
Technical University of Denmark
Center for Electron Nanoscopy

Advanced TEM (16 September 2010)

Microanalysis in the electron microscope

Dr C B Boothroyd

Synopsis

X-ray spectroscopy

1. Electron beam interactions and X-ray generation
2. Recording X-ray spectra: wavelength dispersive X-ray spectroscopy (WDX)
- 3, 4. Recording X-ray spectra: energy dispersive X-ray spectroscopy (EDX)
- 5, 6. Quantitative analysis: Quantification of X-ray spectra, detection limits, standards, spectrum fitting, “ZAF” corrections, rough surfaces

Energy loss spectroscopy (EELS)

7. Electron energy loss spectroscopy (EELS)
8. Energy filtered microscopy (EFTEM)

The aim of this course is to give an introduction to the microanalysis facilities available on electron microscopes.

Lectures 1 to 6 will cover microanalysis using X-ray spectroscopy in both scanning and transmission electron microscopes. Lectures 7 and 8 cover energy loss spectroscopy and energy filtering which are available only on transmission electron microscopes.

Note that this version (DTU 2010) is virtually unchanged from the handout for the 2002 course in Cambridge. Although most of the content of the DTU course is present, there are a number of recent developments (such as EELS very low losses) that are missing from this handout.

References

Scanning electron microscopy and X-ray microanalysis

Jl Goldstein, DE Newbury, P Echlin, DC Joy, AD Romig, CE Lyman, C Fiori and E Lifshin
(Plenum, New York, second edition 1992) (Mat Sci library, first edition 1981, MB199b)

Transmission electron microscopy

DB Williams and CB Carter (Plenum, New York, 1996) (Mat Sci library, Mb309–Mb312)

Scanning electron microscopy

L Reimer (Springer-Verlag, Berlin, 1985) (Mat Sci library, Mb237)

Fundamentals of energy dispersive X-ray analysis

JC Russ (Butterworths, 1984) (Mat Sci library, NbC76)

Electron microprobe analysis

SJB Reed (CUP, 1993) (Mat Sci library, first edition 1975, Mb159)

Quantitative electron-probe microanalysis

VD Scott, G Love and SJB Reed (Wiley, 1994) (Mat Sci library, first edition 1983, Mb225a)

Microscopy of materials

DK Bowen and CR Hall (Macmillan Press, 1975) p 68–112 (Mat Sci library, Mb144a)

Electron energy-loss spectroscopy in the electron microscope

RF Egerton (Plenum, second edition 1996)

1. Electron beam interactions and X-ray generation

1 Purpose of microanalysis

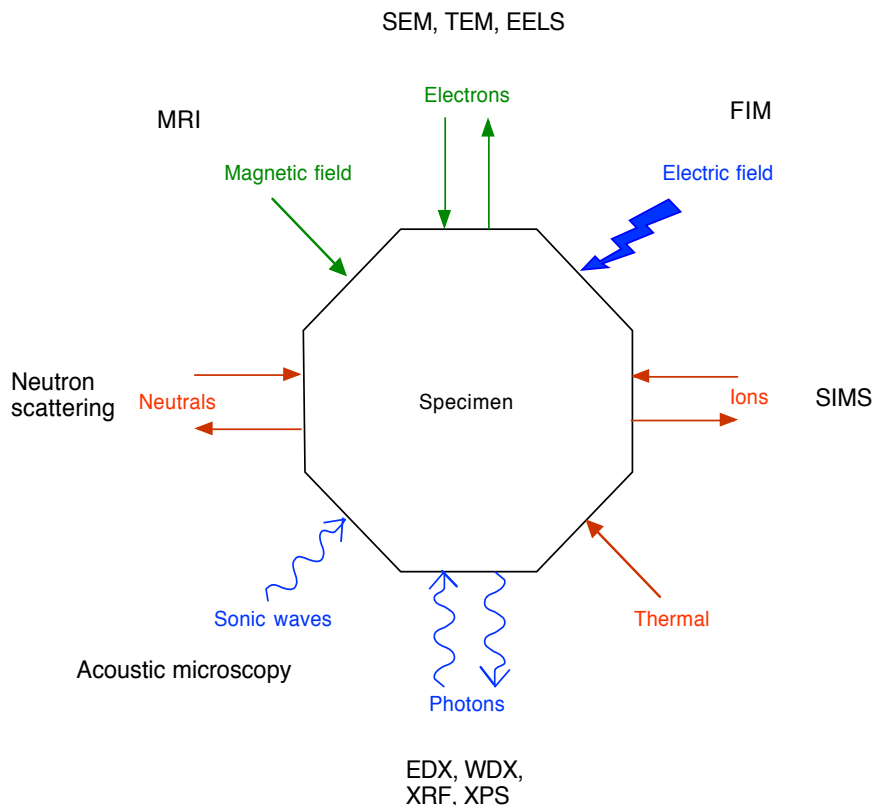
Microanalysis is the determination of the composition of the specimen on a microscopic scale. In an electron microscope this is done by irradiating with a beam of electrons and analysing the radiation emitted.

Why do it? Examination of a material in an electron microscope allows the structure to be seen, but gives little clue as to the composition of the phases present. X-ray detectors allow the elements present in your specimen to be identified and their proportions measured. X-ray analysis is quick and provides an easy check of the composition of your specimen.

Quantitative X-ray analysis, to get accurate elemental concentrations is more difficult and requires an understanding of the mechanism of generation and detection of X-rays.

2 Interaction of the electron beam with the specimen

There are many possible surface analysis techniques. They can be characterised by the radiation used to probe the specimen (electrons, X-rays, etc) and the radiation emitted by the specimen. Some of the different possible input and output signals are shown below along with the corresponding techniques that use those signals.



Surface analysis techniques – input and output signals

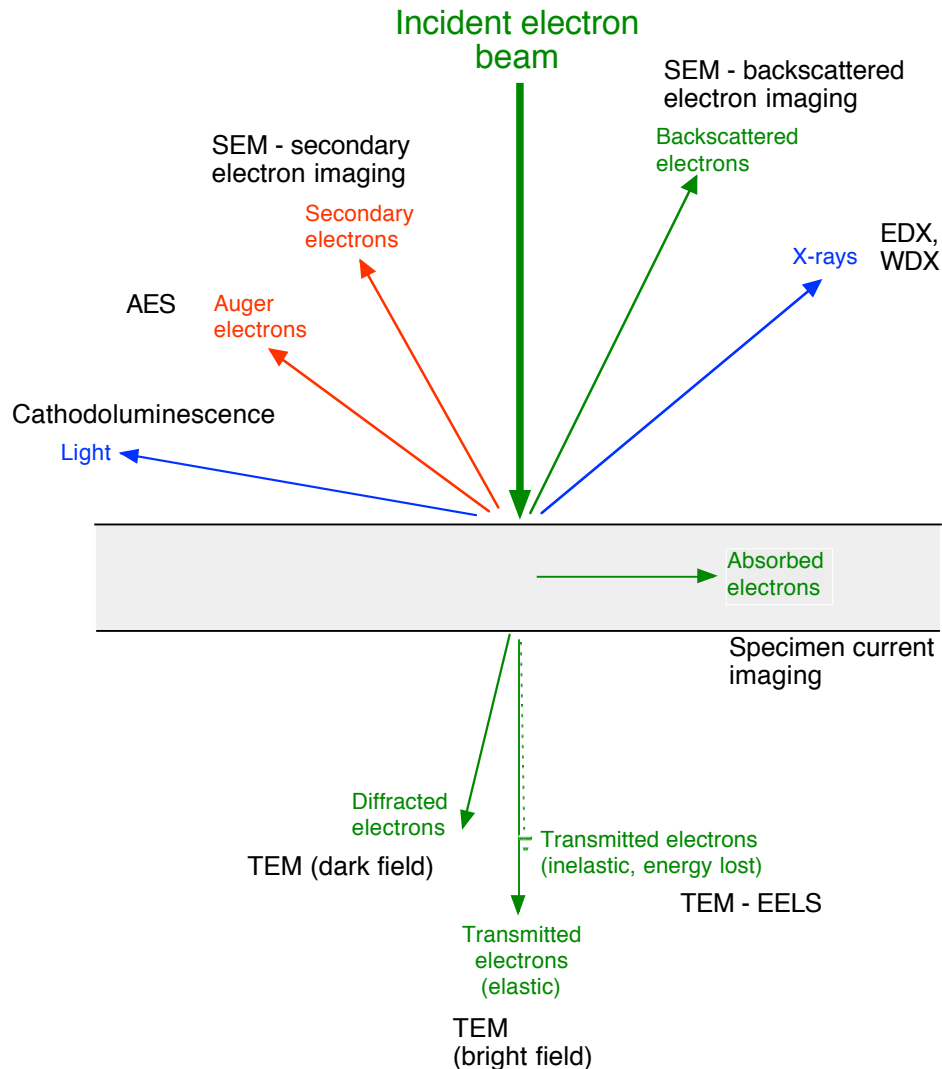
Summary of microanalysis techniques

Technique	Input signal	Output signal	Depth sampled (nm)	Lateral resolution (nm)	Elemental range	Sample	Detection limit (at%)
Wavelength dispersive X-ray spectroscopy (WDX, WDS)	electron	photon (X-ray)	~1000 nm	~1000 nm	$Z \geq 4$	flat polished surface	0.005% (50 ppm)
Energy dispersive X-ray spectroscopy in SEM (EDX, EDS)	electron	photon (X-ray)	~1000 nm	~1000 nm	$Z \geq 5$	flat polished surface	0.1%
Energy dispersive X-ray spectroscopy in TEM or STEM (EDX, EDS)	electron	photon (X-ray)	10 to 100 nm (sample thickness)	2 nm	$Z \geq 5$	TEM foil	0.1%
Electron energy loss spectroscopy in TEM (EELS), EFTEM	electron	electron	10 to 100 nm (sample thickness)	1 nm	$Z \geq 3$	TEM foil	0.1%
X-ray fluorescence (XRF)	photon (X-ray)	photon (X-ray)	~1000 nm	1 to 10 mm	$Z \geq 4$	flat polished surface	0.0002% (2 ppm)
Secondary ion mass spectroscopy (SIMS)	ion	ion	0.1 to 1 nm	1 to 10 μm	all	surface (destructive)	0.00001% (0.1 ppm)
Auger electron spectroscopy (AES)	electron	electron	0.5 to 2 nm	400 nm	$Z \geq 2$	surface	0.1%
X-ray photoelectron spectroscopy (XPS), Electron spectroscopy for chemical analysis (ESCA)	photon (X-ray)	electron	1 to 3 nm	1 to 10 mm	$Z \geq 2$	surface	1%
Field ion microscopy (FIM)	electric field	ion	individual atoms	individual atoms	$Z \geq 1$	fine wire with sharp point	individual atoms
Laser microprobe mass spectrometry (LAMMS)	photon (laser)	ion	~100 nm	~1000 nm	all	polished surface or foil	0.00001% (0.1 ppm)

Other acronyms

SEM	Scanning electron microscope
TEM	Transmission electron microscope (= CTEM, conventional transmission electron microscope)
HREM	High resolution electron microscope
AEM	Analytical electron microscope
EFTEM	Energy filtered transmission electron microscopy
ELNES	Energy loss near edge structure
EXELFS	Extended energy loss fine structure
XANES	X-ray absorption near edge structure
EXAFS	Extended X-ray absorption fine structure
CL	Cathodoluminescence
EBIC	Electron beam induced conductivity
NMR	Nuclear magnetic resonance
MRI	Magnetic resonance imaging (like NMR but without the word "nuclear"?)

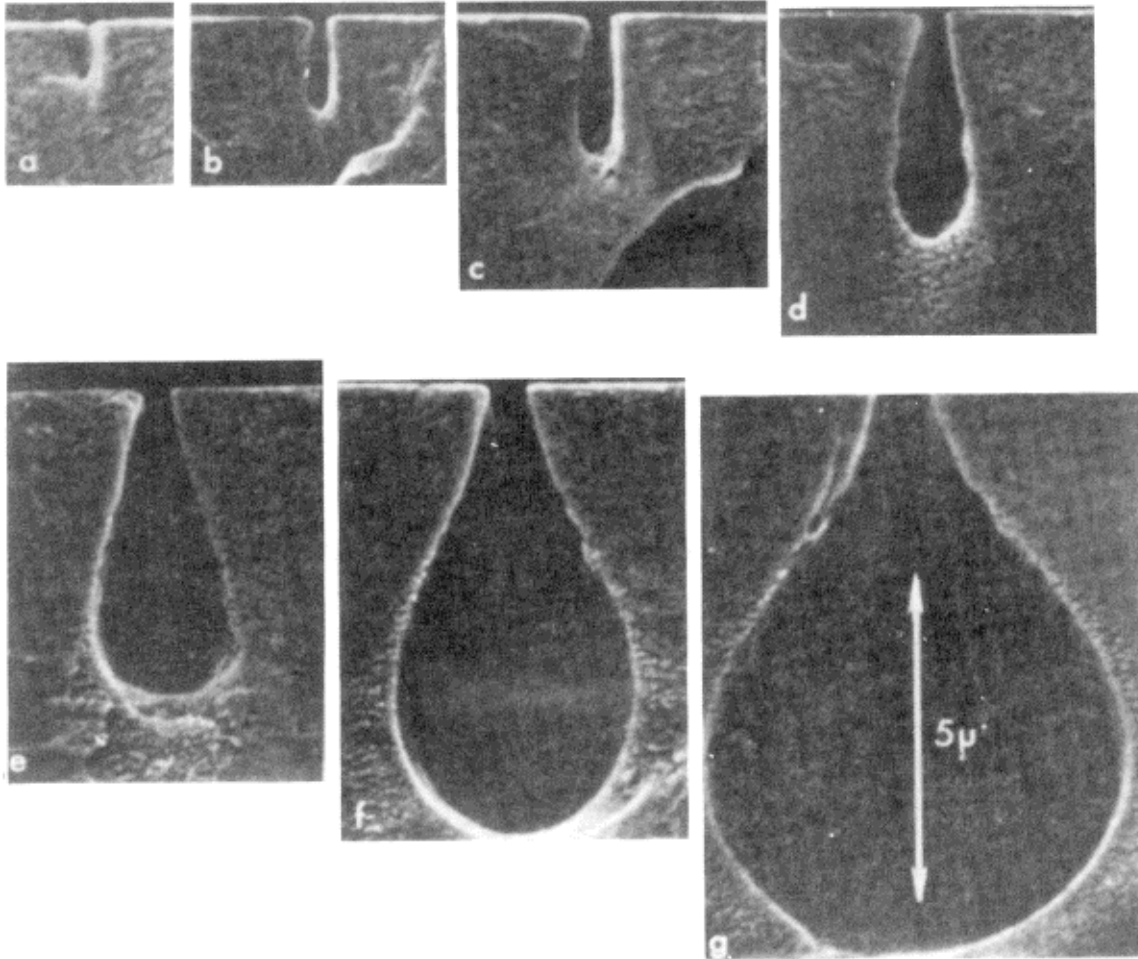
In electron microscopes, only electrons are used to probe the specimen. The diagram below shows the signals emitted when a specimen is bombarded with electrons and the techniques that use these signals.



For a thick specimen, as in SEM, the transmitted electrons are absorbed in the specimen and give rise to a specimen current.

3 Interaction volume

When an electron beam hits a material a characteristic interaction volume is generated by the electrons as they penetrate and slow down.

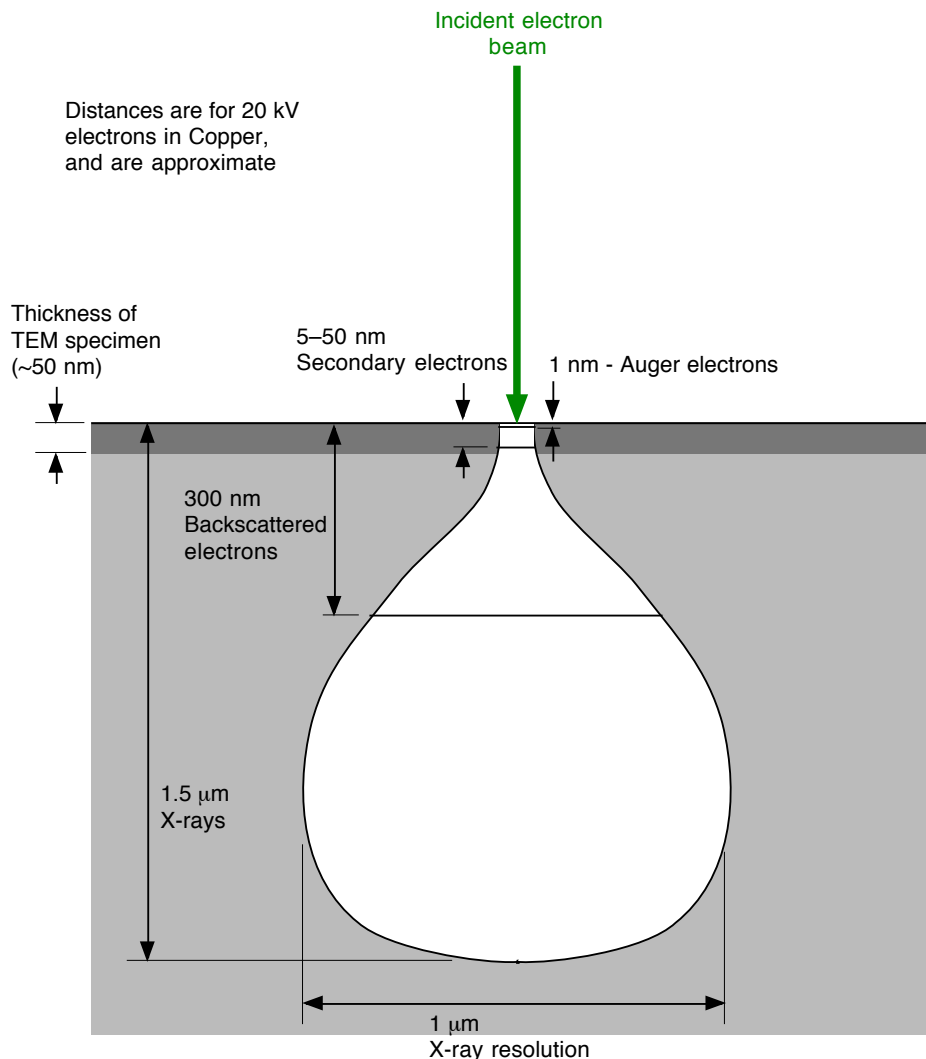


Electron interaction volume in polymethylmethacrylate (PMMA)

In a to g the electron dose is the same, but the etching time is increased to show the contours of decreasing energy deposition.

Goldstein (1st ed) p 60

Where the electrons enter the material they have high energy and so suffer little lateral scattering. Thus they form a narrow neck. As the electrons are slowed down, they are scattered sideways more forming the bulb shape.



Interaction volume for electrons in a bulk sample. Distances are for 20 kV electrons in Cu. For Al, multiply by 3

The distances depend strongly on the material and the beam energy, see Goldstein page 89. Notice that for 20 kV electrons in Cu, as drawn above, the width of the interaction volume is about 1 μm . This width puts a limit on the best resolution obtainable no matter how small the electron beam is.

X-rays are generated throughout the whole interaction volume and are detected from this volume. Backscattered, secondary and Auger electrons are also generated from all of this volume but due to their lower energies have limited ranges in the sample. Backscattered electrons are emitted from relatively large depths, eg 300 nm for Cu, secondaries from the top 5 to 50 nm and Auger electrons from the top 1 nm only. Thus the X-ray and backscattered signals have a much poorer resolution than the secondary or Auger signals.

In the TEM, the sample thickness is typically around 50 nm, as shown above. Most of the electrons go right through the sample with little interaction. Thus the interaction volume does not degrade the resolution like it does in SEM. However the much smaller interaction volume in TEM means a much smaller signal.

The interaction volume varies with the atomic number (Z) of the target, the energy of the electron beam and the angle between the beam and the specimen surface. There are two ways of measuring the electron range in a material (and hence the size of the interaction volume), the Bethe range

(average distance travelled by an electron derived from theory) and Kanaya-Okayama (K-O) range (more empirical). The Kanaya-Okayama range is perhaps the more accurate and more useful:

$$R = \frac{0.0276 A E_0^{1.67}}{Z^{0.889} \rho}$$

Where

A = atomic mass

E_0 = beam energy in keV

Z = atomic number

ρ = density in g/cm³

Below is a table showing the electron ranges calculated using the above formula for three target materials at various electron energies.

Target	5 keV	10 keV	20 keV	30 keV
Aluminium	0.41 μm	1.32	4.2	8.3
Copper	0.15	0.46	1.47	2.89
Gold	0.085	0.27	0.86	1.70

K-O electron ranges (in micrometers)

From Goldstein (1st ed) p 73

The width of the interaction volume, and hence the X-ray resolution, is typically about $2/3$ of the electron range shown above. Thus the X-ray resolution is better at lower accelerating voltages, and is much worse for lower atomic number specimens.

The electron range also depends on the angle the electron beam hits the specimen at (measured as the angle between the beam and the specimen normal). This is given approximately by:

Electron range at angle θ

$$R(\theta) = R \cos(\theta)$$

The thickness of TEM specimens is typically about 50 nm, so most electrons go straight through. Thus for TEM there is very little broadening of the beam. In TEM the X-ray resolution depends mostly on the width of the electron beam and is thus much better than for the bulk specimens in SEM.

4 X-ray generation

When an electron beam hits a material X-rays are formed by two processes giving two types of X-ray, bremsstrahlung and characteristic X-rays.

a. Bremsstrahlung process

This is caused by electrons being decelerated by the nucleus and electrons in the atoms in the material. Energy is lost by the electron on being decelerated and is emitted as X-rays. It contains all energies from 0 eV up to the beam energy. The intensity of the bremsstrahlung is given by Kramers law:

$$I \approx \frac{iZ(E_0 - E)}{E}$$

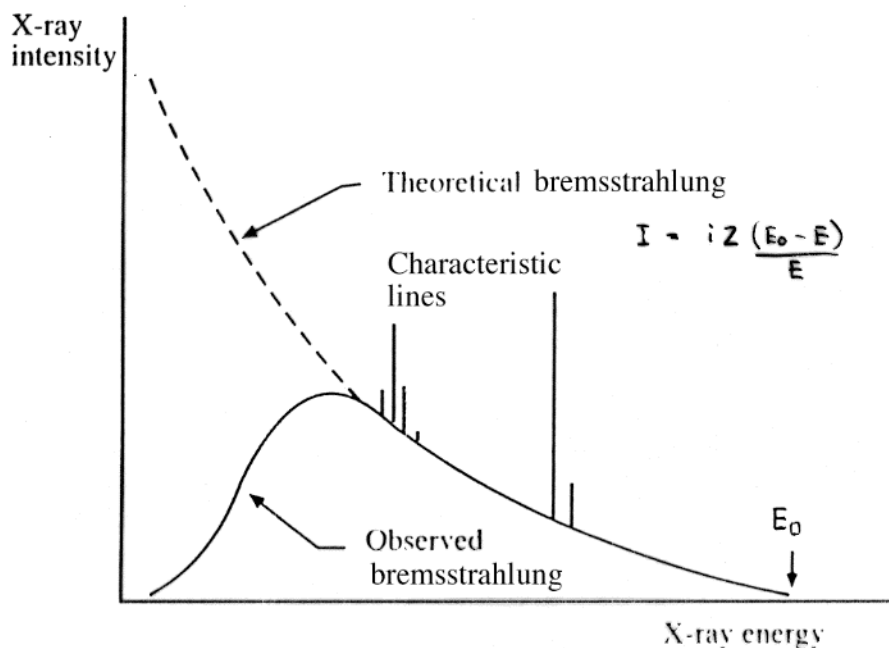
where

i = beam current

Z = average atomic number

E_0 = incident electron beam energy

This is shown as the dotted line in the graph below.



The bremsstrahlung and characteristic X-ray intensity as a function of energy. The generated intensity increases rapidly at low energies but low energy X-rays are absorbed by the specimen and the detector. E_0 is the energy of the electron beam causing the X-ray emission. Two families of characteristic lines are shown superimposed on the bremsstrahlung.

Williams and Carter p 57

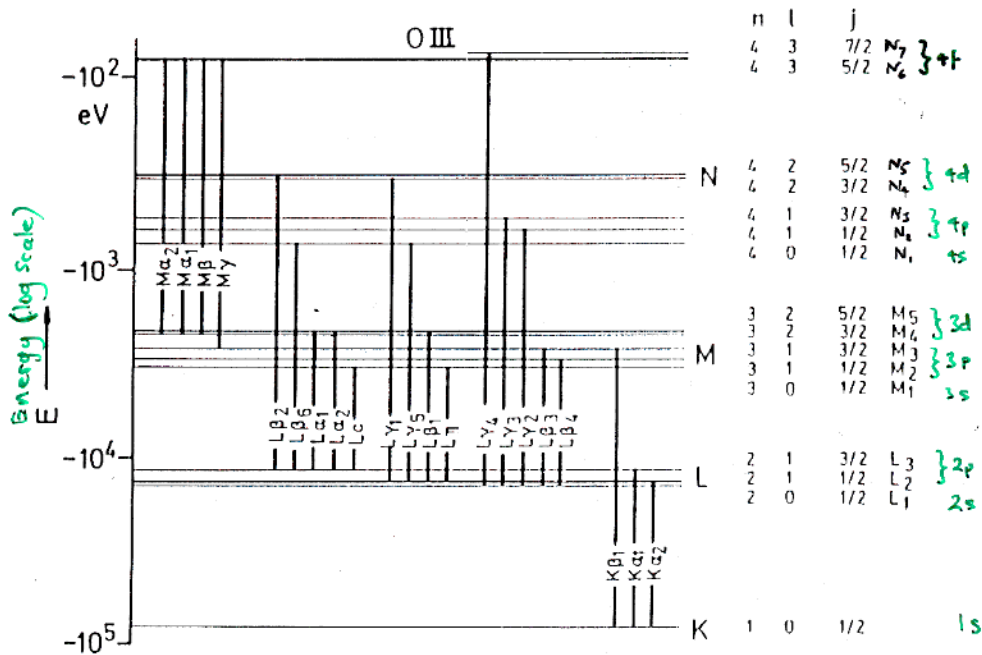
The maximum energy of the bremsstrahlung is E_0 . In practice the bremsstrahlung is cut off at low energies (solid line) because low energy X-rays are absorbed more in the specimen and thus never escape.

The bremsstrahlung X-rays are generally not useful for analysis, they form a background that must be removed.

b. Characteristic X-rays

To produce characteristic X-rays an electron from the beam removes an electron from an inner shell in an atom in the target material leaving a vacancy. The result is an ion in an excited state. For this to happen the beam electron must have energy greater than the shell being ionised. This is called the excitation energy, or the absorption edge energy.

The ion then loses energy by an outer shell electron falling into the inner shell vacancy (in typically 10^{-12} s). The energy is released as either an X-ray or by the ejection of another outer shell electron (an Auger electron).



K, L and M characteristic lines for Au

Reimer p 164

Electrons can fall from any higher shell for which the transition is quantum mechanically allowed. Each shell has 4 quantum numbers, n, l, $j=l+s$ and m. The allowed transitions are:

$$\begin{aligned} \Delta n &\geq 1 \\ \Delta l &= \pm 1 \\ \Delta j &= \pm 1 \text{ or } 0 \end{aligned}$$

Transitions to the K shell are called K, transitions to the L shell are called L, etc. Greek letters are used to denote the shell the transition is from, in approximate order of intensity (α more intense than β etc).

