks and the background be modelled?

Sub-problem 4. How is the spectrum calibrated, and is AZtec different than HyperSpy?

Sub-problem 5. When does the analysis fail, both in AZtec and HyperSpy?

(Question for Ton: Should I have a short paragraph here about the status of EDS analysis today?)

One of the main problems with the existing software for EDS analysis is that they are like black boxes. The manufacturer of EDS sensors, like EDAX, Hitatchi, Thermo Fisher Scientific, and Oxford Instruments **(Question for Ton:**

Cite this?), provide their own software for analyzing EDS data. These software packages are black boxes, because the code is hidden, and the users does not know what is happening inside the software. In other words, the user pushes some buttons to start the analysis of their sample, and then the software does some "magic" to analyze the data and produce some results. The user has few options to change the analysis to fit their needs. Many users tend to accept the results **(Question for Ton: Need to cite something here?)** from the software without questioning them, even though the analyzation "magic" differs between software packages and might be unreliable. The manufacturer Hitachi have trouble with separating Cs from ??, and their solution is to neglect the existence of Cs **(Brynjar: Find citation)**. The manufacturer Oxford Instruments provides the software AZtec which have trouble with ??, and their solution is ?? **(Brynjar: Find something, eg. zero peak)**. Even if the software is not wrong, the user might not understand the results, and the user have no way to change the analysis to fit their needs. Some might say that using e.g. AZtec is EDS analysis for dummies. One could solve this problem with an increase in transparency, because that would make it easier to understand EDS analysis and easier for users to adapt the analysis to their own needs.

(Brynjar: Paragraph about Dispersion, offset, energy resolution?)

(Brynjar: Paragraph about Other parameters of EDS analysis?)

(Brynjar: Paragraph about Improving quantitative EDS analysis?)

This remainder of this report is built up around the main problem statement and the sub-problems. The theory chapter contains the physics of X-rays and empirical adjustments in the analysis, a section on data processing, and a section about the hardware in an EDS setup. The methods chapter explains how the data was collected, while the arguments for and against the different methods are presented in the discussion chapter. The results chapter contains qualitative and quantitative results. The qualitative results are presented as figures of spectras showing the elements in the sample, and also how well different calibrations fit with the theoretical values. The quantitative results are presented as tables with compositional results with different methods and adjustments. The different methods are using AZtec and two approaches in HyperSpy. The different adjustments are results with different calibrations, different background models, **(Brynjar: "and different peak models"?)** The discussion chapter follow the structure of the sub-problems, and discuss both the methods and the results of the analysis consecutively. The conclusion chapter summarizes the report with an answer to the main problem statement, and provides ideas for further work. The appendix contains the code used in the analysis, which is also available on GitHub **(Brynjar: Link to GitHub)**.

Chapter 2

Theory

Energy dispersive X-ray spectroscopy (EDS) is a technique for analyzing the elemental composition of a sample with a spatial resolution, used in SEM and TEM. The technique is based on excitation of core shell electrons, which are bound to the atom with different strengths in different elements, and thus the electron relaxation results in very specific photon energies. EDS can be used to determine both the qualitative and quantitative composition of a sample. This chapter will cover the theoretical formation of characteristic X-rays, the empirical adjustments done due to creation and detection issues, explain quantitative calculations, cover the parameters of a quality control program, and briefly explain the basics of a SEM.

2.1 Theoretical view on characteristic X-rays

This section is primarily based on Hollas [2, Ch. 8.2] and Goldstein [1, Ch. 4.2]. It covers the theoretical physics behind creation of characteristic X-rays.

2.1.1 Formation of characteristic X-rays

The formation of characteristic X-rays is an inelastic quantum mechanical scattering process in two steps. In the following four equations the subscripts are referring to specific electrons in order to distinguish between them, which is also used in FIGURE XX **(Brynjar: make this figure)**. In the first step described in Equation (2.1) electron e_1^- from the incident electron beam eject electron e_2^- from the core orbital of atom A [2, Eq. (8.12)].

$$e_{1_\text{incident}}^- + A \rightarrow e_{1_\text{outgoing}}^- + A^+ + e_{2_\text{ejected}}^- \tag{2.1}$$

The incident electron from the beam loose energy to both breaking the binding energy of the core orbital and to the kinetic energy of the ejected electron. The energy is given by Equation (2.2). The user can control the incident electron energy E_{1_incident} by the acceleration voltage V_acc of the electron gun (and the current I_beam of the electron beam). The

energy of the characteristic X-ray is dependent on the binding energy of the core orbital, $E_{2_{\text{core shell, binding}}}$. In EDS $E_{1_{\text{outgoing}}}$ and $E_{2_{\text{kinetic}}}$ serve no purpose [1, Eq. (4.1)].

$$E_{1_{\text{incident}}} = E_{1_{\text{outgoing}}} + E_{2_{\text{core shell, binding}}} + E_{2_{\text{kinetic}}} \tag{2.2}$$

In the second step electron e_3^- from a higher energy orbital relaxes to the hole in the core orbital of atom A, and the difference in energy is emitted as a photon with a specific energy $h\nu$ called the characteristic X-ray [2, Eq. (8.12)].

$$e_{3_{\text{outer shell}}}^- \rightarrow e_{3_{\text{inner shell}}}^- + h\nu_{\text{X-ray}} \tag{2.3}$$

The energy of the characteristic X-ray is the difference in energy between the ionized orbital and the orbital filling the hole, shown in Equation (2.4). The equation specifies the energy of the X-ray as $h\nu$, but Section 2.2 explains why users of EDS just use the energy directly, usually measured in eV or keV [1, Eq. (4.2b)].

$$h\nu_{\text{X-ray}} = E_{2_{\text{core shell, binding}}} - E_{3_{\text{outer shell, binding}}} \tag{2.4}$$

In the second step in Equation (2.3) it is also a probability that the relaxation energy is used to eject and give kinetic energy to another electron from a higher energy orbital. This process results in two ejected electrons, both the ionized electron from the core orbital and a second ejected electron from a higher energy orbital. The second ejected electron is called an Auger electron. Auger electrons are used for surface studies, because they can penetrate around 2 nm solid material and thus does not escape from inside the sample. The X-rays are emitted in all directions and penetrate typically 4000 nm, and are the signal in EDS. The ratio between the characteristic X-ray photons and Auger electrons are known as the fluorescent (quantum) yield, ω.

$$\omega = \frac{\text{X-ray photons}}{\text{Auger electrons}} \tag{2.5}$$

The fluorescent yield is heavily dependent on the experimental setup and can be approximated, which is covered as one of the empirical factors in Section 2.2.

2.1.2   Naming convention

The transition lines are grouped and named semi-systematic, based on the orbital the vacancy is in, and the orbital the electron is relaxed from. The naming convention is semi-systematic because it is the original empirical system published in Nature by the Swedish physicist Siegbahn in 1916 [4], when they did not have the knowledge we have today. The International Union of Pure and Applied Chemistry made a more systematic naming convention for X-ray lines which is supposed to be the official one [1, Ch. 4.2.4]. However, the Siegbahn notation is used in the X-ray booklet, in HyperSpy, and by the TEM group at NTNU, and thus is used in this thesis.
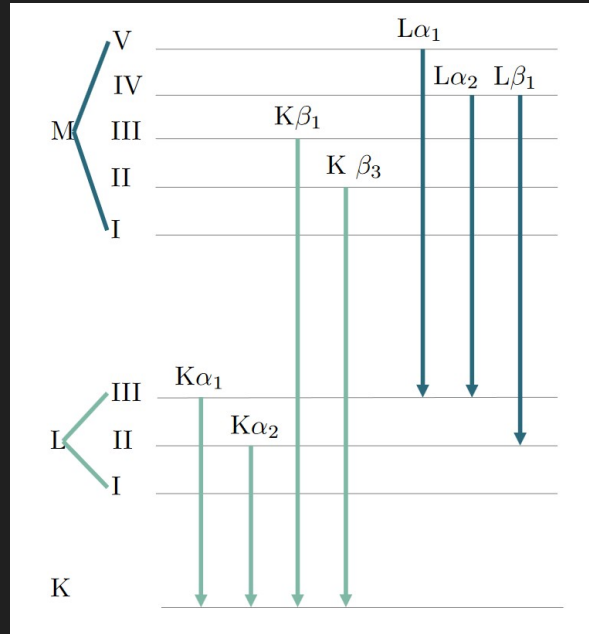
Figure 2.1: Direct copy if Maris Figure 2.9. I will make my own.

The X-rays are first named by which shell in the Bohr model the vacancy is in, i.e. the principal quantum number n of the vacancy orbital. Relaxations to the innermost shell $n = 1$ is named K-transitions, relaxations to $n = 2$ is L-transitions, relaxations to $n = 3$ is M-transitions.

The X-rays are further grouped with Greek letters in families. Orbitals close in energy are usually in the same group, which means that electrons in the same shell usually are in the same family. This naming is non-systematic, but tends to follow a pattern where the transitions labeled α are the lower energy transitions corresponding to the $n + 1$ orbitals, and the transitions labeled β are the higher energy transitions corresponding to the $n + 2$ orbitals. For example, $L \to K$ are α-transitions.

