Contents

Chapter 1 – Introductory Comments ................................................................. 5
  1.1. Microbeam Techniques in Geology ............................................................ 5
  1.2. Units .................................................................................................. 6
Chapter 2 – Historical Background .................................................................. 7
  2.1. Early Work with Cathode Rays ................................................................. 7
  2.2. Discovery of X-rays ........................................................................... 9
  2.3. Further Investigations of X-rays ............................................................. 11
  2.4. Diffraction of X-rays ........................................................................ 12
  2.5. Wave-Particle Duality ....................................................................... 15
  2.6. X-rays and Elements ....................................................................... 16
Chapter 3 – Development of Instrumentation .................................................... 20
  3.1. Magnetic Lenses ............................................................................... 20
  3.2. Scanning Transmission Electron Microscope .......................................... 20
  3.3. Scanning Secondary Electron Microscope .............................................. 21
  3.4. Electron Microprobe .......................................................................... 23
Chapter 4 – Signals Produced by Electron Bombardment .................................... 26
  4.1. Types of Signals ............................................................................... 26
  4.2. Electron Interaction Volume ................................................................ 26
    4.2.1. Introduction ................................................................................ 26
    4.2.2. Monte Carlo Simulations ............................................................... 28
    4.2.3. Detection of Signals ................................................................. 29
  4.3. Electromagnetic Spectrum .................................................................. 29
    4.3.1. Introduction ............................................................................. 29
    4.3.2. Energy-Wavelength Relationship ................................................ 30
  4.4. Continuum X-rays .......................................................................... 30
  4.4. Characteristic X-rays ....................................................................... 34
    4.4.1. Introduction ............................................................................. 34
    4.4.2. Atomic Structure and Characteristic X-rays ................................... 35
    4.4.3. Naming Characteristic X-rays .................................................... 37
    4.4.4. Moseley's Law ......................................................................... 40
    4.4.5. X-ray Intensity ......................................................................... 41
    4.4.6. Satellite Peaks ......................................................................... 42
    4.4.7. Peak Shifts .............................................................................. 43
  4.5. Cathodoluminescence ....................................................................... 44
    4.5.1. Introduction ............................................................................. 44
    4.5.2. Band Theory ............................................................................ 45
    4.5.3. Conductors, Semi-conductors & Insulators ................................... 46
    4.5.4. Types of Semiconductors .......................................................... 47
    4.5.5. Production of Cathodoluminescence ........................................... 47
  4.6. Heat .................................................................................................. 48
Chapter 1 – Introductory Comments

1.1. Microbeam Techniques in Geology

A large number of microbeam analytical techniques have been developed for the analysis of geological materials. These techniques use sharply focused incident beams of particles or energy to determine chemical or isotopic composition in the microscopic scale with diameters ranging from about 100 pm to 50 µm. The table below lists the more familiar techniques; this class will focus on the ones in red.

<table>
<thead>
<tr>
<th>Incident Beam</th>
<th>Technique</th>
<th>Effect or Measured Signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visible light</td>
<td>Reflected light microscopy</td>
<td>Reflected light Transmitted light Transmitted light</td>
</tr>
<tr>
<td></td>
<td>Polarized light microscopy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UV-IR microspectrometry</td>
<td></td>
</tr>
<tr>
<td>Continuous laser light</td>
<td>Micro-Raman spectrometry</td>
<td>Scattered light Heat and released gas</td>
</tr>
<tr>
<td>Pulsed laser light</td>
<td>Laser ablation - Inductively coupled plasma mass spectrometry (LA-ICP-MS)</td>
<td>Volatilized material</td>
</tr>
<tr>
<td>X-rays</td>
<td>X-ray microscopy</td>
<td>Transmitted x-rays Characteristic x-rays</td>
</tr>
<tr>
<td></td>
<td>Micro x-ray fluorescence analysis (Micro XRF)</td>
<td></td>
</tr>
<tr>
<td>Electrons</td>
<td>Transmission electron microscopy (TEM)</td>
<td>Transmitted electrons Backscattered electrons (BSE) Secondary electrons Characteristic X-rays Visible light Auger electrons</td>
</tr>
<tr>
<td></td>
<td>Scanning electron microscopy (SEM)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Electron microprobe analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cathodoluminescence microscopy</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Auger microprobe analysis</td>
<td></td>
</tr>
<tr>
<td>High-energy protons</td>
<td>Proton-excited X-ray emission analysis (PIXE)</td>
<td>Characteristic X-rays</td>
</tr>
<tr>
<td>Low-energy ions</td>
<td>Secondary ion mass spectrometry (SIMS)</td>
<td>Sputtered secondary ions</td>
</tr>
<tr>
<td></td>
<td>Sensitive high resolution ion microprobe (SHRIMP)</td>
<td></td>
</tr>
</tbody>
</table>

Most techniques are non-destructive, but several result in sample damage, usually in the form of pitting where material has been ablated. For example, incident pulsed laser light volatilizes material from the sample in laser-ablation ICP-MS, and low-energy ions ablate the sample leaving holes in SIMS.

Nondestructive techniques are more numerous. Many use characteristic x-rays produced from the sample to analyze composition. As indicated, there are several ways to produce x-rays from the sample, including incident x-rays, electrons, and protons. This course will focus on effects that are produced by electron bombardment (in red above). Using these effects to examine geological materials one may (among other things):

1. Identify minerals; this is especially useful for opaque and micrometer-sized grains;
2. Determine phase compositions, which are required for geothermometry and geobarometry calculations;
3. Document chemical zoning within minerals for petrologic, growth, and diffusion studies; and
4. Locate rare phases, such as zircon, monazite, and badellyite, which often have distinctive chemistries.

1.2. Units

Before proceeding, it will be useful to define the units we will use in the following materials. All units (with one prominent exception) will be in the International System of Units (abbreviated SI from the French language name Le Système international d’unités). The SI system was developed in 1960 from the meter-kilogram-second (mks) system and uses a series of prefixes that are attached to seven base units that are nominally dimensionally independent:

- meter (m), length
- kilogram, (kg), mass
- second (s), time
- ampere (A), electrical current
- kelvin (K), temperature
- mole (mol), amount of substance
- candela (C), luminous intensity

Other units are derived from these seven base units; for example, a pascal (Pa) is 1 N/m², a Newton (N) is 1 kg m/s². Prefixes are attached to base units to denote multiples (or fractions):

- giga (G), 10⁹
- mega (M), 10⁶
- kilo (k), 10³
- centi (c), 10⁻²
- milli (m), 10⁻³
- micro (µ), 10⁻⁶
- pico (p), 10⁻⁹
- nano (n), 10⁻¹²

One non-SI unit very frequently used is the Ångstrom (Å), which is 1 x 10⁻¹⁰ m. X-ray wavelengths are typically 1 to 100 Å (0.1 to 10 nm). Visible light has wavelengths from 400 (violet) to 750 nm (red). Typical distances of interest are µm (sizes of mineral grains, interaction volumes, beam diameters, etc.) and pm (atomic and ionic sizes).
Chapter 2 – Historical Background

2.1. Early Work with Cathode Rays

The English physicist Sir William Crookes (1832-1919) created a vacuum tube around 1875, which he used to study gases. The glass tube contained negative and positive electrodes across which high-voltage electrical currents. Crookes primarily worked with the conductivity of gases, which he placed inside the glass. Although he did not know it, these cathode ray tubes also produce X-rays and Crookes unsuccessfully sought cause of repeated fogging of photographic plates that he had stored nearby.

Crookes also discovered that some minerals would glow if placed in the tube. This is the phenomenon of cathodoluminescence (production of visible light by electron bombardment), which we will consider in detail later. Different minerals yield different colors: calcite (red), apatite (yellow), willemite (bright green), scheelite (bright blue), dolomite (brown), and magnesite (violet).
Later variations on the mineral tube used colored phosphors to produce dramatic displays. In the eponymous Crookes flower tube, copper flowers are covered with different phosphors. Rotating vanes made of mica plates are situated above them. When the tube is activated, the cathode rays turn the vanes, resulting in a moving shadow on the flowers below.

**Flower Tube.** When activated the flowers painted with phosphors glow in different colors. Images source: [http://members.chello.nl/~h.dijkstra19/page7.html](http://members.chello.nl/~h.dijkstra19/page7.html).

In 1858, **Julius Plücker** (1801-1868) and **Johann Wilhelm Hittorf** (1824-1914) discovered that cathode ray beams were deflected by a magnetic field. They used a modified Crookes tube, with an insert consisting of an aluminum sheet covered with phosphor.

**Deflection tube.** A magnet brought near the tube causes the cathode rays to deflect from a straight path. Images source: [http://members.chello.nl/~h.dijkstra19/page7.html](http://members.chello.nl/~h.dijkstra19/page7.html).

In 1869, Hittorf demonstrated that cathode rays go in a straight line and are blocked by metal. He used the Maltese cross arrangement shown above. When the cross was down, the glass face of the tube emitted a green glow, which faded over time. When the cross was up, intercepting the cathode rays, its shadow was visible on the end of the tube.
The German scientist Philipp Lenard (1832-1947) added a thin aluminum window to the basic Crookes tube. The window's foil was thick enough to maintain the vacuum inside the tube, yet thin enough to allow the cathode rays to pass out of it. Lenard concluded from subsequent experiments that cathode rays propagated through air for distances of about 100 cm, but that in a vacuum they traveled for several meters without being weakened. These and other early findings were published by Lenard in 1894. During one of his experiments, the fluorescence from a dissipated ray caused a few pieces of paper that had been soaking a barium platinocyanide solution to glow. Lenard had unknowingly discovered the first evidence for X-rays but failed to investigate the strange phenomenon further. Lenard later claimed that he, rather than Roentgen, should be honored as the discoverer of X-rays. Although he made significant contributions to the field of physics, Lenard was a fervent Nazi and condemned Albert Einstein and other individuals with Jewish backgrounds.

2.2. Discovery of X-rays

A few months later, on November 8, 1895, Wilhelm Conrad Röntgen (re)discovered X-rays at the University of Würzburg in Germany. Röntgen used ~40 kV electrons to bombard inert gas in tubes and noticed that a screen coated with barium platinocyanide (BaPbCN) across room
began to glow. He placed a deck of cards and a two-inch book between the tube and the screen and discovered the rays penetrated these materials. He called them X-rays after the algebraic symbol of the unknown, $x$. Once while holding lead pipe to the rays, he noticed that the bones of his fingers were shadowed on the screen. On December 22nd, Roentgen decided to show his wife what had been occupying his time and took an X-ray image of her hand.

Röntgen with X-ray photographs of this wife Bertha’s hand (left) and of von Kölliker’s hand (right). Sources: http://www.deutsches-museum.de/sammlungen/ausgewaehlte-objekte/meisterwerke-ii/roentgen/ (left and center), http://www.vmas.kitasato-u.ac.jp/radiology/CALS/von-Kolliker.htm (right)

Röntgen also demonstrated that X-rays:

- were produced from the fluorescence part of the wall of the discharge tube;
- traveled in straight lines;
- were not deflected by a magnetic field;
- were absorbed more by denser metals;
- were scattered when passing through a body;
- and could ionize gases.

He submitted the first paper describing his work, Über eine neue Art von Strahlen, to the Würzburg Physical Medical Society on 28 December 1895. He gave a public lecture on 23 January 1896, after which he made a plate of hand of Alfred von Kölliker, a famous anatomist. Kölliker proposed that the newly discovered rays be named Röntgen's Rays; X-rays are still called Röntgenstrahlen in Europe. Röntgen received the first Nobel Prize in Physics in 1901 for his discovery, donating the prize money (then about $40,000) to the University of Würzburg.

In addition to Lenard, there were others who had observed the effects of X-rays before Röntgen. In the United States, A. W.Goodspeed (1860-1943) and William Jennings (1860-1945) made an accidental photograph of coins stacked with photographic plates using X-rays from Crookes tube on 22 February 1890 in Philadelphia. However, unlike Lenard, neither claimed priority for discovery of X-rays, noting that they had ignored the plates until Röntgen’s announcement caused them to review the photographs.
X-rays rapidly became the latest rage. The public was fascinated with this unknown phenomenon; people were amazed at seeing through human flesh; however, they were commonly ill at ease because bones were associated with death. Indeed, people occasionally fainted when first they saw an X-ray image! Thomas Edison capitalized on the American public’s intoxication with the X-rays and announced in March 1896 that he would be the first to photograph the living human brain; however, he never accomplished this feat. American physicians quickly recognized the value of X-rays for examining bone fractures and locating foreign objects in the body. The first diagnostic X-ray was taken at Dartmouth Hitchcock Hospital in 1896; Gilman Frost and his Dartmouth physicist brother, Edwin, used X-ray imaging to help set a boy’s broken arm.