The transitions are very sharp and depend on the element the X-rays are emitted from, and are thus useful for analysis.

The energy (or wavelength) of characteristic X-rays follows Moseley's relation:

$$\text{X-ray wavelength } \lambda = \frac{B}{(Z - C)^2}$$

$$\text{X-ray energy (eV) } E = \frac{hc(Z - C)^2}{eB}$$

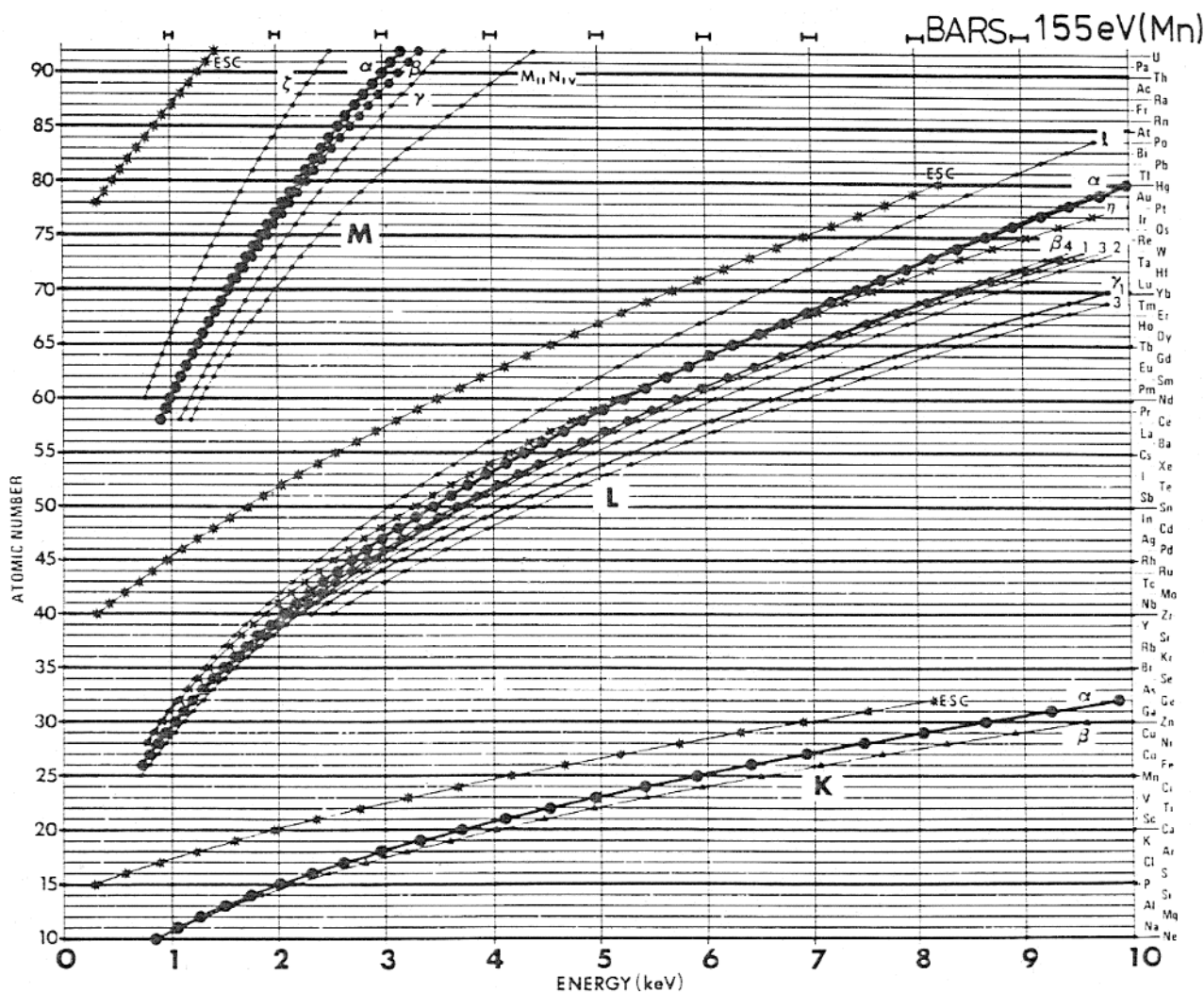
where B, C = constants depending on the X-ray line (K_α, L_α , etc)

Z = atomic number

e = electron charge

c = speed of light

h = Planck's constant



X-ray line energy vs atomic number

Goldstein (1st ed) p 276

The bonding environment of the atom affects the energy of the X-ray lines only slightly. It causes a change in X-ray energy of around 10 eV or less. This is not detectable by EDX and only just detectable by WDX.

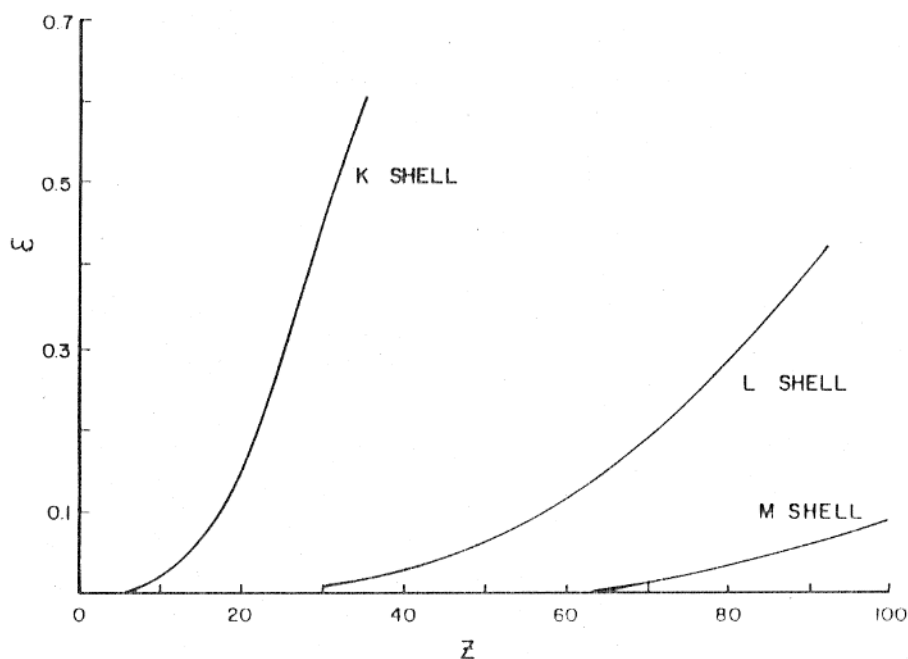
As mentioned above, an ionised atom can release its energy either as an X-ray or an Auger electron. The probability of X-ray emission is called the fluorescent yield, ω . For example, the fluorescence yield for the K shell is defined as:

$$\omega_K = \frac{\text{No. of X - ray photons emitted}}{\text{No. of K shell vacancies}}$$

The fluorescent yield has the following relation to the atomic number, Z :

$$\omega \approx \frac{Z^4}{a + Z^4}$$

For the K shell $a \approx 10^6$. The fluorescent yield is plotted below.



Fluorescence yield (ω) as a function of atomic number for the K, L and M shells

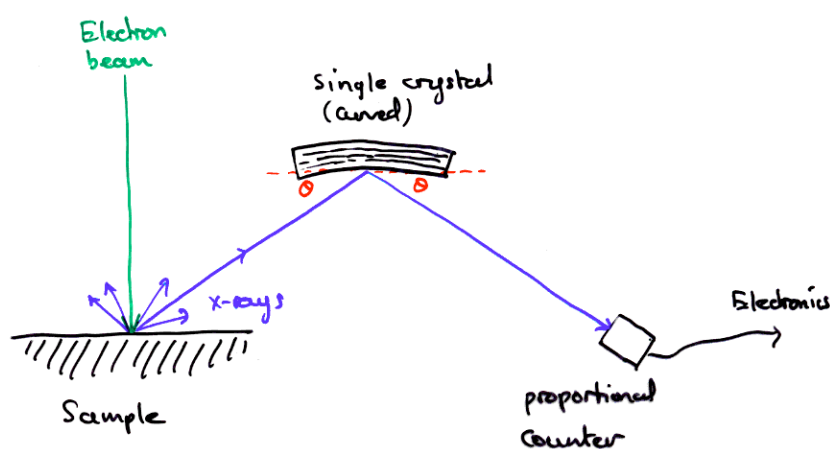
Goldstein (1st ed) p 101

For low Z the fluorescent yield becomes very small, eg for C ($Z=6$) $\omega_K \approx 0.001$ and for Ge ($Z=32$) $\omega_K \approx 0.5$. Thus X-ray emission is very inefficient for low Z , and EELS is a more efficient technique for analysing low Z elements.

2. Recording spectra: wavelength dispersive X-ray spectroscopy (WDX)

To record the X-ray spectrum emitted from a material irradiated by electrons we need to be able to both count the X-rays and measure their energy. In wavelength dispersive X-ray spectroscopy (WDX) the X-rays are separated according to their wavelength by a crystal and detected with a proportional counter. In energy dispersive X-ray spectroscopy (EDX) a semiconductor crystal both separates the X-rays according to their energy and counts them.

1. Basis of wavelength dispersive X-ray spectroscopy (WDX)



Basic diagram of a crystal spectrometer for detecting and measuring X-rays

In WDX a crystal spectrometer is used to separate the X-rays into their different wavelengths. This is a single crystal that diffracts the X-rays and thus monochromates them. The intensity is then measured using a gas proportional counter, plus associated electronics to process the signals. Different X-ray wavelengths are measured by scanning the crystal angle and thus changing the diffraction angle. Because WDX uses diffraction from a crystal it measures the X-ray wavelength, hence its name.

In practice we tend to measure the counts on the X-ray peak of interest (and either side of the peak for a background measurement) rather than scanning the whole spectrum.

I will now look at each part of the detection and processing of X-rays for WDX in turn.

2. Crystal spectrometer

Diffraction from a crystal is used to monochromate X-rays and separate a defined wavelength from the mixed wavelengths coming from the specimen. The wavelength is selected by Bragg's law,

$$n\lambda = 2d \sin \theta$$

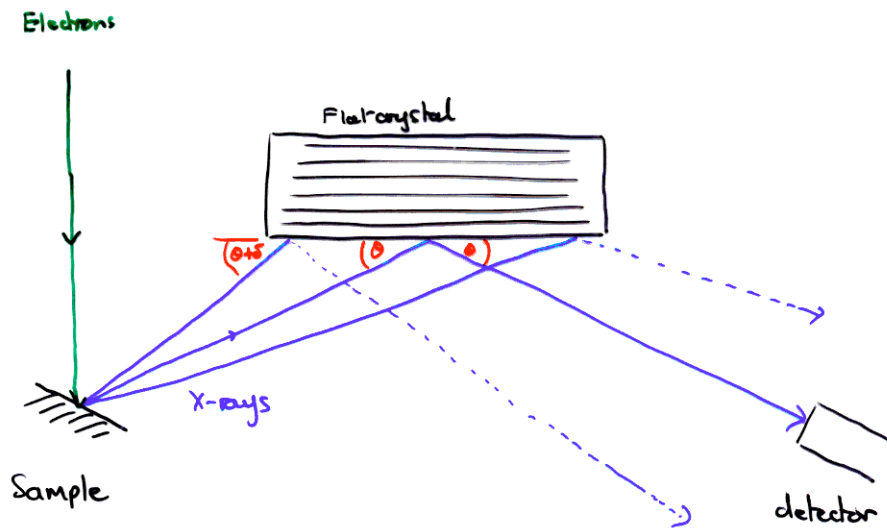
Where

d = spacing of crystal plane (eg LiF)

θ = Bragg diffracting angle (as in diagram)

n = order of reflection

Unfortunately a simple flat crystal is very inefficient, as can be seen from the diagram below.

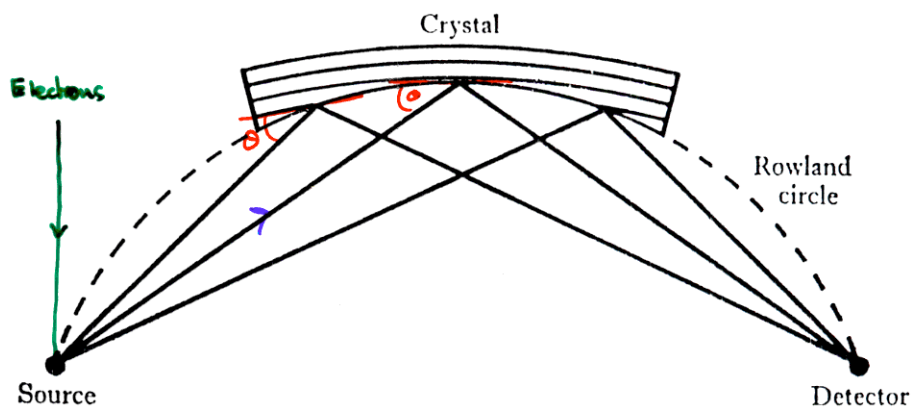


X-ray diffraction from a simple flat crystal. Most electrons hit the crystal at the wrong angle for diffraction and miss the detector

Only those electrons hitting the middle of the crystal are diffracted into the detector. Electrons hitting the ends *both* miss the detector *and* hit the crystal at the wrong angle for diffraction.

For these reasons a focusing crystal is used.

We need to have *both* the X-rays hitting the crystal planes at a constant (Bragg) angle *and* the crystal curved to focus all the X-rays onto detector so as to maximise the collection efficiency.

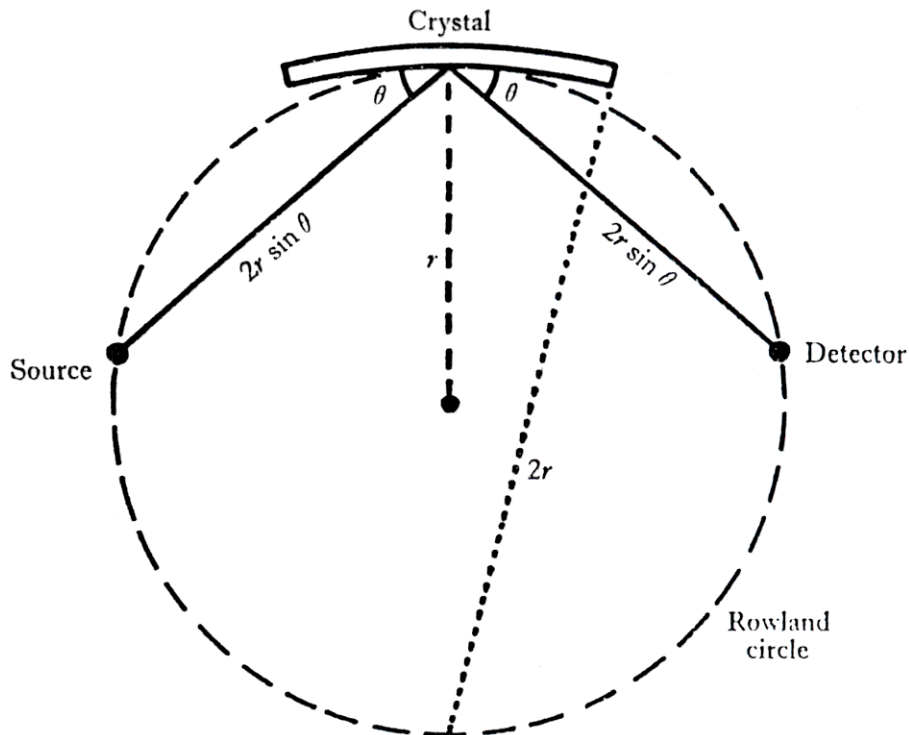


X-ray diffraction from a curved and ground crystal (Johansson focusing spectrometer). The focusing crystal minimises signal loss

Reed p 78

To ensure that the X-rays hit the crystal at constant Bragg angle the crystal must be bent to a radius of $2r$, where r is the radius of the Rowland circle, as shown below. However, accurate focusing requires the front surface of the crystal to lie on the Rowland circle and thus have a radius r .

Thus ideally the crystal is bent to a radius of $2r$ and the front surface ground to a radius of r . In practice grinding introduces defects and it is often better to leave the crystal with a radius of $2r$.

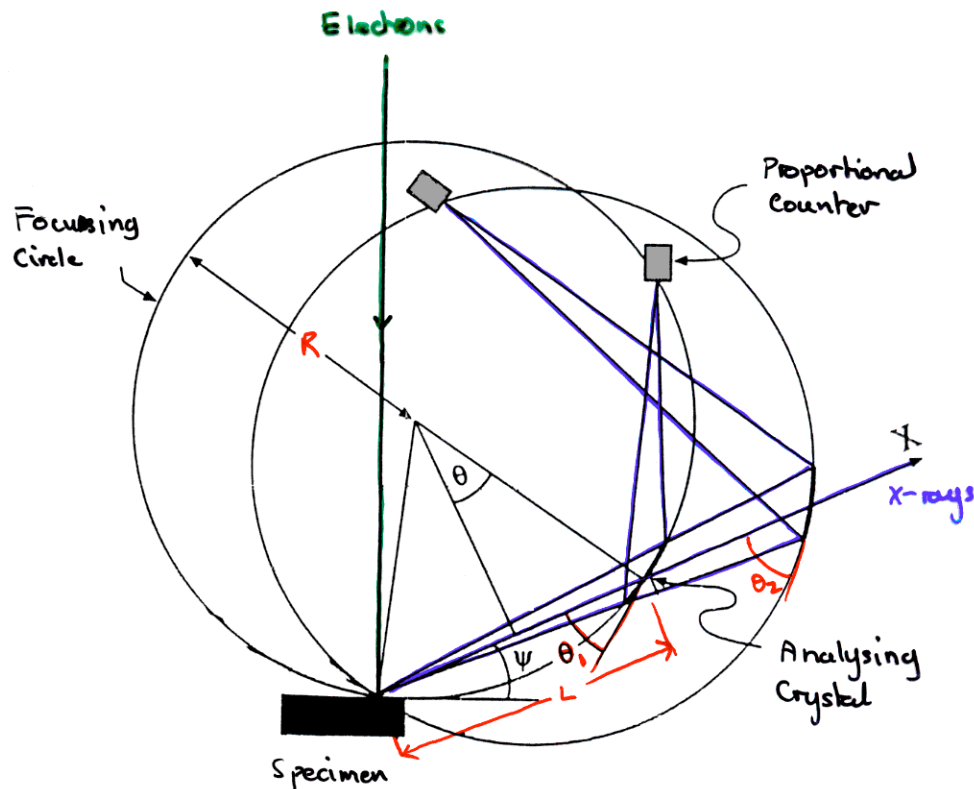


Geometry of a focusing spectrometer

Reed p 76

The above geometry works for focusing onto the detector all the X-rays of one wavelength emitted from the sample. To record a spectrum, we need to be able to change the wavelength recorded so as to be able to scan through a range of wavelengths.

The wavelength can be changed by changing θ , the Bragg angle, and as a consequence the position of the detector must change to keep the focusing geometry. At the same time, the take off angle must remain constant as X-ray emission changes with take-off angle from the specimen. The resulting arrangement is called linear focusing geometry.



Linear focusing spectrometer geometry. The spectrometer is shown in two positions corresponding to analysing two wavelengths. The crystal is moved linearly away from the sample and the crystal and detector are constrained to move on the focusing (Rowland) circle.

Williams and Carter p 570

To maintain a constant radius of the focusing circle (and hence constant radius of bent crystal) the crystal must move linearly away from the specimen. This gives rise to the result that the wavelength detected is proportional to the distance of the crystal from the specimen:

$$\lambda = \frac{dL}{R}$$

where

d = crystal plane spacing

R = focusing circle radius

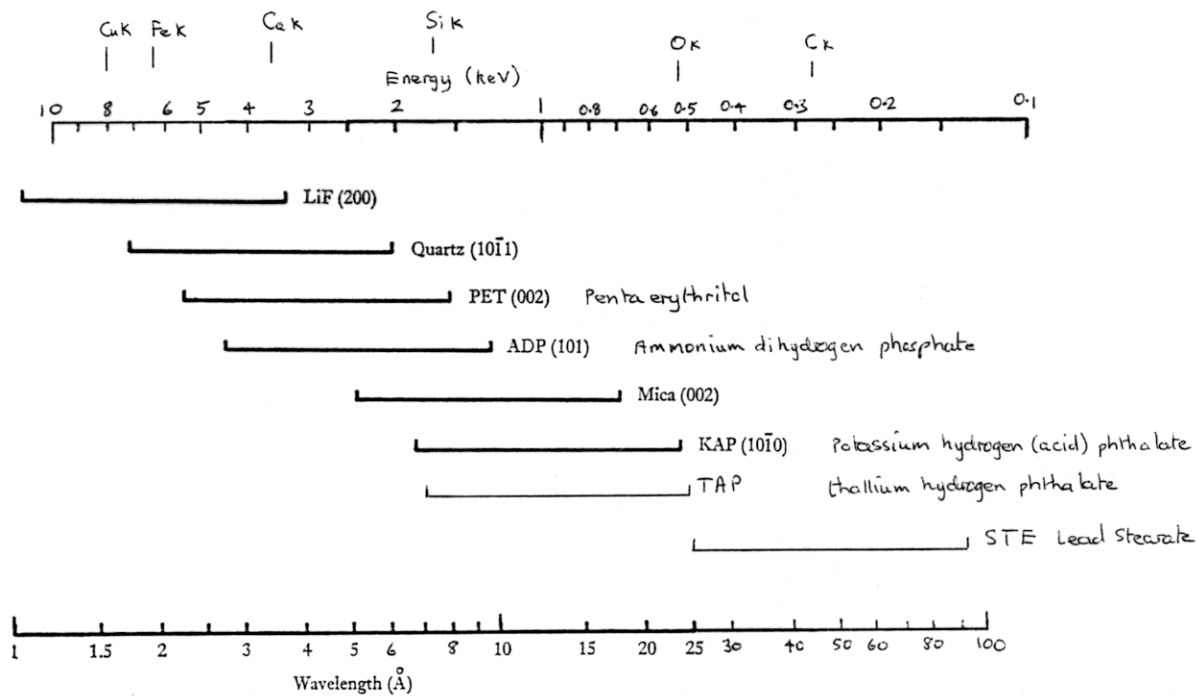
L = distance from specimen to crystal

The geometry limits the range of θ from typically

$$15^\circ < \theta < 65^\circ,$$

giving $100 \text{ mm} < L < 350 \text{ mm}$

and thus limits the wavelength range a given crystal can analyse. Hence more than one crystal must be used to analyse a large range of wavelengths.



Wavelength range of first order reflections for common crystals

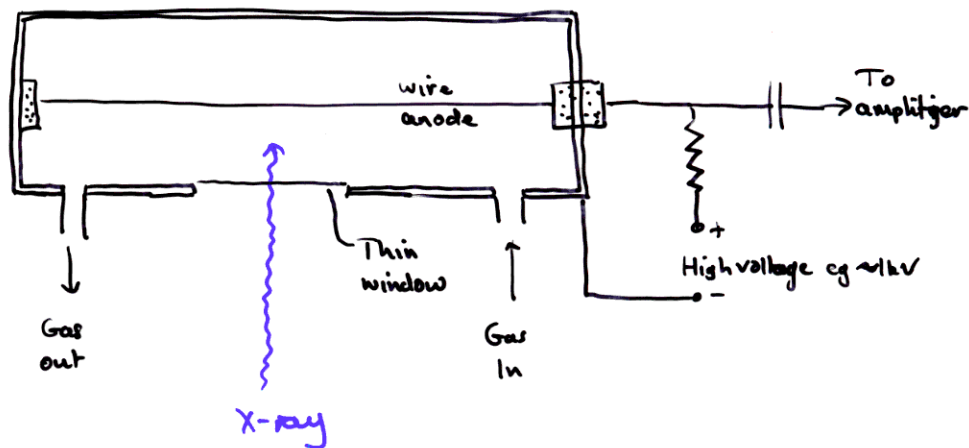
Reed p 83

The choice of which crystal to use is governed by:

- d spacing and hence wavelength,
- highest reflectivity for lowest 2θ ,
- second order reflections ($n\lambda = 2d \sin \theta$),
- intrinsic reflectivity of the crystal (eg LiF and PET have high reflectivity, mica has low reflectivity),
- intrinsic resolution of crystal (eg quartz and mica have high resolution),
- X-ray fluorescence, eg Fe and Ni can fluoresce P X-rays in KAP.

3. Detectors and proportional counters

The X-ray detector needs to amplify the signal from a single X-ray and produce a signal that is proportional to the X-ray energy. This is fed into the electronics for counting. X-ray detection is done by a proportional counter, which may be sealed or more normally a gas flow proportional counter.



Gas flow proportional counter

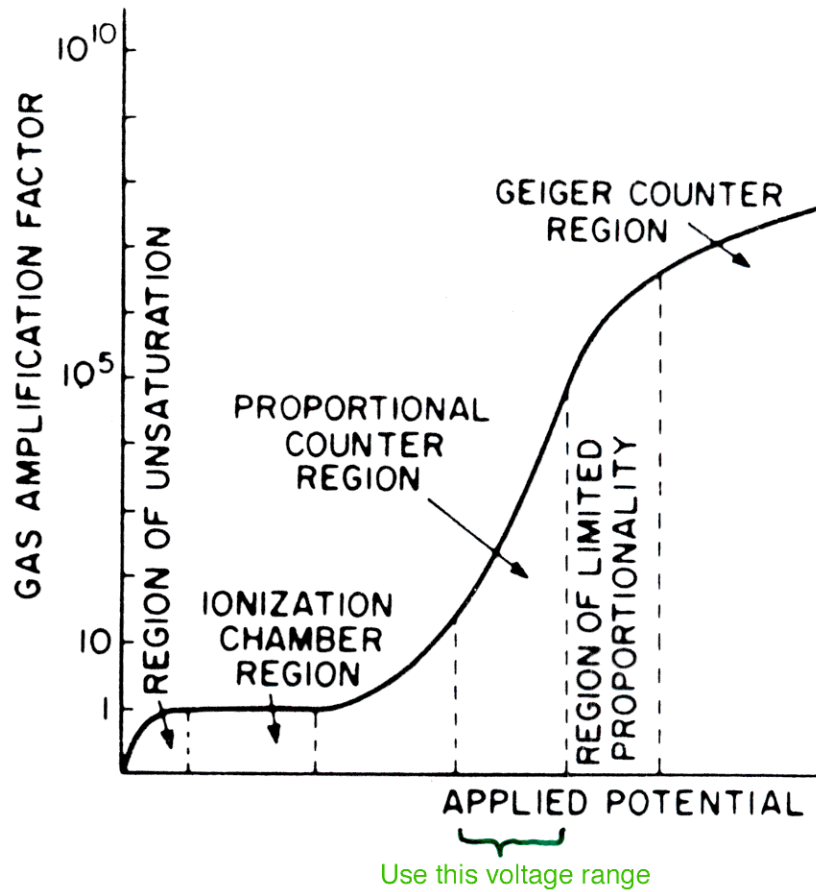
Important parts of the proportional counter are:

Window: this is Be for sealed counters (no gas flow), but Be absorbs low energy X-rays. Organic windows (eg 12 μm mylar, transmits 2-3% of Na K or 4 μm polypropylene, transmits 50% of Na K) transmit more low energy X-rays but are permeable to gas, hence the need for continuously flowing gas.

Gas: typically Ar plus 2.5% CO_2 or 10% methane.

Voltage: typically 1 to 3 kV. The voltage must be in proportional region as defined below.