In addition, the lines in an X-ray group are labeled with subscript numbers which generally start with the highest intensity. This is a splitting of the lines due to different energy levels of the orbitals in the same shell. The different energy levels are due to the spin-orbit coupling, which is the interaction between the electron spin and the orbital angular momentum. The spin-orbit coupling increase with increased Z, which separates the lines more and more. The splitting of the α family to α_1 and α_2 are usually first resolvable in EDS for elements heavier than tin with $Z = 50$ [2, Ch. 8.2.2.3].

Putting these three naming conventions together, we name the transition $L_3 \to K_1$ as $K\alpha_1$, and $L_2 \to K_1$ as $K\alpha_2$, with more examples in Figure 2.1. The transition $L_1 \to K_1$ has $\Delta l = 0$ and is thus forbidden by the selection rules, see Equation (2.6). In gallium the $K\alpha_1 = 9251.74$ eV and $K\alpha_2 = 9224.82$ eV [6] are coupled, but as shown in Figure 2.5 this energy difference of $\Delta E = 26.92$ eV is too low to be resolved in EDS.

2.1.3 Energy and intensity

The energy of the characteristic X-ray depends on which orbital the vacancy is in, which orbital the electron is relaxed from, and the amount of protons in the core of atom A. Higher Z means higher energy of the characteristic X-ray line, because the energy difference between the relaxation orbital and the ionized orbital is larger. Atoms with higher Z have more possible transitions, because they have more electrons and orbitals.

The selection rules, which govern the allowed transitions for the formation of characteristic X-rays, are based on the Pauli exclusion principle and the spin-orbit coupling. See Figure 2.2 for illustration of the quantum number n and l which are relevant for the selection rules, while the quantum number l is dependent on the non-geomertically dependent spin and l. The selection rules in Equation (2.6) is for the electron which relaxes to the vacancy in the core orbital of atom A [2, Sec. 8.2.2.2].

$$\Delta n \geq 1; \qquad \Delta l = \pm 1; \qquad \Delta j = 0, \pm 1 \qquad (2.6)$$

Even though heavy atoms like gold have more than 30 possible transitions, only a few are detectable in EDS. Lines which are undetectable have low abundance, are too close to other lines, or are forbidden by the selection rules [1, Ch. 4.2.3]. Lines which are detectable have an intensity dependent on the amount of the element in the sample, because a higher amount gives more counts in the detector. The counts in EDS the number of X-ray photons detected in a specific energy range. The energy range is typically around 10 eV. In theory, the ratio of the atomic concentration between two elements are proportional to the ratio of the corresponding lines from the elements. However, there are many factors which affect the intensity of the lines, which are covered in Section 2.2.

2.2 Empirical view on characteristic X-rays

On top of the theoretical physics there are many experimental processes affecting both the creation and the detection of characteristic X-rays. The scientific community uses an empirically influenced approach in EDS analysis. This approach includes empirical equations to deal with the imperfect beam, scattering in the chamber, secondary scattering (?), imperfect detectors, …

(Question for Ton: What do I do with the thin film assumptions? How do I mention it?)

2.2.1 From lines to peaks

A striking difference between the theoretical and the empirical view is that the lines are not lines, but peaks with a varying width. This is illustrated both in Figure 2.4 and in Figure 2.5, where we see that the two K_α lines of Ga is merged into one broad peak. The peaks in the EDS spectrums have a Gaussian shape, which is a result of the broadening of the lines. The broadening of the lines is due to …**(Question for Ton: What is the cause of the broadening? And why is it Gaussian?)**. Figure 2.3 shows that lower energy lines are narrower than higher energy lines. This increase in
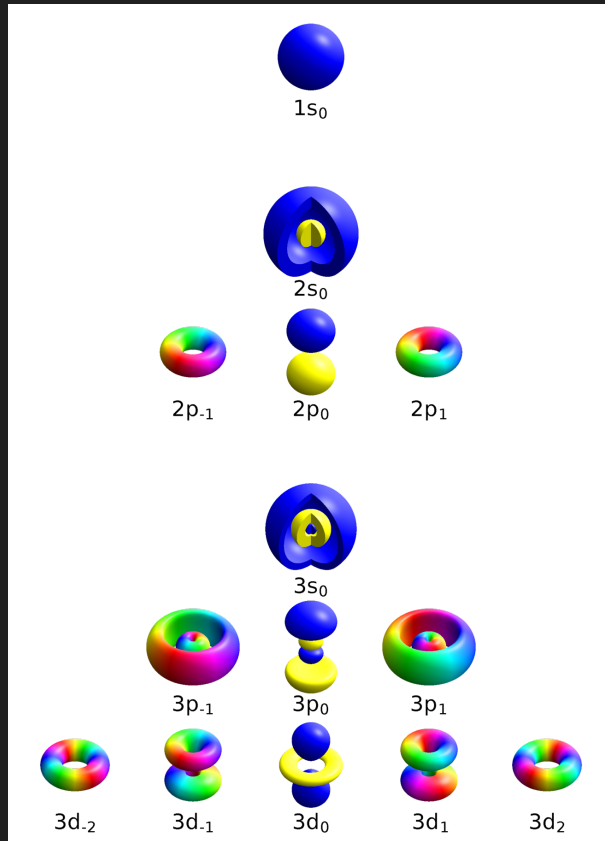
Figure 2.2: The quantum numbers n, l, and m in a hydrogen-like atom. The principal quantum number is shown as the block with values $n = 1, 2, 3$. The Azimuthal quantum number is the rows as $l = s, p, d$. The magnetic quantum number is the columns as $m = -2, -1, 0, 1, 2$. The spin quantum number, s, is not geometrically dependent and thus not shown. The total angular momentum quantum, l number is the sum of l and s. The figure is copied from the article "Quantum number" on Wikipedia, made by Geek3 - Own work, Created with hydrogen 1.1, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=67681892.
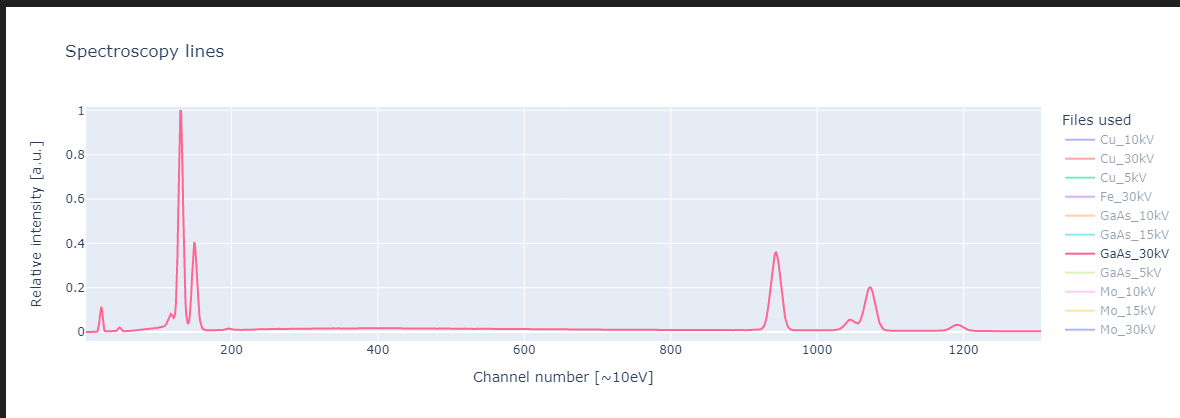
Figure 2.3: The whole GaAs spectrum. The next two figures zoom in on different parts of this plot. Data from the SEM Apreo at NTNU NanoLab, 30 kV.
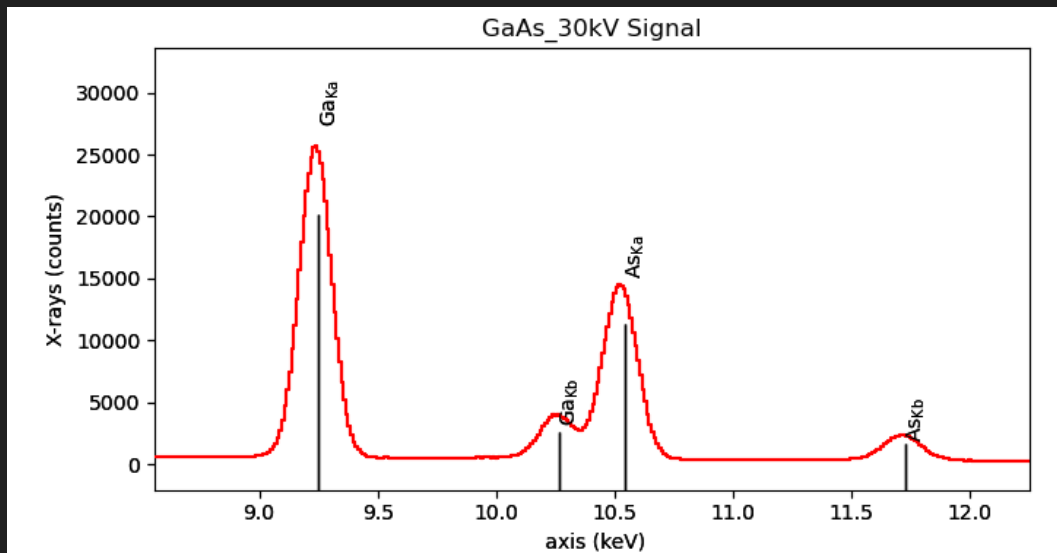
Figure 2.4: Plot produced by HyperSpy of the X-ray lines in GaAs from 8 to 12.5 keV. The theoretical lines are marked with black lines, but in reality the lines become peaks with Gaussian shapes. The height of the black lines are empirically estimated and are available in HyperSpy. Notice also how the center of the peak and the black line are slightly right shifted, which can be solved by calibrating the energy scale, see Section 2.3. Data from the SEM Apreo at NTNU NanoLab, 30kV.
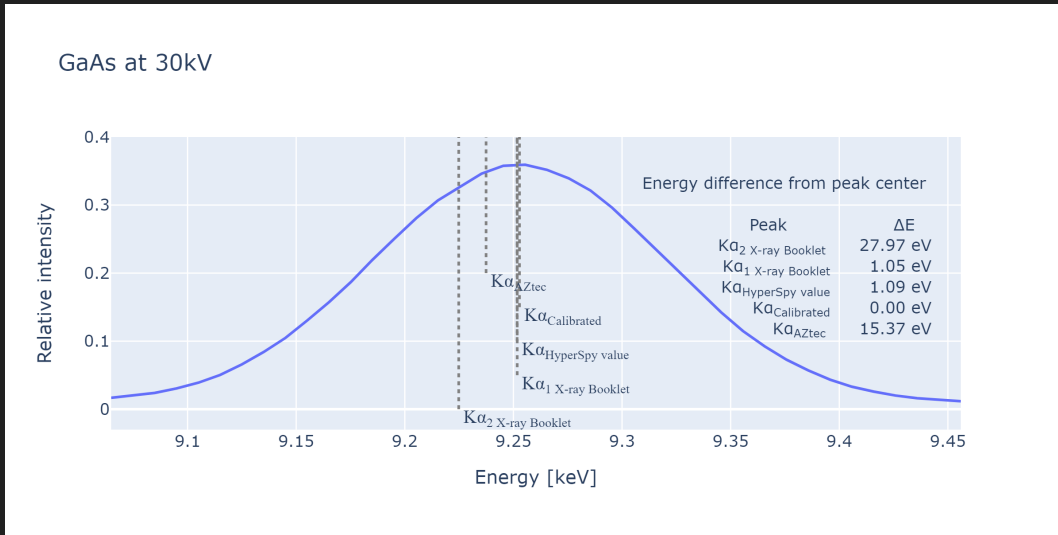
Figure 2.5: The theoretical lines $K\alpha_1$ = 9.25174 and $K\alpha_2$ = 9.22482 from the X-ray booklet with $K\alpha$ = 9.251 from HyperSpy and the fitted peak $K\alpha$. The spectrum is calibrated to Ga $L\alpha$ and As $K\alpha$, where the energy of the peaks is taken from HyperSpy. HyperSpy gives Ga $K\alpha$ almost directly on top of the theoretical $K\alpha_1$. This figure show that the lines which are separated with small energy differences are so close that they make one single peak in the EDS spectrum. Data from the SEM Apreo at NTNU NanoLab, 30kV. **(Question for Ton: This is more results than theory I would say. But I do remember you saying something about having this in the theory section.)**