2.3. Further Investigations of X-rays

Another major figure in the investigation of the properties of X-rays was Charles Glover Barkla (1877-1944). He demonstrated in 1905 that X-rays could be polarized, thus having properties
identical to visible light. He thus eliminated the possibility that X-rays might be a type of longitudinal wave (like sound) rather than a transverse wave.

In 1909, Barkla discovered what are now termed **characteristic X-rays**. He found that under certain conditions, the emergent X-rays contained one strong homogeneous component with a constant absorption coefficient. Barkla documented that the absorption coefficient decreased with increasing atomic weight of the anode material. Since no one could yet measure the wavelengths (or frequencies) of X-rays, Barkla measured the absorption of the secondary radiation by directing the secondary radiation through a 0.01 cm-thick layer of aluminum and measuring how much of the beam was absorbed. Barkla next discovered that his homogeneous X-rays were, in fact, heterogeneous. He wrote: "... the radiations from Sn, Sb, I ... emit ... radiation of variable penetrating power."

Plotting his results yielded two monotonic curves, one for the lighter elements and one for the heavier ones, with some elements showing two absorption values. At first, he labeled the two curves with by letters B and A, but in 1911 changed his notation remarking that, "[t]he letters K and L are, however, preferable as it is highly probable that series of radiations both more absorbable and more penetrating exist."

In 1912-13, R. T. Beatty demonstrated that electron bombardment of the anode produced two types of x-rays: the characteristic X-rays described by Barkla and continuum radiation, which is also called **bremsstrahlung**. He noted that characteristic X-rays required a threshold electron energy, below which only continuum X-rays were produced. Beatty also showed that depth of x-rays production by electron bombardment was very small (<10 mm).

### 2.4. Diffraction of X-rays

Max von Laue (1879-1960), a junior colleague of Röntgen, considered passage of waves of light through crystalline arrangement of particles and realized that short wavelength electromagnetic rays, such as X-rays, should cause diffraction or interference phenomena in crystals. In 1912, he and his lab assistant, Walter Friedrich and doctoral student, E. Paul...
Knipping, confirmed diffraction of X-rays by systematic crystal. Von Laue worked out a mathematical formulation for this behavior and published this discovery.

Von Laue took diffraction as positive proof that X-rays are electromagnetic radiation (waves), not particles. However, he was forced to suggest that the incident X-rays must contain only certain wavelengths to account for missing diffracted beams from the crystal. He received Nobel Prize in 1914 for his work. X-ray diffraction work by Rosalind Franklin (1920-1958) was key in determining the structure of DNA.

In 1896, when William Henry Bragg learned of W. K. Röntgen's discovery of X-rays, he set about producing the new radiation. On 13 June 1896, he photographed his son's broken elbow using the primitive equipment. W. H. Bragg was proponent of the idea that X-rays were particles, not waves. He supported this view with the evidence that X-rays produced ionization. Bragg engaged in controversy with Barkla, who advocated the wave theory; however, in retrospect, the source of their differences was that he was studying high energy (hard) X-rays, whereas Barkla was studied low-energy (soft) X-rays. Bragg developed an X-ray spectrometer that allowed many different types of crystals to be analyzed by precisely measuring the diffraction.
angles, but perhaps Bragg's greatest accomplishment was developing an X-ray detector based on ionization of gas. Detectors of this type are still used today in many applications.

In August 1912, William Lawrence Bragg (1890-1971), WHB’s son who had just graduated from Cambridge, realized that the pattern of spots in the Laue diffractogram could be explained by reflection of waves from crystal planes. L. W. Bragg used his father's spectrometer to observe the diffraction of Pt-Lα X-rays by NaCl crystal. He published his results in 1913, *The diffraction of short electromagnetic waves by a crystal*, in the Proceedings of the Cambridge Philosophical Society. In it, WLB concluded that salt consisted of a three-dimensional lattice of Na⁺ and Cl⁻ ions. For some time after, chemists refused believe that NaCl contains no NaCl molecules (just alternating array of Na⁺ and Cl⁻ ions! The collaboration between father and son led many to believe that WHB had initiated the research, a fact that upset WLB and continued to haunt him throughout his life. The Braggs shared the 1915 Nobel Prize in Physics.
While studying scattering of X-rays by crystals in 1913, the Braggs noticed a different distinctive pattern of peaks appeared for each of the different anodes used to produce X-rays. These occurred when there was constructive interference of the characteristic X-rays. This method is used today in wavelength-dispersive spectrometers to examine elements of interest.

Moseley used a simple flat diffracting crystal, but it was later shown that curved crystals would provide higher intensities by better "focusing" the diffracted X-rays. In 1931, H. H. Johann showed that bending the crystal would produce a better focus, however, this results in some broadening and asymmetry of the focus point. In 1933, T. Johannson proposed grinding the surface of a bent crystal to achieve a perfect focusing. However, the machining difficulties of such a procedure are tremendous and most modern microprobe use Johann crystals.

2.5. Wave-Particle Duality

The apparent conflict between those who considered X-rays particles and those who considered them waves reflects the way they were studied. All electromagnetic energy and particles depending upon the method of examination will show either wave-like behavior (diffraction) or particle like behavior (discrete energies, ionization, scattering, photoelectric effect). In 1924, Louis-Victor de Broglie (1892-1987) formulated the de Broglie hypothesis, claiming that matter, not just light, has a wave-like nature. He related wavelength, \( \lambda \), and momentum, \( p \):

\[
\lambda = \frac{h}{p} = \frac{h}{m \nu} \sqrt{1 - \frac{\nu^2}{c^2}}
\]

where \( \lambda \) = the particle's wavelength, \( h \) = Planck's constant, \( p \) = the particle's momentum, \( m \) = the particle's rest mass, \( \nu \) = the particle's velocity, and \( c \) = the speed of light in a vacuum. For example, consider the wavelength of a baseball (0.15 kg), thrown at 90 mph (40.2 m/s):

\[
\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34} \text{ kg m}^2\text{s}^{-1}}{0.15 \text{ kg} \times 40.2 \text{ m s}^{-1}} = 1.095 \times 10^{-34} \text{ m}
\]

This wavelength is considerably smaller than the diameter of a proton (about \( 10^{-15} \text{ m} \)) and approaches the Planck length (\( 1.61 \times 10^{-35} \text{ m} \)). The wave-like properties of this baseball are too small to be observable. In contrast, electrons have much smaller mass and move much faster, so much faster that relativistic effects (mass increase) must be considered. The wavelength of 10 kV electrons is \( 12.3 \times 10^{-12} \text{ m} \), permitting diffraction to be observed.
Electron Interference. Electrons are sent through the slits one at a time, but still build up an interference pattern like that expected of waves (time elapses $a \rightarrow b \rightarrow c \rightarrow d$). Image sources: (left) [http://www.blacklightpower.com/theory/DoubleSlit/Fig_37-1_Two_Slit_Particles.jpg](http://www.blacklightpower.com/theory/DoubleSlit/Fig_37-1_Two_Slit_Particles.jpg), (right) [http://www.hqrd.hitachi.co.jp/em/emgif/fig2.gif](http://www.hqrd.hitachi.co.jp/em/emgif/fig2.gif).

Wikipedia reports that the diffraction of C\textsubscript{60} fullerenes was reported by researchers from the University of Vienna in 1999. Fullerenes have an atomic mass of about 720 amu; their de Broglie wavelength is $2.5 \times 10^{-12}$ m. The diameter of the molecule is about 400 times larger. As of 2005, this is the largest object for which wave-like properties have been directly observed.

Image source: [http://abyss.uoregon.edu/~js/images/wave_particle.gif](http://abyss.uoregon.edu/~js/images/wave_particle.gif)

### 2.6. X-rays and Elements

In late 1913, Henry Gwyn Jefferys Moseley constructed X-ray spectrometer using a potassium ferrocyanide, K\textsubscript{4}Fe(CN)\textsubscript{6}·3H\textsubscript{2}O, diffracting crystal. Moseley had been working on radioactivity since 1910 but decided to study X-ray diffraction since it was the hottest new field in physics. With his new spectrometer, he measured the frequency of characteristic K-series X-rays produced from tubes with different anode materials (Ca to Zn). His results showed that the ordering of the wavelengths of the X-ray emissions of the elements coincided with the ordering of the elements by atomic number. This relationship is called Moseley's Law and will be discussed in detail later in this course.
Prior to Moseley's work, atomic numbers were considered semi-arbitrary, based on atomic masses but altered when necessary to put an element in the appropriate place in the periodic table. For example, cobalt and nickel had been assigned atomic numbers of 27 and 28, respectively, based on their chemical properties, since they have nearly identical atomic mass. Moseley's experiments showed that the order of these elements should be reversed. Moseley showed that there were gaps in the atomic number sequence at numbers 43, 61, 72, and 75 and suggested that these would be filled by undiscovered elements. After this work, which was done in Manchester, Moseley moved to Oxford in 1914 and continued his experiments by studying the L-spectra of heavier elements.
Moseley's Results. The linear relationship between atomic number (y-axis) and the square root of frequency (x-axis) is Moseley's Law. Notice in this diagram that the lighter elements have two distinctive characteristic X-ray lines plotted and the heavier have up to three. Image source: http://photos.aip.org/history/Thumbnails/moseley_henry_d9.jpg.

However, with the onset of war, Moseley enlisted in the Royal Engineers and was killed on the Gallipoli Peninsula (Turkey), shot through the head while telephoning an order on 10 August 1915. He was just 27. Belatedly, recognizing this great loss to science, the British army changed its policy in World War II, no longer allowing scientists to enlist for combat.

In 1923, Georg Charles von Hevesy proposed using X-rays to excite characteristic X-rays from a sample and use diffraction to identify its constituent elements. This is termed X-ray fluorescence. Later he introduced a method of activation analysis based on neutron
bombardment (neutron activation); this method yields better detection limits than X-ray analysis with fluorescent X-rays. In 1923, he and Dirk Coster discovered the element hafnium in X-ray spectrum emitted from zircon. He received the 1943 Nobel Prize in Chemistry for his work on the use of isotopes as tracers in studying chemical processes. Hevesy was the first to apply the radioactive tracer technique to biology, and he later used it in medical research.

Chapter 3 – Development of Instrumentation

3.1. Magnetic Lenses

Hans Busch showed theoretically in a 1927 paper that coaxial magnetic field produced by an electric coil could be expected to focus a beam of electrons. He also predicted that the focal length of such a magnetic electron lens could be changed continuously by varying the coil current. This theory was confirmed in 1929 by Ernst Ruska (1906-1988) at the High Voltage Institute, Berlin, under the direction of Max Knoll (1897-1969). In 1931, Ruska constructed a magnetic lens of the type that has been used in all magnetic high-resolution electron microscopes since then. Further work, conducted with Knoll, led to the construction in 1933 of an electron microscope that for the first time gave better definition than a light microscope. Ruska's 1934 Ph. D. thesis investigated the properties of electron lenses with short focal lengths.

Ruska and Knoll's Electron Microscope. The first electron microscope was a transmitted electron instrument shown here in a wonderful mad-scientist photograph (Knoll is at left). Image source: http://www.microscopy.ethz.ch/history.htm.

Major limitation of their microscope was that electrons are unable to pass through thick specimens. Thus, it was impossible to utilize the instrument to its full capacity until the diamond knife and ultra-microtome were invented in 1951. Ruska was awarded half of 1986 Nobel Prize for Physics; the other half was divided between Heinrich Rohrer and Gerd Binnig for their invention of the scanning tunneling microscope (STM).

3.2. Scanning Transmission Electron Microscope

The earliest known paper presenting the concept of a scanning electron microscope was by Knoll (1935). Subsequently, in 1938, Manfred von Ardenne (1907-1997) constructed a scanning transmission electron microscope (STEM) by adding scan coils to a transmission electron microscope. The first micrograph was of a ZnO crystal imaged at an operating voltage of 23 kV.
at a magnification of 8000 times. The spatial resolution was between 50 and 100 nm. The micrograph contained 400 x 400 scan lines and took 20 min to record, because the film was mechanically scanned in synchronization with the beam. The instrument also had a viewing CRT, but it was not used to record the image.

![Von Ardenne and the Scanning TEM](http://www.g.eng.cam.ac.uk/125/achievements/mcmullan/images/fig4.jpg) ![Von Ardenne and the Scanning TEM](http://www.dhm.de/lemo/objekte/pict/f64_1025/200.jpg)

**3.3. Scanning Secondary Electron Microscope**

The first, true scanning electron microscope (SEM) was developed and described in 1942 by Russian immigrant **Vladimir Kosmo Zworykin** (1889-1982), J. Hillier and R. L. Snyder, working in the RCA Laboratories in the United States. Zworykin also invented the cathode-ray tube called the kinescope in 1929 the precursor of modern television picture tubes; he is sometimes called the "father of television."
The instrument consisted of an inverted column (electron gun at the bottom), three electrostatic lenses and electromagnetic scan coils placed between the second and third lenses. A photomultiplier tube detected the scintillations on a phosphor screen caused by the secondary electron emissions. This detector was an early version of the combination of phosphor and photomultiplier that Everhart and Thornley used nearly twenty years later. The electron gun was located at the bottom, so the specimen chamber was at a comfortable height for the operator. However, this arrangement had the slight disadvantage that the specimen could fall down the electron column! The instrument achieved a resolution of about 50 nm, but this figure was considered unexciting compared to results achieved by the rapidly developing TEM, and further development lapsed.