Each ionisation event requires about 30 eV, so for example for a Cu K_α X-ray, which has an energy of 8 keV, about 300 electrons would be produced. This is too small to detect on its own, but if the potential of the wire is at the correct voltage amplification occurs.



Gas amplification in a proportional counter

4. Pulse processing electronics

The stages in processing pulses from the proportional counter are:

Preamplifier and amplifier: These amplify and shape the pulses from the proportional counter into gaussian shapes. The width of the gaussian depends on the time constant. A long time constant gives a high amplitude signal and thus less noise, but means more overlapping pulses.

Discriminator (pulse height analyser): X-rays of the same energy should give rise to gaussian pulses of the same shape and height. The discriminator selects pulses of a given height and produces a uniform output from these, rejecting all other pulses. Other pulses present due to:

noise

second order reflections (higher energy thus larger pulse)

sum peaks (two X-rays arriving simultaneously)

overlapping pulses

escape peaks – when an Ar K X-ray is lost from the counter so too little charge is collected.

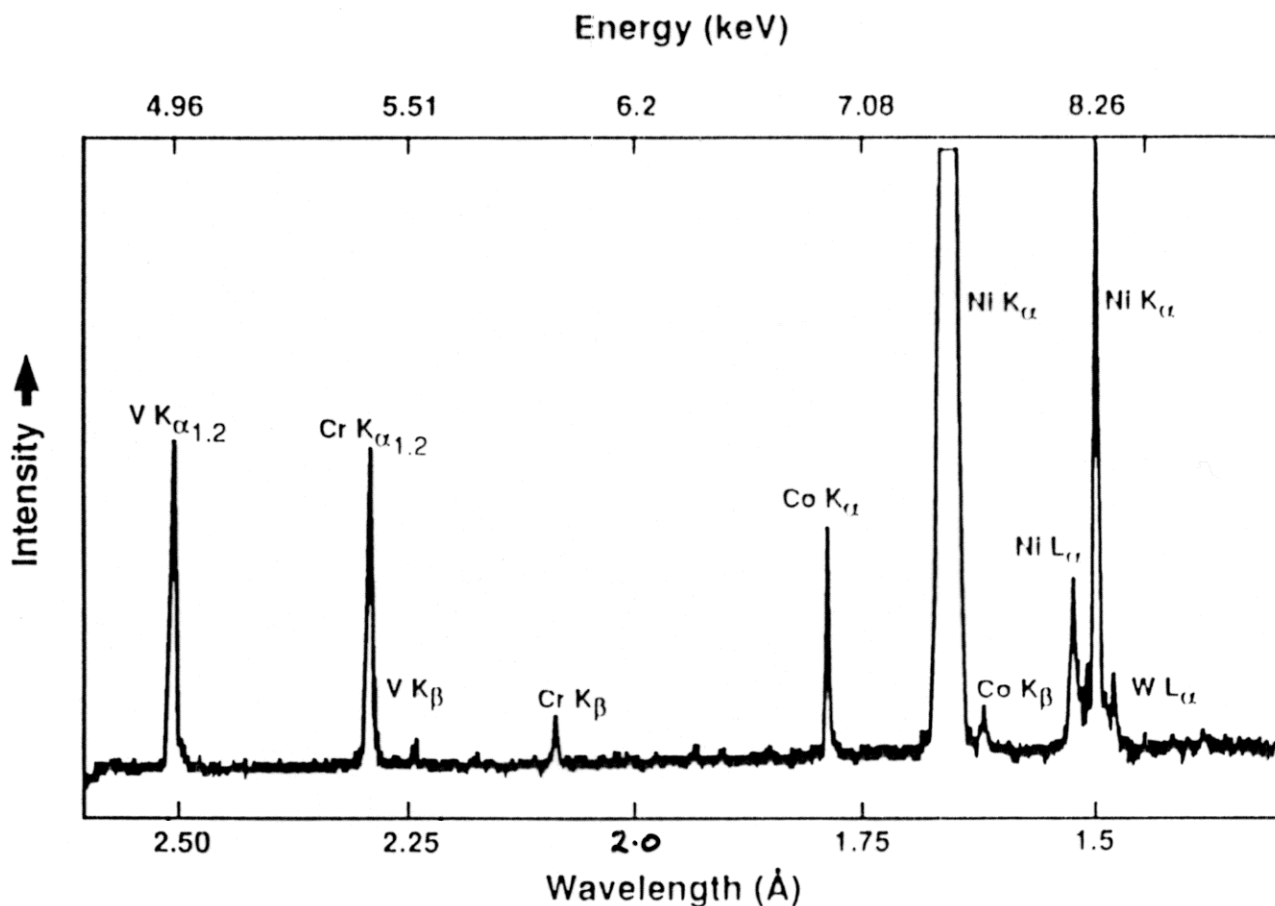
Computer: counts pulses and displays spectrum as a function of wavelength (detector position).

5. Dead time

The detector and electronics takes a short time to process a pulse. Thus after a pulse arrives there is a short time while the pulse is processed during which the system cannot respond to further pulses. This is the dead time and is typically about $2 \mu\text{s}$. X-rays arriving in the dead time are lost. The dead time can be corrected for.

If n_{obs} is the observed count rate and τ is the dead time then in 1 sec the system is dead for $n_{obs} \cdot \tau$ sec and live for $1 - n_{obs} \cdot \tau$ sec. Thus

$$\text{true counts} = \frac{n_{obs}}{1 - n_{obs}\tau}$$



Typical WDX spectrum. Ni base superalloy using a LiF crystal

Goldstein p 290

6. WDX advantages

WDX spectrometers are now only used on dedicated SEM based microanalysers. These often contain 4 or 5 crystal spectrometers to allow measurement of many elements simultaneously. TEMs and most SEMs (such as those in our department) use EDX detectors.

WDX spectrometers do have some advantages over EDX:

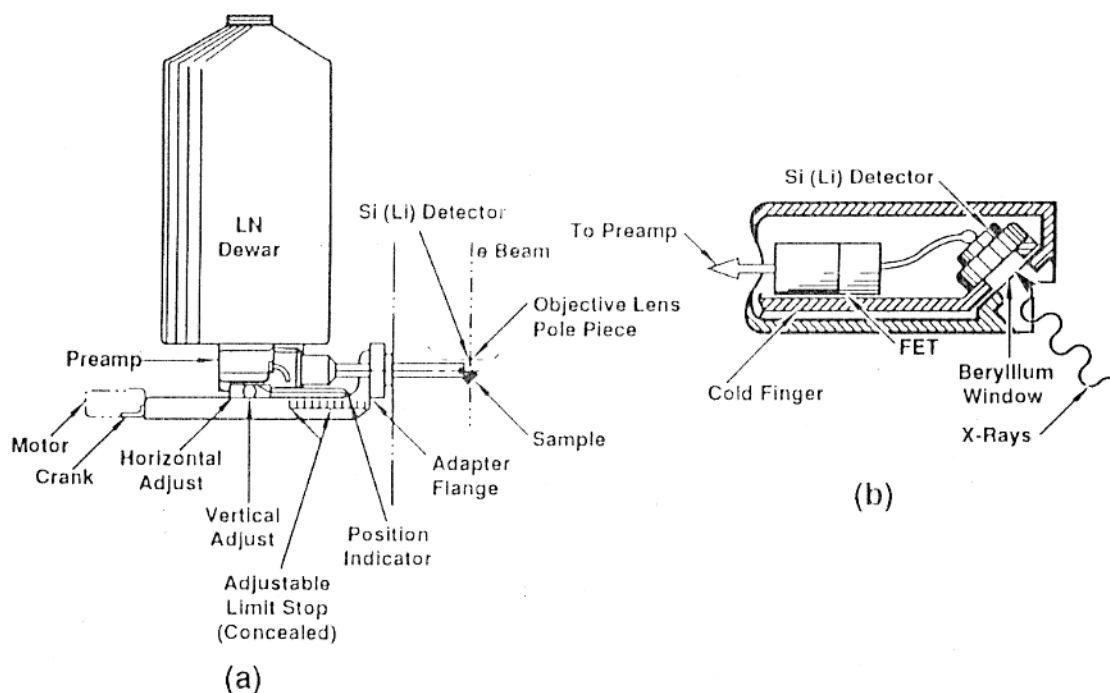
- Better energy resolution (~ 10 eV) so fewer peak overlaps. It is even possible to detect chemical shifts
- Better peak to background ratio allowing smaller amounts of elements to be detected
- Better detection of light elements
- No artefacts from detection electronics. The only artefact is higher order lines from Bragg equation.
- Higher count rates possible.

3. Recording spectra: energy dispersive X-ray spectroscopy (EDX)

In this section I will look at how an EDX spectrometer works and how it detects and processes X-rays.

1. Basis of energy dispersive X-ray spectroscopy (EDX)

In EDX a Silicon (or sometimes germanium) chip is used as both an X-ray monochromator and a proportional counter. When an X-ray hits the Si chip it creates electron-hole pairs. These electron-hole pairs can be counted by an amplifier, resulting in a pulse whose height is proportional to the X-ray energy. The pulses are counted as a function of energy and displayed on a computer. Below is a diagram of a typical EDX detector.



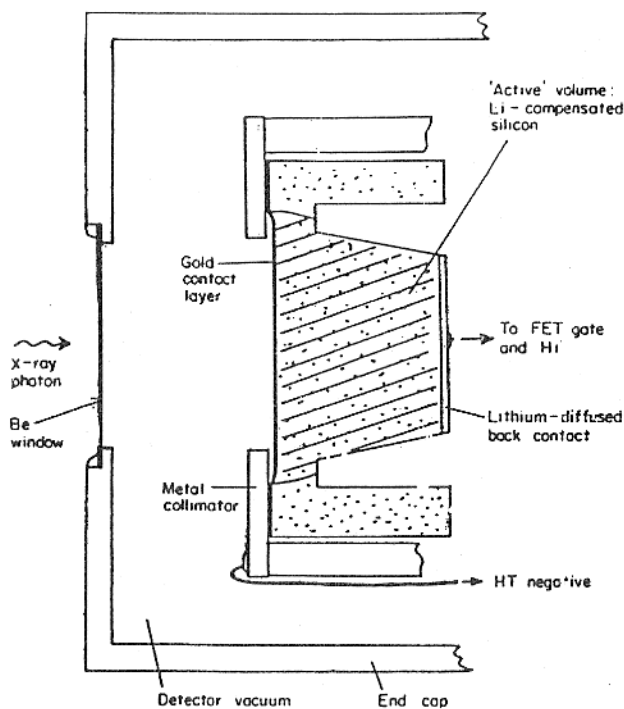
Energy dispersive X-ray spectrometer
 (a) Physical appearance of EDX detector
 (b) Detail of Si(Li) detector assembly

Goldstein p 295

2. Hardware

a. Detector

The Si crystal that detects the X-rays is a reversed biased p-i-n (p-type, intrinsic, n-type) detector. This forms a diode. The p-type layer is at the front of the crystal and the n-type layer at the back. Contacts are made to the front and back surfaces of the crystal. Adjacent to the contacts is a “dead layer” where charge is not completely detected.



Cross-section of Si(Li) crystal and end of EDX detector

The detection of X-rays takes place in the intrinsic layer in the middle of the detector. In intrinsic silicon all the electrons are in the valence band with none in the conduction band. An X-ray ejects a photoelectron in the Si which loses its energy by promoting electrons from the valence band to the conduction band forming electron-hole pairs. A reversed bias of 500 to 1000 volts is used to ensure most of the electron-hole pairs are collected.

It is not possible to make silicon pure enough to be intrinsic, it usually contains acceptor impurities and thus acts as a p-type semiconductor. The acceptor sites are thus “filled” with Li, and the Si is then called “lithium drifted silicon” or Si(Li).

The energy gap for Si is 1.1 eV but normally on average 3.8 eV is required to form an electron-hole pair as not all the energy creates electron-hole pairs. Thus for example a Cu K_{α} X-ray will create $8040/3.8 = 2300$ electron-hole pairs or about 10^{-16} C. This is a very small signal to measure.

The signal thus needs to be amplified by about 10^{10} , hence there is a need to minimise noise. The detector and preamplifier are thus cooled with liquid nitrogen in order to:

reduce thermally generated electron-hole pairs

prevent the Li atoms from diffusing which would otherwise occur under the applied bias

reduce noise in the FET preamplifier.

b. Window

Cooling the detector causes ice and hydrocarbons from the microscope vacuum to condense on the detector thus reducing its efficiency. This can be prevented by separating the detector from the microscope vacuum with a window. There are 3 distinct types of window:

Be window

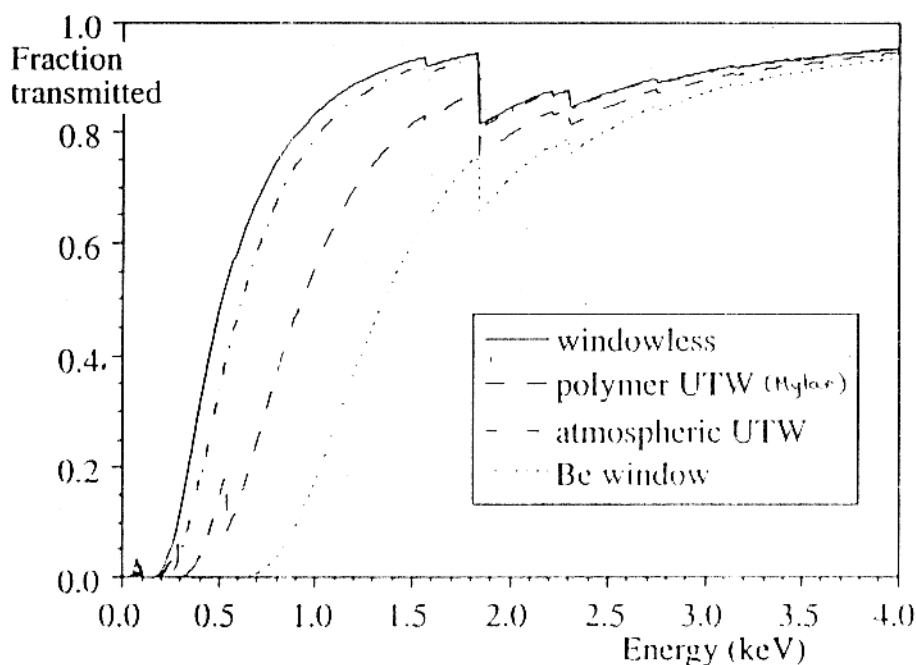
A thin sheet of Be ideally around $7\ \mu\text{m}$ but more usually up to $12\ \mu\text{m}$ thick which is able to withstand atmospheric pressure (as when air is let into the microscope chamber). This has the disadvantage that it absorbs the lower energy X-rays and thus Be window detectors cannot detect elements with Z less than 11 (Na)

Ultrathin window

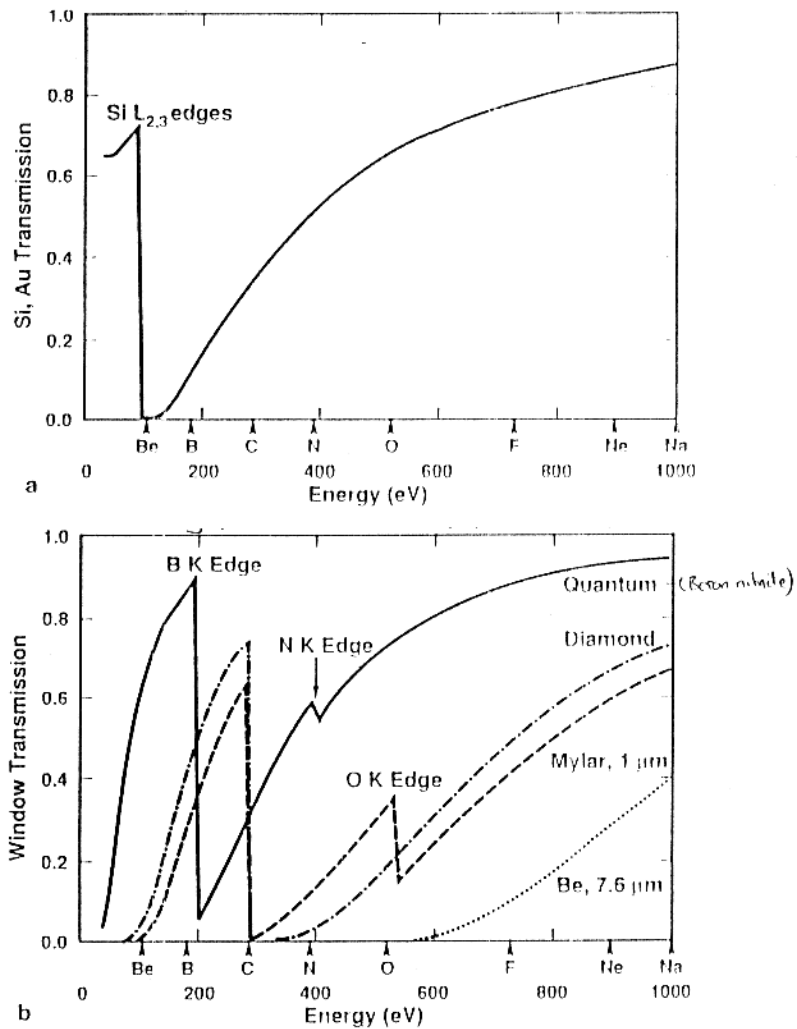
These are thin ($<100\ \text{nm}$) films of polymer, diamond, boron nitride or silicon nitride all of which can withstand atmospheric pressure. Early ultrathin window detectors used a thin polymer film that could not withstand atmospheric pressure and thus had to be withdrawn behind a vacuum valve whenever a specimen was exchanged. Ultrathin windows allow through more low energy X-rays thus allowing lighter elements to be detected. Each different window material absorbs X-rays differently, eg carbon containing windows absorb N K X-rays strongly. Most ultrathin window detectors can detect elements with Z from 6 (C) upwards.

Windowless

Windowless detectors have the highest collection efficiency and are typically able to collect elements with Z from 5 (B) upwards and at their best can detect Be. However they need a good ($<10^{-8}\ \text{Pa}$ ($<10^{-6}$ torr or mB)) vacuum free of water vapour and must be withdrawn when changing specimen. They are generally only suited to UHV TEMs and STEMs.



X-ray detector efficiency

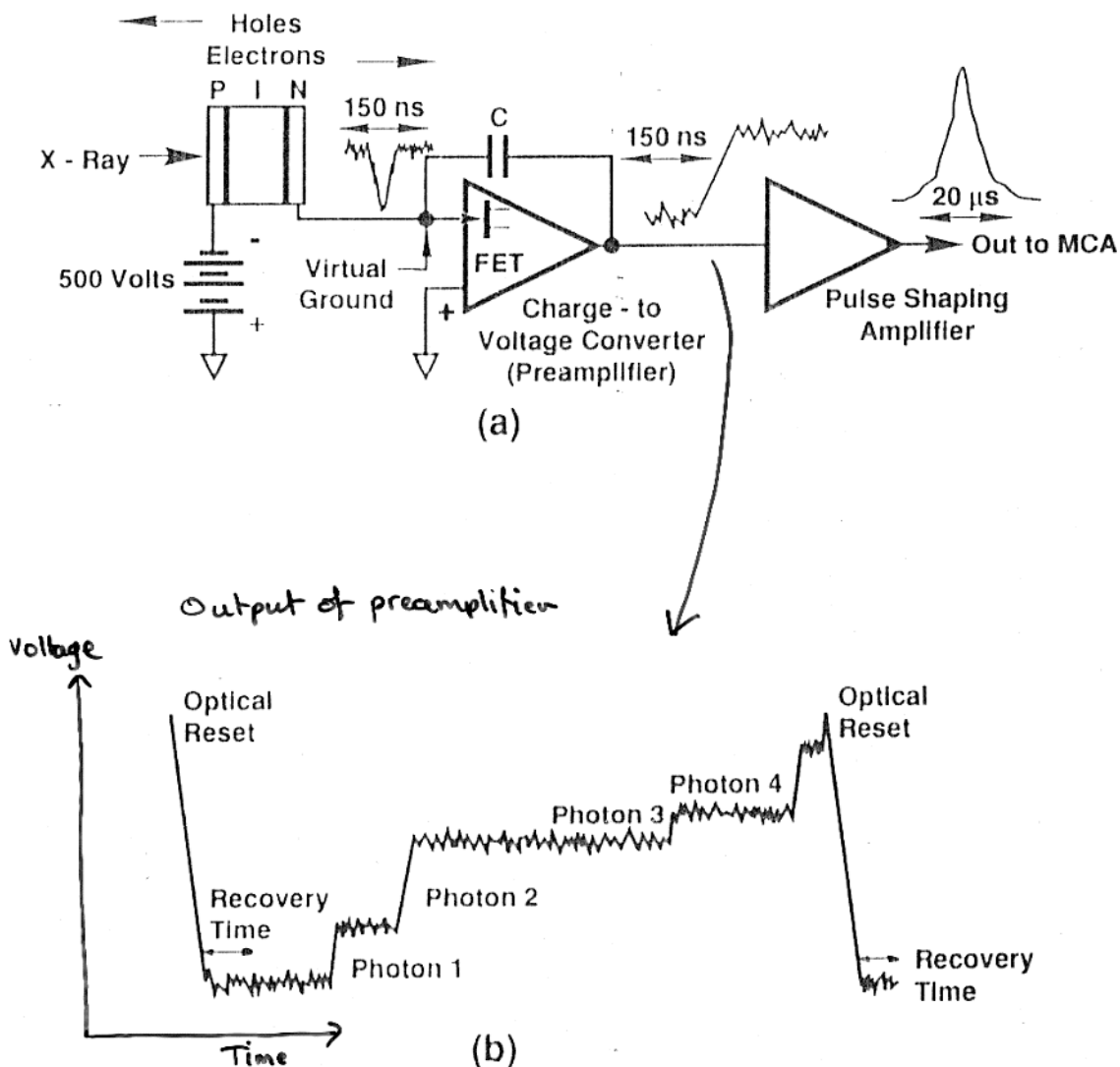


(a) Transmission curve for a "windowless" detector. This measures the transmission of the 10 nm Au contact and 100 nm Si dead layer.

(b) Transmission curves for several window materials. The actual transmission of the detector is the combination of (a) and (b).

c. Electronics

The Si detector has no internal gain so the electronics that process the pulses must be very low noise.



EDX electronics – signal processing

(a) Detector charge to voltage converter and pulse shape linear amplifier

(b) Output of charge-to-voltage converter after a series of photons

Goldstein p 297

The stages in processing are:

Preamplifier

This consists of a field effect transistor (FET) mounted close to the Si crystal and cooled to $\sim 100\text{K}$ along with the Si crystal to minimise stray capacitance and noise pick up. The Si detector produces a short current pulse proportional to the incident X-ray energy. The preamp acts as a charge to voltage converter and integrates this signal. This means its output rises with time and must be reset. A resistor is too noisy so use “pulsed opto-feedback” to reset when voltage gets too high. Typical pulse width at this stage is 150 ns.

Amplifier

This converts the stepped output of the preamp of few millivolts into shaped pulses of several volts height and around 20 μs width.

The time constant of this amplifier determines the trade off between X-ray energy resolution and maximum count rate. Short time constants (eg 10 μs) allow higher count rates before the detector is saturated, while long time constants (eg 40 μs) allow better energy resolution. This time constant can usually be set manually.

Processing

There are various stages involved in processing the pulses, including:

Pulse pile up rejecter: this rejects overlapping peaks by terminating processing if another pulse arrives while one is being processed. For this a second amplifier is used with a short time constant. If 2 pulses arrive too close together then both are ignored.

Noise peak or zero energy strobe: the natural electronic noise can be allowed through into the final spectrum. This allows the zero energy to be calibrated accurately and the electronic noise to be measured.

Analogue to digital converter (ADC): this converts pulse height (proportional to X-ray energy) into a digital signal. For accurate digitisation this may take up to 10 μs per pulse.

Multi channel analyser (MCA) or computer: Typically 1024 channels of 20 eV width to cover 0 to 20 keV. Each pulse is digitised and 1 added to the corresponding bin in the MCA. For a given X-ray pulse the X-ray energy it is assigned to depends on the gain of the amplifier which can drift over time. Thus the system must be calibrated by acquiring a spectrum from a known element such as Co or Cu. Most EDX software includes calibration using the zero strobe (or an L line if no strobe is present) and a known element when the system is started.

4. Recording spectra: energy dispersive X-ray spectroscopy (EDX) (continued)

3. Energy resolution

EDX spectra typically have a resolution of about 120 eV. It is important to understand the reasons for this resolution and how it can be improved. Also an accurate measure of the detector resolution is needed to analyse spectra quantitatively. The resolution depends on a number of factors.

a. Ionisation statistics

Ideally each X-ray photon of the same energy would give the same number of electron hole pairs

$$N = \frac{E}{\varepsilon}$$

where ε = electron-hole pair energy = 3.8 eV for Si

E = X-ray energy

If the ionisation were ideal as above then the standard deviation width for N pairs would be

$$\sigma_N = 0$$

Alternatively if the electron-hole pair formation is random (Poisson) then

$$\sigma_N = \sqrt{N}$$

In practice electron-hole pair formation is somewhere between these extremes

$$\sigma_N = \sqrt{NF}$$

where F = Fano factor, typically around 0.125 for Si

Converting to energy

$$\sigma_E = \varepsilon \sigma_N = \sqrt{EF} \varepsilon$$

If the peak is gaussian, then its full width at half maximum (fwhm) is

$$\Delta E_{\text{ionisation}} = \sqrt{8 \ln(2)} \sigma_E = 2.355 \sqrt{EF} \varepsilon$$

Thus $\Delta E_{\text{ionisation}} \rightarrow 0$ as $E \rightarrow 0$

For a Si X-ray detector, $\Delta E_{\text{ionisation}} \approx 125$ eV at the energy of Mn K_{α} , 5.898 keV.

This is the fundamental limit of resolution. The only way to improve it is to reduce ε , ie reduce the energy per electron-hole pair and thus increase number of electron-hole pairs. For example by using a Ge detector instead of Si.