Table 2.1: Ga lines from HyperSpy cite ?? and X-ray booklet [6]. ** HyperSpy operates with a single line for each K and L shell, but the X-ray booklet lists two lines for each shell. * HyperSpy lists Ln, Ll and Lb_3 as separate lines, but I am not sure what they correspond to in the X-ray booklet. TODO: Figure out this, eventually just dropping the additional lines.

	X-ray booklet Energy [keV]	HyperSpy Energy [keV]	HyperSpy Weight, unitless
$K\alpha_1$	9.25174	9.2517	1.0
$K\alpha_2$**	9.22482		
$K\beta_1$	10.2642	10.2642	0.1287
$L\alpha_1$	1.09792	1.098	1.0
$L\alpha_2$**	1.09792		
$L\beta_1$	1.1248	1.1249	0.16704
Ln*		0.9843	0.02509
Ll*		0.0544	0.9573
Lb_3*		0.0461	1.1948

broadening with higher energy is due to …**(Question for Ton: What is the cause of the broadening with higher E again?)**

2.2.2 Intensity

The intensity, or weight, of a line is dependent on multiple empirical and physical factors. This subsection will briefly cover fluorescent yield, critical ionization energy, and empirical weights in HyperSpy. In practice, the weight of the lines are included as a part of the k-factors or k-ratios, which are presented in Section 2.2.3. Only the strongest lines are listed in the X-ray booklet, where the theoretical values of the characteristic X-ray lines are listed. The list include: $K\alpha_1$, $K\alpha_2$, $K\beta_2$, $L\alpha_1$, $L\alpha_2$, $L\beta_1$, $L\beta_2$, $L\gamma_1$, $M\alpha_1$ [6].

(Brynjar: This info about fluorescent yield is interesting, but will I actually use it?) The fluorescent yield ω is non-linearly dependent on Z. Figure (4.3) in Goldstein [1] **(Question for Ton: Reproduce the figure? Would take like 2h to use the Crawford data properly I guess)** shows the fluorescent yield for the first 90 elements (which is based on Crawford 2011). For K- and L-shell fluorescent yield, ω is strictly increasing. For M-shell fluorescent yield, ω is strictly increasing till around $Z = 80$, and then it starts to decrease. The figure also show that for the same element $\omega_K > \omega_L > \omega_M$ **(Brynjar: But L-peaks are higher in the spectra. Explain this in the discussion?)**. When dealing with thin samples, the fluorescent yield can be approximated by an empirical formula based on tin. The formula is given in Equation (2.7), where $a = 10^6$ for K-shell. **(Brynjar: What is it for L and M? Also, is it really relevant here?)** The two key takeaways from this formula is that ω is dependent on Z, but kinda similar for close elements like Ga $Z = 31$ and As $Z = 33$.

(Brynjar: Find the source for this formula. Ton said it is from Williams and Carter. Is it valid for bulk? The K-shell have the same shape as Figure 4.3 in Goldstein.)

(Question for Ton: I'm not sure if I will actually use this formula, should I just remove it? And I don't get how this formula just becomes a part of the k-factor. Except from the key takeaways mention above.)

$$\omega = \frac{Z^4}{a + Z^4} \tag{2.7}$$

The critical ionization energy, E_C, is the energy the incident beam need to ionize a core electron in an atom. If the energy of the incident beam is lower than E_C, the core electron is not ionized and no peak can be detected. The critical ionization energy is dependent on the atomic number, Z, of the atom. Higher Z means higher E_C, because the core electrons are bound stronger to the nucleus. When the incident beam has an energy higher than E_C and continues to increase, the amount of ionization is not constant and not linear. The amount of ionization with varying energy above E_C is dependent on the ionization cross section and the overvoltage, which again are dependent on what shell the ionization happens in. The solution to this is empirically estimations and using k-factors where all factors like this is either cancelled out or included in the k-factor correction.

(Question for Ton: Is this good enough? If not, what do I write about the ionization cross section and

overvoltage? What Mari wrote is for thin samples, I think. Why include it if I'm not going to use any equations about the ionization cross section or overvoltage? You said something about just write that intensity it is lower for low energy because of many reasons (absorption, efficiency), with the same reason that background is lower for low energy. But I kinda need to explain some of the reasons to use them in the discussion. Another question: in the k-factor all these other factors fall out, so EDS-people does not use these equations. Am I wrong?)

HyperSpy have an integrated list of the characteristic X-rays, with both the energy and the weight of the lines. **(Brynjar: I guess the weights are empirically estimated, but I have not found any information about how they are estimated. In addition: I'm not sure if HyperSpy adjust the line height to the spectrum, or if it knows that Ga Ka are higher than As Ka as in Figure 2.4.)** Goldstein uses different intensity weights for isolated atoms, thin foils and bulk samples [1, Ch. 4.2.6], and all are dependent on the atomic number and ionized shell. One could venture down the rabbit hole of finding the theoretical weights for different lines, but that will not be done in this thesis. Examples of the weights from HyperSpy are given in Table 2.1. **(Question for Ton: OK?)**

2.2.3 K-factors and k-ratios

About em k-factors and k-ratios.

2.2.4 Detection system

Detector efficiency, detector resolution, dead time, angle/placement of detector, beam issues, stray: secondary excitations in the sample, Si stray, holder / chamber stray detection.

2.2.5 Sample thickness

Thin vs bulk samples.

2.2.6 Bremsstrahlung - the background radiation

Why linear and why sixt order polynomial. Which are better. **(Brynjar: I might use background removal as some results, but I'm not sure. That would be comparing the different methods, and looking at smoothing of the data before removing the background.)**

2.3 Calibration of the spectrum

E_1 and E_2 is the energy of the two characteristic X-rays, and c_1 and c_2 is the channel number of the two characteristic X-rays.

$$\text{Dispersion} = \frac{E_2 - E_1}{C_2 - C_1} \tag{2.8}$$

Get as little extrapolation as possible by selecting the longest possible distance between the two characteristic X-rays, while still using peaks with good signal to noise ratio. **(Brynjar: Quantify signal to noise ratio?)** Assume calibration on one spectrum and use it on all spectra for the same instrument.