![Vladimir Zworkyin and a Schematic of the first SEM.](image1.jpg)

The early history of the SEM is very well described at the Cambridge University, Department of Engineering SEM site. The following summary is based on this information.

In the late 1940s, Charles Oatley at the Engineering Laboratories of the University of Cambridge decided that another look at the SEM might be worthwhile. He had decided that "Zworykin ... had shown that the scanning principle was basically sound" and believed that the improvements in electronics that had resulted from work during the war would allow better results. Specifically, Oatley thought that the RCA detector had a too low efficiency and thus that the images were noisy in spite of the long recording time.
Oatley selected Dennis McMullan to build an SEM as his Ph.D. project. McMullan first finished building a 40 kV electrostatically focused TEM, which had been begun by another PhD student, converting it into an STEM. Next, he converted the STEM into an SEM by adding scan coils, an electron multiplier detector, and a long persistence cathode-ray tube.

Many improvements were incorporated in subsequent SEM2, SEM3, and SEM4 models. In 1952, K. C. A. Smith developed a way to efficiently detect low-energy secondary electrons (previously images were made using high-energy electrons), allowing vastly improved surface imaging. In 1955, T. E. Everhart devised a new electron detector; the "Everhart-Thornley" detector is still used today. Magnetic lenses were added to the SEM4 in 1961. Eventually, Oatley persuaded Associated Electrical Industries (AEI), a company that manufactured both transmission electron instruments and electron probe microanalyzers, to take an interest in the SEM. The first four production models, sold under the trade name "Stereoscan", were delivered in 1965.

3.4. Electron Microprobe

In 1944, James Hillier (1915-2007) and Baker at the Radio Corporation of America (RCA) Labs at Princeton, New Jersey, built an electron microprobe, combining an electron microscope and an energy-loss spectrometer. Electron energy-loss spectrometry is very good for light element analysis and they obtained spectra of C-Kα, N-Kα and O-Kα radiation. In 1947, Hillier patented the idea of using an electron beam to produce analytical X-rays, but never constructed a
A working model. His proposed design used Bragg diffraction from a flat crystal to select specific X-ray wavelengths and a photographic plate as a detector.

**Hillier's Electron Microprobe.** a. Schematic of microprobe constructed; b. Patented design of the unconstructed microprobe. Note that the electron gun is at the bottom of the column in both designs. Image sources: (left) [http://www.geology.wisc.edu/~johnf/g777/Hillier-EELS.jpg](http://www.geology.wisc.edu/~johnf/g777/Hillier-EELS.jpg), (center) [http://www.geology.wisc.edu/~johnf/g777/Hillier_xtal.jpg](http://www.geology.wisc.edu/~johnf/g777/Hillier_xtal.jpg), (right) [http://www.cedmagic.com/mem/whos-who/hillier-james.html](http://www.cedmagic.com/mem/whos-who/hillier-james.html).

In 1948-1950, Raimond Castaing (1921-1999), supervised by André Guinier built the first electron microprobe ("microsonde electronique") at the University of Paris. He was apparently unaware of Hillier's ideas and patents and worked independently. The resulting instrument produced an electron beam diameter of 1-3 mm with beam current of ~10 nA and used a Geiger counter as a detector. In 1950, Castaing added a fully focused Johannson quartz crystal between sample and detector to permit wavelength discrimination and an optical microscope to viewpoint of beam impact.

Castaing's 1951 Ph.D. thesis laid the foundations of the theory and application of quantitative analysis by electron microprobe. Castaing recognizing that unknown X-ray intensities measured relative to a pure element could be used as a first approximation for quantifying the chemical composition of the specimen being bombarded. He soon realized that X-ray generation in a multi-element specimen was complex. For example, the interaction volume varies with composition, as does the path for the x-rays leaving the specimen. These complications are known as atomic number (Z) and absorption (A) effects. The interaction is further compounded by a secondary fluorescence effect (F), which addresses the generation of secondary x-rays by absorption of primary x-rays within the specimen. Castaing also discussed effects such as instrumental drift and X-ray background. Because of the enormous breadth of his contributions to the field, Castaing is often described as the "father" of electron microprobe analysis.
Cameca (France) produced the first commercial microprobe, MS85, in 1956; based on Castaing’s design. Later in his distinguished career, Castaing was involved in the development of the ion beam microprobe.
Chapter 4 – Signals Produced by Electron Bombardment

4.1. Types of Signals

In the SEM and electron microprobe, the sample is bombarded by a focused beam of electrons. Most incident electrons, rather than penetrating the sample in a linear fashion, interact with specimen atoms and are scattered. They follow complicated twisting paths through the sample material, losing energy as they interact. The scattering events are of two types either elastic or inelastic. In elastic scattering, the electron trajectory changes, but its kinetic energy and velocity remain essentially constant. This is because of the large difference between the mass of the electron and atomic nuclei. Alternatively, in inelastic scattering, the trajectory of the incident electron may be only slightly perturbed, but energy is lost through interaction with the orbital electrons of the atoms in the specimen. Inelastic interactions produce a number of effects:

- Secondary electrons (SE)
- Backscattered electrons (BSE)
- Cathodoluminescence (CL)
- Continuum x-ray radiation (bremsstrahlung)
- Characteristic x-ray radiation
- Phonons (heat)

Interaction Effects. Some the interaction effects due to electron bombardment emerge from the sample. Some, such as sample heating (not shown) stay within the sample. The lines within the interaction volume delineate regions where the effect indicated predominates (see figure below). For example, only x-rays emerge from the sample from the deeper parts of the volume.

4.2. Electron Interaction Volume

4.2.1. Introduction

The volume of material analyzed by an electron microprobe (or SEM equipped with X-ray spectrometers) depends upon many factors. The interaction volume is limited by energy loss
through inelastic interactions and electron loss or backscattering through essentially elastic interactions. The probability of beam electrons interacting (scattering) with an atom or ion in the sample is specified by the interaction cross-section. However, interaction cross-sections are very small, corresponding to very low probabilities of interaction with a given atom. Cross-sections are measured in barns (1 barn = \(10^{-24} \text{ cm}^2 = 10^{-4} \text{ pm}^2\)). However, there are a very large number of atoms in a small volume of a material, making the total probability of interaction unity. A single electron may be scattered many times.

The depth of electron penetration of an electron beam and the volume of sample with which it interacts are a function of its angle of incidence, the magnitude of its current, the accelerating voltage, and the average atomic number (Z) of the sample. The resulting excitation volume is a hemispherical to jug-shaped region with the neck of the jug at the specimen surface. The analyst must remember that the interaction volume penetrates a significant depth into the sample.

Electron penetration generally ranges from 1-5 µm with the beam incident perpendicular to the sample. The depth of electron penetration, \(x\), is approximately (Potts, 1987, p. 336):

\[
x(\mu m) = \frac{0.1E_0^{1.5}}{\rho}
\]

where \(E_0\) = accelerating voltage (kV) and \(\rho\) = density (g/cm³). For example, bombarding a material with a density of 2.5 g/cm³, about the minimum density for silicate minerals, with \(E_0 = 15\) kV, gives \(x = 2.3\) µm. The width of the excited volume, \(y\), can be approximated by (Potts, 1987, p. 337):

\[
y(\mu m) = \frac{0.077E_0^{1.5}}{\rho}
\]

Both of these are empirical expressions. A theoretical expression for the range of an electron, \(r\), the straight-line distance between where an electron enters and its final resting place, for a given \(E_0\) is (Kanaya and Okayama, 1972):
\[ r(\mu m) = \frac{2.76 \times 10^{-2} AE_{0}^{1.67}}{\rho Z^{0.89}} \]

where \( A \) = average atomic weight (g/mole), \( E_{0} \) = accelerating voltage (kV), \( Z \) = average atomic number, \( \rho \) = density (g/cm³)

### 4.2.2. Monte Carlo Simulations

The volume of electron interaction may be modeled from first principles using a Monte Carlo method, where the paths of a series of incident electrons are modeled probabilistically with equations for elastic and inelastic scattering determining the scattering angles, mean free-paths, and the rate of energy-loss. Each electron trajectory is simulated in stepwise fashion.

![Monte Carlo Simulation of Electron Paths](image1)

**Monte Carlo Simulation of Electron Paths.** This simulation is of 15 KV electrons in fayalite (Fe₂SiO₄). Distances are given in nanometers (1000 nm = 1 µm). Paths of backscattered electrons are in red; those of absorbed electrons in blue. One should remember that this slice through a three-dimensional volume. This model was run using the Casino software described at http://www.gel.usherbrooke.ca/casino/What.html.

Monte Carlo results may be used to models not just the electron paths, but distribution of the resulting effects and implanted energy. This documents the strong dependence of depth of interaction on atomic number and density.

![Monte Carlo Simulation of Energy Distributions](image2)

**Monte Carlo Simulation of Energy Distributions.** Energy distributions assuming 15 kV electrons and the materials indicated. The contour lines correspond to energy amounts of 5% (pale blue), 10% (red), 25% (green), 50% (yellow), 75% (dark blue), and 90% (purple). The models were run using the Casino software described at http://www.gel.usherbrooke.ca/casino/What.html.
4.2.3. Detection of Signals

Two factors control which effects can be detected from the interaction volume. Firstly, some effects are only produced from certain parts of the interaction volume. Beam electrons lose energy during interactions within the sample. Electron energy generally decreases with depth; thus, if a certain amount of energy is required to produce an effect, it will not be possible to produce it from the deeper portions of the volume. Secondly, the degree to which an effect, once produced, can be observed is controlled by how strongly it is diminished by absorption and scattering in the sample. For example, secondary electrons are produced throughout the interaction volume, but have very low energies and can only escape from a thin layer near the sample's surface. Similarly, soft x-rays, which are absorbed more easily than hard x-rays, will escape more readily from the upper portions of the interaction volume. Absorption is a very important phenomenon and is discussed in detail below.

4.3. Electromagnetic Spectrum

4.3.1. Introduction

Of all the effects produced from the interaction volume, characteristic X-rays provide the greatest amount of information, reflecting as they do the composition of the material. X-rays represent an energetic portion of electromagnetic spectrum with $\lambda$ from $\sim$1 nm to 1000 nm (0.1 to 100 Å). X-rays with $\lambda$ from $\sim$50 to 100 Å are termed "soft" X-rays; those with shorter wavelengths are called "hard" X-rays and overlap with $\gamma$-rays.
4.3.2. Energy-Wavelength Relationship

X-rays are specified either by their energies (in electron volts) or wavelengths (in Å). It is very useful to be able to convert readily between these units. For all electromagnetic radiation:

\[ E = h \nu \]

where \( h \) = Planck’s constant \((6.6260 \times 10^{-34} \text{ Joule-second})\) and \( \nu \) = frequency in cycles/second (Hertz).

For electromagnetic radiation,

\[ \nu = \frac{c}{\lambda} \]

where \( c \) = speed of light \((2.99782 \times 10^8 \text{ m/sec})\) and \( \lambda \) = wavelength (m).

Substitution yields

\[ E = \frac{hc}{\lambda} \]

and plugging in appropriate values gives

\[ E = \frac{1.98636 \times 10^{-25}}{\lambda} \]

where \( \lambda \) = wavelength (m) and \( E \) = energy (Joules).

Conversion to nanometers and electron volts \((1 \text{ eV} = 1.6021 \times 10^{-19} \text{ Joule})\) yields the Duane-Hunt equation:

\[ E = \frac{1239.84}{\lambda} \]

where \( \lambda \) is in picometers (pm) and \( E \) is in electron volts (eV).

If a conversion from Ångstroms to eV is desired, the appropriate conversion factor is 12398.4.

4.4. Continuum X-rays
Continuum X-rays are produced when incident beam electrons are slowed to varying degrees by the strong electromagnetic field of atomic nuclei in the sample. All degrees of electron braking are possible and, thus, the resulting X-rays have a continuous range of all energies. Each incident electron potentially can undergo many such interactions in a solid. Continuum radiation is also called bremsstrahlung, German for "braking radiation".