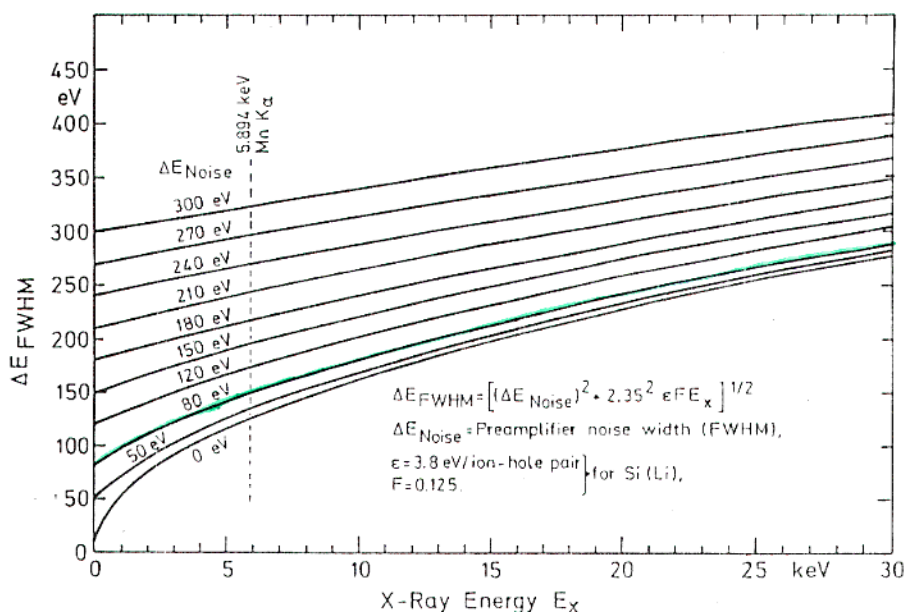
b. Noise

Noise is caused by a DC leakage current in the Si detector, thermal noise in the Si and the FET and electronic noise in amplifiers. It can be measured by the width of the strobe peak. For a typical EDX system $\Delta E_{noise} \approx 80$ to 85 eV.

c. Overall resolution

The overall resolution of the detector depends on the sum of both the ionisation statistics and detector noise resolution limits discussed above.

$$\begin{aligned} \text{Resolution } \Delta E_{fwhm} &= \sqrt{\Delta E_{noise}^2 + \Delta E_{ionisation}^2} \\ &= \sqrt{\Delta E_{noise}^2 + 2.355^2 EF\epsilon} \end{aligned}$$



EDX resolution (ΔE_{fwhm}) measured as the full width half maximum of the Gaussian peak of an X-ray line of energy E_x for different levels of amplifier noise

Reimer, SEM p 205

For a modern detector the overall resolution at the energy of Mn K_α (5.898 keV) is 130 to 150 eV.

d. Consequences

1. EDX peaks are broad (~ 130 eV) cf WDX (~ 10 eV). This gives a lower peak to background in EDX than in WDX and the detection limit in EDX is $\sim 0.1\%$
2. The above equation can be used to predict the peak width for curve fitting in quantitative analysis.
3. The energy resolution improves as E decreases, but the separation of K lines of adjacent elements decreases as E decreases. Hence it is possible to separate elements of adjacent atomic number more easily for high Z elements than for low Z .

4. Dead time

The time constant τ is the time that the detector takes to process a pulse. It is typically from a few to 40 μs and can usually be changed by the user.

Short τ allows higher count rates to be processed. Long τ gives better energy resolution as the electronics has more time to process the pulse.

The detector cannot process further pulses while one is being processed hence it is “dead” for this time. If a further pulse arrives while the first is being processed then the dead time is increased.

Intervals between pulses obey the Poisson distribution so If R_{in} is the rate X-rays (of all energy) are hitting the detector then the observed count rate (always smaller than R_{in}) is given by

$$R_{out} = R_{in} \exp(-R_{in} \tau)$$

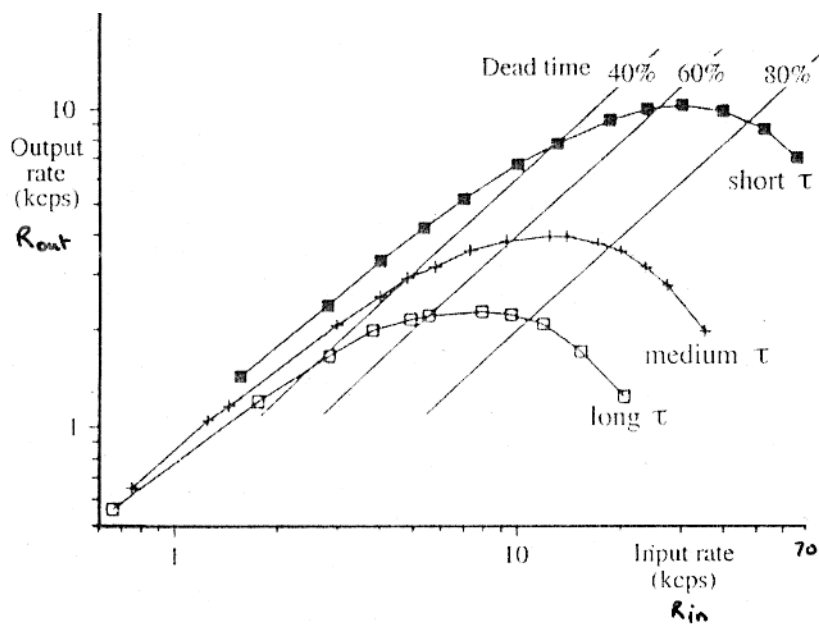
The dead time is usually quoted as a percentage:

$$\begin{aligned} R_{out} &= \left(1 - \frac{R_{out}}{R_{in}}\right) \times 100\% \\ &= 1 - \exp(-R_{in} \tau) \times 100\% \\ &= \frac{\text{time} - \text{live time}}{\text{time}} \times 100\% \end{aligned}$$

where time = actual time taken to collect the spectrum

live time = time the detector is live for, always less than time

Generally a spectrum is collected for a given live time (eg 100 s), so the actual time to take the spectrum will always be larger and will depend on the dead time and thus the input count rate, R_{in} .



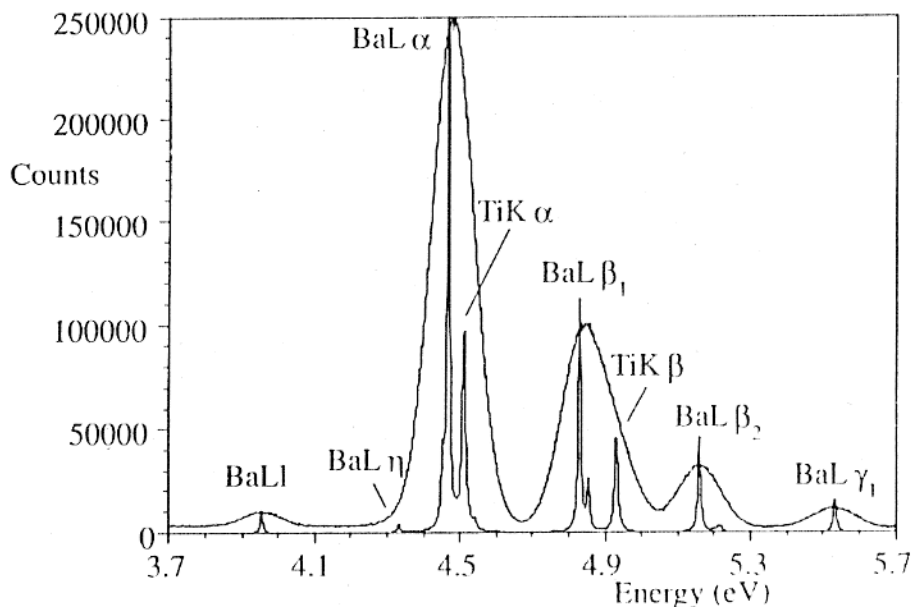
Relation between output count rate and input count rate. The output count rate takes account of the dead time

The maximum processing efficiency is reached at about 50% dead time. Much higher input count rates saturate the detector (100% dead time) In practice it is good to work with between 30 and 50% dead time.

5. Comparison of WDX and EDX

Approximate typical values of parameters for WDX and EDX spectrometers

	WDX	EDX
Energy resolution	10 eV	140 eV
Max count rate	50,000 counts/sec on one peak	3000 counts/sec over all spectrum
Spectrum acquisition time	30 min	1 min
Detection limit	0.005%	0.1%
Peak to background ratio	1000	50
Artefacts	High order lines	Sum peaks Escape peaks Si K internal fluorescence etc

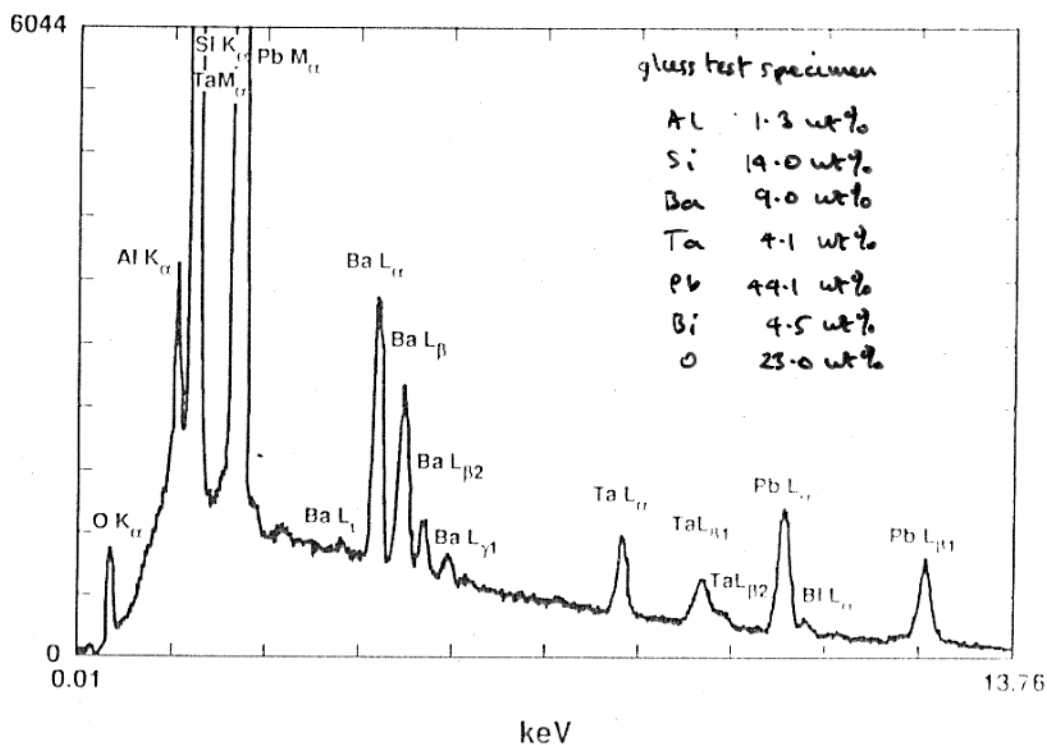
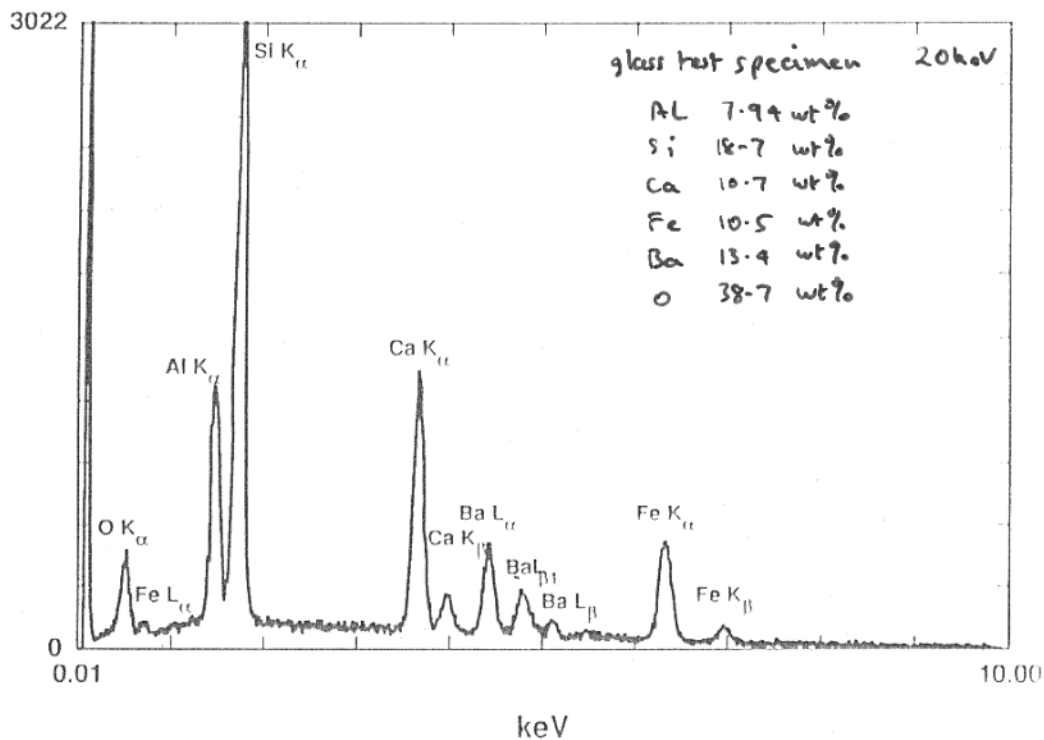


Comparison of WDX and EDX spectra from BaTiO₃. The broad peaks are from EDX and the sharp peaks from WDX.

Williams and Carter p 571

Notice that the Ti and Ba peaks cannot be separated using EDX.

6. Interpretation of X-ray spectra



EDX spectra from glass test specimens of two different compositions

Goldstein, top: p 356, bottom: p 358

Steps to interpret an EDX spectrum of an unknown material, for example the spectra above.

1. Ensure incident beam energy is high enough, for quantitative work need beam energy > 2 × highest peak energy.

2. Ensure spectrum is reliable:

- repeat spectrum from same area or similar area
- ensure there are enough counts so that peaks are not lost in the noise of the background
- ensure count rate is not too high, <3000 counts/sec and dead time <50% to minimise sum and escape peaks.

3. Use prior knowledge of sample to know which elements are likely to be present and which are not likely to be present.

4. Confirm elements by looking for other peaks from that element.

5. Work from high energy to low energy identifying peaks. At high energy there are fewer peaks and better resolution of neighbouring peaks.

6. Peak shapes and energies:

For energy range 0 to 20kV (typical of most spectrometers)

- *K series* – elements B (Z=4) to Ru (Z=44)
if $Z > \sim 16$ (S) look for K_{β} peak, intensity 10 to 20% of k_{α}
 $K_{\alpha 1}$ and $K_{\alpha 2}$ not resolvable
peak shape gaussian with shoulder when K_{β} peak just resolvable

- *L series* – elements Cl (Z=17) up
If $Z > \sim 42$ (Mo) look for L_{β} and others.

Intensities:

L_1	4%
$L_{\alpha 1}$	100%
$L_{\beta 1}$	70%
$L_{\beta 2}$	20%
$L_{\gamma 1}$	8%
$L_{\gamma 3}$	3%

- *M series* – elements ~Ag (Z=47) up. Highest M peak (U, Z=92) at 3.2kV
M alpha and M beta not resolvable.
Peaks gaussian with small shoulder for highest Z
- Confirm presence of element by looking for peaks from all series (K, L, M)

7. Beware of peak overlaps, eg

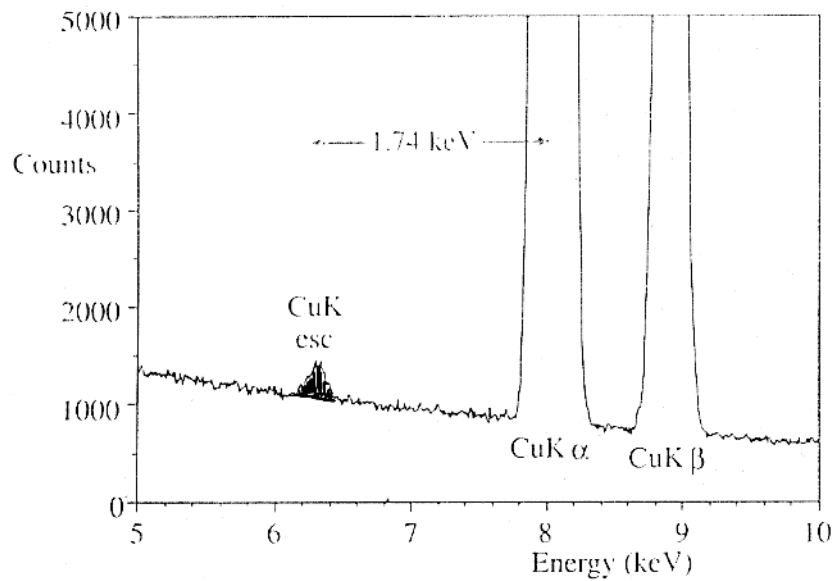
S K 2.31 keV	Mo L 2.29 keV	Pb M 2.35 keV
N K 0.39 keV	Ti L 0.45 keV	

7. Artefacts

There are many possible artefacts in EDX spectra, some of which can be mistaken for peaks. Below is a list of the main artefacts to beware of.

1. *Escape peaks*

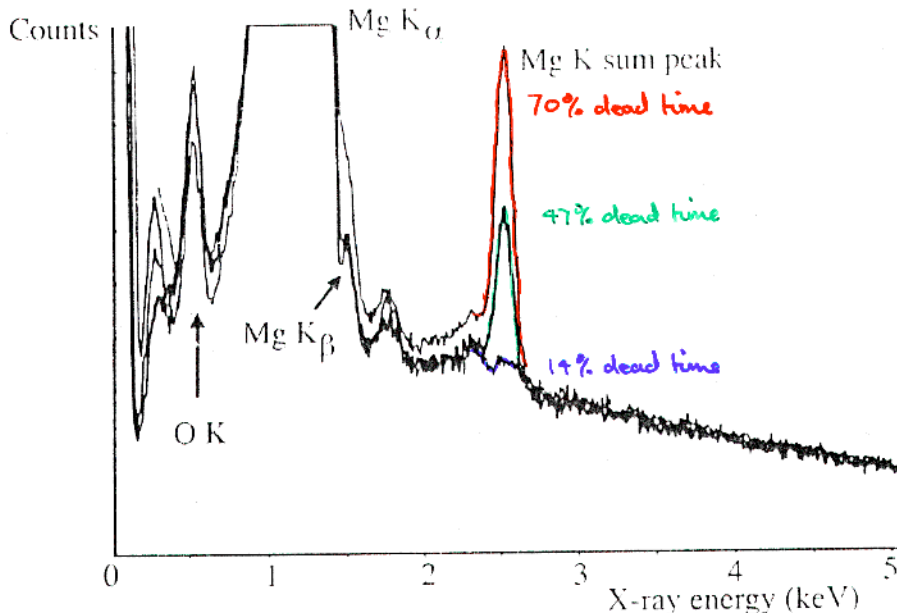
Escape peaks occur when a Si K X-ray escapes from the Si detector. The energy recorded is lower by the energy of a SiK X-ray. You see a peak 1.74 keV below any intense peak. The intensity is around 0.2% to 2% of the main peak. They are most often seen for elements between P K (2.0 keV) and Zn K_{α} (8.6 keV)



Escape peak in Cu, 1.74 keV below the Cu K_{α} peak

2. Sum peaks

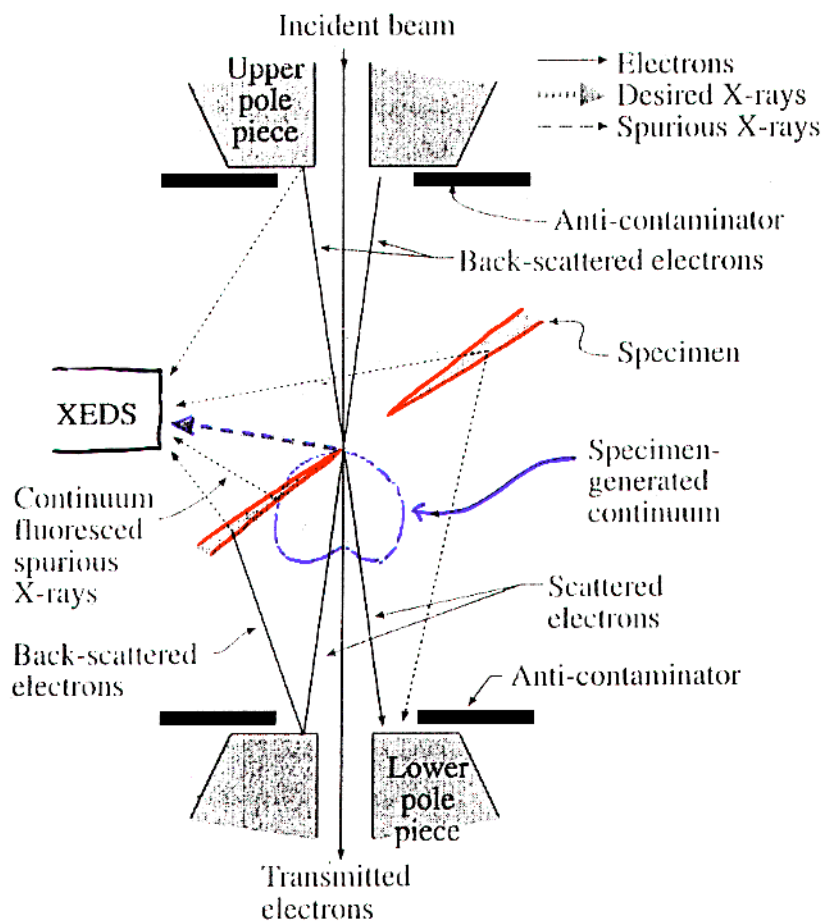
Sum peaks appear when 2 X-rays arrive at the same time. The electronics cannot distinguish them from one X-ray. You see a peak at double the energy of strong peaks. This happens for high count rates and thus high dead times (typically >50%)



Sum peak in Mg for various dead times. No sum peak is present for 14% dead time and below

Williams and Carter p 569

3. Stray radiation peaks, secondary fluorescence



Sources of stray X-rays generated in TEM when the electron beam is scattered by a tilted specimen.

There are many possible sources of stray radiation. X-rays can be collected from various parts of the microscope chamber, detector, sample holder or parts of the sample away from the area of interest. They can be created due to backscattered electrons and/or X-rays hitting parts of the specimen and/or chamber and exciting secondary X-rays. This is especially a problem for TEM where there is much less space. For this reason it is important to remove the objective aperture, it is particularly good at scattering electrons back onto the specimen.

Typical stray radiation peaks are:

Fe, Cr, Ni from stainless steel

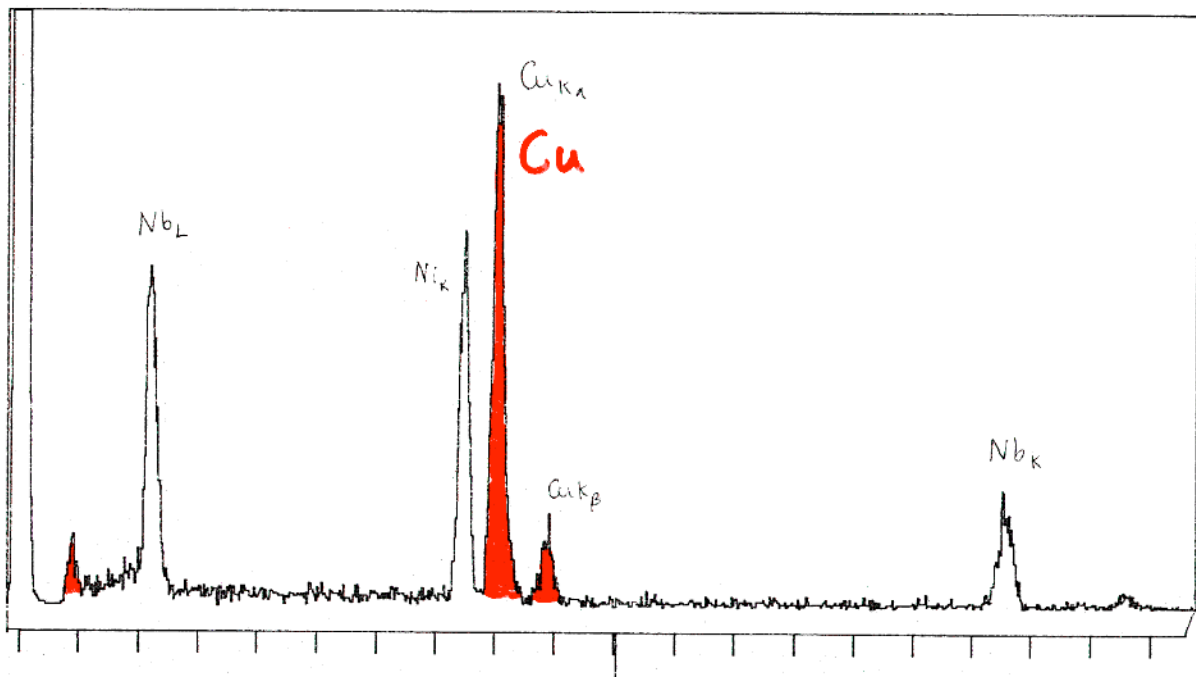
Cu, Zn from brass specimen holders

Al from SEM stubs

Cu from Cu grids supporting specimens (especially TEM)

Mo/S

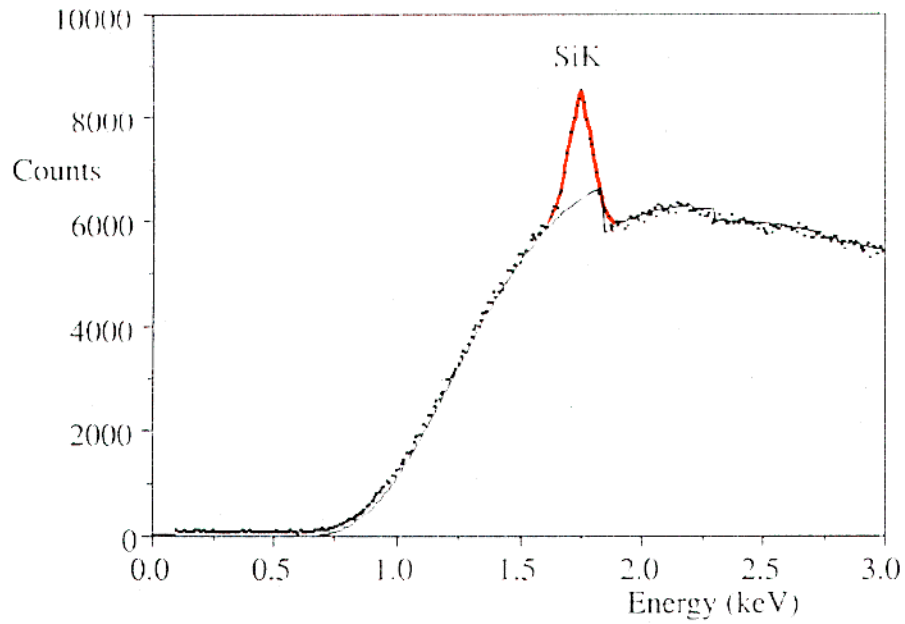
Pt or Mo from objective aperture in TEM.



Amorphous Ni-Nb film on a Cu TEM grid. Note the large Cu peak present even though the beam passes through the middle of a grid square. The Cu X-rays come from secondary fluorescence and stray radiation generated in the microscope subsequently illuminating the whole of the specimen.