Zero-offset in channels is:

$$\text{Zero-offset in channels} = C_1 - E_1 \cdot \text{Dispersion} \tag{2.9}$$

2.4 SEM

- how (e-beam, vacuum, ...) - hardware (scanning coils, astigmatism) - imaging (contrast, SE, BSE)

Chapter 3

# Method / experimental design

3.1 Introduction

The data needed in this project was SEM EDS spectra. The spectra were acquired on a SEM APREO at NTNU NanoLab. The spectra were from a sample containing different sections with different elements to get different spectra to work with. Thanks to Mari Skomedal [5] and Martin Lundeby [3] who made manuals for easily extracting data from AZtec. The relevant material for data extraction is available on the GitHub repository for this project (ref). The data is also available on the GitHub repository for this project, under the folder `data/2022-09-06_EDS-Apreo`. This method chapter describes the material used, the microscope and detector, the analysis in AZtec and HyperSpy, and the post data treatment.

3.2 Materials

EDS data was collected on one sample with different sections containing different materials. A half 2″ Si wafer was mounted with Cu tape on an Al FIB stub. On the Si wafer a smaller piece of a GaAs wafer was mounted with Cu tape. A Mo disk was mounted to the Si wafer with Ag paint. In the middle of the Mo disk there was a hole where a TEM grid was mounted. The TEM grid had GaAs nanowires on it, and the grid was made of Mo with C film. Growth of the GaAs nanowires is described in [5, Ch. 3]. A schematic of the sample is shown in Figure 3.1.

3.3 The microscope and the detector

The data in this project was collected with the SEM Apreo from FEI with an Oxford EDX detector at NTNU NanoLab. The detector is an EDX Oxford Xmax 80 mm2 Solid angle detector, with reported energy resolution of 127 eV (Cite: ntnu.norfab.no on the instrument). The acceleration voltage, V_{acc}, can be set from 0.2 to 30 kV. The beam current,
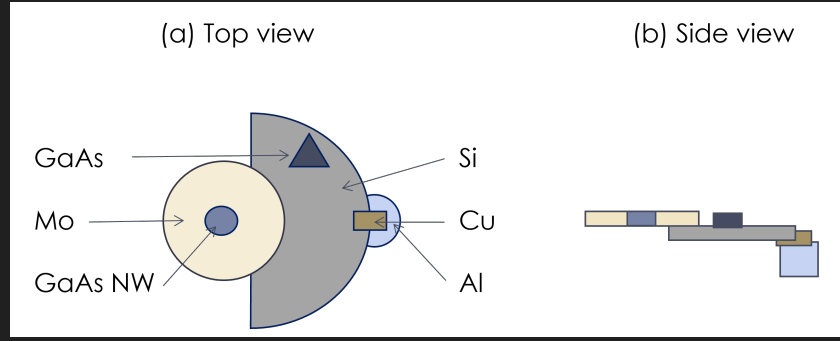
Figure 3.1: The sample used for the EDS data collection in the SEM Apreo. (a) is top view and (b) is side view. GaAs is a piece of a GaAs wafer. Mo is a Mo disk. GaAs NW is the TEM grid of Mo with C film and GaAs nanowires. Si is the Si wafer. Cu is the Cu tape. Al is the Al FIB stub. Four spectra with different V_{acc} was collected on the GaAs, Mo, Si, and GaAs NW. Three spectra was taken on the Cu tape and one spectrum on the Al FIB stub.

I_{beam}, has a maximum value of 400 nA. The SEM has an adjustable working distance. The SEM is equipped with a BSE and a SE detector.

The data collection was done on 5, 10, 15, and 30 kV. The beam current was 0.2, 0.4, 0.8, or 1.6 nA, depending on the dead time of the detector, trying to get the dead time around 30%. Dead time at around 30% was recommended by the supervisor of this project, Antonius T. J. von Helvoort, and is well below the problematic dead time of >60%. The dead time on the nano wire area was closer to 20%, which is what Goldstein [1, page 223] recommends. All samples were collected on 2048 channels ranging from 0 to 20 keV. Working distance was 10 mm. Processing time was set to 5 and each spectrum was collected with 2 minutes time live. Some spectra were stopped early due to high dead time and thus very long sampling. At leas one SE picture was taken for each sampling area. Counts per second in and out was noted down, but not used in this project. See table 3.1 for the settings used for the different spectra.

3.4   Analysis in AZtec and data extraction

The results from AZtec was taken as described in [3, Appendix A]. Both qualitative and quantitative results were acquired. The quantitative results were acquired after selecting the elements in the sample and noted as atomic percent. The k-factors calculated theoretically by AZtec was noted down, as they are needed for the quantitative analysis in HyperSpy. **(Brynjar: Uncertainties? Also get this data properly from AZtec.)**

3.5   Analysis in HyperSpy

The data was analyzed in HyperSpy, both qualitative and quantitative. The qualitative analysis was done by loading the spectra and plotting with lines marking the theoretical peak center and empirical weight. The quantitative analysis was done in different ways, and just on the GaAs bulk spectra. All the quantitative analysis was done with the k-factors

from AZtec. The first way was to use the area under the peaks in the raw spectra, where the background was removed linearly. The second way was to make a fitted model of the background and the peaks as Gaussian curves, and use the area under the peaks in the model. Then the x-axis spectrum was calibrated, and the two ways of quantification was used again. The different calibrations used was a model fitted calibration from HyperSpy, a self made calibration and the calibration from AZtec. The calibration values from AZtec is reffered to as uncalibrated. **(Brynjar: Uncertainties from HyperSpy?)**

3.6   Data treatment

In this project the data was also analyzed with new code, which is available on the GitHub profile of the author, `https://github.com/brynjarmorka/`. The end goal of this code is to improve understanding of the process. The code is written in Python and uses Jupyter notebooks. HyperSpy 1.7.1 was used for the analysis. NumPy 1.22.4 is used for calculations, SciPy 1.9.0 for fitting and peak finding, and Plotly 5.10.0 for plotting **(Question for Ton: do I reference these packages?)**. The repository "eds-analysis"[1] contains the code developed throughout the semester, and the repository "eds-analysis-final"[2] contains the final code. The final code is the code which is intended to enhance a users understanding of the analysis steps. **(Brynjar: Make the "eds-analysis-final" repository public.)**

3.6.1   Normalization

Since the total amount of counts in a spectrum differs with V_{acc}, I_{beam}, DT, etc., the spectra had to be normalized to be able to compare them. Initially the spectra were normalized to the highest peak in each spectrum, $\text{Intensity}_{\text{relative to max}} = \text{Counts}_{\text{raw}}/\text{Counts}_{\text{max}}$. Later other normalization methods were explored, but since the quantitative analysis is based on the area under the peaks, it does not matter how the spectra are normalized. However, for the SciPy function `curve_fit` to work, the spectra had to be normalized and the raw channel numbers used as x-axis. This was probably due to the fact that the function `curve_fit` is not designed for spectra with a large dynamic range. Whatever normalization method is used, the quantitative results would be the same because of the use of peak ratios. **(Question for Ton: This is getting argumentative, but it is the base for the decision and not really anything with the end results. It is just to reproduce. OK?)**

3.6.2   Gaussian fitting and peak finding

The fitting of a model to the spectra was done progessively more and more advanced. All the peaks was fitted as Gaussian curves, and the fitting itself was done with the SciPy function `curve_fit`. The first model fitting was just making a Gaussian at two specified peaks. The second model was fitting a Gaussian at all the peaks in the spectrum, but still with a user input of the peak positions. The third model was fitting a Gaussian at all the peaks in the spectrum, but with the

---

[1] `https://github.com/brynjarmorka/eds-analysis/`
[2] `https://github.com/brynjarmorka/eds-analysis-final`

peak positions found by the SciPy function `find_peaks`. The problem with these models was that one of the peaks was usually moved to compensate for the background. Thus the forth model was made, where the background was fitted as a sixth order polynomial. The background was fitted after removing the peaks, and then the peaks were fitted on top of the background. The fifth and final model was fitting the peaks and background in one go. The background was fitted as a n-th order polynomial, and different orders were tested.

3.6.3 Calibration

A third repository, "spectroscopy-channel-calibration"[3], was made specifically for calibration of spectra, which was used in the course "TFY4255 - Materials Physics" at NTNU, October 2022. The model used in this calibration is the first model from 3.6.2, and thus is not as advanced as the model used in the final code. The calibration is done with a spectrum of known elements, where the user inputs the energy of the peaks. The user further specifies the channel value of two peaks. The energy of the peaks are available through HyperSpy, or can be set manually from e.g. the X-ray booklet. The code makes a Gaussian fit to the two peaks, to find the true peak center. A plot of the spectrum and the fit are shown, and the user can decide if the fit is good enough. The code then calculates the dispersion, and the zero-offset. In the end the code plots the spectrum with the calibration, and the user can decide if the calibration is good enough. The same calibration principle is implemented in "eds-analysis-final", but without the plots and the user interaction. **(Brynjar: Implement the sentence above.)** The final calibration is using the fifth model with both background and peaks in one fit. However, since the calibration only needs the distance between two peaks, using the first model could be sufficient.