The highest energy X-ray that can be produced by electrostatic braking has an energy equivalent to the accelerating voltage of the electron beam. This cutoff is sometimes described as the short wavelength limit and may be written as

$$\lambda_{swl} = \frac{hc}{E_0},$$

Where $E_0 =$ accelerating voltage. The wavelength of maximum continuum intensity ($\lambda_{max}$) occurs at approximately 1.5 $\lambda_{swl}$. Increasing the accelerating voltage shifts $\lambda_{max}$ toward $\lambda_{swl}$, while $\lambda_{swl}$ moves to shorter wavelength and the overall X-ray output of the continuum increases. Increasing the beam current, increases the overall X-ray output of the continuum, but $\lambda_{swl}$ and $\lambda_{max}$ remain the same. Continuum intensity is a function of three variables: atomic number ($Z$), beam current ($i_b$) and the accelerating voltage ($E_0$). The relationship between intensity, $I$, and energy of interest, $E$, is given by Kramer's equation, which is derived from classical theory:

$$I_c = i_b Z \frac{(E_0 - E)}{E}$$

This relationship yields an exponential curve with $I_c = 0$ at $E = E_0$ with $I_c \to \infty$ as $E \to 0$. However, the observed X-ray continuum drops off at very low energies (long wavelengths), because very soft X-rays are easily absorbed. Increasing either the accelerating voltage, average atomic number of the sample, or the beam current produces a higher continuum.
Continuum X-rays. (left) Effect of changing accelerating voltage and beam current; (right) Continuum observed from graphite enclosed in Fe-Ni metal. The spectrum was accumulated using an EDS system, using a 15 kV accelerating voltage. Low channel numbers correspond to low energies. The small peak near channel 622 is due to secondary fluorescence producing characteristic X-rays in the surrounding Fe-Ni metal.

It should be noted that the continuum is nonlinear. The continuum radiation limits minimum detectable amount of an element and its curvature may present difficulties in insuring that the backgrounds of characteristic X-ray peaks are correctly determined. Backgrounds are usually measured on each side of a peak of interest assuming a linear background between them. However, if the continuum is nonlinear and the backgrounds are measured too far from peak assuming linear background, incorrect peak intensity will be determined. The peak intensity will be under measured if there is a slight upward curvature in the continuum; conversely, intensity will be overestimated if there is a downward curvature. These problems are only significant when analyzing for trace elements.

Only a very small fraction of the incident energy from the electron beam is emitted as continuum X-rays. The fraction of beam energy, $p$, may be approximated by

$$p = 1.1 \times 10^{-6} Z E_0,$$

where $Z$ = the average atomic number, and $E_0$ is the accelerating voltage (energy) of the incident electrons. For example, 15 kV electrons bombarding an iron sample ($Z = 26$) yield only 0.04% of their energy as continuum X-rays. Most energy is dissipated as heat in sample. The low efficiency of continuum production makes it necessary to cool the X-ray tubes used in XRF and X-ray diffraction (XRD) analysis.

Continuum radiation is peculiar to particle bombardment (electrons or protons), resulting as it does from electrostatic charges; bombardment of a sample with X-rays, as in X-ray fluorescence (XRF) analysis will not produce continuum radiation, yielding much lower background counts. Consequently, element detection limits are much lower in XRF analysis than electron beam analysis. Particle mass also plays a role in determining the magnitude of the continuum radiation. The intensity of continuum radiation, $I$, emitted by a decelerating particle is

$$I \propto a^2 \propto \left( \frac{t}{m} \right)^2,$$
where $a$ = the acceleration of the particle, $f$ = the decelerating force, and $m$ = the mass of the particle. The intensity of continuum produced by electron bombardment is greater than that produced by comparably accelerated protons because of the mass difference. Similar nuclear coulombic forces ($f$) are involved, but protons are 1836 times more massive than electrons. As a consequence, proton-induced X-ray emission (PIXE) has markedly lower backgrounds than those produced during electron bombardment, allowing detection of trace elements.

Spectrum observed from proton bombardment. PIXE spectrum of a spin-valve type recording head, which has many thin layers composed of similar mass elements. Note the accelerating voltage of 2.3 MeV and the essentially flat continuum background! Image source: http://www.almaden.ibm.com/st/scientific_services/materials_analysis/ib_surface/pixe/PIXEData.gif.
4.4. Characteristic X-rays

4.4.1. Introduction

The beam electrons knock a small fraction of electrons out of inner shells orbitals, in a process called *inner-shell ionization*. An atom remains ionized for only ~10^-14 second before inner-shell vacancies are filled by outer-shell electrons, emitting a characteristic X-ray. Each atom interacts with an incident electron every ~10^-12 seconds, thus, it is possible for a given atom to be repeatedly ionized. Approximately 0.1% of the beam electrons cause K-shell ionizations.

This process requires critical threshold energy to occur; beam electrons with lower energies cannot displace inner-shell electrons. This threshold energy is called the *excitation potential* or *critical ionization potential* (E_c). Excitation potentials decrease with distance from the nuclei reflecting the diminishing electrostatic forces holding the orbital electrons. Excitation potentials are always larger than the energy of the associated X-ray. Incident electrons that are not energetic enough to excite Kα radiation may still have sufficient energy can excite Lα or Mα radiation. For example, E_c for Fe-Kα is 7.11 kV; whereas, E_c for Fe-Lα is only 0.71 kV. X-ray analysts use Lα (and β) lines for elements with atomic numbers from about 30 to 70 and Mα lines for elements with higher atomic numbers.

![Production of Characteristic X-rays](http://nobelprize.org/educational_games/physics/x-rays/what-3.html); (right) [http://ehs.unc.edu/training/self_study/xray/9.shtml](http://ehs.unc.edu/training/self_study/xray/9.shtml)

The spectrum of characteristic X-ray peaks produced during neutralization of inner-shell vacancies is superimposed on the continuum X-ray spectrum.
**EDS X-ray Spectrum.** The characteristic peaks of gold (Au) and copper (Cu) are superimposed upon the background of the continuum. Although the peaks have distinct quantized energies, their detection is a statistical process, which broadens the peaks. Image source: http://www.trincoll.edu/~alehman/_images/Engr232_02/Au_latex_042302.gif.

### 4.4.2. Atomic Structure and Characteristic X-rays

The Bohr model of the atom provides a convenient way to understand the nature of characteristic X-rays. One may consider an atom to be a positively charged nucleus surrounded by shells of negatively charged electrons, labeled K, L, M, N... (starting from innermost, most strongly bound shell). Each electron in an atom is uniquely defined by several quantum numbers.

<table>
<thead>
<tr>
<th>Name</th>
<th>Values</th>
<th>Defines</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>positive integers (1, 2, 3...)</td>
<td>electron shell (1=K, 2=L, 3=M...)</td>
<td>Principle binding energy</td>
</tr>
<tr>
<td>l</td>
<td>integers from 0 to (n-1)</td>
<td>electron cloud shape (0=sphere, 1=dumbbell...)</td>
<td>Orbital angular momentum. Chemists follow optical spectroscopy conventions using letters rather than numbers: sharp (l = 0), principal (l = 1), diffuse (l = 2), and fundamental (l = 3)</td>
</tr>
<tr>
<td>m</td>
<td>-l to +l (including 0)</td>
<td>electron orientation in magnetic field</td>
<td>Not significant in absence of an external magnetic field</td>
</tr>
<tr>
<td>s</td>
<td>± ½</td>
<td>electron spin direction</td>
<td>Clockwise or counterclockwise (up or down)</td>
</tr>
<tr>
<td>j</td>
<td>± ½, (j ≠ 0 - ½)</td>
<td>total angular momentum vector</td>
<td>Determines which transitions are permitted between electron shells. For s orbitals (l = 0), j can only be ±½ (vector sum always positive)</td>
</tr>
</tbody>
</table>

The Pauli Exclusion Principle was proposed by Wolfgang Pauli to explain the arrangement of electrons in atoms. It has been generalized to assert that no two fermions of the same type (electrons are fermions) can exist in the same state, i.e., have the same quantum numbers, at the same place and time. Thus, the maximum number of electrons per level is $2n^2$. The correspondence between first four electron shell names and their quantum numbers is given in the following table. Note that the j values increase systematically within a given shell.
Characteristic X-rays are produced by transitions between electron shells. However, not all transitions are possible. Permissible transitions are specified by what are termed quantum selection rules:

1. The change in $n$ must be $\geq 1$ ($\Delta n \neq 0$)
2. The change in $l$ can only be $\pm 1$
3. The change in $j$ can only be $\pm 1$ or 0

Consider the element scandium, Sc, which has an electronic configuration of $1s^22s^22p^63s^23p^63d^24s^2$. We can represent this configuration in $n$-$l$ space:

Now we can restate the first two rules as: (1) no transitions between shells in same row, and (2) no transitions can occur between shells in the same column or that skip columns (e.g., from $l = 3$ to $l = 1$).
The transitions permitted according to these two rules are:

<table>
<thead>
<tr>
<th>n=1</th>
<th>1s</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n=2</td>
<td>2s</td>
<td>2p</td>
<td></td>
</tr>
<tr>
<td>n=3</td>
<td>3s</td>
<td>3p</td>
<td>3d</td>
</tr>
<tr>
<td>n=4</td>
<td>4s</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Finally, we must apply Rule 3 ($\Delta j = 0, \pm 1$). There are four possible cases to consider for the $p \to d$ transitions ($\pm \frac{1}{2}$ spin in the $p$ orbital $\to \pm \frac{1}{2}$ spin in the $d$ orbital).

<table>
<thead>
<tr>
<th>3d</th>
<th>2p</th>
<th>(\Delta j)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(l)</td>
<td>(s)</td>
<td>(j)</td>
</tr>
<tr>
<td>2</td>
<td>+(\frac{1}{2})</td>
<td>2(\frac{1}{2})</td>
</tr>
<tr>
<td>2</td>
<td>+(\frac{1}{2})</td>
<td>2(\frac{1}{2})</td>
</tr>
<tr>
<td>2</td>
<td>-(\frac{1}{2})</td>
<td>1(\frac{1}{2})</td>
</tr>
<tr>
<td>2</td>
<td>-(\frac{1}{2})</td>
<td>1(\frac{1}{2})</td>
</tr>
</tbody>
</table>

Note that the second transition in the list is not possible because \(\Delta j = 2\). There are only two potential transitions to consider when the $s$ orbitals are involved because $j$ cannot be negative: $\pm \frac{1}{2}$ spin (non-$s$ orbital) $\to +\frac{1}{2}$ spin ($s$ orbital). Both are possible.

<table>
<thead>
<tr>
<th>2p</th>
<th>1s</th>
<th>(\Delta j)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(l)</td>
<td>(s)</td>
<td>(j)</td>
</tr>
<tr>
<td>1</td>
<td>+(\frac{1}{2})</td>
<td>1(\frac{1}{2})</td>
</tr>
<tr>
<td>1</td>
<td>-(\frac{1}{2})</td>
<td>1(\frac{1}{2})</td>
</tr>
</tbody>
</table>

All the other transition possibilities can be evaluated in the same manner...

4.4.3. Naming Characteristic X-rays

In 1923, Manne Siegbahn (1886-1978) introduced a notation for identifying X-rays in his book “The Spectroscopy of X-rays.” His nomenclature, now called the Siegbahn notation, was based on the relative intensity of lines from different series but provided no information about specific origin of X-ray lines. Beta (β) and gamma (γ) lines were discriminated based on intensities and some inconsistencies resulted. Siegbahn did such good enough job labeling the lines that, after the fact, generalizations can be made about the relationship between the names of the spectral lines and the processes that produced them:

1. X-ray lines are labeled according to the element that they represent (Si, Ca, Fe, etc.)
2. Lines are further labeled (K, L, M...) based on the electron shell that was ionized and subsequently filled.
3. Greek letters are added based on the "distance" of shell that provided the neutralizing electron (\(\alpha\) indicates that the electron came from the next shell out, \(\beta\) indicates two shells out, \(\gamma\) indicates three shells, etc.)

4. If needed, sublines are identified by numerical subscripts.

Thus, Ca-K\(\alpha\) is a calcium X-ray produced when an electron fills the K shell from the L shell; Ca-K\(\beta\) indicates the electron came from the M shell. Siegbahn got Nobel Prize in Physics in 1924 for this work.

**Siegahn Notation.** The diagram shows both notations used for the electron shells and examples of transitions. the Image source: http://xdb.lbl.gov/Section1/Sec_1-2.html

The table below summarizes all significant characteristic X-rays and gives the line intensities relative to the lines of each series. Note that the intensities of most lines are very low, making them of no use for analytical work. Analysts use \(\alpha\) lines almost exclusively, with \(\beta\) lines used for analyses of the rare-earth elements where there are significant line overlaps on \(\alpha\) lines.
Recognizing that there were inconsistencies in the Siegbahn notation, the International Union of Pure and Applied Chemistry proposed a new system for identifying X-ray lines. The IUPAC system labels X-rays using:

1. A level symbol for final state (K, L, M ...), i.e., the shell being filled;
2. A level symbol for initial state (K, L, M ...), i.e., the source of the neutralizing electron; and
3. Arabic numerals rather than Roman numerals for identifying subshells (L2 and L3 instead of LII and LIII, etc.)