4. Si fluorescence peak

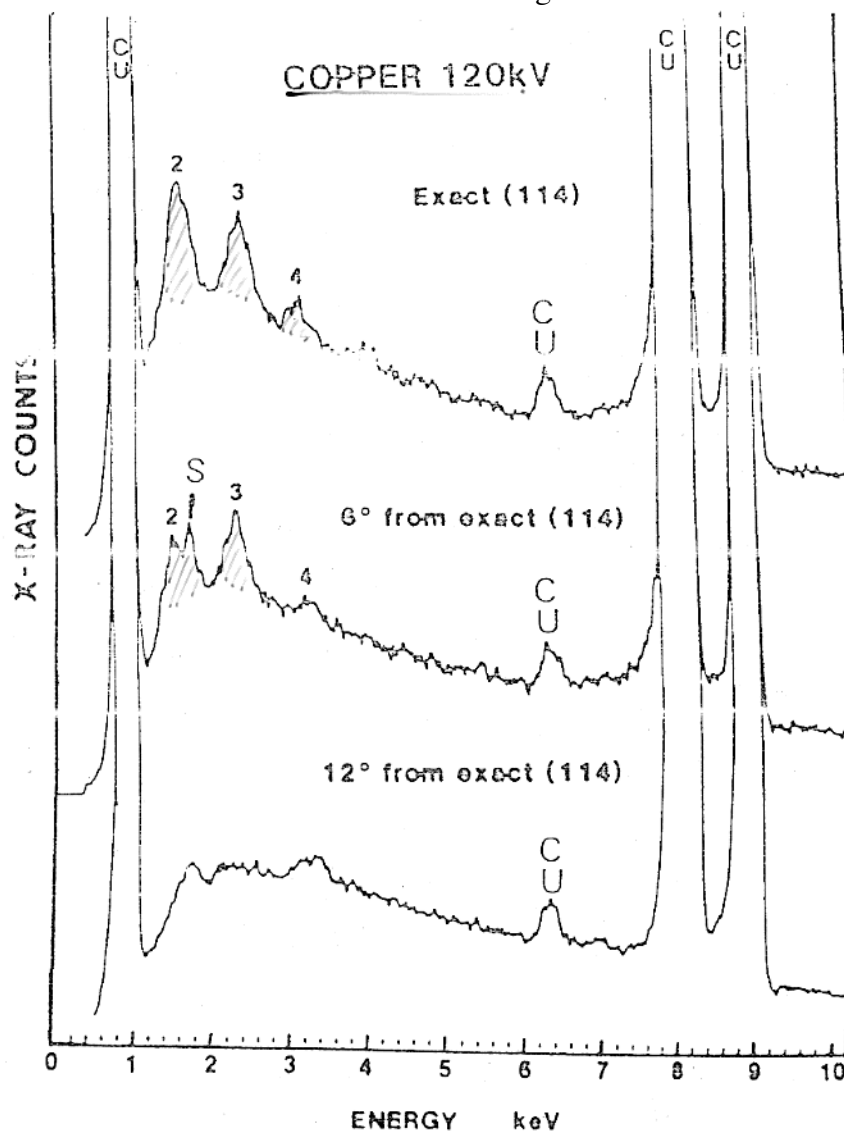
This peak is due to fluorescence of the Si dead layer in front of the detector. You see a small Si peak even when no Si present.



Si internal fluorescence peak in a spectrum from pure C obtained with a Si(Li) detector.

5. Coherent bremsstrahlung

Coherent bremsstrahlung can produce 1 to 3 small peaks at low energies. It comes from crystalline materials near to or at a zone axis. It can be mistaken for eg Si or S or similar elements.



Coherent bremsstrahlung. Two or three small peaks at low energy which can be mistaken for low Z elements, eg Si or S. Coherent bremsstrahlung peaks appear in crystalline materials when the beam is down a zone axis.

6. Rough or bent specimens

If the surface is rough then the angle the beam hits the specimen at will be different from the angle the specimen is tilted to. This will change the absorption length and if very rough (SEM) or bent (TEM) may obscure the direct line from where the beam hits the specimen to the detector.

5. Quantitative analysis

In *qualitative* analysis we are just interested in finding the elements present in our sample and getting a rough idea of how much of each. This can be done by just looking at and identifying the peaks in the X-ray spectrum.

In *quantitative* analysis we want to measure the proportions of each element present as accurately as possible, by measuring the areas of the X-ray peaks.

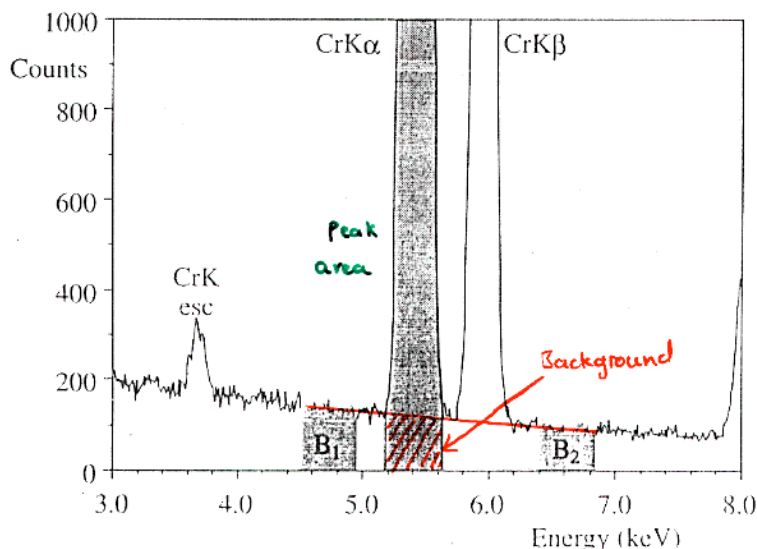
To perform quantitative analysis first we need to separate the X-ray counts in each peak from the background and separate overlapping X-ray peaks. Then we need to convert the number of X-ray counts measured from each element to an atomic (or weight) fraction.

1. Background subtraction and peak integration

There are a number of approaches to background subtraction and measurement of peak areas.

a. Measure background either side of peak

This is the simplest and obvious method. Measure the background intensity from a window either side of the peak and linearly interpolate under the peak to find the background.



Background subtraction by averaging background measured in two windows on either side of the peak. There must be no other peaks near the background windows. Most useful for WDX.

Williams and Carter p 602

This relies on the background being linear (often not the case due to absorption and the background is often curved at low energies). It also relies on there being some background with no other peaks nearby and no peak overlap.

It is simple and reliable and can be done by hand.

b. Background modelling

Use Kramers relation to model the background.

$$I = \frac{kZ(E_0 - E)}{E}$$

where

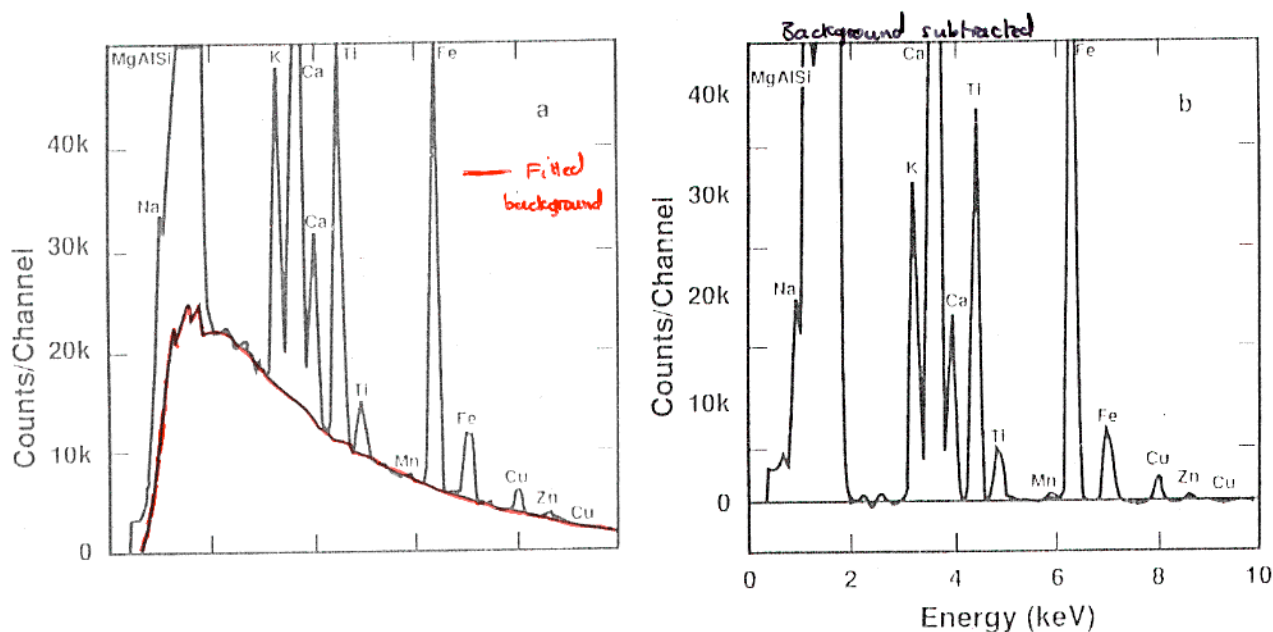
E = X-ray energy

E_0 = beam energy

Z = atomic number

k = constant

In practice a better model is needed to account for absorption of low energy X-rays in the specimen and the detector. The background found using Kramers relation can be subtracted from the spectrum peak and peak areas can then be found easily.



Background subtraction by background modelling of bremsstrahlung plus absorption by specimen.

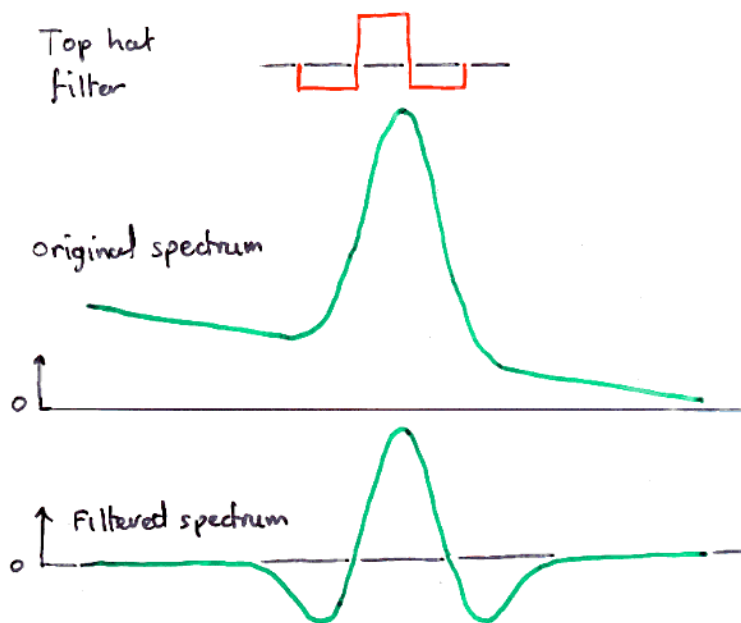
Goldstein p 372

c. Fourier filtering

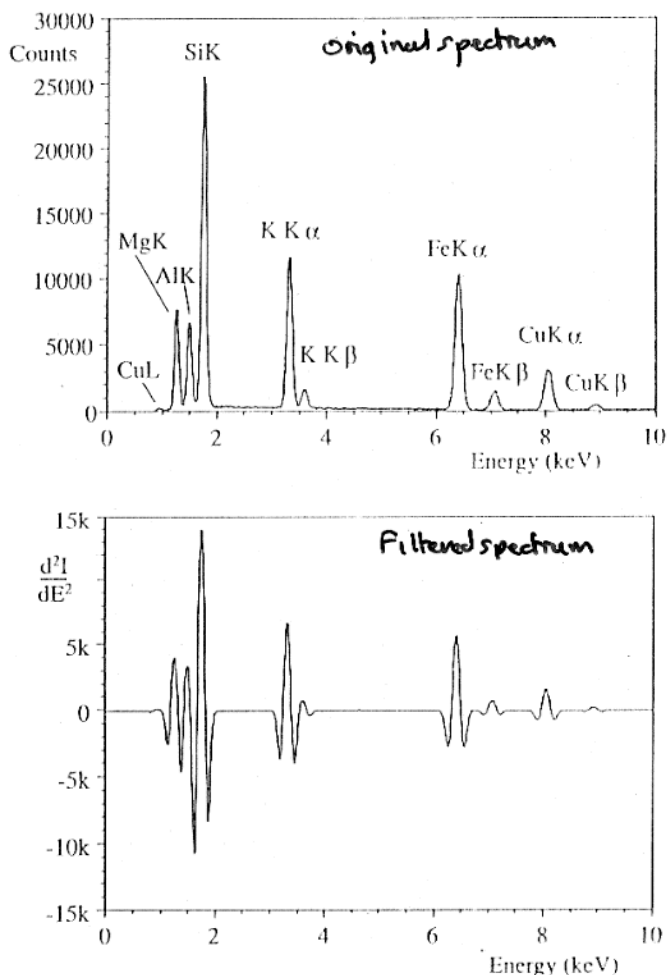
The background is slowly varying and thus contains mostly low frequencies. The peaks are sharp and thus contain high frequencies. So Fourier transform the spectrum, remove the low frequencies and back transform. The result is a spectrum without the background.

d. Top hat filtering

This is similar to Fourier filtering. The method is to convolute the spectrum with a top hat filter. This gives approximately the second differential of the spectrum, with some smoothing. This will remove the background completely if it is linear and sloping but not if it is curved. This method works well and is used in some commercial X-ray programs such as Link (now Oxford) spectrometers.



Background subtraction by top hat filtering. Original spectrum is convoluted with top hat filter.



Spectrum before and after top hat filtering.

To find peak areas standard peaks from each element are needed which have been filtered in the same way. A multiple least squares fitting method is then used to decompose the experimental filtered spectrum into fractions of each filtered standard peak.

2. Converting peak areas to concentrations: standards (Castiang)

Given that the background has been subtracted and the areas of the X-ray peaks measured it is now necessary to convert these into elemental concentrations.

Historically this was first done by Castiang. It is difficult to predict the X-ray intensity that will be emitted from a given alloy directly so Castiang used a standard of known composition.

If I is the X-ray intensity and C_i the concentration of element i then

$$\frac{C_i}{C_{i(std)}} = K \frac{I_i}{I_{i(std)}}$$

where K is a sensitivity factor that takes in to account the difference between the generated and measured X-ray intensities for the standard and the unknown specimen.

K depends on:

- Z the atomic number of all the elements in the specimen
- A the absorption of X-rays in the specimen
- F the amount of fluorescence within the specimen.

As a result these are known as "ZAF" corrections.

The original method as described suffers from the need for standard specimens of each element of interest.

3. Converting peak areas to concentrations: Cliff-Lorimer ratio technique

This technique was originally developed for TEM where specimens are thin and absorption and fluorescence corrections are small. Nowadays the absorption and fluorescence corrections are included as well.

Here the above equation is rewritten

$$\frac{C_A}{C_B} = k_{AB} \frac{I_A}{I_B}$$

Where A and B are 2 elements in the unknown alloy. k_{AB} is called the Cliff-Lorimer k factor or just the "k-factor". It is not a constant. k_{AB} is related to the atomic number correction factor (Z). Note that no standards are needed.

Since measuring k factors for every pair of elements would be tedious, they are measured with respect to one element, usually Si. Si was originally chosen because many minerals contain Si and its K edge has high enough energy to be detected on the Be window detectors then available (cf oxygen, the other obvious choice). Then

$$k_{AB} = \frac{k_{ASi}}{k_{BSi}}$$

4. ZAF corrections: Z – atomic number

ZAF corrections are an alternative way of converting peak areas to concentrations used mostly on EDX systems on SEMs.

The Z correction is a combination of 2 effects, backscattering, R (the proportion of electrons remaining in sample), and stopping power, S (the rate of energy loss in sample).

a. Back scattering, R

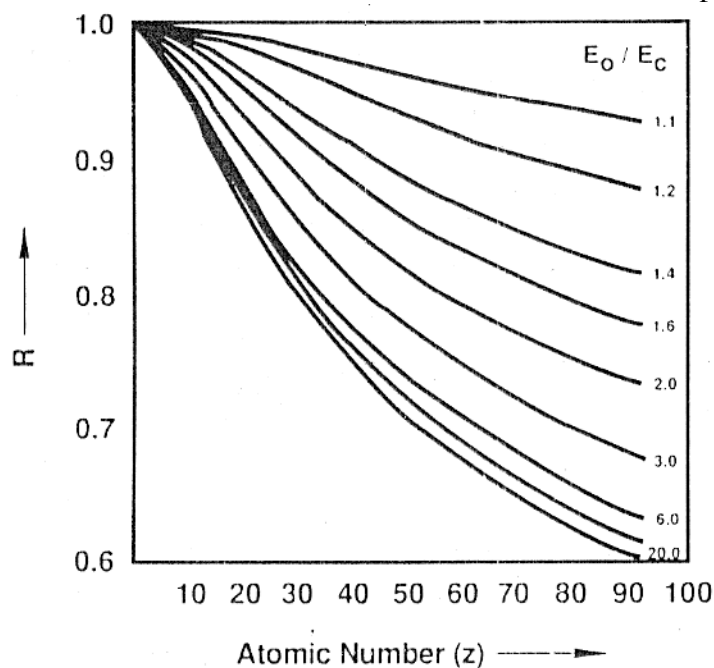
This is the proportion of electrons remaining in the sample

$$R = \frac{\text{number of ionisations in the specimen}}{\text{number of ionisations if no electrons escaped}}$$

$$= 1 - \eta$$

where η = backscatter coefficient

As Z increases more electrons are backscattered and so fewer remain in sample to produce X-rays.



R – fraction of ionisations remaining in a specimen of atomic number Z after the loss of electrons from the beam due to backscattering.

Goldstein p 421

b. Stopping power, S

Stopping power is the rate of energy loss of the electron in the sample.

$$S = \frac{1}{\rho} \frac{dE}{dx}$$

where ρ = density

Low Z elements are better at removing energy from electrons (and hence producing X-rays) than high Z elements. Therefore low Z elements have higher stopping power, S .

5. ZAF corrections: A – absorption

X-rays generated in the sample are absorbed on their way out from the sample. The distance the X-rays have to pass through depends on the take-off angle, typically 45 to 70 deg.

The intensity of X-rays after passing through a distance x of a material is

$$\text{Intensity} \quad I = I_0 e^{-\mu x}$$

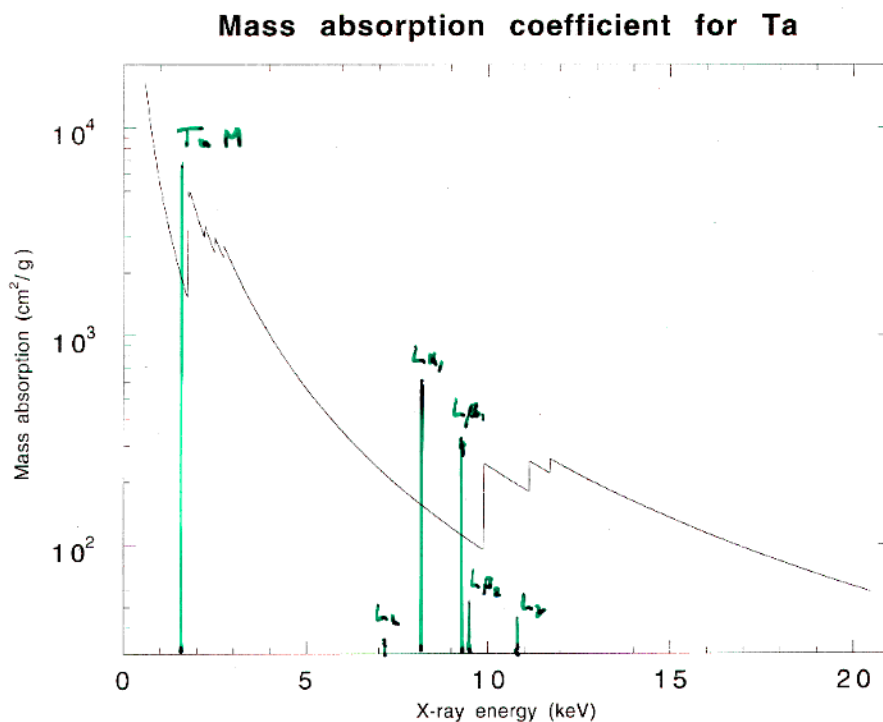
where I_0 = initial X-ray intensity

μ = absorption coefficient

In practice the mass absorption coefficient ($\frac{\mu}{\rho}$) is used, so

$$I = I_0 e^{-\left(\frac{\mu}{\rho}\right)\rho x}$$

Absorption depends on both the X-ray trying to get out and the sample it is passing through in a non-linear way.



Mass absorption coefficient for Ta as a function of X-ray energy.

Note that the absorption edge energy is above the energy of the corresponding X-ray line.

Absorption is greatest when the X-ray energy is just above the absorption edge of one of the major constituents of the sample. eg Fe K_{α} (6.40 keV) is heavily absorbed by Cr (absorption edge 5.99 keV). Element Z tends to absorb element $Z+1$.

6. F - Fluorescence

An X-ray on passing through the sample may excite X-rays of lower energy, typically from elements of lower atomic number (with the original X-ray being absorbed). This enhances the X-ray signal from these lighter elements.

The extent of fluorescence depends on:

Value of fluorescent yield. This is small for light elements.

Ionisation cross section.

As for absorption this is most likely when the energy of the original X-ray is just above the absorption edge of one of the constituents of the sample. In the example above Fe K_{α} would fluoresce Cr K_{α}

6. Quantitative analysis (continued)

7. Doing quantitative analysis

```

CO calibration,                               FILENAME: C096CU
CO REF 1 1950                                LIVETIME= 100
  ENERGY   RES     AREA
-    .4     87.59   68225
  6924.6   158.12  68366
TOTAL AREA= 173352 GF= 50.006
    
```

} gain calibration using Co K

```

Specimen spectrum,                           FILENAME: C097CU
SPEC 3.1                                     LIVETIME(spec.)= 100
  ENERGY   RES     AREA
-    .4     87.81   67590
TOTAL AREA= 233256
    
```

} noise peak

```

.....
Peak at    .74 keV omitted?   Co L
Peak at    .92 keV omitted?   Cu L
Peak at    3.44 keV omitted?  Si K sum peak
    
```

} peak areas

```

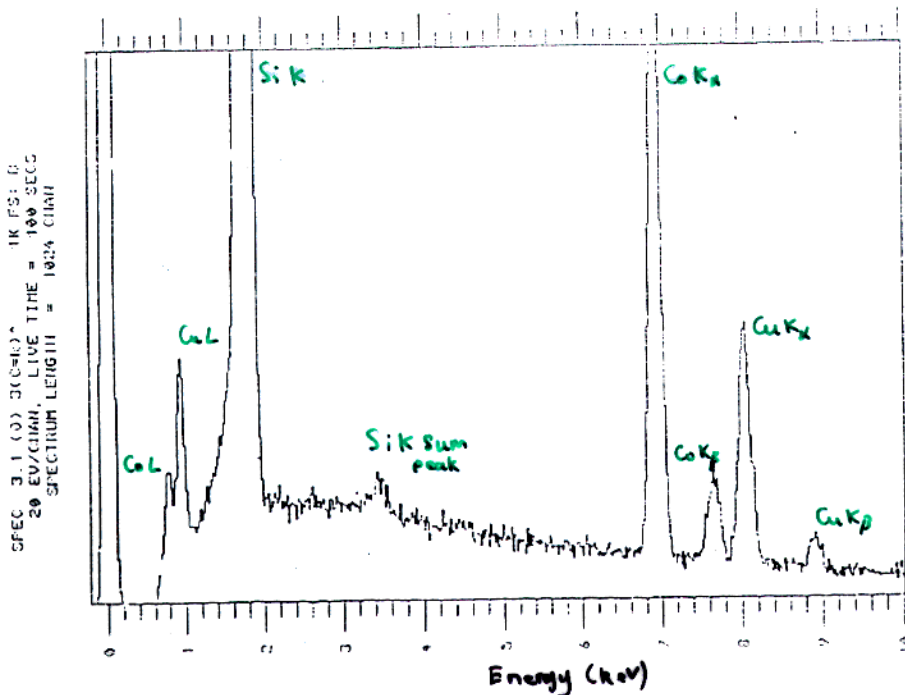
FIT INDEX= 2.46

ELNT  APP. CONC  ERROR(WT%)
Si     59.524    .219
Co     15.447    .295
Cu      7.307    .290
... 2 ZAF/SJ
    
```

} ZAF correction

```

20.00 kV TILT=45.00 ELEV= .00 AZIM= .00 COSINE=1.000
Spectrum: SPEC 3.1                               Co/Cu; Si 4911
All elmts analysed
          wt
ELNT  ZAF  %ELMT  ATOM.%
SiK: 1  .844  71.236  55.302
CoK: 1  .877  17.629  10.062
CuK: 1  .935   9.756   4.635
TOTAL  .977  97.622 100.000
    
```



Steps in quantitative analysis on a SEM

To work out the ZAF corrections (and background modelling) the sample composition needs to be known. But this is what we are trying to determine. Hence an iterative approach is needed. We use the raw peak areas (and k-factors in TEM) to determine an initial guess at the sample composition. We then use this composition to determine Z, A and F corrections to get a better estimate of the composition and repeat the process.

An example of the steps in quantitative analysis performed on a (rather old!) EDX system are shown on the previous page.

a. Fit index

This is a measure of the fit used by the software on Link (Oxford) EDX spectrometers. Other manufacturers may use different ways of measuring how good the fit is. The fit index represents the goodness of fit relative to the amount of noise present. A lower value means a better fit. It is a normalised χ^2 .

$$\chi^2 \propto \sum_{\text{all channels}} \frac{(\text{actual counts} - \text{fitted counts})^2}{(\text{actual counts})^2}$$

Normalised such that $\chi^2 = 1$ for an ideal fit (noise only)

Thus a value of 1 would be obtained for a perfect fit on a noisy spectrum, ie the only contribution to the misfit is noise.

The fit index should be around 2 or less. If it is much more than 2, there is a problem with the fit.

The fit index will increase as the number of counts increases, suggesting it is better to use fewer counts - this is not the case. All it means is that the contribution to the misfit from noise is decreasing.

b. Standards

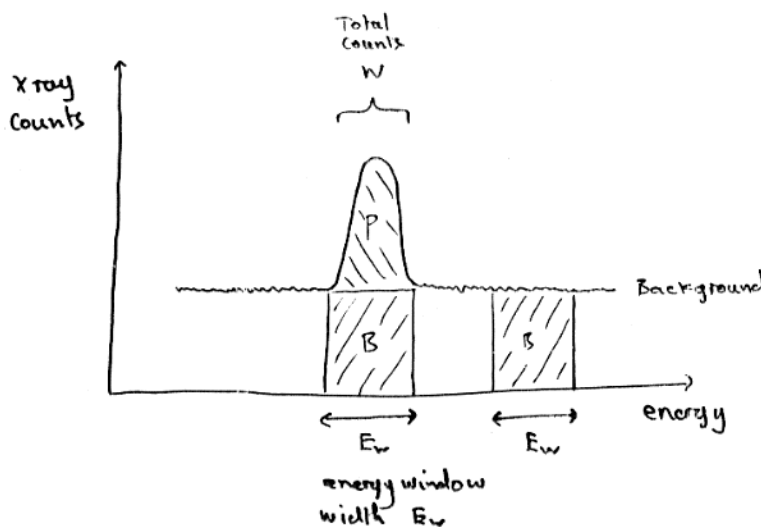
The standard spectra store all peaks from a series, ie K_α and K_β or all the peaks of an L series. This increases the statistics a little as all counts not just those in 1 peak are used.

c. Quantitative analysis works only if:

- The sample is of *uniform* composition over the area the beam spreads to, plus adjacent areas that X-rays pass through
- The sample is flat and oriented at the angle given to the program (ie the surface is not rough)
- All the elements in the specimen are considered by the program
- The program only looks at 1 peak (eg K or L or M) per element.

8. Errors in peak area

Random errors



Want to find peak area, P , and error in peak area, σ_p .

Measure counts in window containing peak and background, $W = P + B$

Measure background from adjacent region using same window, B .

As measuring X-ray counts assume Poisson statistics so:

Error in window area, W

$$\sigma_W = \sqrt{W}$$

Error in background, B

$$\sigma_B = \sqrt{B}$$

Error in peak area, $P = W - B$

$$\begin{aligned} \sigma_P &= \sqrt{\sigma_W^2 + \sigma_B^2} \\ &= \sqrt{W + B} \\ &= \sqrt{P + 2B} \end{aligned}$$

Lowest possible error is if measure background level over a large area so that σ_B is small or tends to zero. Then:

Error in peak area

$$\begin{aligned} \sigma_P &= \sqrt{W} \\ \sigma_P &= \sqrt{P + B} \end{aligned}$$

Note that the error in the peak area depends on the number of counts in the peak plus the background. Thus for small peaks on a large background the error in the measurement of the peak area comes mostly from the background.