3.6.4 Area under the peaks

A function finding the area under the peaks was made as a simple first step in a quantitative analysis. This simple analysis was used to quantify if different calibrations gave different results. **(Brynjar: Do this.)** Implementing the Cliff-Lorimer method for quantification was not done, but could be done in the future. The author started to look at implementation of a factorless quantification method, but did not have time to dive deep enough into that in this project.

---

[3] `https://github.com/brynjarmorka/spectroscopy-channel-calibration`

Table 3.1: Table giving the detector settings for the different spectra. The other detector settings are the same for all the measurements, and they were the following. Working distance was 10 mm. Range was 0 to 20 keV, with 2048 channels. Processing time was set to 5 and each spectrum was collected with 2 minutes time live. Since the dead time on the GaAs bulk sample were high, the sampling was manually cut short.

| Sample	V_{acc} [kV]	I_{beam} [nA]	Dead time [%]
GaAs	5	1.6	33
GaAs	10	1.6	65
GaAs	15	0.8	58
GaAs	30	0.2	30
NW	5	1.6	5
NW	10	1.6	18
NW	15	1.6	18
NW	30	1.6	12
Mo	5	1.6	30
Mo	10	1.6	66
Mo	15	0.4	42
Mo	30	0.2	35
Si	5	1.6	27
Si	10	1.6	65
Si	15	0.2	25
Si	30	0.2	40
Cu	5	0.2	10
Cu	10	0.2	30
Cu	30	0.4	15
Al	30	0.1	22

Chapter 4

Results

4.1 Introduction

The results are presented in this chapter. First qualitative then quantitative results are presented. All the spectra taken were qualitatively analyzed. Only the GaAs bulk spectra was quantitatively analyzed. The spectrum from the GaAs bulk wafer taken on 30 kV is shown in Figure 4.1. This plot was made with HyperSpy, which utilize Matplotlib for plotting. The plotting method in HyperSpy can add where the theoretical peak centers are. The lines added also show an estimate of the weight of the peak.

4.2 Qualitative results

Figure 4.2 to Figure 4.7 shows the spectra for the six different areas of the sample plotted with PLotly. The plots are available as an interactive HTML plots on the GitHub repository. **(Brynjar: Upload the HTML files to the GitHub repository.)** The calibration used in these spectra is based on the calibration of Ga Lα and As Kα from the GaAs bulk wafer. The y-axis is normalized to the highest peak value in each spectrum.