For example, the Ka line, which is produced by the transition from the LIII to the K shell, is labeled K-L3. When a single spectral line results from two transitions, this is indicated by duplicating the symbol for the source shells. For example, the KB2 corresponds to K-N2N3. When X-ray lines are unresolved, subscripts are used to indicate this. For example, the recommended IUPAC notation for the Ka1,2 X-ray lines, which are unresolved in the light elements, are labeled K-L2,3 (Ka2 is K-L2 and Ka1 is K-L3).

<table>
<thead>
<tr>
<th>Source Shell</th>
<th>Shell Filled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
</tr>
<tr>
<td>L_{I}</td>
<td></td>
</tr>
<tr>
<td>L_{II}</td>
<td>K_{αα} (50)</td>
</tr>
<tr>
<td>L_{III}</td>
<td>K_{αγ} (100)</td>
</tr>
<tr>
<td>M_{I}</td>
<td>K_{βα} (1)</td>
</tr>
<tr>
<td>M_{II}</td>
<td>K_{βα} (20)</td>
</tr>
<tr>
<td>M_{III}</td>
<td>K_{βα} (50)</td>
</tr>
<tr>
<td>M_{IV}</td>
<td>K_{βα} (1)</td>
</tr>
<tr>
<td>M_{V}</td>
<td>K_{βα} (100)</td>
</tr>
<tr>
<td>N_{I}</td>
<td>K_{βα} (5)</td>
</tr>
<tr>
<td>N_{II}</td>
<td>K_{βα} (20)</td>
</tr>
<tr>
<td>N_{III}</td>
<td>K_{βα} (50)</td>
</tr>
<tr>
<td>N_{IV}</td>
<td>K_{βα} (100)</td>
</tr>
<tr>
<td>N_{V}</td>
<td>K_{βα} (1)</td>
</tr>
<tr>
<td>N_{VI}</td>
<td>K_{βα} (100)</td>
</tr>
<tr>
<td>N_{VII}</td>
<td></td>
</tr>
<tr>
<td>O_{I}</td>
<td>K_{δα} (0.1)</td>
</tr>
<tr>
<td>O_{II}</td>
<td>K_{δα} (0.1)</td>
</tr>
<tr>
<td>O_{III}</td>
<td>L_{η} (1)</td>
</tr>
<tr>
<td>O_{IV}</td>
<td>L_{η} (1)</td>
</tr>
<tr>
<td>O_{V}</td>
<td>L_{η} (1)</td>
</tr>
</tbody>
</table>
The IUPAC notation is consistent with notation used for Auger electron spectroscopy; however, it is cumbersome, and the X-ray analytical community largely uses the Siegbahn notation.

4.4.4. Moseley’s Law

The relationship between atomic number and characteristic X-ray wavelength was established by Henry G.J. Moseley in 1914:

$$\lambda = \frac{K}{(Z-\sigma)^2},$$

where $K$ and $\sigma$ are constants for given X-ray spectral line. Sigma ($\sigma$) is called the shielding constant and compensates for the effect of electron shells within the ionized shell; it is approximately 1 for K-lines, but 7.4 for the more shielded L-lines. Moseley derived his formula empirically by plotting the square root of X-ray frequencies against a line representing atomic number. However, it was almost immediately noted in 1914 that his formulae could be explained in terms of the 1913 Bohr model of the atom if certain reasonable extra assumptions about atomic structure in other elements were made.
Moseley's Law may be rewritten to relate atomic number and X-ray frequency:

$$\sqrt{\nu} = K(Z - \sigma),$$

where K and σ are constants for given X-ray spectral line (but not the same constants as in the previous equation).

For the Kα and Lα lines, these equations are:

$$\nu = 2.47 \times 10^{15}(Z - 1)^2$$
$$\nu = 4.57 \times 10^{14}(Z - 7.4)^2$$

### 4.4.5. X-ray Intensity

X-ray line intensity is function of the transition probability and rate of ionization. The probability of ionization depends on the target-atom ionization cross-section, Q. Typical capture cross-sections for K shell are about $10^{-20} \text{ cm}^2$ (~1/100 atomic diameter); there is a very low probability of K-shell ionization! The capture cross-section depends on the overvoltage, $U$, etc.
which is the ratio accelerating voltage, $E_0$, to the critical ionization potential, $E_c$. For the K-shell the capture cross-section is

$$Q = 6.51 \times 10^{-20} \left( \frac{n_s b_s}{U E_c^2} \right) \ln C_s U$$

where $n_s$ = number of electrons; $b_s$ and $C_s$ = constants; $E_0$ = accelerating voltage; $E_c$ = critical ionization potential for shell of interest (kV); and $U$ = overvoltage ($E_0/E_c$).

The intensity of a given spectral line, $I$, may be described by the following empirical relationship:

$$I = C i_b (E_0 - E_c)^p ,$$

where $C$ = constant; $i_b$ = beam current; $E_c$ = critical excitation potential of the line of interest; $E_0$ = accelerating voltage; and $p = 1.7$ for $E_0 < 1.7E_c$ (smaller for higher values of $E_c$).

The only ways an analyst can increase a specific line's intensity is to optimize the over voltage (as above) or increase the beam current, $i_b$.

4.4.6. Satellite Peaks

At high overvoltages more than one electron may be ejected simultaneously from an atom. This causes the overall structure of the electron shells to change, resulting in the production of X-rays with slightly lower energies than those produced during single electron ionization. These multi-ionization X-rays appear as small satellite peaks near the characteristic X-ray peaks. Another important source of satellite peaks is the Auger process described below.

Satellite peaks are generally not important, except when there is a very large overvoltage. An accelerating voltage, $E_0$, of 15 kV is about optimum for exciting Fe-Kα X-rays ($U = 15 / 7.112 = 2.1$), but results in an eight-fold overvoltage for Si-Kα ($U = 15 / 1.838 = 8.2$). Production of satellite peaks may decrease peak-to-background ratios.
4.4.7. Peak Shifts

There are small, but detectable, changes in the wavelength (energy) of an X-ray peak produced from a pure element compared with that produced from a compound of the element. This occurs because the configuration of the inner shells of an atom is influenced by the outer valence electrons. The orbitals of the valence electrons reflect (1) the manner in which an element is bonded (which is related to its oxidation state) and (2) the coordination of an atom inside a crystal. Peak shifts most significant when the X-ray-producing inner shells are less shielded from the valence electrons. This is the case for the outermost transitions (M- and N-lines) in most elements and in low-Z elements (B through F for Kα, Al through Cl for Kβ).

The shifts are most apparent comparing metals with oxides and halides (figure below). Consequently, it is not advisable to use a metallic standard material for analysis of oxide compounds. In the case of energy-dispersive analysis, the peak shifts are undetectable, and metals can be used as reference standards for quantification.

As a detailed example, consider Al as a metal and as an oxide, Al₂O₃ (figure below). In Al metal, X-rays are produced by transitions from the 3s and 3p levels into the 1s level; the 3s and 3p electrons are not involved in chemical bonding. The result is a relatively narrow X-ray peak. In contrast, in Al₂O₃, the valence electrons are combined with oxygen 2s and 2p electrons and span wider range of energies. As a result, the Kβ emission becomes broader, asymmetric, and separates into Kβ and a satellite Kβ' peak.
People have attempted to use observed wavelength shifts to determine oxidation state. For example, iron in minerals exhibits two common valences, 2+ and 3+. In theory, there should be a detectable energy change between Fe-Lα and Fe-Lβ, however, this is very difficult to observe. The only apparently workable method uses the shape of the Fe-Lβ peak to evaluate oxidation state. A readily observable shift occurs for the S-Kβ peak when comparing sulfides (S^2-) with sulfates (S^{6+}).

The coordination of an atom or ion in a mineral will also influence the electron cloud configuration. In essence, while equal numbers of protons pull on the electron clouds, variable numbers of coordinating anions pull out. Higher coordination (more anions pulling out) results in the electron shells being pulled farther from nucleus, yielding higher energy transitions. As with bonding, coordination effects are most easily observed in light elements. For example, the shift in Al-Kα lines relative to line from Al metal for kaolinite (Al^{VI}) is about twice that observed for feldspar (Al^{IV}).

### 4.5. Cathodoluminescence

#### 4.5.1. Introduction

Luminescence is the emission of visible light from a solid when it is excited by some form of energy. Luminescence can be subdivided into two subtypes: **fluorescence**, which ceases immediately after withdrawal of the exciting source (persisting <10^8 seconds) and **phosphorescence**, which persists for some time after removal of excitation (persisting >10^8 seconds). However, the distinction between these is somewhat arbitrary and confusing. Alternatively, perhaps more usefully, luminescence may be subdivided based upon the type of exciting energy. Luminescence produced by incident X-rays or γ-rays is radioluminescence, that...
produced by visible and UV light is photoluminescence, that produced by heat is thermoluminescence (TL) and that generated by incident electrons is **cathodoluminescence** (CL).

Cathodoluminescence was first studied by William Crookes in 1879. He discovered that scheelite, CaWO₄, placed opposite the cathode in a vacuum discharge tube yielded spectacular fluorescence. He concluded that some particle was traveling from cathode to the target causing the fluorescence well before J. J. Thomson discovered the electron in 1897. Crookes also studied luminescence from many other minerals (diamond, ruby, sapphire, zircon, etc.).

![Cathodoluminescence in Carbonate Cement. Note the multiple generations of cementation. Image source: http://www.univ-lille1.fr/geosciences/umr_pbds/techniques/cathodo.html#exemples.](image)

### 4.5.2. Band Theory

Cathodoluminescence is explained using the band theory of solids. Additionally, band theory will be used in the discussion of semiconductor X-ray detectors later in the course; thus, it will be covered in some detail here. Solids consist of an essentially infinite number of atoms, and it is incorrect to consider each atom individually when considering the behavior of electrons in them. Band theory considers the structure of the solid as a whole and provides a way to describe metals and other solids to explain their unique chemical and physical properties.

In basic molecular orbital theory, one assumes that when atoms are brought together, they form bonding, non-bonding and antibonding orbitals of different energies. For \( n \) atomic orbitals in a molecule, \( n \) molecular orbitals are produced. Thus, a molecule with three atoms (assuming 1 atomic orbital for each) forms 3 molecular orbitals. As the number of molecular orbitals increases, the energy difference between the lowest bonding and the highest antibonding orbitals increases, while the space between each individual orbital decreases. With an essentially infinite number of atoms, the spacing between the lowest bonding and highest antibonding orbital reaches a maximum and there are so many molecular orbitals so close together that they blur into one another forming a **band**.
Origin of a Band. Schematic diagram of the formation of a band by ever increasing numbers of molecular orbitals. Diagram redrafted from a figure at: http://www.chembio.uoguelph.ca/educmat/chm729/band/concept.htm

Electrons fill the molecular orbitals within a band starting with the lowest energy orbital and follow the Pauli Exclusion Principle. No more than two electrons may occupy a single orbital and if two do occupy a single orbital, then their spins must be spin paired. Additionally, Hunds rule must not be broken and when more than one orbital has the same energy, electrons must occupy separate orbitals and have parallel spins. If each molecular orbital contains only one electron, the band is only half full \((n \text{ electrons for } n \text{ orbitals})\), leaving empty levels in the band. Conversely, if each atomic orbital contains 2 electrons, the band is full \((2n \text{ electrons for } n \text{ molecular orbitals})\). Whenever a band is half full, the energy of the uppermost filled molecular orbital at 0 K is called the Fermi energy. At \(T >0\) K, the energy rises above the Fermi energy level because the electrons start occupying higher states due to thermal excitation.

4.5.3. Conductors, Semi-conductors & Insulators

The nature of a material can be characterized by the relationship between the valance band (low energy electrons) and the conduction band (higher energy electrons). Most solids are insulators, which means that there is a large band gap between energies of the valence and conduction bands. All electrons are confined to the valence band, and there is no electrical conduction.

Some materials are semiconductors, in which the valence band is separated from the conduction band by a small energy gap, \(E_g\). There are no electrons in the conduction band at 0 K, but with the addition of energy, electrons can be promoted into it and an electrical current can flow. Semiconductors may be either intrinsic or extrinsic. In the latter case a material contains small amounts of impurities that have roughly the same atomic size, but more or fewer valence electrons than the host material. Artificial addition of impurities is termed doping. Impurities can transform a non-cathodoluminescent material into one that is luminescent with the nature of the impurity controlling the CL color.

The valence and conduction bands overlap in conductors, such as metals. There is no real difference between the valence and conduction bands and electrons can move through the material easily, yielding electrical conduction. This free-floating electron distribution is also described as metallic bonding.

The electrical properties of all types of materials depend strongly on temperature. The conductivity of an insulator or semiconductor increases with increasing temperature, whereas conductors become less conductive.