Fractional error in peak area

$$\text{Fractional error in peak area} \quad \frac{\sigma_P}{P} = \frac{\sqrt{P + B}}{P}$$

$$\text{For large peaks (ie small background, } B) \quad \frac{\sigma_P}{P} \approx \frac{1}{\sqrt{P}}$$

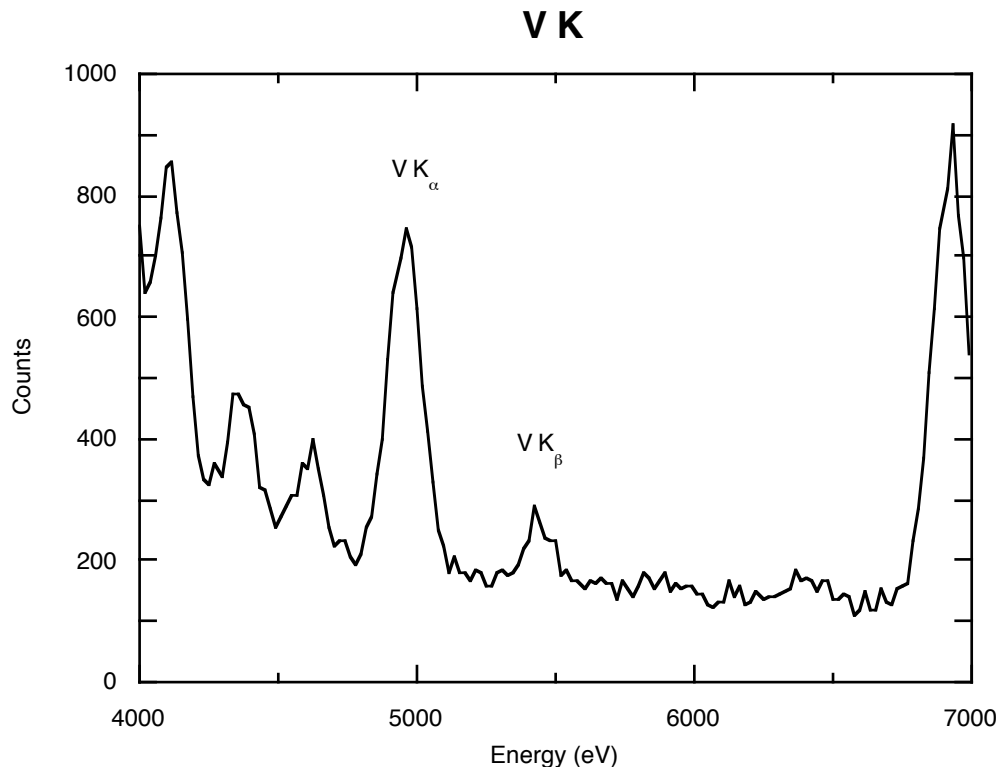
So to get 1% error in peak area need 10000 counts in peak.

For small peaks, peak is 95% certain of being real if $P > 2\sigma_p$

ie $P > 2\sqrt{P+B}$

For small P and large B $P > 2\sqrt{B}$

Example: V K_β peak



Total counts in window around V K_β peak, $W = 2433$

Counts in background, $B = 1955$

Peak area, $P = W - B = 478$

Error in peak area (assuming no error in background counts, B)

$$\sigma_p = \sqrt{P+B} = 49$$

Fractional error $\frac{\sigma_p}{P} = \frac{\sqrt{P+B}}{P} = 10\%$

To be significant need $P > 2\sqrt{P+B}$

ie $478 > 98$

So peak is significant.

Generally random errors are around ± 1 at%. At best if use long counting times get ± 0.1 at%.

Hence you can ignore all (except possibly the first) decimal places in the compositions.

Random errors can be estimated as shown above.

To get an idea of the random errors analyse same area a few times and analyse similar areas in other parts of specimen.

Systematic errors

Systematic errors cannot be estimated. Examples of systematic errors are:

- wrong k-factors because k-factors (Z correction) are calculated not measured
- wrong absorption and fluorescence correction because specimen surface is rough or not uniform.

Systematic errors are usually greater than random errors.

To overcome systematic errors compare the composition of your unknown area with an adjacent area of known composition. Such comparisons are much more accurate than absolute measurements as many systematic errors cancel.

Errors in concentrations calculated from peak counts

Given a spectrum with three peaks from three different elements A, B, C

Peak areas: P_A, P_B, P_C

Errors in peak areas: $\sigma_A, \sigma_B, \sigma_C$

k factors: k_A, k_B, k_C

Want to work out the concentration, C , of each element and the error in these concentrations, δC .

$$\text{Concentration of element A, } C_A = \frac{k_A P_A}{k_A P_A + k_B P_B + k_C P_C}$$

$$\text{and likewise concentration of B } C_B = \frac{k_B P_B}{k_A P_A + k_B P_B + k_C P_C}$$

etc.

The error in the concentration is less straightforward. It is *not* correct to say

$$\delta C_A = \frac{\sigma_A \cdot C_A}{P_A}$$

this over-estimates the error. Instead use

$$\delta C_A = \frac{\sqrt{(k_A \sigma_A)^2 [k_B P_B + k_C P_C]^2 + (k_A P_A)^2 [(k_B \sigma_B)^2 + (k_C \sigma_C)^2]}}{(k_A P_A + k_B P_B + k_C P_C)^2}$$

And similarly for δC_B and δC_C .

For example, if:

Peak areas: $P_A, P_B, P_C = 5000, 4800, 200$ counts

Errors in peak areas: $\sigma_A, \sigma_B, \sigma_C = 78, 76, 35$ counts (assumes a background of 1000 for each)

k factors: $k_A, k_B, k_C = 1, 1, 1$ (for simplicity)

Then:

$$\text{Concentration of element A, } C_A = 50\%$$

$$\text{Concentration of element B, } C_B = 48\%$$

Concentration of element C, $C_C = 2\%$

and

error in concentrations $\delta C_A = 0.57\%$

$\delta C_B = 0.57\%$

$\delta C_C = 0.34\%$

ie $C_A = 50.0 \pm 0.6 \%$, $C_B = 48.0 \pm 0.6 \%$ and $C_C = 2.0 \pm 0.3 \%$.

9. Practical limitations

a. Probe size

The probe size can be bigger than it looks, especially for small probes with large condenser apertures where spherical aberration creates a bright ring.

Resolution is made worse by contamination.

b. Secondary fluorescence

Mentioned already, but can be a big problem, especially in TEM. Warning signs are large peaks from unexpected elements eg Cu from Cu grids, large contribution from matrix when analysing precipitate.

c. X-rays and ions in column

Some condenser apertures are thin and allow X-rays from the gun to pass down the column. Ensure the condenser apertures are thick and designed for EDX. Microscopes with bad vacuums can also have ions passing down the column. Check by measuring a hole count.

d. Crystalline specimens

We saw earlier that close to a zone axis we can get coherent bremsstrahlung peaks. In addition X-ray peak intensities change close to strong diffracting conditions by few percent. Keep away from strong diffraction. This is especially a problem when analysing boundaries.

7. Electron energy loss spectroscopy (EELS)

References

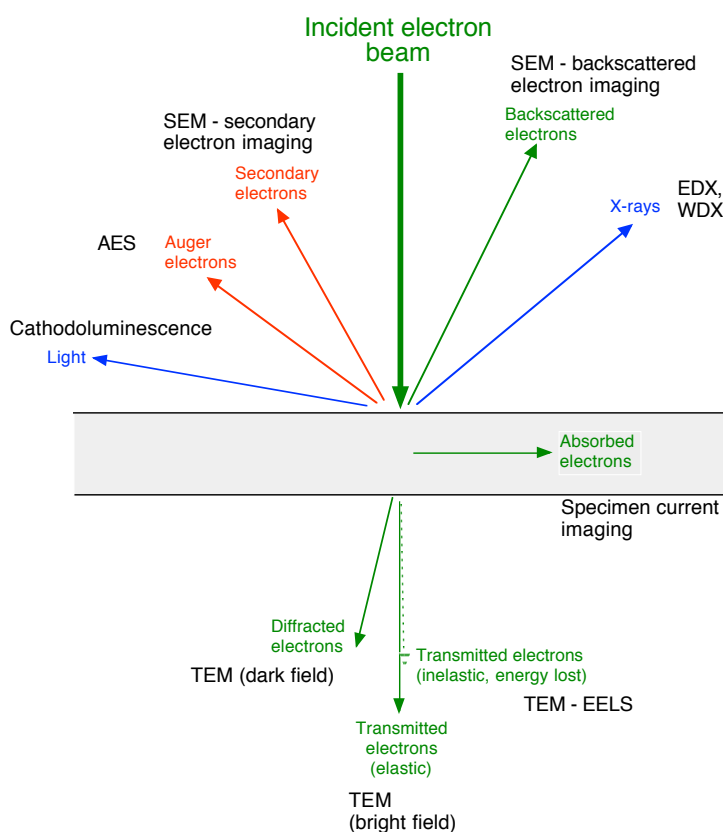
Transmission electron microscopy

DB Williams and CB Carter (Plenum, New York, 1996) chapters 37 to 40

Electron energy-loss spectroscopy in the electron microscope

RF Egerton (Plenum, second edition 1996)

1. Introduction to EELS



Output signals generated by electron bombardment of a thin specimen

When an electron creates an X-ray or an Auger electron or any other process in the sample the electron loses energy. If the electron is allowed to exit the sample and its energy spectrum measured then this spectrum contains information on all the interactions the beam has made with the specimen. The requirement for the electrons to exit the specimen means that EELS can only be done in a transmission electron microscope.

EELS can be used for compositional analysis: X-rays have a low fluorescent yield and poor collection efficiency for light elements. EELS is much more efficient for light elements.

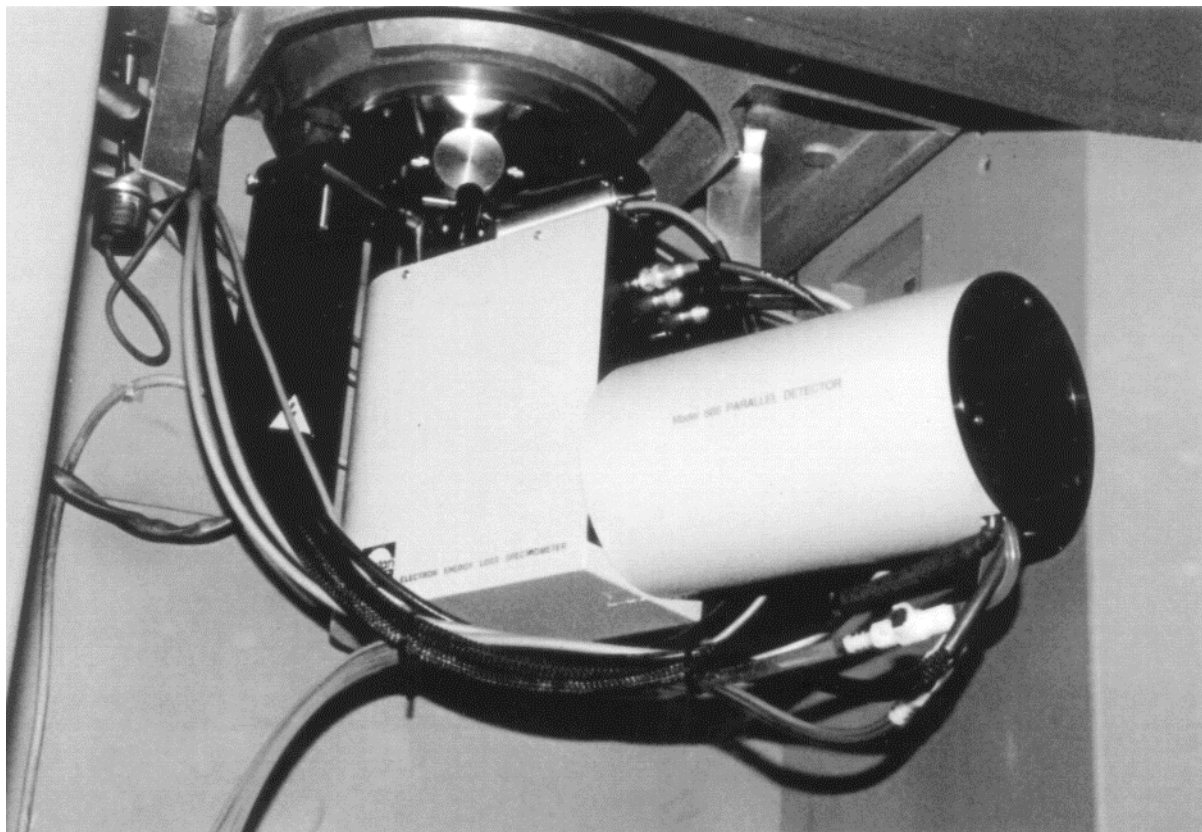
EELS can also be used for chemical analysis: shapes of energy loss edges depend on local bonding and oxidation state.

2. EELS spectrometer

Most electrons go straight through the specimen and are scattered through only very small angles, $<1^\circ$. Therefore the energy loss spectrometer has to be at the bottom of the microscope, under the viewing screen, as shown below.



Transmission electron microscope with electron energy loss spectrometer below column and control electronics on desk to left.

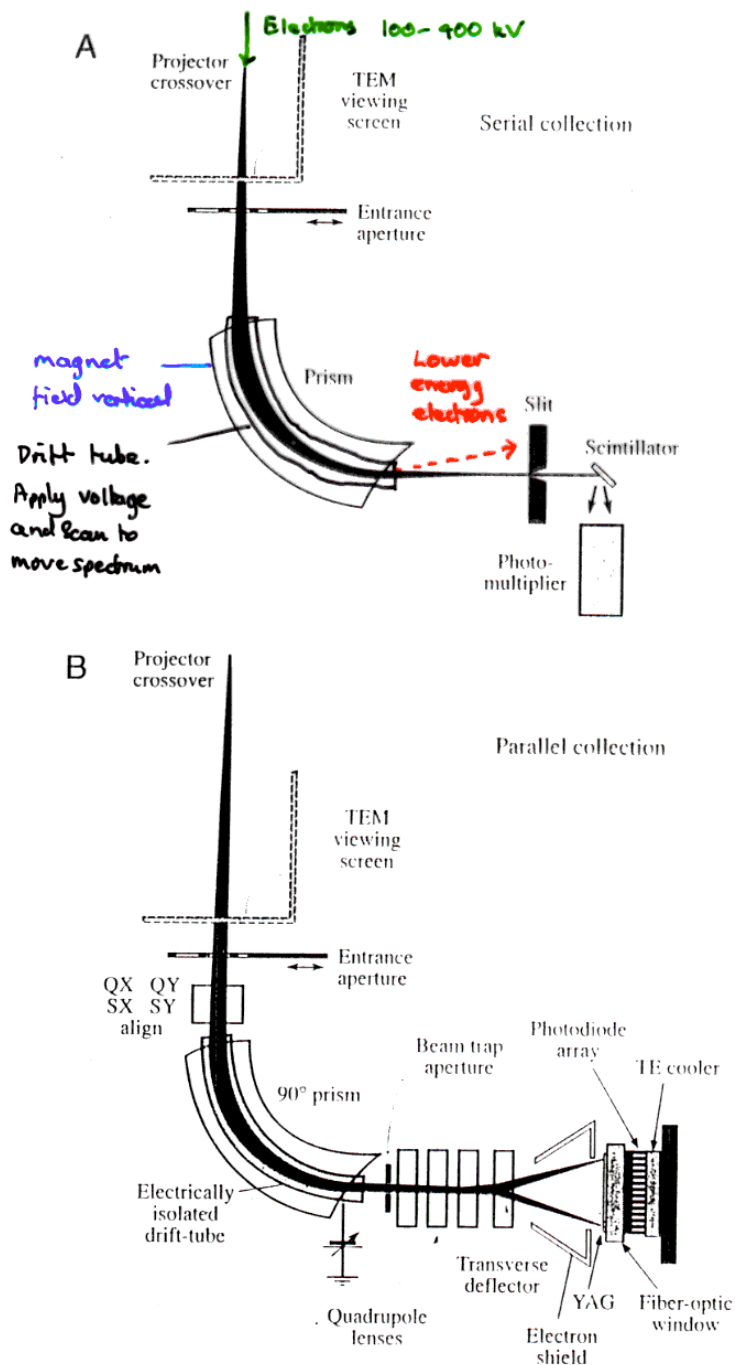


Close up of electron energy loss spectrometer below viewing screen of microscope.

A typical transmission electron microscope voltage is 100 to 400 kV but typical energy losses are generally less than 1 kV. We need a spectrometer that can separate these.

An energy loss spectrometer bends the electrons through 90° with a magnetic prism. It separates different energy losses just like a glass prism. The lower energy electrons are bent through a smaller radius than the higher energy electrons.

Curved faces of the magnet are used to focus the electrons. Additional quadrupole and sextupole lenses are used to fine tune the focus.



Diagrams of EELS spectrometers. A Serial spectrometer, spectrum is scanned across a slit and recorded serially on a photomultiplier. B Parallel spectrometer, spectrum is recorded in parallel on a photodiode array.

There are two ways of recording energy loss spectra.

a. Serial detection

This is the old way and is not used much now. There is a slit at the focus of the spectrometer. The electrons pass through the slit onto a phosphor screen, which gives off light. A photomultiplier is used to record the light intensity.

To collect electrons of different energies either the magnet strength is varied or the voltage on the drift tube is varied to scan the spectrum across the slit. We normally scan from low energy to high to avoid phosphorescence problems with the intense zero loss electrons.

Serial detection is easy to do but has a poor collection efficiency.

b. Parallel detection

Typically a diode array with 1024 or more diodes is used to collect the whole spectrum in parallel. On more recent spectrometers and energy filters a two-dimensional 1024 square charge coupled diode (CCD) detector is used.

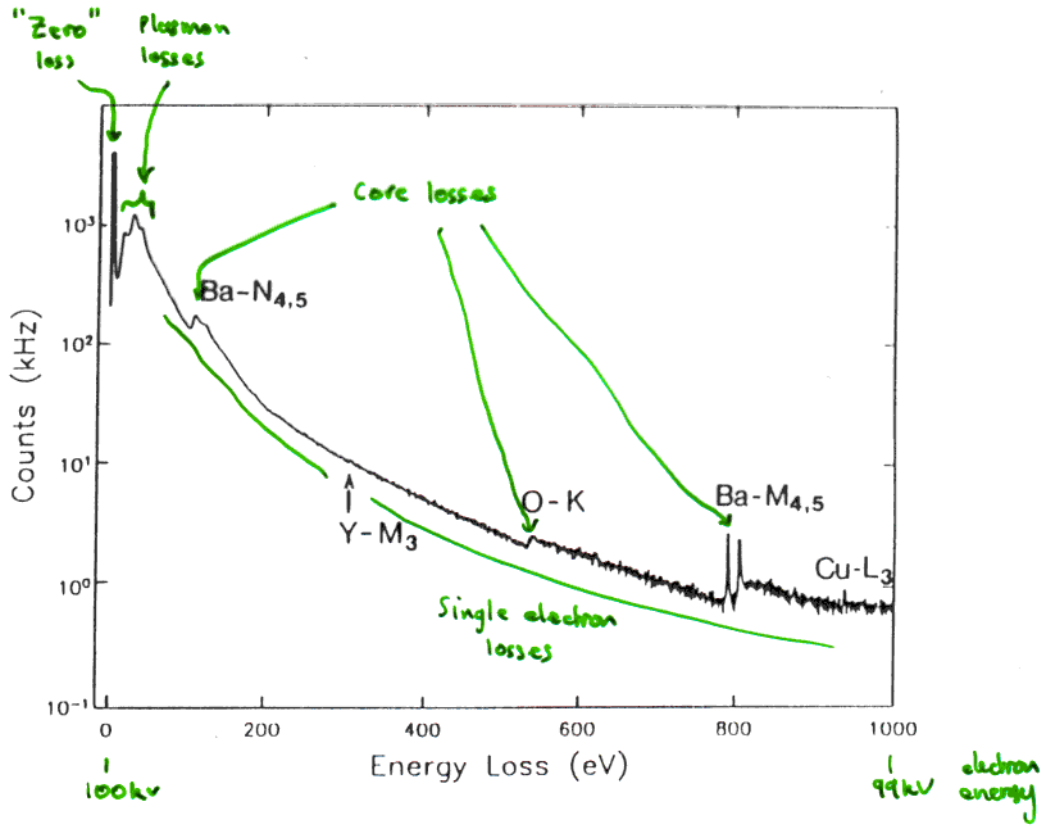
The electrons hit a YAG (yttrium aluminium garnet) scintillator which converts them into light. The light is then channelled via fibre optics to the diode or CCD chip. The diode chip is cooled by a thermo-electric cooler to about -35°C .

Parallel collection is much more efficient than serial collection, but has more artefacts.

Some systems (Gatan imaging filter, GIF) have the ability to form filtered images. These will be covered later.

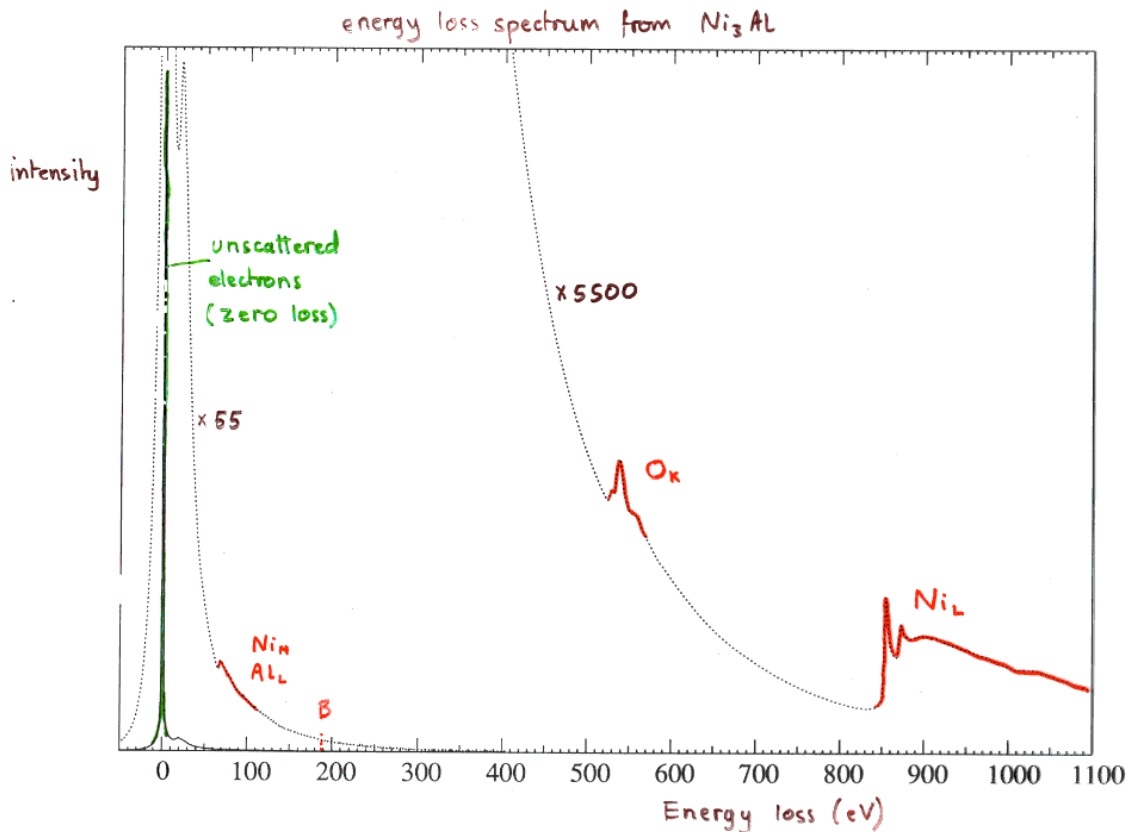
3. Energy loss spectra

Below are two examples of typical energy loss spectra.



Electron energy loss spectrum of $\text{YBa}_2\text{Cu}_3\text{O}_7$ with intensity on log scale.

Egerton p 7



Electron energy loss spectrum from Ni_3Al with intensity on linear scale.

Notice that there is an enormous variation in the intensity of the spectrum – typically of the order of 10^4 from the bright zero loss peak to the highest energy losses. As a result spectra often have to be displayed on multiple scales, as in the lower spectrum, to show all the detail.

Energy loss spectra can be separated into 4 regions, the zero loss peak, the low loss region, the core losses and the background.

a. Zero loss peak

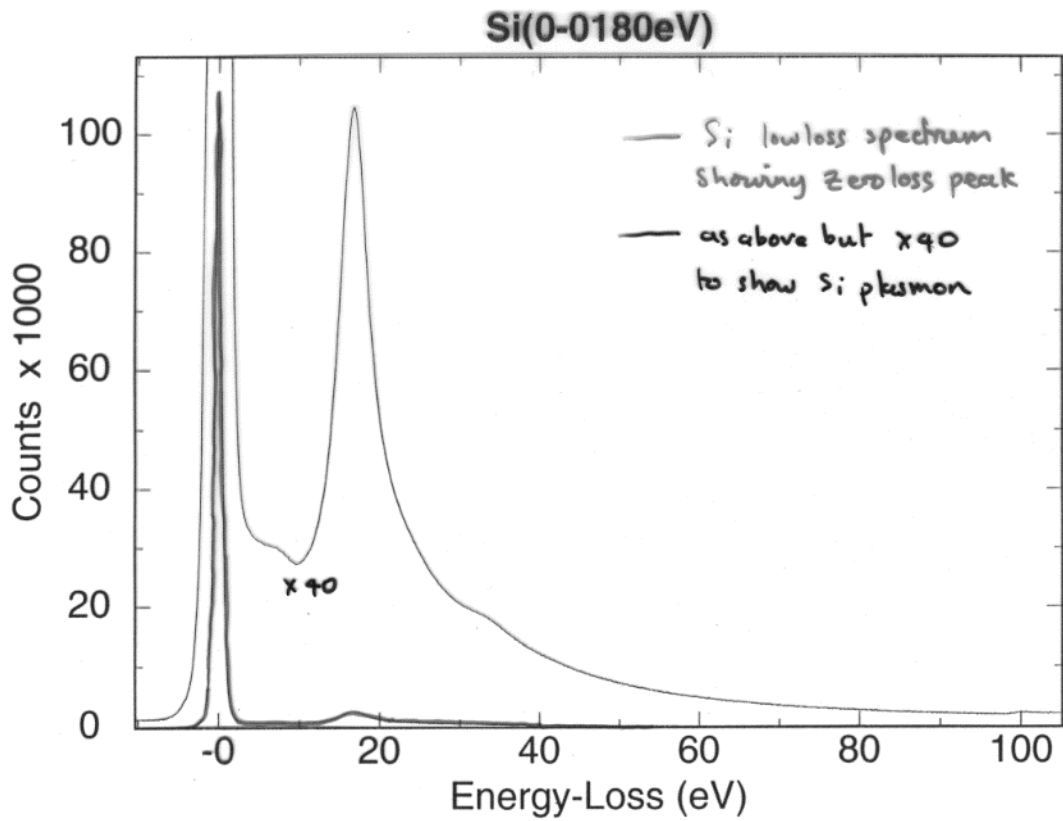
The zero loss peak contains elastically and phonon scattered electrons.