4.2.1 Peaks and background

All the spectra have peaks with high peak-to-background ratio. The zero peak from the Oxford detector is visible in all the spectra as the first peak, with a center at 0.00924 keV. The GaAs, NW and Mo spectra show clearly that the lines broadens with higher E. All the highest peaks are below 5 keV. When doing the qualitative analysis, it became clear that the FIB stub was not made of Fe as expected, but rather of Al with a peak at 1.48 keV. The FIB stub have a small peak at 1.25 keV, which could be from Mg Kα at 1.253 keV. The Al FIB stub spectrum also have a peak at 5.90 keV, which could be from Mn Kα at 5.899 keV. Another discovery was that the Cu-tape does not give a good Cu signal. The high peak in the Cu-tape spectra at 0.260 keV are from C and not the Cu Lα peak. Only the Cu-tape taken at 30 kV has a Cu peak,
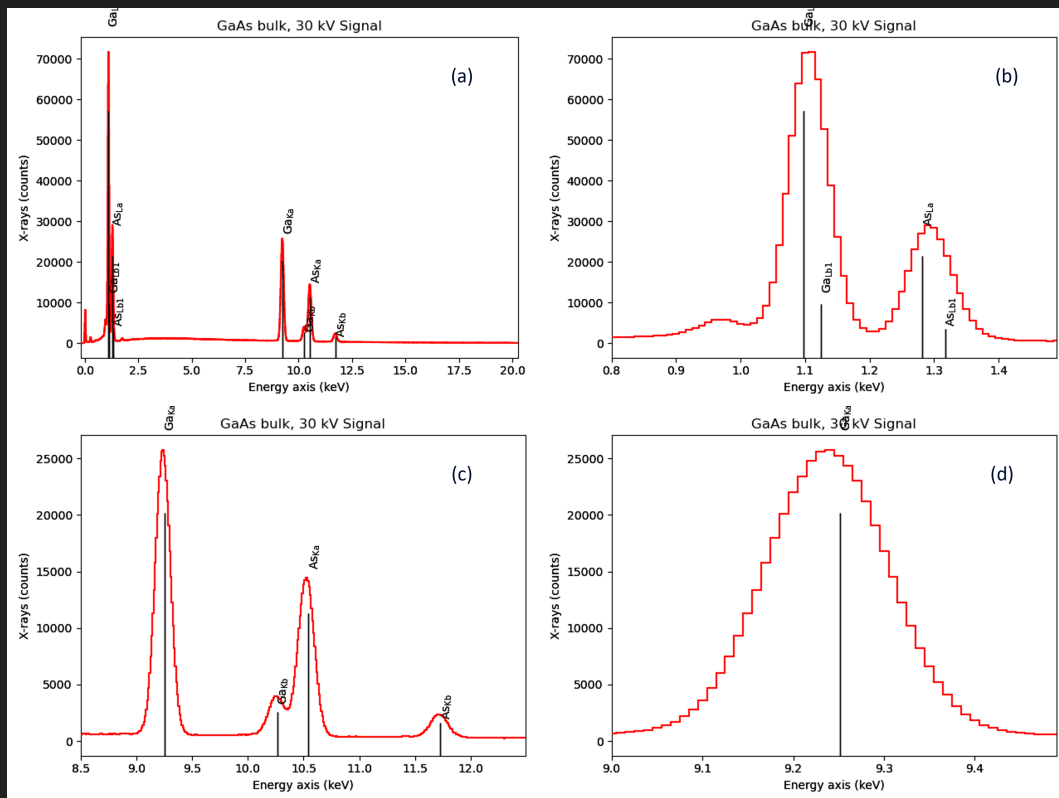
Figure 4.1: The GaAs spectrum taken at 30 kV. This plot was made with HyperSpy, which use Matplotlib. The theoretical peak centers are added as lines. (a) is the whole spectrum. (b) is the zoomed in on the L-peaks. (c) is zoomed in on the K-peaks. (d) is zoomed in on the Ga Kα peak. This plot has the calibration from AZtec, and it is clear that the line position is deviating from the center of the peaks.
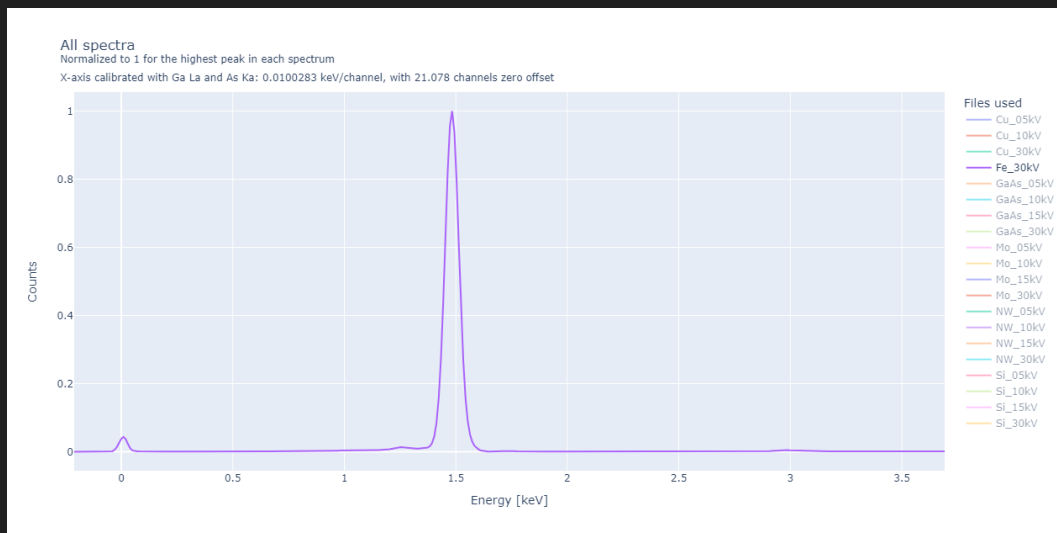
Figure 4.2: The spectra of the Al sample part. This was expected to be Fe, thus the label is wrong. The peak at 1.48 keV is the Al Kα peak.
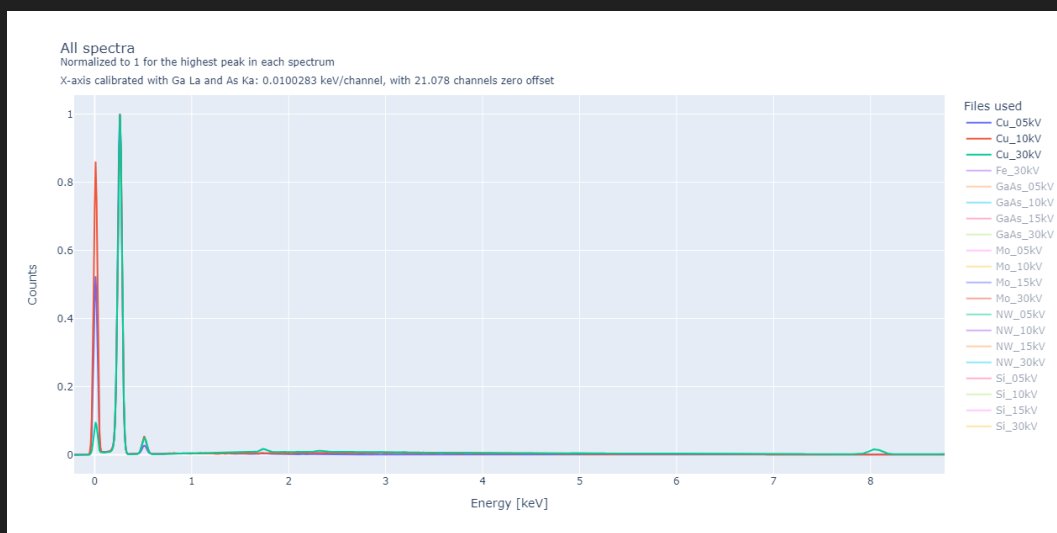
Figure 4.3: The spectra of the Cu sample part. The Cu sample was Cu-tape from the lab, but the Cu Kβ peak is only barely visible at the 30 kV spectrum. The highest peak in all three spectra is at 0.260 keV, which is the C Kα peak, slightly off from the expected 0.277 keV.
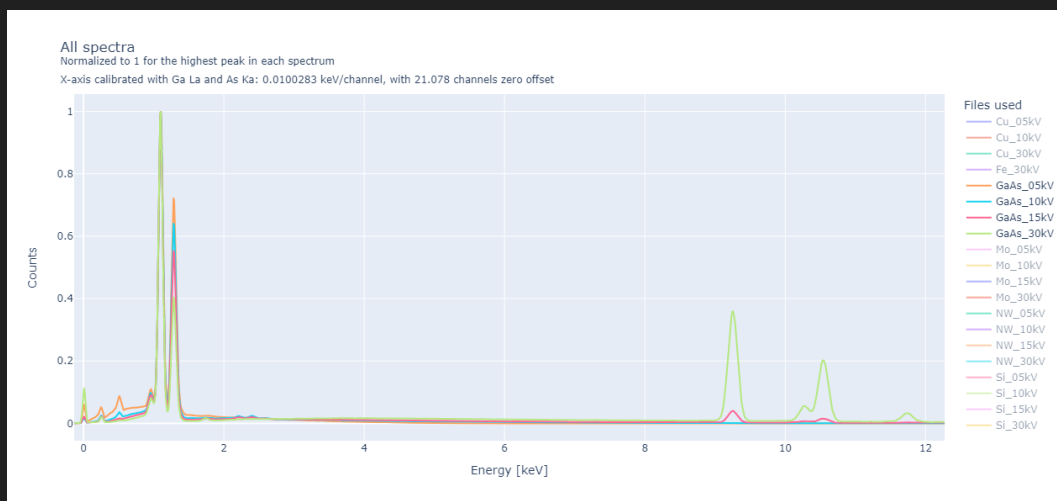
Figure 4.4: The spectra of the GaAs sample part. Both the K-peaks and the L-peaks of Ga and As are visible. There is a peak at 0.511 keV, which is the O Kα peak. There is a peak at 0.260 keV, which is the C Kα peak.
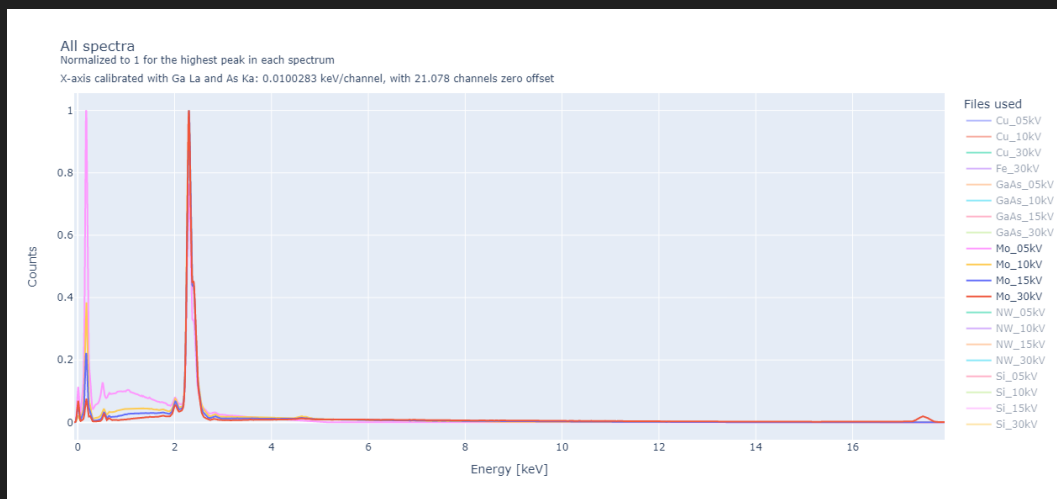
Figure 4.5: The spectra of the Mo sample part. The Mo Kα peak at 17.47 keV is barely visible at the 30 kV spectrum, and has a high noise level. The high dobble peak is Mo Lα at 2.29 keV and Mo Lβ at 2.39 keV.
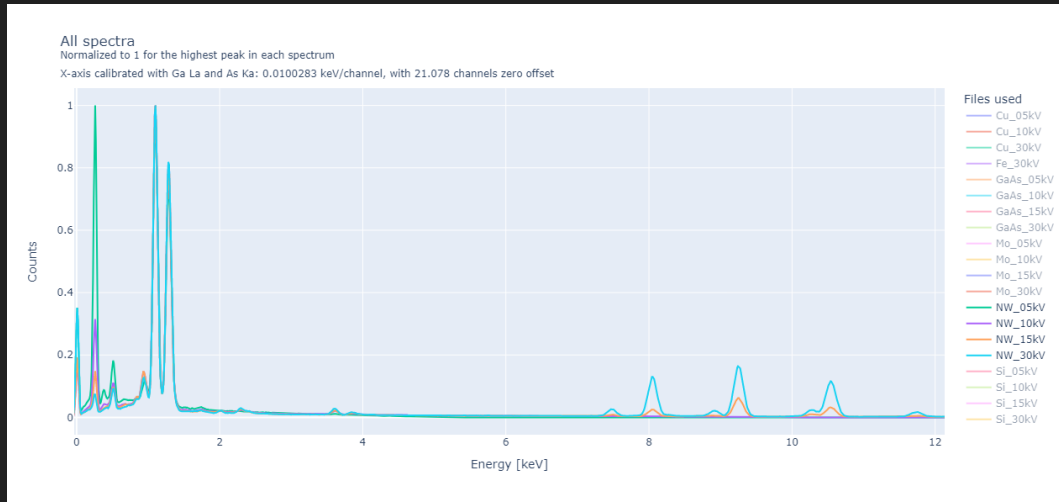
Figure 4.6: The spectra of the nanowire sample part. This spectrum have the most peaks, and contains Ga, As, Cu, Sb (? at 3.6 keV), Mo, C, and O.
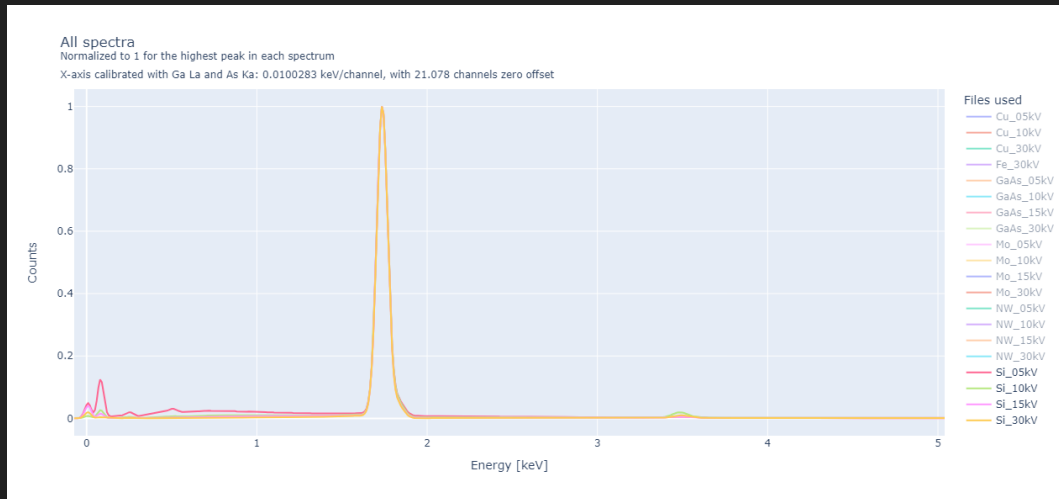
Figure 4.7: The spectra of the pure Si wafer sample part. All four spectra have one large peak at 1.73 keV, which is the Si Kα peak. The four spectra also have a small peak at 3.48 keV, which is not identified.

but it is very small. Even though the Cu spectra at 30 kV has the Cu Kα peak, the Cu Lα peak is completely missing.

Some of the peaks in the spectras are overlapping, which shifts the shape of the peaks. An example of this is the As Kα peak and the Ga Kβ peak in the GaAs bulk wafer spectra. These peaks are overlapping, but also far enough apart that the peaks are still distinguishable. Another example is the Mo Lα peak and the Mo Lβ_1, but these are overlapping so much that they are harder to distinguish. Since they are harder to distinguish, the peak fitting makes one Gaussian for the two peaks, which is off on both peak centers. Overlapping peaks makes counting the signal from a spesific peak harder.