4.5.4. Types of Semiconductors

There are two types of semiconductors. A **p-type** semiconductor is formed when the dopant has fewer valence electrons than the host (e.g., Al$^{3+}$ into Si$^{4+}$). The impurity atoms occupy sites in the structure but contribute fewer electrons to the valence band than the host atoms. This produces "holes" in the lower-energy valence band, which allow electrons to move from one orbital to another within the valence band with a small input of energy (smaller than required for the semiconductor without the doping). Alternately, in an **n-type** semiconductor, the impurity has more valence electrons than the host (e.g., P$^{5+}$ into Si$^{4+}$) and contributes extra electrons to the valence band. However, because the valence band is already filled, the extra electrons must go into the higher-energy (conduction) band. These electrons partially fill the conduction band, can move easily between the orbitals, and thus throughout the solid.

4.5.5. Production of Cathodoluminescence

Cathodoluminescence is produced in materials that have semiconductor properties. Beam electrons provide the energy to knock photoelectrons from the valence band into conduction band. This leaves behind a positively charged valence band “hole”. The average energy required to produce an electron-hole (e/h) pair (eV) is:

$$E_{e/h\,\text{pair}} = 2.8E_{\text{gap}} + M,$$

where $M$ is a constant ($0 < M < 1$) and $E_g$ is the size of the band gap. We may thus calculate the number of electron-hole pairs per beam electron as approximately:

$$G = \frac{E_0(1-\eta)}{E_{e/h\,\text{pair}}}.$$
where \( \eta = \) backscattered electron coefficient. Backscattered coefficients in silicates are generally 0.05-0.20, so each incident electron potentially can produce a large number of electron-hole pairs. For example, a single electron with \( E_g = 30 \text{ kV} \) could produces 1300 electron-hole pairs in zircon (\( \eta = 0.27, E_g = 5.4 \text{ eV} \)). Free electrons recombine with the valence band holes, radiating light with ultraviolet to infrared wavelengths (~160-2000 nm). The resulting CL can be polarized by the crystal structure.

It should be emphasized that the processes that yield CL are, except in rare cases, different than those that produce mineral colors. Mineral coloration reflects the absorption of incident white light. This generally occurs by electron transfers in \( d \) suborbitals produced by crystal field splitting or by intervalence charge transfer. The band properties of a mineral can produce color by absorption, but this is a relatively uncommon because most minerals are insulators. Generally, the gaps are too large for visible light to promote an electron to the conduction band and no visible light is absorbed, yielding a white or colorless mineral (diamond, C). In minerals that act as semiconductors, light photons with \( E > E_g \) can promote electrons to the conduction band. In the cases where \( E_g \) is less than energy of visible light, a mineral will appear black or gray (e.g., galena, PbS). Larger gaps may only absorb short wavelengths of visible light, leaving a red mineral (e.g., cinnabar, HgS). Minerals with overlapping conduction and valence bands absorb the entire range of colors because electrons can be promoted to conduction band by any wavelength; the result is an opaque mineral.

### 4.6. Heat

#### 4.6.1. Introduction

Significant amounts of heat are produced within a sample because electron excitation of X-rays is not very efficient. Most low energy continuum photons and low-energy inelastically scattered electrons do not escape the sample and their energy is transformed into higher vibrational energies of the bonds (heat).

A crystal lattice consists of bonded atoms, which cannot vibrate independently. Instead, the entire lattice vibrates in a coordinated fashion, with vibrations taking the form of collective modes, which propagate through the material. These harmonic vibrations can be decomposed into elementary vibrations called phonons. The total number of phonons in a vibrating crystal is
related to its temperature of the system. At higher temperatures, vibration of an object is stronger and the number of phonons larger. As every phonon carries a quantum of vibrational energy, this means that the internal energy of the object is also larger.

![Rayleigh phonon on the (100) face of a fcc solid (amplitude exaggerated 10x).](http://www.fhi-berlin.mpg.de/th/personal/hermann/pictures.html)

**4.6.2. Quantification of Heating Effects**

Castaing (1951) demonstrated that the maximum temperature rise (°C) for a material can be expressed as

\[
\Delta T = \frac{4.8E_0b_i}{C_t d_0},
\]

where \( E_0 \) = accelerating voltage (kV), \( b_i \) = beam current (µA), \( C_t \) = thermal conductivity (W/cm·K), and \( d_0 \) = beam diameter (µm).

Consider an accelerating voltage of 15 kV, a specimen current of 0.05 µA (= 50 nA), and a beam diameter of 1 µm. A material like copper with a high thermal conductivity (\( C_t = 3.97 \text{ W/cm·K} \)) has \( \Delta T = 0.9 \) °C, whereas zircon (\( C_t = 0.042 \)) has \( \Delta T = 86 ^\circ\text{C} \). Note that going to a beam diameter of 10 µm would drop this latter number by a factor of 10 to 8.6 °C. Actually, epoxy will decompose ("burn") long before attaining \( \Delta T \) in excess of 150 °C. In contrast, most minerals can survive temperatures well in excess of 200 °C. Values of \( \Delta T \) at 15 kV for some materials are given below.

<table>
<thead>
<tr>
<th>Material</th>
<th>( C_t ) (W/cm·K)</th>
<th>1 µm diameter spot</th>
<th>5 µm diameter spot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>10 nA 25 nA 50 nA</td>
<td>10 nA 25 nA 50 nA</td>
</tr>
<tr>
<td>Epoxy</td>
<td>0.002</td>
<td>360 900 1800</td>
<td>72 180 360</td>
</tr>
<tr>
<td>Mica</td>
<td>0.007</td>
<td>103 257 514</td>
<td>21 51 103</td>
</tr>
<tr>
<td>Calcite</td>
<td>~0.04</td>
<td>18 45 90</td>
<td>4 9 18</td>
</tr>
<tr>
<td>Quartz</td>
<td>~0.09</td>
<td>8 20 40</td>
<td>2 4 8</td>
</tr>
<tr>
<td>Diamond</td>
<td>~5</td>
<td>0.1 0.4 0.7</td>
<td>0 0.1 0.1</td>
</tr>
</tbody>
</table>

Epoxy decomposes ("burns") long before attaining \( \Delta T > 150 \) °C. However, most minerals easily survive temperatures >200 °C. Thermal conductivity is a strongly vectorial property, meaning
that it varies with crystallographic orientation. For example, quartz ranges from 0.065 along the $a$ axis to 0.12 along $c$ axis. The table below summarizes the sparse data available for minerals.

<table>
<thead>
<tr>
<th>Material</th>
<th>$C_v$ (W/cm·K)</th>
<th>Material</th>
<th>$C_v$ (W/cm·K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalusite</td>
<td>0.11</td>
<td>Kyanite</td>
<td>0.17</td>
</tr>
<tr>
<td>Barite</td>
<td>0.017</td>
<td>Magnetite</td>
<td>0.05</td>
</tr>
<tr>
<td>Calcite</td>
<td>0.038 ($</td>
<td>a</td>
<td>$), 0.044 ($</td>
</tr>
<tr>
<td>Cordierite</td>
<td>0.022</td>
<td>Obsidian</td>
<td>0.014</td>
</tr>
<tr>
<td>Diamond</td>
<td>5.4, 20 (synthetic)</td>
<td>Orthoclase</td>
<td>0.042</td>
</tr>
<tr>
<td>Dolomite</td>
<td>0.029</td>
<td>Periclase</td>
<td>0.69</td>
</tr>
<tr>
<td>Epoxy</td>
<td>0.002</td>
<td>Quartz</td>
<td>0.065 ($</td>
</tr>
<tr>
<td>Fluorite</td>
<td>0.01</td>
<td>Rutile</td>
<td>0.088 ($</td>
</tr>
<tr>
<td>Forsterite</td>
<td>0.059</td>
<td>Sapphire</td>
<td>0.34</td>
</tr>
<tr>
<td>Graphite</td>
<td>1.7</td>
<td>Sylvite</td>
<td>0.07</td>
</tr>
<tr>
<td>Gypsum</td>
<td>0.013</td>
<td>Spinel</td>
<td>0.14</td>
</tr>
<tr>
<td>Hematite</td>
<td>0.13</td>
<td>Zircon</td>
<td>0.042</td>
</tr>
</tbody>
</table>

**4.7. Secondary Electrons**

**4.7.1. Introduction**

The strongest region in the electron energy spectrum is due to secondary electrons (SE), which are defined as those emitted with energies less than 50 eV.

![Electron Energy Spectrum](image)

*Electron Energy Spectrum.* Secondary electrons (SE) form a large low-energy peak. Auger electrons (AE) produce relatively small peaks on the backscattered electron (BSE) distribution. Figure after Goldstein et al. 1981.

**4.7.2. Production of Secondary Electrons**

Secondary electrons are produced when an incident electron excites an electron in the sample and loses some of its energy in the process. The excited electron moves towards the surface of
the sample undergoing elastic and inelastic collisions until it reaches the surface. Here it can escape if its energy exceeds the surface work function, $E_w$, which defines the amount of energy needed to remove electrons from the surface of a material. One of the major reasons for coating a non-conductive specimen with a conductive material is to increase the number of secondary electrons that will be emitted from the sample (decrease $E_w$).

The mean free path length of secondary electrons in many materials is ~1 nm (10 Å). Thus, although electrons are generated throughout the region excited by the incident beam, only those electrons that originate less than 1 nm deep in the sample escape to be detected as secondary electrons. This volume of production is very small compared with BSE and X-rays. Therefore, the resolution using SE is better than either of these, and is effectively the same as the electron beam size. The shallow depth of production of detected secondary electrons makes them very sensitive to topography and they are used for scanning electron microscopy (SEM).

**4.7.3. Secondary Electron Yield**

The average number of SE produced per primary electron is called the *secondary-electron yield*, $\delta$, and is typically in the range 0.1 to 10 (varying between different materials). For a given sample material, $\delta$ decreases with increase in incident energy $E_0$ since the probability of inelastic scattering of a primary electron within the escape depth decreases. SE yield depends on the angle of tilt of the specimen relative to the primary-electron beam, $\phi$. The value is lowest for perpendicular incidence ($\phi = 0$) and increases with increasing angle between the primary beam and the surface normal. This effect may make cracks appear bright in SE images.


The edge effect is modified by the position of the detector relative to the sample. SE detectors are positioned to one side of the sample enhancing topographic effects. Faces oriented towards the detector will be brighter, whereas those in the opposite direction will be dark.
4.7.4. Secondary Electron Components

From the above discussion one might conclude that SE production will be independent of the electron beam accelerating voltage. However, only part of the SE signal, the SE1 component, comes from the sample surface. Other components arise from SE produced by backscattered electrons as they exit the specimen (SE2) and when BSE strike the walls of the specimen chamber (SE3).

The SE2 component depends on a sample's backscattered coefficient, η, which reflects chemical differences well below the surface. This effect increases with increasing penetration depth, so the sample will appear more "transparent" at higher $E_0$, and less so at low $E_0$. Finer surface structure images can generally be obtained with lower accelerating voltages. At high $E_0$, the SE2 and SE3 signals are larger, reducing image contrast and veiling fine surface structures.

4.8. Backscattered Electrons

4.8.1. Introduction

Backscattered electrons (BSE) are high energy primary electrons that suffer large angle (> 90°) scattering and re-emerge from the entry surface of a specimen. Most BSE have energies slightly lower than that of the primary electron beam, $E_0$, but may have energies as low as ~50 eV (the upper cut-off for secondary electrons). The fraction of beam electrons backscattered from a sample, $n_b$ (also symbolized $\eta$), depends strongly on the sample’s average atomic number, $Z$, reflecting the increasing charge of the atomic nuclei. Individual scattering events are generally elastic, where a negligible amount of energy is lost by the primary electron in the process. The direction of the electron may be altered, but its energy remains essentially the same. However, an electron that has undergone inelastic scattering (having excited a plasmon, phonon, caused inner shell ionization, or interacted with an electron in the valence band) may subsequently escape the sample surface as a BSE.

4.8.2. Backscattered Coefficient

A good approximation for the backscattered coefficient (Love & Scott, 1978) is:

$$n_b = n_{b_0} \left[ 1 + a \ln \left( \frac{E_0}{20} \right) \right],$$

where

$$a = -0.11128 + 3.0289 \times 10^{-3} \tilde{Z} - 1.5498 \times 10^{-5} \tilde{Z}^2,$$

$$n_{b_0} = -5.23791 \times 10^{-3} + 1.5048371 \times 10^{-2} \tilde{Z} - 1.67373 \times 10^{-4} \tilde{Z}^2 + 7.16 \times 10^{-7} \tilde{Z}^3,$$

and

$$\tilde{Z} = w_1 Z_1 + w_2 Z_2 \ldots + w_n Z_n$$
where \( w_i \) = weight factor of element \( i \) and \( Z_i \) = atomic number of element, \( i \). The energy of a backscattered electron depends upon the number of interactions that it has undergone before escaping the sample surface.

**Backscattered Electron Energies.** Distribution of BSE energies with an iron target. Note that the energy peak of the BSE for this material occurs \( \sim 12 \) keV. The model was run using the Casino software described at [http://www.gel.usherbrooke.ca/casino/What.html](http://www.gel.usherbrooke.ca/casino/What.html).