Elastic scattered electrons are those electrons that pass through the specimen without losing any energy, ie an energy loss of 0 eV. They are the electrons that either do not interact with the specimen or are diffracted. The width of the zero loss peak is determined by the energy spread of the microscope, typically 1 eV and a minimum of 0.3 eV if the microscope has a field emission gun (FEG). The typical angular scattering for diffraction, eg for Si 111, is 12 mrad (0.7°).

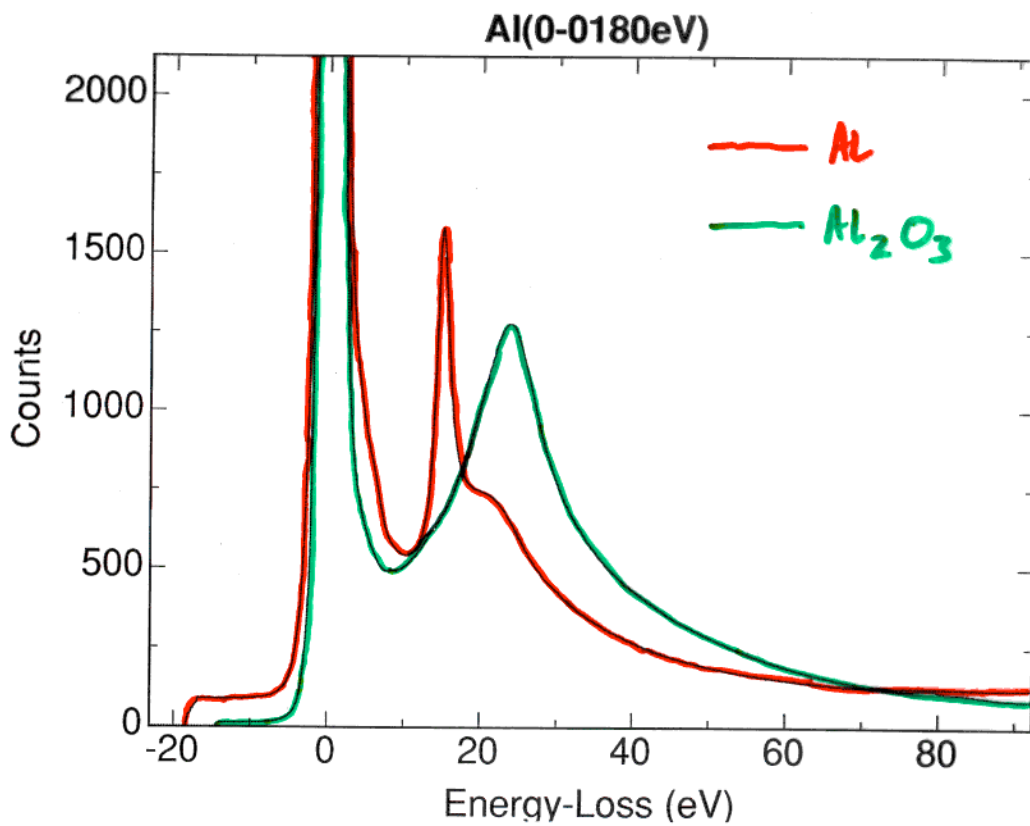
Phonon scattering is where an electron creates one or more phonons, or scatters off a phonon (thermal diffuse scattering). For each phonon created approximately kT (~ 0.025 eV) of energy is lost. This energy loss is too small to measure. However the angular scattering due to phonon scattering is significant, eg up to 50 mrad. Phonon scattering is one source of Kikuchi lines in diffraction patterns.

b. Low loss region ~5 to ~50 eV

The low loss region contains mostly plasmon scattering. This is when an electron excites a plasmon, which is an oscillation of the conduction band electrons (in the same way that a phonon is an oscillation of the crystal lattice). Each compound has characteristic plasmon energy, typically between 10 and 25 eV. For Si the plasmon energy is 16.7 eV and for Al the plasmon energy is 15 eV.



Si plasmon peak.

Al and Al₂O₃ plasmon peaks.

The angular scattering for plasmon losses is around a few mrad.

For thicker specimens multiple scattering gives a series of peaks at multiples of the plasmon energy.

For thin specimens, the first plasmon intensity

$$I_p = I_0 e^{\left(\frac{-t}{\lambda_p}\right)}$$

where

I_0 = zero loss intensity

t = specimen thickness

λ_p = plasmon mean free path

This equation provides a convenient way of measuring specimen thickness, provided the plasmon mean free path, λ_p , is known. If λ_p is not known then it provides a way of measuring relative thicknesses.

Plasmon scattering is delocalised over about 1 nm. This means for example we see one plasmon peak from multilayers of period 1 nm, not two peaks, one from each material.

c. Background - single electron excitations ~30 eV up

For energies above 30 eV energy loss is mainly by collision and ejection of single electrons from the valence band. These electrons become the secondary electrons you use for SEM imaging. In EELS they form a background.

The background has a characteristic shape:

$$I = AE^{-r}$$

where

I = Intensity

E = Energy loss

r = constant, usually between 2 and 6

A = constant

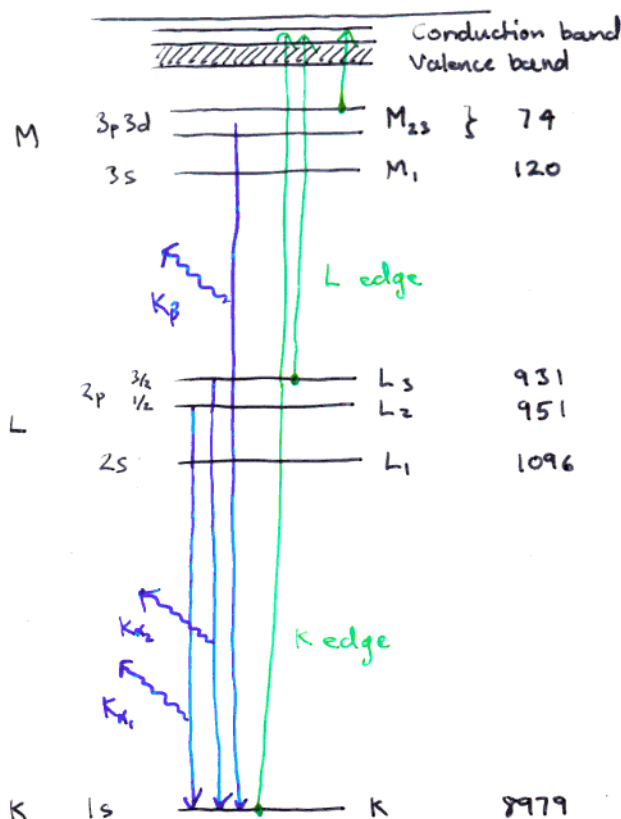
This formula enables the background to be modelled and subtracted. It is only valid above about 50 eV loss.

d. Core losses

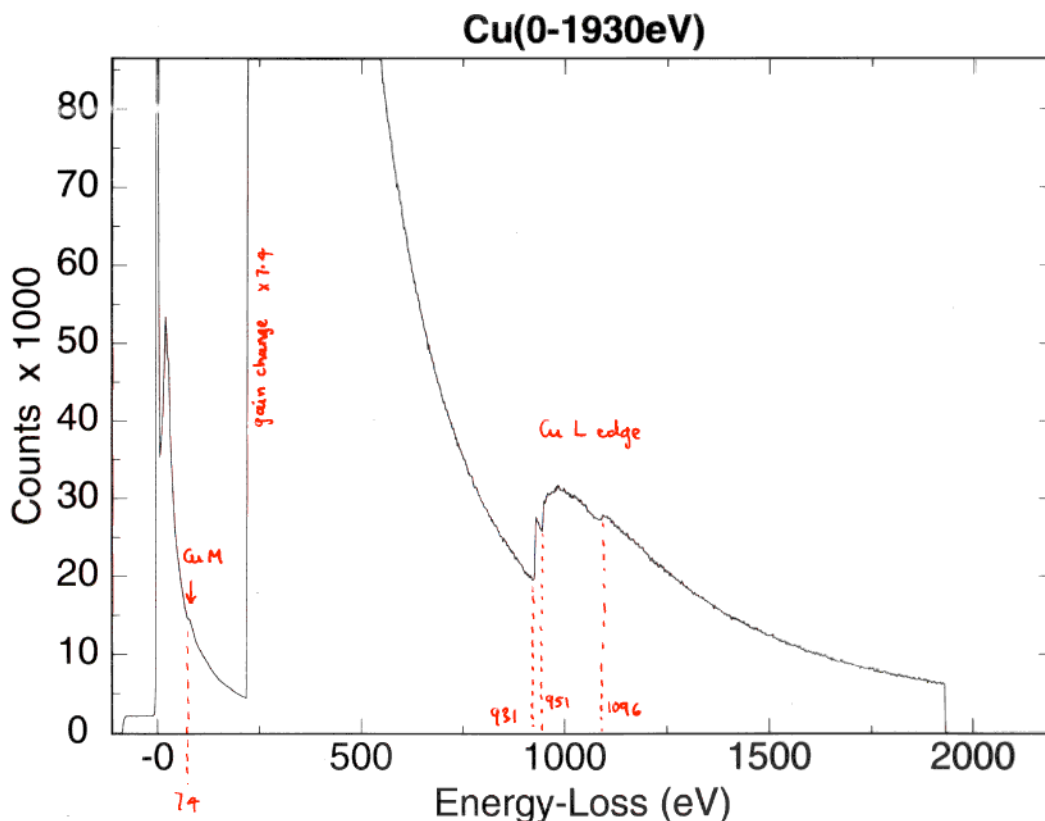
Core losses are formed by the excitation of a core level electron leaving an ionised atom. The core level electron can be excited just to the conduction band (the lowest empty state) or to higher energies. Hence an “edge” is seen in the spectrum, rather than a sharp peak. Each shell has its own edge.

Energy loss spectra are similar to X-ray absorption spectra but can be obtained from a much smaller volume.

It is the filling of the hole in the core level that causes X-ray and Auger electron emission.



Atomic energy levels for Cu.



Core loss EELS spectrum from Cu.

4. Quantifying energy loss spectra

Quantifying energy loss spectra is more difficult than for X-ray spectra mainly because the EELS edges extend over a range of energy, unlike the sharp X-ray peaks. In addition a thin specimen is

needed otherwise multiple scattering obscures the edges. Ideally the plasmon peak should be less than $1/10$ of the zero loss peak.

The shape of EELS edges means the intensity is spread over large range of losses rather than being concentrated in a sharp peak. Hence the edge area is difficult to measure. It is only possible to fit the background above the edge.

The steps in quantifying an EELS edge are:

1. Fit the background using the equation $I = AE^{-r}$ using a window above the edge and extrapolate this background under the edge.
2. Subtract the background from the edge.
3. Calculate the edge area within a window.
4. Calculate the ionisation cross-section σ and integrate over same window. To calculate the cross-section we need to know the collection angle (ie the objective aperture size) and the microscope voltage. This calculation is done by the EELS software (EL/P).
5. Repeat this process for other edges and calculate compositions from:

$$\frac{X_A}{X_B} = \frac{C_A / \sigma_A}{C_B / \sigma_B}$$

where

X = fraction of element

C = counts under edge as determined in step 3 above

σ = ionisation cross-section as determined in step 4 above

Ionisation cross-sections can be treated like EDX k-factors.

The cross-sections are not particularly accurate, especially for thick specimens. Thus expect errors of $\pm 10\%$.

The cross-sections are larger for the lowest energy losses, hence EELS is better for low Z elements. However the background is also larger for low energy losses.

5. Comparison of EDX and EELS for composition determination

EDX

High signal to noise

Compositions to 1 at% or better

EELS

Detection limit as low as 0.05 at% for some elements

Can detect low Z elements better, can see Li and He has been measured.

Better spatial resolution

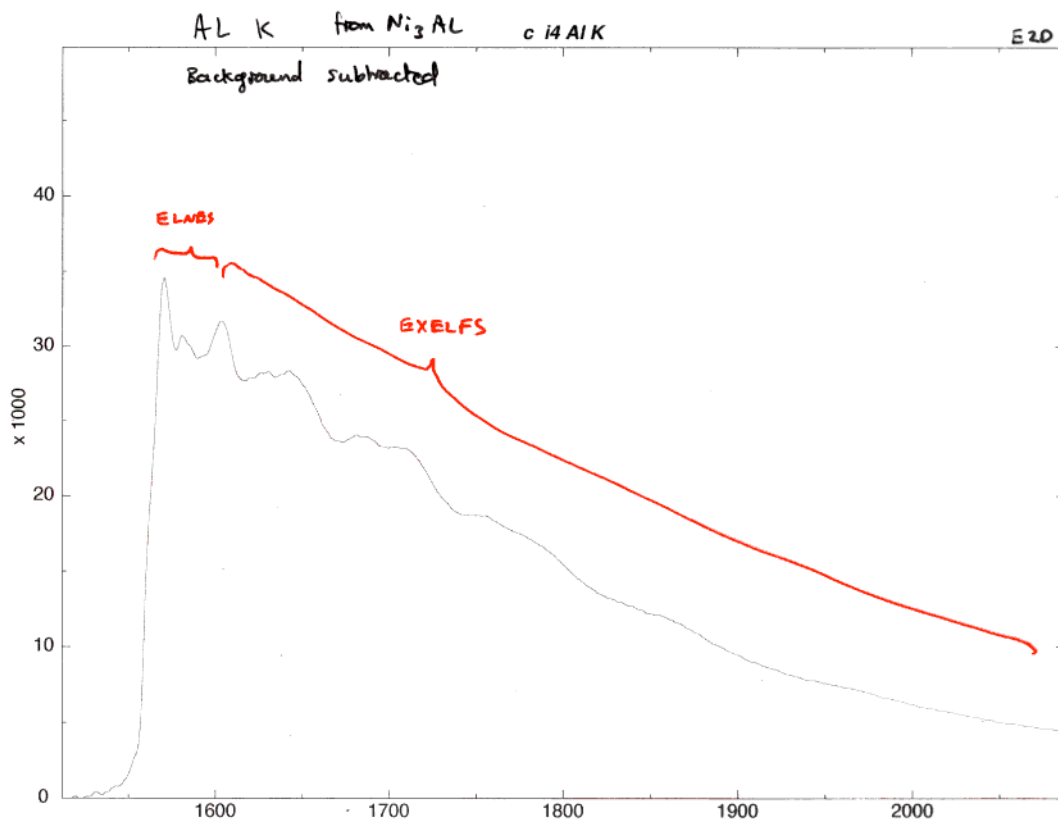
No stray scattering/fluorescence problems.

Poor compositional accuracy, $\pm 10\%$

8. Electron energy loss spectroscopy (EELS) (continued), Energy filtered microscopy (EFTEM)

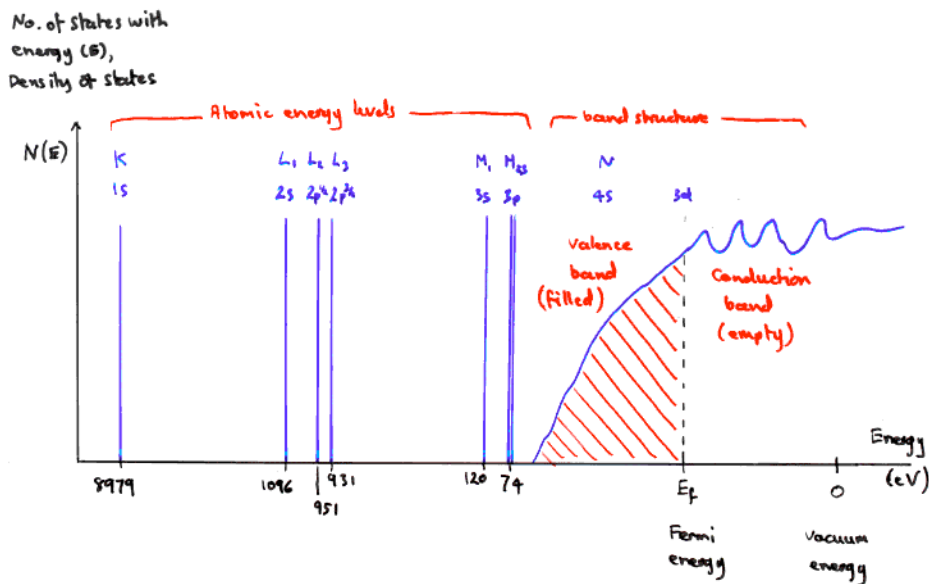
6. Energy loss near edge structure (ELNES)

Near edge structure refers to the fine structure within 50 eV of an EELS edge. It depends on the conduction band density of states.



ELNES and EXELFS (extended energy loss fine structure) in an Al K edge from Ni_3Al .

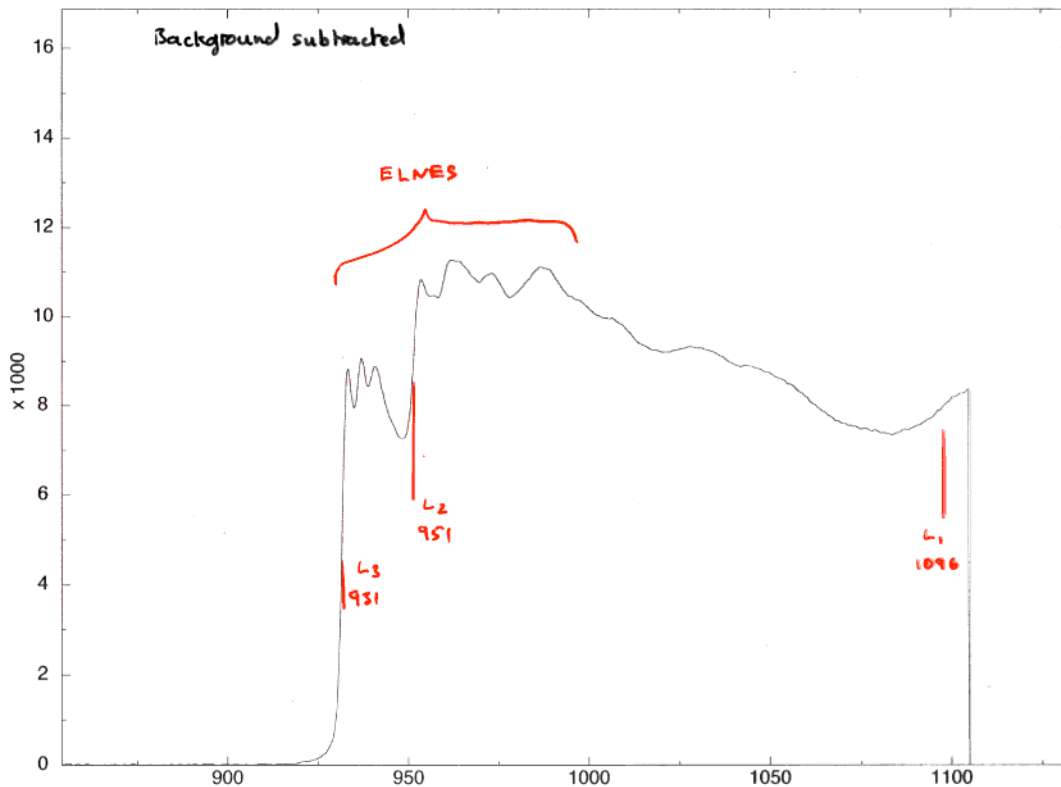
An electron is excited from a core level, eg L, to the first empty state in the atom, the conduction band. The probability of the electron ending up at a particular energy level in the conduction band depends on the conduction band density of states. So peaks near an EELS edge onset should show the conduction band density of states.



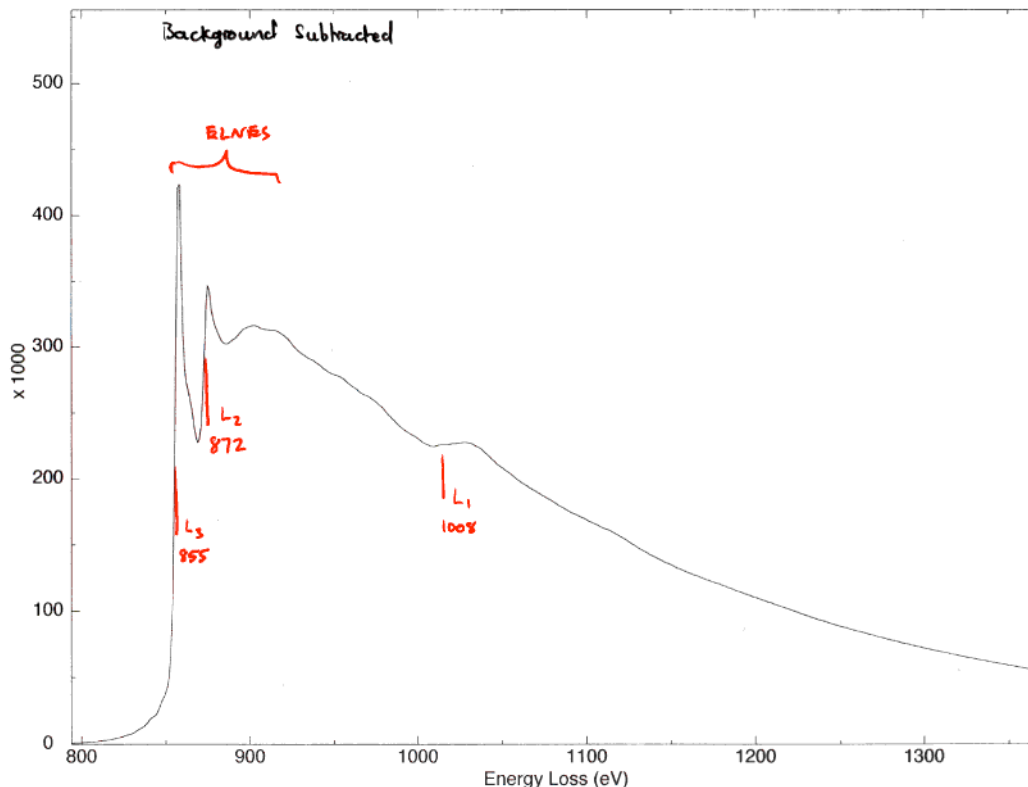
Electron configuration for Cu 2 8 18 1
 i.e. 3d full, 4s half full.

Electron configuration for Ni 2 8 16 2
 i.e. 3d not full, 4s full.

Energy levels in Cu.



Part of EELS spectrum from Cu showing the Cu L edge.

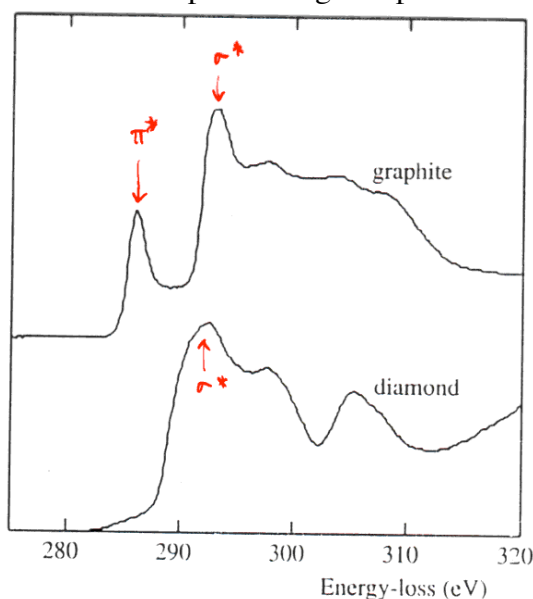


Part of an EELS spectrum from Ni₃Al showing the Ni edge.

In practice the process is more complicated. The quantum mechanical selection rules say $\Delta j = \pm 1$, ie $2p \rightarrow 3s$ or $3d$, but not $3p$. So the near edge structure doesn't show the full density of states, only a partial density of state.

The conduction band density of states depends on the bonding state of the atom, for example the C K edge shapes for graphite and diamond are different.

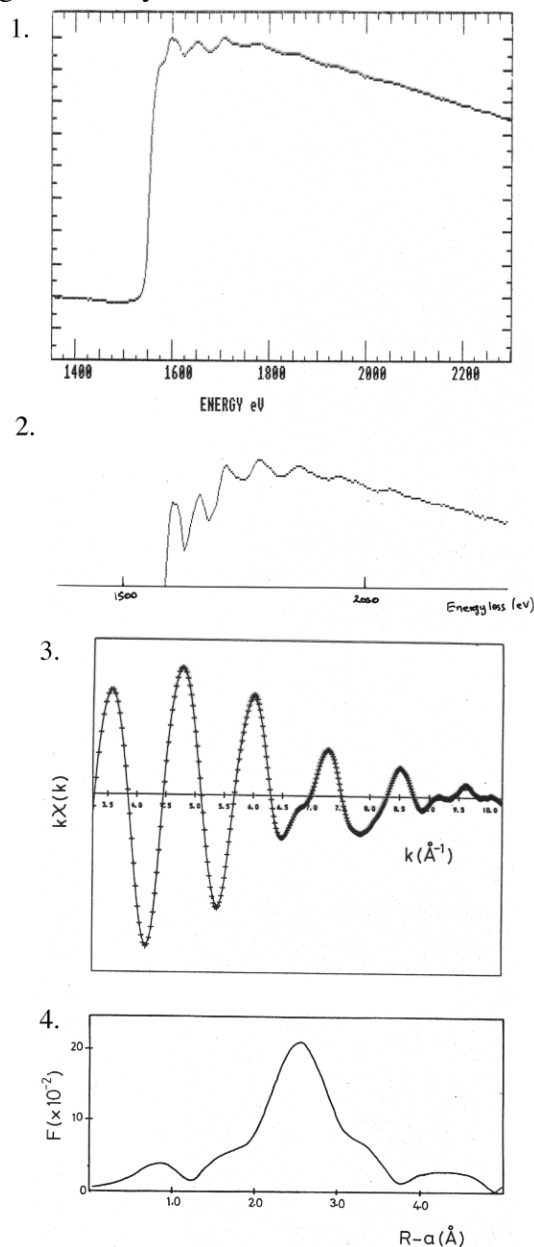
Usually complex calculations are needed to predict edge shapes.



Carbon K edge from graphite and diamond. The conduction band density of states and hence the EELS edge shape depends on the bonding of the C atom.

7. Extended energy loss fine structure (EXELFS)

In addition to near edge structure, EELS edges have structure extending from about 50 eV to about 200 eV beyond the edge. This structure depends on the local arrangement of atoms around the excited atom. EXELFS is similar to EXAFS, extended X-ray absorption fine structure. It can be used to get nearest neighbour distances and radial distribution functions. However low noise spectra are needed and the processing is not easy. Below is an illustration of the steps involved.



Determining the radial distribution function from EXELFS in Al.

Stages in processing EXELFS

1. Original spectrum, Al K edge with EXELFS.
2. Subtract background from edge, subtract smoothed calculated edge shape to leave just EXELFS.
3. Weight and convert from energy to wavelength⁻¹ scale.
4. Fourier transform to give radial distribution function.

8. Energy filtered microscopy (EFTEM)

Energy filtering is an extension of energy loss spectroscopy. In energy filtering we are able to collect an image from a given energy loss. There are two ways of doing this:

STEM: Focus a small probe on the specimen. Collect an energy loss spectrum through a slit at the desired energy loss. Scan the probe to create an image. Effectively we are collecting spectrum at each image point.

TEM: Use extra lenses after the energy selecting slit to allow the original image to be reformed. The extra lenses form part of an imaging filter (GIF, Gatan imaging filter). Effectively we are collecting an image at each energy loss.