The signal from the background is another factor which makes counting more difficult. In general, background in the aquired spectra is low, but with different shapes. In the GaAs, Si and Mo the height of the background decrease with higher acceleration voltage. In the Cu spectra the bacgrkoud increase with higher acceleration voltage. In the NW spectra the background is fairly similar, except for the 5 kV spectrum where the background is cut off at 5 keV. All the 5 kV spectra have decrease linearly from 1 to 5 keV. The background is very low and almost flat above the highest peak in the spectra. For example, the Cu 10 kV spectrum have its highest peak at 0.5 keV, and the background is almost flat above 0.5 keV. The background in the Si 30 kV spectrum increase a lot up to the high Si Kα peak at 1.7 keV, but the drop off significantly after the peak. The Mo 30 kV spectrum have the same behavior as the Si 30 kV spectrum. In general, the background singals are low, but their different shapes makes it harder to fit the peaks. **(Question for Ton: The last sentence is meant to be a transition/finishing sentence, but might be too much discussion.)**

Some of the peaks in the spectra are artifacts. All the spectra have a Si peak at 1.74 keV. Some of the spectra have peaks from areas outside the main beam, like the Mo, Sb and Cu peak in the NW spectra. The Sb peaks in the NW spectra are at 3.60, 3.85, 4.10 and 4.35 keV, being the Lα_1, Lβ_1, Lβ_2 and Lγ_1 peaks. NW have clear C Kα and O Kα peaks ar 0.260 and 0.515 keV. GaAs and Mo have some C and O signal. The C and O signal is higher at lower acceleration voltage. All four Mo spectra have a peak at 0.175 keV, which match with B Kα or Be Kα.

Another aritfact present in most spectra is sum peaks. The Si spectra have a sum peak at 3.49 keV, which is the sum of Si Kα at 1.74 keV. The Al spectrum have its highest peak at 1.48 keV and a sum peak at 2.98 keV The Mo spectrum have a sum peak at 4.65 keV, which is the sum of Mo Lα at 2.293 keV and Mo Lβ_1 at 2.395 keV. The GaAs spectrum at 30 kV have two small sum peak signals at 18.5 and 19.5 keV, while the NW spectrum at 30 kV with lower DT does not have these peaks.

4.2.2   Calibration

Different calibrations were explored. The initial calibration is the one from AZtec, and is the one used in the spectra in Figure 4.1. This calibration has a left shift for the L-peaks and a right shift for the K-peaks. The second type of calibration is the one given by the model fit in HyperSpy. The third type is from the self made model fit, using the distance between two high intensity and far apart peaks to calibrate the energy scale. Values for the three calibrations are given in Table 4.1. The deviations are a few percent, and the accuracy of the different calibrations give on spesific

peaks are given in Table 4.2. Here accuracy is the difference between the theoretical peak position and the peak center in the spectrum, given in percent.

Table 4.1: Different calibration values. The AZtec calibration is reffered to as the uncalibrated value. The dispersion is calculated with Equation (2.8). The offset is calculated with Equation (2.9). The own calibration was done on Ga L_α and As K_α from the 30 kV measurement on the GaAs wafer. The HyperSpy cailbration was done by making a model and fitting it to the data on the 30 kV GaAs spectrum.

Calibration method	Dispersion, [keV/channel]	Zero offset [channels]
AZtec	0.01	20
HyperSpy	0.010028	21.0787
Own calibration	0.010030	21.127

4.3 Quantitative results

Its okay to only do the GaAs here. Could also look at NW, but not nessesary.

To do the quantitative analysis, HyperSpy needs k-factors. The k-factors for Ga and As are given in Table 4.3. These k-factors are from the GaAs bulk wafer, and HyperSpy have estimated them theoretically. **(Question for Ton: Shall I list the other k-factors for the other sample areas? I do not think I will use them, since I've only quantified the GaAs bulk wafer. But the other k-factors are results too. Eventually including NW data too, but I do not know that ratio.)**

Table 4.2: Peak accuracy of the different calibration methods. The accuracy here is the deviation from the theoretical peak to the measured peak. The measured peak is the Gaussian fitted center of the peak. All results are from the 30 kV measurements. The own calibration was done on Ga L$_\alpha$ and As K$_\alpha$ from the 30 kV measurement on the GaAs wafer. The HyperSpy cailbration was done by making a model and fitting it to the data on the 30 kV GaAs spectrum. AZ is short for AZtec. HS is short for HyperSpy. All deviations are in percentage difference from the theoretical peak value.

Peak and line	Theoretical [keV]	AZ deviation [%]	HS deviation [%]	Own deviation [%]
As Lα	1.2819	1.000	0.439	0.422
As Kα	10.5436	-0.202	-0.025	-0.010
Ga Lα	1.098	1.044	0.342	0.318
Ga Kα	9.2517	-0.153	0.009	0.024
Cu Lα	0.9295	1.767	0.888	0.857
Cu Kα	8.0478	-0.114	0.031	0.045
Mo Kα	17.4793	-0.325	-0.108	-0.090
Mo Lα	2.2932	1.047	0.859	0.858
Si Kα	1.7397	0.167	-0.175	-0.182
Al Kα	1.4865	0.200	-0.247	-0.259
Cu Kα	8.0478	-0.116	0.029	0.043
C Kα	0.2774	-2.955	-6.583	-6.738

Table 4.3: K-factors for Ga and As, extracted from AZtec. All the k-factors are theoretically estimated. AZtec provides either the k-factor for the Lα or the Kα line, and selects automatically based on the energy of the incident electrons.

V_{acc} [kV]	Element	Line	K-factor
5	Ga	Lα	1.086
5	As	Lα	1.210
10	Ga	Lα	1.223
10	As	Lα	1.310
15	Ga	Lα	1.259
15	Ga	Lα	1.331
30	Ga	Kα	3.268
30	As	Kα	4.191

The initial quantification was done on the data from the GaAs wafer in AZtec and in HyperSpy as out-of-the-box as possible. The results are presented in Table 4.4. The wafer is a 1:1 alloy of gallium and arsenic, so the atomic percent of Ga and As should be 50% and 50% respectively.

Table 4.4: Initial quantification of the GaAs wafer. The ratio in the wafer is 1:1, so the correct ratio is 50% and 50%, because the results are in atomic percent. **(Brynjar: Put in the actual results here. Use both HyperSpy linear and model fitted results?)**

V_{acc}	AZtec		HyperSpy	
	Ga	As	Ga	As
5 kV	50 %	50 %	50 %	50 %
10 kV	50 %	50 %	50 %	50 %
15 kV	50 %	50 %	50 %	50 %
30 kV	50 %	50 %	50 %	50 %

One of the adjustments explored was the affect of the calibration on the quantification. Using different the calibrations in 4.1 gave different quantification results when using Cliff-Lorimer in HyperSpy. The results are presented in **??**.

Chapter 5

Discussion

5.1 Introduction

The discussion is presented in this chapter. Producing code from scratch is a time consuming process, but it is also a learning process. While developing the code, the author learned a lot about EDS analysis and new ideas emerged. The sections below follow the structure of the sub-problems of the main problem statement formulated in Chapter 1.

5.2 Analysis steps in HyperSpy

(Question for Ton: Is this interesting to write about? The problem here is that the text is both method, some results and kind of discussion. What do I do with that? I want to keep it, but also restructuring it.)

The next sub-problem was to find out what is done with the data at the different steps in the analysis when using HyperSpy. In these steps it is assumed that the user have done qualitative analysis and want to do quantitative analysis on a set of elements. The analysis in AZtec is done as a black box, so it is not possible to see what is done with the data at the different steps. All variables inside croccodile need to be set by the user, e.g. `<element_list>` would be set to `['Ga', 'As']` for the GaAs wafer. An example notebook with quantification of the GaAs wafer is attached in APPENDIX. **(Brynjar: Make a notebook with GaAs quantification in HyperSpy, with the data somehow.)**

5.2.1 Loading the data and specifying the elements

```
s = hs.load(<filepath>, signal="EDS_TEM")
s.set_elements(<element_list>)
```

The first step in the analysis is to load the data as a HyperSpy `signal` type, and specifying the signal as TEM. The `signal` type is a class in HyperSpy that contains the data and the metadata, and it has methods for analysis. The

`signal` type must be specified as TEM, because the `signal` type for SEM is very limited and does not have a method for quantification. When using .emsa files from AZtec, as is done in this project, the metadata contains some relevant and some irrelevant information. The information relevant later in this project is: acceleration voltage, dispersion, zero offset, energy resolution Mn $K\alpha$, After loading, it is possible to plot the data with `s.plot()`.