**Backscattered Electron Energies.** Distribution of BSE energies with an aluminum target. Note that the energy peak of the BSE for this material occurs \( \sim 10 \) keV, reflecting generally greater depths of electron penetration than for iron (above). The models were run using the Casino software described at [http://www.gel.usherbrooke.ca/casino/What.html](http://www.gel.usherbrooke.ca/casino/What.html).
Chapter 5 – Absorption of Electromagnetic Radiation

5.1. Introduction

All forms of electromagnetic radiation (X-rays, light, etc.) are attenuated by absorption from elements when it passes through a material. This results from two processes:

- Compton scattering ($\sigma$), in which photons are diverted in directions different from that of the primary beam, and
- Photoelectric absorption ($\tau$), which produces fluorescence or Auger electrons from a material

These two effects can be combined into a single mass absorption coefficient ($\mu$):

$$\mu = \tau + \sigma$$

Throughout this discussion we will use the notation, $\mu$, for the mass attenuation coefficient. However, more accurately these coefficients are $\mu/\rho$, and are tabulated as such. The mass absorption coefficient depends upon both wavelength (energy) of the photon of interest and the composition of the material. In geological materials photoelectric absorption comprises ~95% of the total absorption and Compton scattering can be largely ignored (it is not significant at $\lambda < 1$ Å).

### Selected Mass Absorption Coefficients for Kα Radiation

<table>
<thead>
<tr>
<th>Emitter</th>
<th>Na</th>
<th>Mg</th>
<th>Al</th>
<th>Si</th>
<th>P</th>
<th>S</th>
<th>Cl</th>
<th>K</th>
<th>Ca</th>
<th>Ti</th>
<th>Cr</th>
<th>Mn</th>
<th>Fe</th>
<th>Ni</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>3653</td>
<td>2169</td>
<td>1345</td>
<td>805</td>
<td>574</td>
<td>391</td>
<td>273</td>
<td>141</td>
<td>105</td>
<td>60</td>
<td>36</td>
<td>28</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>Na</td>
<td>562</td>
<td>5409</td>
<td>3365</td>
<td>2170</td>
<td>1444</td>
<td>987</td>
<td>691</td>
<td>359</td>
<td>286</td>
<td>152</td>
<td>92</td>
<td>72</td>
<td>57</td>
<td>37</td>
</tr>
<tr>
<td>Mg</td>
<td>768</td>
<td>463</td>
<td>4287</td>
<td>2768</td>
<td>1843</td>
<td>1261</td>
<td>883</td>
<td>460</td>
<td>341</td>
<td>195</td>
<td>116</td>
<td>93</td>
<td>74</td>
<td>46</td>
</tr>
<tr>
<td>Al</td>
<td>1019</td>
<td>614</td>
<td>366</td>
<td>3451</td>
<td>2300</td>
<td>1575</td>
<td>1104</td>
<td>576</td>
<td>427</td>
<td>245</td>
<td>148</td>
<td>116</td>
<td>93</td>
<td>60</td>
</tr>
<tr>
<td>Si</td>
<td>1319</td>
<td>794</td>
<td>499</td>
<td>525</td>
<td>2816</td>
<td>1929</td>
<td>1353</td>
<td>707</td>
<td>524</td>
<td>301</td>
<td>182</td>
<td>143</td>
<td>114</td>
<td>74</td>
</tr>
<tr>
<td>P</td>
<td>1670</td>
<td>1006</td>
<td>632</td>
<td>411</td>
<td>276</td>
<td>2325</td>
<td>1832</td>
<td>854</td>
<td>633</td>
<td>364</td>
<td>220</td>
<td>174</td>
<td>138</td>
<td>90</td>
</tr>
<tr>
<td>S</td>
<td>2076</td>
<td>1251</td>
<td>786</td>
<td>511</td>
<td>343</td>
<td>238</td>
<td>1940</td>
<td>1016</td>
<td>754</td>
<td>434</td>
<td>262</td>
<td>207</td>
<td>165</td>
<td>108</td>
</tr>
<tr>
<td>Cl</td>
<td>2542</td>
<td>1551</td>
<td>962</td>
<td>826</td>
<td>420</td>
<td>289</td>
<td>204</td>
<td>1194</td>
<td>587</td>
<td>511</td>
<td>306</td>
<td>244</td>
<td>195</td>
<td>127</td>
</tr>
<tr>
<td>K</td>
<td>3583</td>
<td>2207</td>
<td>1386</td>
<td>902</td>
<td>605</td>
<td>417</td>
<td>294</td>
<td>155</td>
<td>1188</td>
<td>686</td>
<td>415</td>
<td>329</td>
<td>262</td>
<td>171</td>
</tr>
<tr>
<td>Ca</td>
<td>4326</td>
<td>2606</td>
<td>1637</td>
<td>1066</td>
<td>715</td>
<td>493</td>
<td>347</td>
<td>183</td>
<td>136</td>
<td>783</td>
<td>475</td>
<td>376</td>
<td>300</td>
<td>196</td>
</tr>
<tr>
<td>Ti</td>
<td>5871</td>
<td>3571</td>
<td>2222</td>
<td>1446</td>
<td>970</td>
<td>668</td>
<td>471</td>
<td>249</td>
<td>185</td>
<td>107</td>
<td>607</td>
<td>480</td>
<td>384</td>
<td>251</td>
</tr>
<tr>
<td>Cr</td>
<td>7728</td>
<td>4655</td>
<td>2924</td>
<td>1903</td>
<td>1277</td>
<td>889</td>
<td>620</td>
<td>327</td>
<td>244</td>
<td>141</td>
<td>86</td>
<td>68</td>
<td>478</td>
<td>313</td>
</tr>
<tr>
<td>Mn</td>
<td>8780</td>
<td>5289</td>
<td>3323</td>
<td>2163</td>
<td>1451</td>
<td>1000</td>
<td>705</td>
<td>372</td>
<td>277</td>
<td>160</td>
<td>97</td>
<td>77</td>
<td>62</td>
<td>347</td>
</tr>
<tr>
<td>Fe</td>
<td>9918</td>
<td>5975</td>
<td>3755</td>
<td>2443</td>
<td>1639</td>
<td>1129</td>
<td>796</td>
<td>420</td>
<td>313</td>
<td>181</td>
<td>110</td>
<td>87</td>
<td>70</td>
<td>382</td>
</tr>
<tr>
<td>Ni</td>
<td>12464</td>
<td>7509</td>
<td>4717</td>
<td>3070</td>
<td>2060</td>
<td>1419</td>
<td>1001</td>
<td>528</td>
<td>393</td>
<td>228</td>
<td>138</td>
<td>110</td>
<td>88</td>
<td>57</td>
</tr>
</tbody>
</table>

5.2. Lambert's Law

Lambert's law describes the amount of attenuation (absorption) that radiation undergoes passing through a thickness of material:

$$I = I_0 e^{-\mu pt}$$
where \( I_0 \) = initial intensity, \( I \) = final intensity, \( \rho \) = density (g/cm\(^3\)), \( t \) = thickness (cm), and \( \mu \) = mass absorption coefficient (cm\(^2\)/g). For our purposes, intensities will be measured in X-ray counts per second (cps).

In a multi-element material (with n elements present) one multiplies the mass absorption coefficients at the wavelength of interest, \( \mu_i \), by the mass fraction each element, \( w_i \), and sums the result to yield a bulk mass absorption coefficient:

\[
\mu_{bulk} = w_1 \mu_1 + w_2 \mu_2 + \cdots + w_n \mu_n
\]

### 5.3. Half-thickness

Absorption in a material may be characterized by a **half-thickness**, \( t_{1/2} \), i.e., the thickness required to halve the intensity of the radiation. The expression for half-thickness is derived from Lambert's Law, using \( I = \frac{1}{2} I_0 \) we can write:

\[
\frac{1}{2} = e^{-\mu \rho t_{1/2}}.
\]

Taking the natural log of both sides, we get:

\[
\ln \frac{1}{2} = -\mu \rho t_{1/2}.
\]

Rearranging yields:

\[
t_{1/2} = \frac{\ln 0.5}{-\mu \rho}.
\]

The natural log of 0.5 = -0.6931, so the final equation is:

\[
t_{1/2} = \frac{0.6931}{\mu \rho}.
\]

The concept of a half-thickness can be applied to the penetration and any energy deposition by photons (x-ray, \( \gamma \)-ray, bremsstrahlung, etc.) in biological materials, solids, and complex materials as long as mass absorption coefficients are available. The NIST site, [http://www.physics.nist.gov/PhysRefData/XrayMassCoef/cover.html](http://www.physics.nist.gov/PhysRefData/XrayMassCoef/cover.html), provides a compilation of absorption coefficients from 1 kV to 100 MeV and a detailed discussion of mass attenuation coefficients.
As an example, consider the absorption of Ni-Kα in forsterite, Mg$_2$SiO$_4$ ($\rho = 3.25$ g/cm$^3$). The weight fractions of elements in forsterite are: Mg = 0.3455, Si = 0.1996, and O = 0.4549. The values of $\mu$ for Ni-Kα radiation for these elements are 48, 74 and 14, respectively, so $\mu_{\text{Ni-K}\alpha} = 0.3455 \times 48 + 0.1996 \times 74 + 0.4549 \times 14 = 37.7$ cm$^2$/g. The half-thickness is 0.0057 cm or 56 µm. Given that X-ray production in a mineral is only ~3 µm deep, absorption of Ni-Kα X-rays will not be a major problem in forsterite. Using Lambert’s Law, we see that the observed intensity will be 96.4% of the original assuming a 3 µm path length.

### 5.4. Absorption Edges

Bulk mass absorption coefficients decrease with increasing X-ray energy; more energetic radiation is absorbed less than less energetic radiation. However, there is strong absorption if the radiation is energetic enough to cause inner shell ionization of another element (i.e., $E_{\text{photon}} > E_c$). This marked increase in absorption produces “teeth” on a curve of absorption vs. photon energy that are called absorption edges.

![Absorption coefficient versus energy for Sm](image)

*Absorption coefficient versus energy for Sm.* Although the absorption in Sm generally decreases with increasing X-ray energy, there are several absorption edges that correspond to values of $E_c$ (identities and energies indicated) for different Sm X-ray lines.

### 5.5. Secondary Fluorescence

Absorption edges can produce dramatic and unexpected effects. Consider the absorption of Ni-Kα and Co-Kα in pyrite, which consists of 0.4655 Fe and 0.5345 S (element weight fractions). The bulk mass absorption coefficients are:

\[
\mu_{\text{Ni-K}\alpha} = 0.5345 \times 108 + 0.4655 \times 382 = 236 \\
\mu_{\text{Co-K}\alpha} = 0.5345 \times 133 + 0.4655 \times 56 = 97
\]
Ni-Kα is more energetic than Co-Kα (7.48 kV vs. 6.91 kV) but is absorbed over twice as much. Why is this? The Ni-Kα radiation is energetic enough to excite Fe-Kα radiation ($E_c = 7.112$ kV) from the pyrite, a process called secondary fluorescence (figure below). Note that the result is enhanced Fe-Kα intensity. Some elements undergo self-absorption, in which the K-line X-rays of an element may be absorbed to produce L-line X-rays in the same element.

Absorption coefficient versus energy for Fe, Co and Ni. Values of $E_c$ are indicated for Kα X-rays of each of the elements.
Chapter 6 – Instrumentation

6.1. Introduction

All electron beam instruments are built around an electron column, which produces a stable electron beam, controls beam current, beam size and beam shape, and raster the beam. Electron optics are a very close analog to light optics, and most of the principles of an electron beam column can be understood by thinking of the electrons as rays of light and the electron optical components as simply their optical counterparts.

There have been great technical advances in microprobes and scanning electron microscopes during the last 30 years (mostly in automation and electronics), but all instruments are built around an electron column, which produces a stable electron beam, controls beam current, beam size and beam shape, and raster the beam for SEM work. Scanning electron microscopes and electron microprobes have very similar electron columns. Electron optics are a very close analog to light optics, and most of the principles of an electron beam column can be understood by thinking of the electrons as rays of light and the electron optical components as simply their optical counterparts.