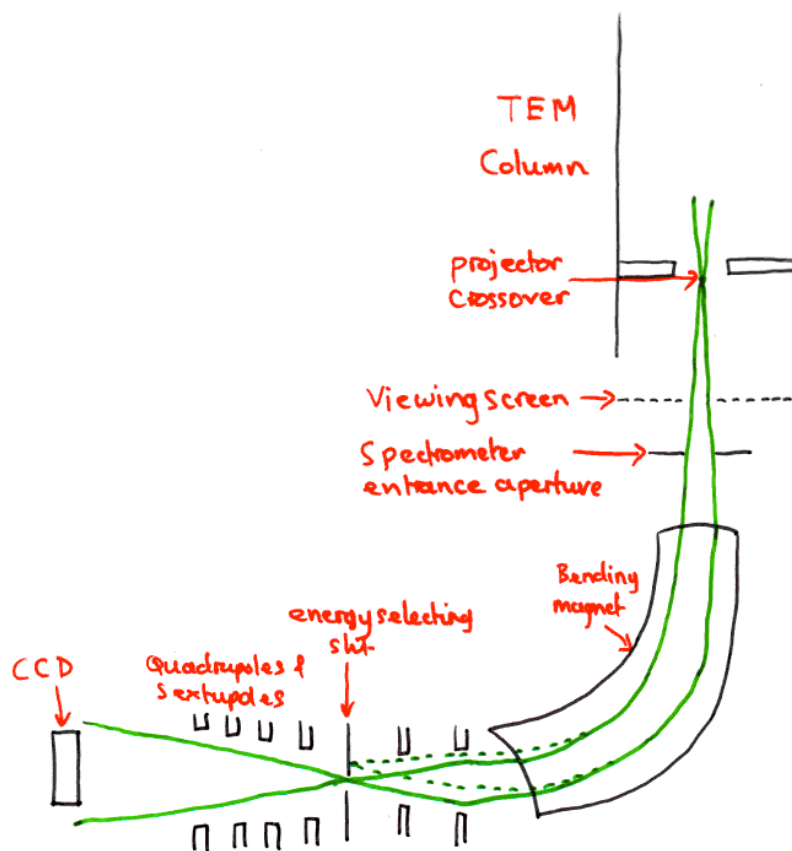


Diagram of an imaging filter.

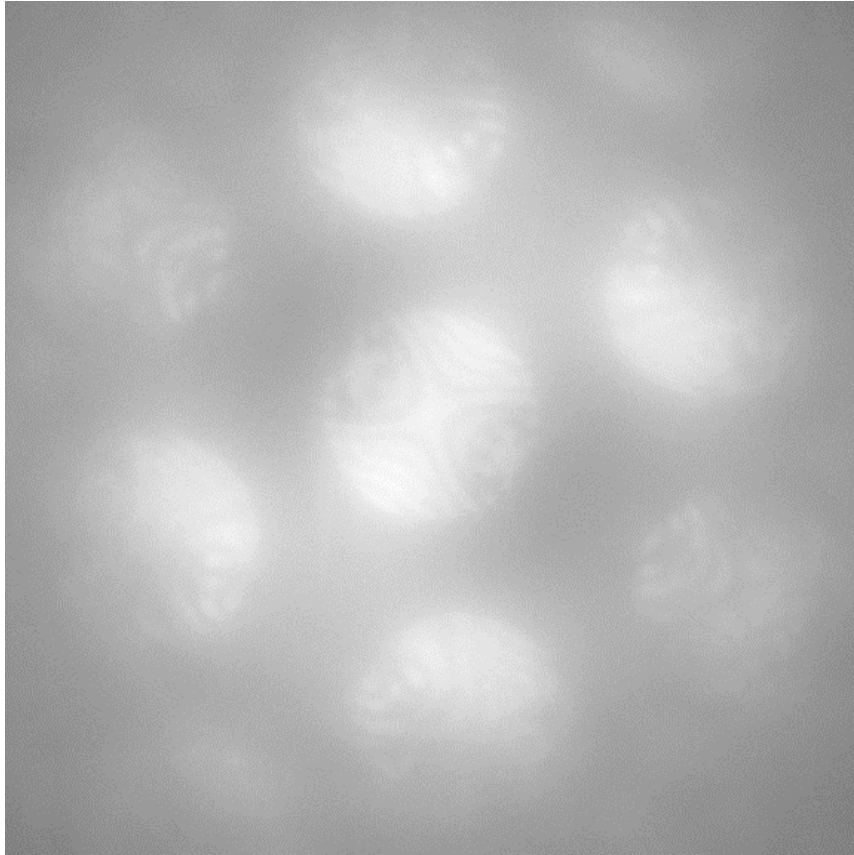
There are a number of uses of energy filtering.

a. Zero loss filtering

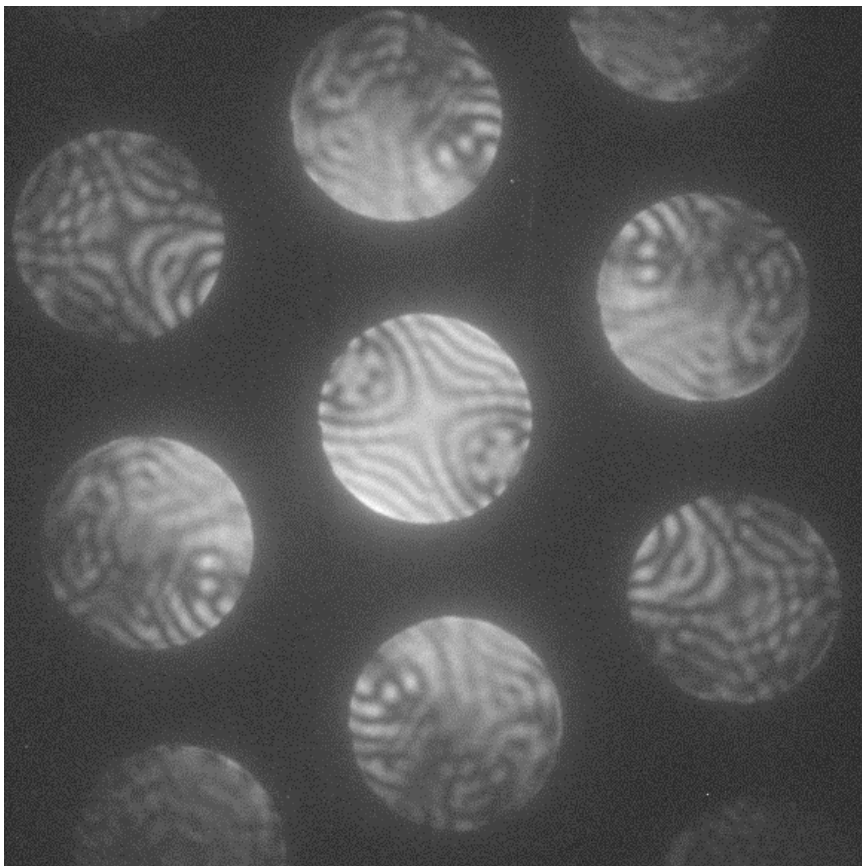
In zero loss filtering the slit is positioned so as to collect only electrons that have lost no energy, i.e. those in the zero loss peak. This means that all the inelastically scattered electrons are prevented from reaching the image.

This allows thicker regions to be examined (loss electrons suffer chromatic aberration & thus blur image). It also allows quantitative work, since most calculations assume elastic scattering only. An example is the improvement in visibility of convergent beam patterns from thick specimens.

Effect of energy filtering: Si 110 convergent beam pattern taken at 100kV (VG HB501 STEM + GIF).



Unfiltered, elastically and inelastically scattered electrons.



Zero loss filtered, elastically scattered electrons only.

b. Thickness determination

The specimen thickness can be determined by a method similar to that used to determine thickness in EELS from the plasmon peaks. The zero loss intensity given by

$$I_0 = I_{unf} e^{-\frac{t}{\lambda}}$$

so

$$\frac{t}{\lambda} = \ln\left(\frac{I_{unf}}{I_0}\right) = \ln\left(1 + \frac{I_{loss}}{I_0}\right)$$

where I_{unf} = intensity of unfiltered spectrum = $I_0 + I_{loss}$

I_0 = zero loss intensity

I_{loss} = loss electrons (ie all except zero loss)

λ = total inelastic mean free path

Thus to get a thickness map of an area we need to collect a zero loss image (I_0) and either an unfiltered image (I_{unf}) or a loss image (I_{loss}) and calculate t/λ from the above formula. Such a map is often called a “ t/λ ” map. Note that this map depends on λ as well as on the thickness. Thus it doesn't give a true thickness map for specimens of non-uniform composition.

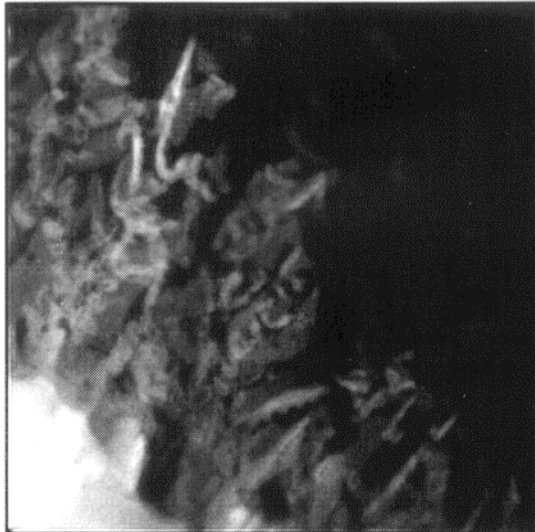
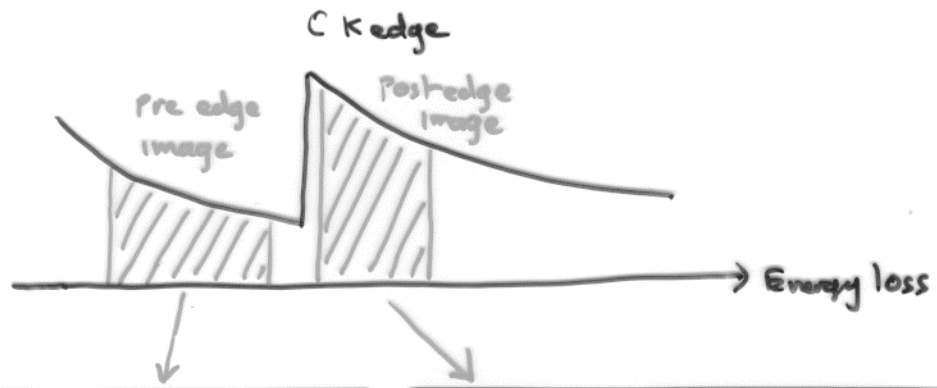
c. Core loss mapping

The aim of core loss mapping is to collect a map of a particular core loss edge, and thus also an elemental map. Core loss mapping is difficult because of need to remove the background from under the core loss edge. There are various methods of calculating core loss maps:

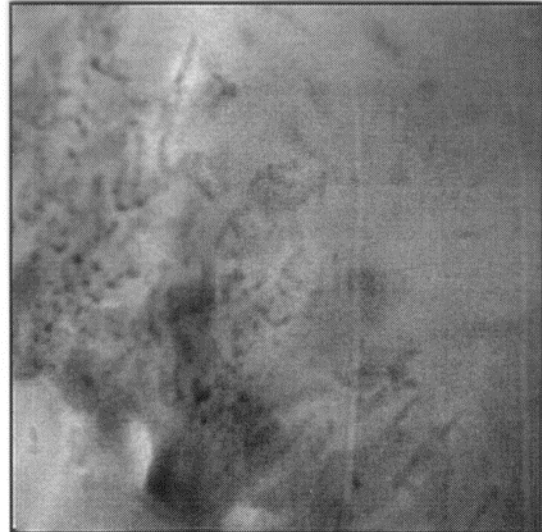
Jump ratio:

To form a jump ratio image we collect one image just before the edge (pre-edge image) and one image on the edge (post-edge image). The jump ratio image is the ratio of these images.

A jump ratio image only gives a approximate idea of the elemental concentration. Its advantage is that thickness variations are removed (at least approximately).

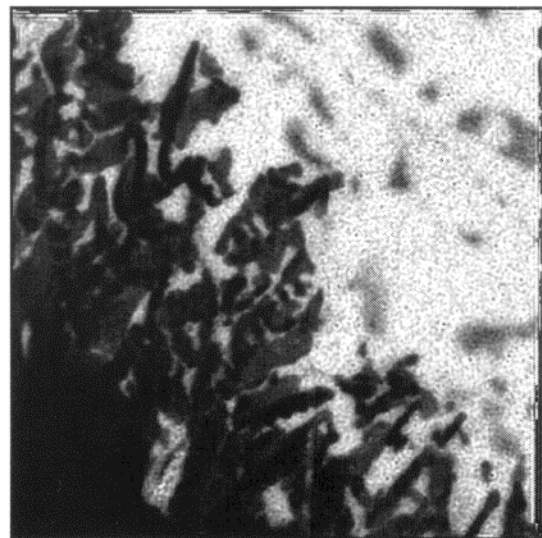


pre-edge image
 $\Delta E = 260\text{eV}$



post-edge image
 $\Delta E = 280\text{eV}$

Divide images

C jump ratio image

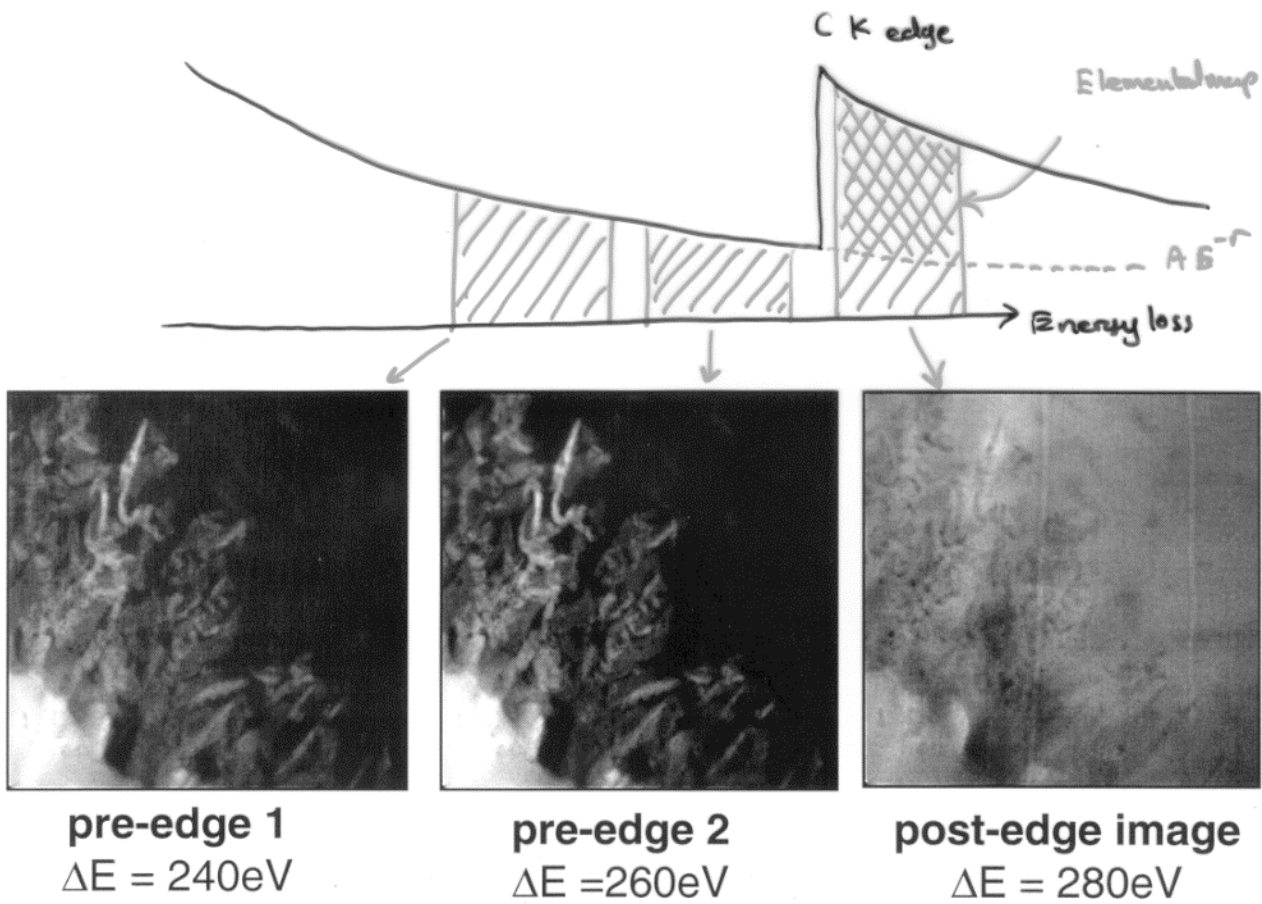
An example of a jump ratio image for carbon.

3 window method:

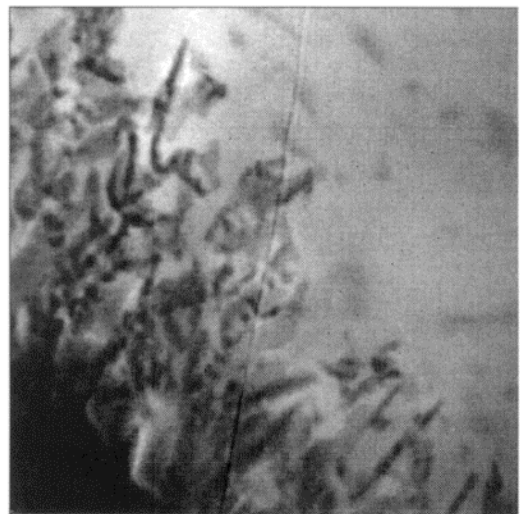
To form a 3-window image we collect 2 pre-edge images and one post-edge image.

The 2 pre-edge images are used to estimate the background under the post-edge image using the formula $I = AE^{-r}$. This background can then be subtracted from post edge image.

The method is equivalent to normal method of finding EELS edge areas from energy loss spectra and thus can be quantified. The 3-window image is the sum of the edge area over the thickness of the specimen. Hence to get true elemental maps the specimen thickness variations must be compensated for.

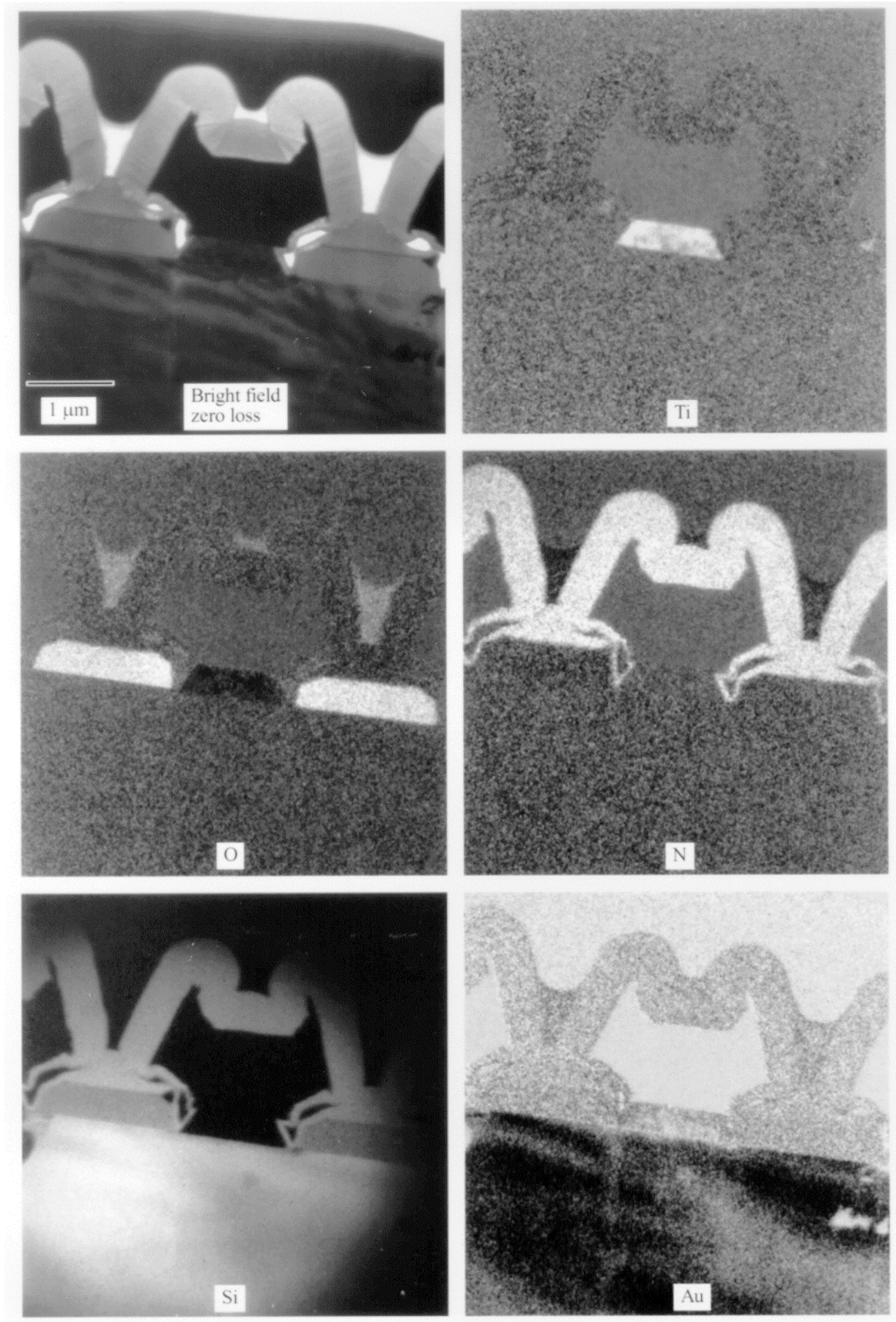


Calculate and remove
AE^{-r} background



C elemental map

An example of a 3-window elemental map for carbon.



Energy filtered mapping of a semiconductor device.

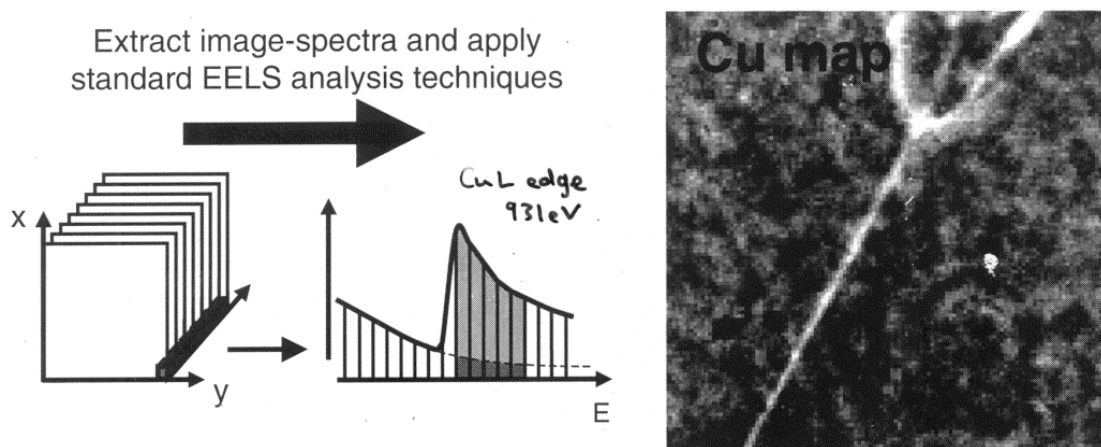
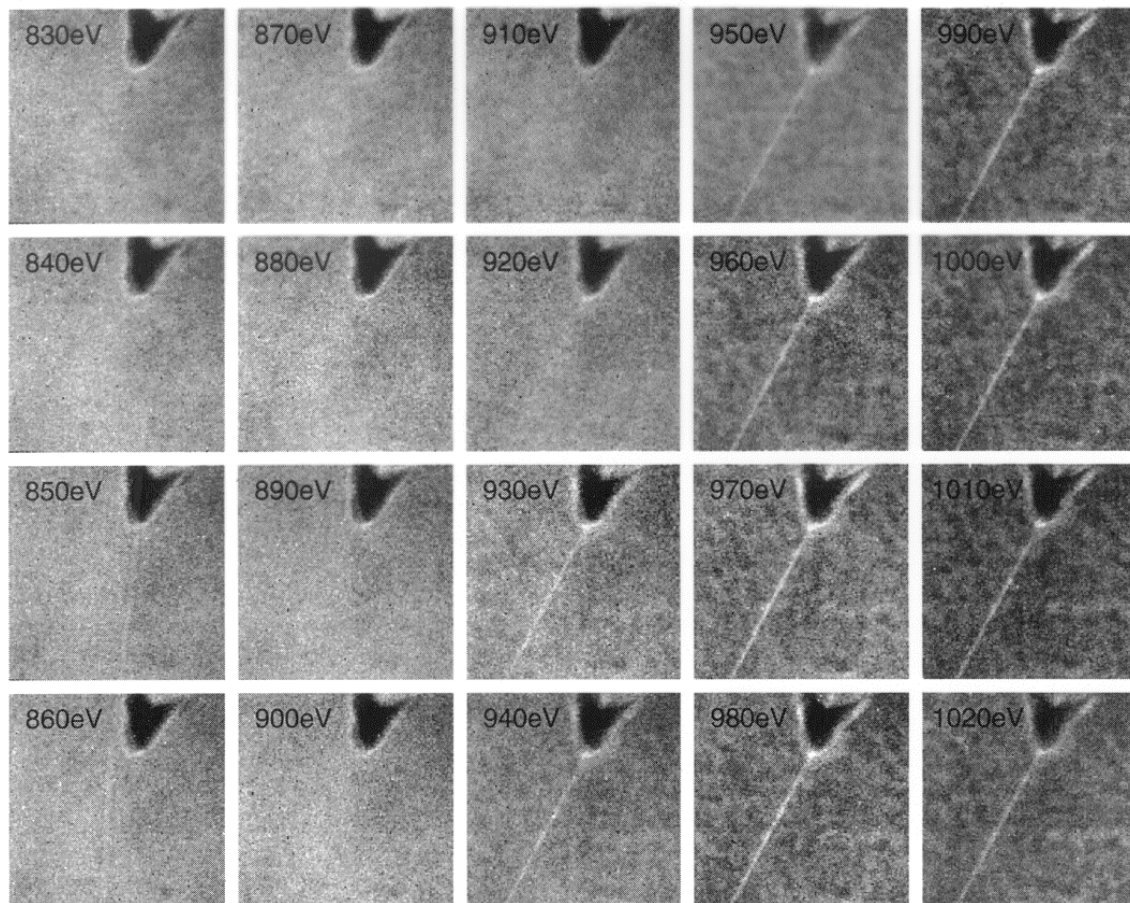
The edges used are: Ti K 456 eV, O K 535 eV, N K 400 eV, Si K 1839 eV, Au O 54 eV.

Image spectroscopy:

A big improvement in determining the background can be made using image spectroscopy. A series of images is collected both before and after the edge of interest. It is possible to fit a much better background to many images than just two.

We are effectively collecting a spectrum at each point in the image.

Image spectroscopy gives a better signal to noise ratio and is more quantitative than the 3 window or the jump ratio methods. Unfortunately it takes longer and requires a larger dose of electrons.



Example of elemental mapping using image-spectroscopy. The map is a map of Cu segregation in Al-Zn-Mg-Cu