5.2.2 Removing the background linearly

```
bw = s1.estimate_background_windows(windows_width=<number>)

iw =  s1.estimate_integration_windows(windows_width=<number>)
```

The next step is to remove the background, which with the above code is done by a linear fit. The background can be removed through model fitting, which is covered in Section 5.2.4. The variable `windows_width` sets how wide the windows are for the background and integration, measured in FWHMs. A good starting value for `windows_width` is 2, but it should be tested by the user with a plot to see if the background will be removed correctly. The estimated windows can be plottet with:

```
s.plot(xray_lines=True, background_windows=bw, integration_windows=iw)
```

5.2.3 Quantification after linear background removal

```
s_i = s.get_lines_intensity(background_windows=bw, integration_windows=iw)

k_factors = [<k-factor 1>, <k-factor 2>]

quant = s.quantification(s_i, method='CL', factors=k_factors)

print(f'E1: {quant[0].data[0]:.2f} \%, E1: {quant[1].data[0]:.2f} \%')
```

The quantification is done with the four lines of code above, where the last one prints the results. The first line gets the intensity of the peak corresponding to the lines of the specified element. HyperSpy selects automatically which lines to use for quantification. To see which lines are used, the `s_i` variable can be printed. The second line sets the k-factors. The k-factors in this project have been the one from AZtec, which are theoretically estimated. The third line does the quantification, where the method is specified. The method is the Cliff-Lorimer method, described in detail in Mari Skomedal's master thesis [5, Sec. 2.2.3]. HyperSpy has a method for quantification with the zeta factor method. The zeta method requires the value for the beam current, which was not measured in this project. [1]

5.2.4 Removing the background with model fitting

Another way to remove the background is to fit a model to the data. This step would be done right after loading the data. If the raw data contains a zero peak, as is the case for most Oxford instrument EDS detectors, the zero peak needs

---

[1]Results from the zeta methon can be converted to the cross section method, see the "EDS Quantification" documentation in HyperSpy.

to be removed before fitting the model. The zero peak is removed by skipping the first n channels, where n=30 works well with the data from the GaAs wafer. The model fitting is done with the following code:

```
s = s.isig[<zero_peak>:]

m = s.create_model(auto_background=False)

m.add_polynomial_background(order=12)

m.add_family_lines(<list_of_element_lines>)

m.plot()
```

The lines above removes the zero peak, create a model from the `signal` s, adds a 12th order polynomial, add the lines of the elements in the `signal`, and plot the model. This model is not fitted, it is just a generated spectrum with the lines of the elements. Eventually, the method `create_model()` can take the boolean argument `auto_add_lines=True`, which will automatically detect the elements in the sample. The model consists of a number of components, which can be accessed with `m.components`. The components are all the gaussian peaks in the spectrum, in addition to the background as a 12th order polynomial. The order of the polynomial can be changed, but it should be tested by the user to see if it is a good fit. Further, the model must be fitted.

```
m.fit()

m.plot()
```

The first line fits the model to the data to the components and the second line plots the model. HyperSpy have a own option for fitting only the background. Since the background is one of the components in m, it is fittet with the code line above.

5.2.5   Quantification after model fitting

```
m_i = m.get_lines_intensity()

k_factors = [<k-factor 1>, <k-factor 2>]

quant = s.quantification(s_i, method='CL', factors=k_factors)

print(f'E1: {quant[0].data[0]:.2f} \%, E1: {quant[1].data[0]:.2f} \%')
```

The quantification after model fitting is done in the same way as in Section 5.2.3, but with intensity from the model instead of the signal. When modelling GaAs, the model can add the intensity from both K-lines and L-lines. Since AZtec only gives the k-factors for either the K-lines or the L-lines, the user must remove the lines without k-factors before quantification.

5.2.6   Calibrating the spectrum with the HyperSpy model

```
m.calibrate_energy_axis(calibrate='scale')

m.calibrate_energy_axis(calibrate='offset')
```

The two lines above calibrates the spectrum with the HyperSpy model and updates the dispersion and zero offset. The metadata in the `signal` s is updated with the new calibration. Thus, doing the previous step with quantification after model fitting can give a more correct quantification.

5.3   Peak and background modelling

The next sub-problem was to find out how the peaks and the background are modelled in a way that is easy to understand. The model was built without HyperSpy, with the idea of making every step easier to understand. The model was used to be able to remove the background and be able to calibrate the spectrum. The model was compared to the HyperSpy model. The model could be used to quantify the elements in the sample, but this was not done in this project. **(Brynjar: Do I want to do this?)**

The first step in creating a model is to identify the peaks. The peaks are assumed to be gaussian curves. The initial way of identifying peaks was that the user manualy identified the peaks. Later the peaks were identified with the function `find_peaks()` from the `scipy.signal` package. Different peak prominence were tested, and the peak prominence of 0.01 gave the best results.

The second step is to make a gaussian in each peak and one polynomial for the background. To do the fitting, the components need an initial guess. The background needs a coefficient for each order of the polynomial. Each gaussian need to have a mean, a standard deviation, and a height. The mean is the peak position. The standard deviation is the width of the peak, where FWHM $= std * 2 * \sqrt{2 * \ln 2}$ [2]. The height is the amplitude of the peak. The easiest way to get the initial guesses for the gaussians is to normalize the data and set all three parameters to 1. In the normalization the highest peak was set to 1, and the rest of the peaks were scaled accordingly. The best way to get the initial guesses for the background is to clip out the peaks with linear interpolation and fit a polynomial. The initial guesses for the background is then the coefficients of the polynomial. With the initial guesses, the whole model is ready to be fitted.

The third step is to fit the model to the data. Using the `curve_fit()` function from the `scipy.optimize` package, the model is fitted to the data. The function `curve_fit()` uses the Levenberg-Marquardt algorithm to fit the model to the data. The function `curve_fit()` returns the optimal parameters for the model. Fitting both the gaussians and the background at the same time makes the fitting more stable. One of the first iterations, where the user manually inputted the peaks, the fitting tended to partially fail. The issue was that the fitting only was done on the peaks. To minimize the error in the fitting, one of the gaussian curves with a low amplitude was moved and got

---

[2]FWHM defined at: `https://en.wikipedia.org/wiki/Full_width_at_half_maximum`

a huge standard deviation, which compensated the background. This was fixed by fitting both the gaussians and the background at the same time. Doing this made the fitting both better and it failed less often.

(Brynjar: Issue: fitting e.g. Mo with two clear peaks, but not with enough prominence to be found by the peak finder.)

5.4   Calibration

The next sub-problem was to calibrate the data with a self produced Python script. With a fitted model of the spectrum, the calibration can be done. Calibration can both be done on raw data with channels on the x-axis and on poorly calibrated data with energy on the x-axis. The dispersion is calculated with Equation (2.8). Table Table 4.1 shows calibration from AZtec, HyperSpy, and the self produced Python script.

5.5   Background models

The next sub-problem was to find out how different background models affect the quantitative analysis done in Hyper-Spy, and how well different order polynomials fit the background. The background models were tested on the spectrum of GaAs, and later also on *(Brynjar: TODO: other spectra. Also make a table here with results)*. The background was modelled as a polynomial of different orders. To quantify the different background models, the residuals were calculated. The residuals are the difference between the data and the model. *(Brynjar: use root-mean-square error?)* The TABLE XXXX *(Brynjar: make table)* shows the residuals for the different order background models. The best orders were visually inspected. A later idea was to model the background as a spline, which is a piecewise polynomial. The spline is a piecewise polynomial with a smooth transition between the pieces. The spline was not tested in this project, but it could be a good alternative to the polynomial background model.

5.6   Analysis failure

The next sub-problem was to find out when the analysis fails, both in AZtec and HyperSpy.

(Question for Ton: Section about normalization too?)

5.7   Calibration decision

Why I selected the Ga La and As Ka peaks for calibration. Less extrapolation. PEaks are far apart. The peaks need a good Gaussian fit. Need a nice curve. High peak to background. Should be a sample that is easily available. Cu tape wpuld be nice, since it is in all labs and samples. Mo is far apart, but looks bad. Mo is also harder to fit automatically, because of the close double peak.

5.8 Choises in HyperSpy

Commersial packages does quantify SEM signal, and they are kinda good at it. Should be more accurate on NW sample, but is it? Reference the discussion on GitHub?

Bibliography

[1] Joseph I. Goldstein, Dale E. Newbury, Joseph R. Michael, Nicholas W.M. Ritchie, John Henry J. Scott, and David C. Joy. *Scanning Electron Microscopy and X-Ray Microanalysis*. Springer New York, New York, NY, 2018.

[2] J. Michael Hollas. *Modern spectroscopy*. J. Wiley, Chichester ; Hoboken, NJ, 4th ed edition, 2004.

[3] Daniel Lundeby. Improving the accuracy of TEM-EDX quantification by implementing the zeta-factor method. Master's thesis, Norwegian University of Science and Technology, 2019.

[4] Manne Siegbahn. Relations between the K and L Series of the High-Frequency Spectra. *Nature*, 96(2416):676–676, February 1916.

[5] Mari Sofie Skomedal. Improving quantitative EDS of III-V heterostructure semiconductors in low voltage STEM. Master's thesis, Norwegian University of Science and Technology, 2022.

[6] Albert C Thompson, David T Attwood, Malcolm R Howells, Jeffrey B Kortright, Arthur L Robinson, James H Underwood, Kwang-Je Kim, Janos Kirz, Ingolf Lindau, Piero Pianetta, Herman Winick, Gwyn P Williams, and James H Scofield. *X-Ray Data Booklet*. Lawrence Berkeley National Laboratory, University of California, 2nd ed edition, 2004.