[Cameca Camebax MBX Microprobe overview image]

Cameca Camebax MBX Microprobe overview.
6.2. Vacuum System

6.2.1. Introduction

A pumping system must be employed to remove air from the electron column. In most microprobes and scanning electron microscopes, vacuum is achieved through a combination of mechanical and diffusion pumps. Various valves allow sequential pumping of the electron column and sample chamber. The two main units used to measure pressure (vacuum) are the torr and Pascal. Standard atmospheric pressure (STD) is 760 torr or 1.01 x 10^5 Pascal (Pa). (1 torr = 133.32 Pa; 1 Pa = 0.0075 torr). The electron column and wavelength dispersive spectrometers must be operated under high vacuum for five reasons:

1. To produce a mean-free path for electrons greater than the length of the electron column. This corresponds to a vacuum of better than about 10^-4 torr (<0.1 Pa);
2. To avoid arcing between the cathode (filament) and the anode plate. There is very high voltage between these two components and stray air or gas molecules can cause electrical arcing between them. The dielectric strength of air depends strongly on pressure. To maintain a voltage of 20 kV between the Wehnelt and anode plate at a pressure of 10^-4 torr, requires a gap of about 2 mm; higher voltages require better vacuums.
3. To avoid collisions between electrons of the beam and stray molecules. These collisions can result in spreading or diffusing of the electron beam or, more seriously, can result in volatilization event if the molecule is organic in nature (for example, vacuum oil). Volatilizations can severely contaminate the microscope column, especially apertures, and degrade the image;

4. To avoid damaging the filament. The volume around the electron gun must be kept free of gas molecules especially oxygen, which will greatly shorten the filament life.

5. To prevent absorption of X-rays produced from the sample by air molecules. At high vacuum, even soft X-rays (such as B-Kα) are transmitted without loss.

The gas in a vacuum system can be in a viscous state, in a molecular state or in a state that is intermediate between these two. Evacuation of the sample chamber from atmospheric pressure to operating vacuums proceeds from the viscous flow regime into the molecular flow regime. The type of flow regime reflects the mean free path of the air molecules. The mean free path, \( l_{MFP} \), for air at 25 °C can be approximated by

\[
l_{MFP} = \frac{4.5 \times 10^{-3}}{P},
\]

where \( P \) = pressure (torr) and \( l_{MFP} \) = the mean free path length (cm).
Thus, $l_{\text{MFP}}$ of gas molecules is very small at atmospheric pressure ($6 \times 10^{-6}$ cm) and the flow of the gas is limited by its viscosity. Under these conditions, gas molecules collide more readily with one another than with the chamber walls and will move as a viscous mass in the general direction of low pressure. However, when the pressure drops to the point where the MFP exceeds the size of the chamber, the gas molecules do not collide with one another and no longer exhibit viscous flow. At this pressure, gas movement is independent of any pressure gradient and molecules may "backstream" from the pump into the sample chamber, moving against the pressure gradient. Backstreaming of pump oil vapor can be an important issue with diffusion pumps. Usually, water-cooled plates or baffles are positioned above the diffusion pump to promote condensation of pump oil before it enters the electron column.

**Flow regimes.** Note that the type of flow depends strongly on the diameter of the volume. Redrafted from image at: http://www.rdmag.com/images/0508/vactech_lrg.jpg.

### 6.2.2. Vacuum Systems

Electron microprobes and scanning electron microscopes have similar pumping systems. These usually include mechanical (rotary) pumps and oil diffusion pumps. When an ultra-clean vacuum environment is required, a turbomolecular pump often replaces the oil diffusion pump. A mechanical pump is used to evacuate the sample chamber from atmospheric pressure to a moderate vacuum. Once the chamber pressure low enough ($\sim 10^{-3}$ torr), a gate valve is opened, and high vacuum achieved using an oil diffusion pump. Cameca’s SX-50 and SX-100 designs use an additional ion pump to help maintain vacuum around the electron gun.

Cameca microprobes have high- and low-vacuum sections. The wavelength-dispersive spectrometers are isolated from the electron column by thin windows made of polypropylene or Mylar. The spectrometers are evacuated by a separate mechanical pump and kept at a pressure of about $10^{-3}$ torr = 1 Pa, whereas the column is kept at a pressure of about $10^{-6}$ torr. Polypropylene windows are used on the light-element spectrometer(s) because they absorb
fewer X-rays; however, they are more fragile than Mylar windows. This arrangement using windows allows a smaller volume (the electron column) to be kept at high vacuum. Additionally, they help keep the column clean by limiting out-gasses from gear and fitting oils and detector gas leaking from the spectrometers.

Schematic diagram of the MBX vacuum system. (1) Electron gun isolation valve; (2) Window separating spectrometers from column; (3) Sample change airlock; (4) Ballast tank; (5) Mechanical pumps; (D) Oil diffusion pump.

6.2.3. Pump-down Sequence

The pump-down sequences for a microprobe and scanning electron microscope are very similar. Initially, a mechanical pump used to evacuate the column and sample chamber. Pressure is monitored using a thermocouple gauge. Thermocouple gauges work down to a pressure of about 10^{-3} torr. Once the vacuum is sufficiently good, the "gate" valve that isolates an oil diffusion pump from the column is opened. This allows the diffusion pump to evacuate the column further. There is a water-cooled baffle located above the pump to prevent oil vapor from entering the electron column. A cold-cathode (ion) vacuum gauge, which operates at pressures down to 10^{-6} torr, is used to monitor the pressure at higher vacuums. A cold-cathode gauge gets dirty with time and reads a better vacuum than actually exists.
Schematic diagram of the MBX vacuum system. During initial pumping (green and blue) the column is evacuated using mechanical pump 1. Once a vacuum of about 1 torr is achieved the gate valve opens to access the oil diffusion pump (green and red), isolating mechanical pump 1. The diffusion pump is kept at operating vacuum by mechanical pump 2. The ballast tank serves as a vacuum reservoir and allows the diffusion pump to function for short intervals with pumps 1 and 2 turned off. This minimizes mechanical vibration and permits high magnification operation.

6.2.4. Vacuum Pumps

In general, high vacuum pumps can be divided into capture and momentum transfer pumps. Capture pumps, such as ion-getter and cryopumps, operate by sequestering air molecules onto a surface where they are held it either temporarily (cryopump) or permanently (ion-getter). Momentum transfer pumps, which include oil diffusion, and turbomolecular pumps, move gas by compressing it using some form of mechanical impact, and exhausting it at a higher pressure into another volume at lower pressure. These pumps require initial pumping (“roughing”) to reach operating vacuum and continued pumping to maintain a low exhaust pressure (“backing”).

Water vapor is a major problem in vacuum systems because it desorbs slowly from the internal surfaces of the sample chamber. It is the major residual gas after a chamber is pumped below $10^{-3}$ torr. Because water is released so slowly from surfaces, microprobes are constantly pumped. It may take many hours to reach operating pressures after an electron column has been vented to atmosphere because of the attendant admission of water vapor into the instrument.
6.2.4.1. Mechanical Pump

There are two mechanical (or "sealed-oil") pumps on the Cameca microprobe and JEOL scanning electron microscope. One is dedicated to pumping on ("backing") the diffusion pump; the other serves to evacuate the column, sample chamber and spectrometers during initial pump-down. Mechanical pumps work at a rate of ~12 m³/h and can achieve a final vacuum of ~10⁻² torr.

![Schematic diagram of a mechanical rotary pump. After Potts 1987.](https://example.com)

Mechanical pumps operate by rotation of an off-center cylinder. Longitudinal vanes move out of the cylinder and make contact with the sides of the pump housing. The vanes are either spring-loaded (E) or extended by the centripetal force of the spinning cylinder. Pumping proceeds as follows:

1. The pump takes in air from the volume to be pumped through the inlet. Air molecules are sucked into the space between a pair of vanes (C and D) by the depressurization caused by rotation of the cylinder;

2. The isolated packet of air is compressed by the off-center rotation of the cylinder (E) until it has sufficiently high pressure high to force open the exhaust valve;

3. The air molecules exit through the oil to the outside.
The outlet valve (F) and moving parts of the pump are bathed in oil that serves to both lubricate the moving parts of the pump but also to trap air molecules. The process operates until there are too few air molecules to push open the exhaust valve. Gas molecules in the pressure range of a mechanical pump pressure range move via viscous flow.

6.2.4.2. Oil Diffusion Pump

The diffusion pump is used to produce and maintain low pressures in the electron column and the sample chamber (and in the Oxford Spectrometer on the SEM). An oil diffusion pump requires an initial vacuum of about 0.05 torr (5 to 10 Pa) to operate. Therefore, it is isolated from the column and sample chamber by a "gate" valve until the mechanical pump has produced a sufficiently good vacuum in the column as described above. Diffusion pumps are so named because rather than viscously "pull" air molecules out of a volume, they wait for the molecules to diffuse into the active part of the pump where they are then trapped and removed.

A diffusion pump operates by boiling oil at the base of the pump. Oil molecules escape to produce a supersonic vapor (400-400 m/s) that is directed into a funnel-shaped set of baffles (known as a “chimney stack”) and jetted towards the sides. Air molecules that have diffused into the vicinity are compressed against the jets and entrained. The upper part of the pump wall is cooled by circulating cold water through coils wrapped around it. When the oil vapor strikes the upper pump wall, it condenses into liquid and sinks to the bottom of the pump taking the adsorbed air molecules with it. At the bottom, the heater re-boils the oil, releasing the air molecules, producing an increased concentration of air molecules in the lower region. An attached mechanical pump removes these air molecules. The diffusion pump produces a working pressure in the sample chamber of $10^{-5}$ to $10^{-6}$ torr.
6.2.4.3. Turbomolecular Pump

On some microprobes the oil diffusion pump is replaced by a turbomolecular pump. Turbomolecular pumps use no oil and operate like jet engines with multiple, angled blades rotating at very high speed. Most turbomolecular pumps employ a series of rotor/stator pairs mounted in series. As gas molecules enter through the inlet, the rotor, consisting of a number of angled blades, strikes the molecules, imparting energy to them. Gas molecules are propelled into gas transfer holes in a plate below the blades called a “stator”. Gas captured by the first rotor is sent into the lower rotor/stator pairs and is successively compressed until it reaches a pressure where it can be removed by a mechanical backing pump. The blades are made as thin as possible and slightly bent for max compression. Turbo pumps can reach $10^{-7}$ to $10^{-10}$ torr. Unfortunately, because of the high rotation speeds (10-20,000 rpm), turbomolecular pumps have shorter lifespans than oil diffusion or sealed-oil mechanical pumps.
6.2.4.4. Sputter Ion Pump

Sputter ion pumps are a type of capture pump, wherein air molecules are removed from the chamber by plating (gettering) them onto a surface. The gettering can be accomplished by making the pump wall very cold (cryopumps) or by ionizing gas within a magnetically confined cold cathode discharge (ion pumps). The ultimate pressure achieved by an ion pump is generally in the region of $2 \times 10^{-11}$ Torr. Ion pumps require an initial starting pressure is $5 \times 10^{-3}$ Torr or lower. Operating an ion pump at high pressure for extended periods shortens the pump life.

Ion pumps consist of short stainless-steel cylinders (anodes) sandwiched between two metal (Ti, or Ti and Ta) plates (cathodes), all situated within a strong magnetic field aligned parallel to the cylinder axes. A high voltage is applied between the anodes and cathodes, and the electrons produced from the cathodes move in long helical trajectories through the anode tubes. The long electron paths increase the probability of collision with and ionization of gas molecules. In the main pumping mechanism, the ionized molecules are accelerated toward one of the cathodes where they are buried by Ti atoms. Ion impacts also sputter titanium from the cathode and the resulting Ti atoms acts as a getter for reactive gases (O, N, CO and H) and are deposited as stable oxides, carbides, nitrides and hydrides elsewhere in the pump.
6.2.5. Vacuum Gauges

No single gauge can measure pressure from atmosphere to high vacuum, so different gauges are used to monitor the level of vacuum present within the electron column and spectrometers at different pressures. Microprobes and scanning electron microscopes are usually fitted with thermocouple (1 to 10⁻³ torr range) and cold cathode gauges (10⁻³ to 10⁻¹⁰ torr).

6.2.5.1. Thermocouple Gauge

A thermocouple (or T/C) gauge is used to monitor column pressure during initial pump-down (roughing) and the fore-line pressure of the high vacuum pumps. It operates by measuring the thermal conductivity of a gas. The T/C tube contains a filament heated with a constant current and a thermocouple in contact with the filament. As the pressure decreases, the filament becomes hotter because the number of gas molecules hitting the wire and conducting heat away from the wire decreases. As temperature rises, the thermocouple voltage increases and is measured by a sensitive meter that has previously been calibrated against a manometer to determine the pressure. Each type of T/C tube has its own calibration curve.

Sputter Ion pump. (left) Diode ion pump, which has two cathodes. Image source: http://www.cae2k.com/howto.html. (right) Mechanism of operation of an ion pump. Image source: unknown
6.2.5.2. Cold Cathode Gauge

A cold cathode (Penning) gauge is used to monitor column vacuum during operation. Cold cathode ionization vacuum gauges contain just two unheated electrodes, a cathode and an anode. A "cold" discharge is initiated by an external event (cosmic ray, radioactive decay) and maintained between the cathode and anode by a voltage of ~2 kV. A strong magnetic field makes the electron paths long enough so that they collide with enough gas molecules to maintain the discharge. The magnetic field is oriented so that the field lines cross the electric field lines, confining electrons to a helical path. The resulting current depends on pressure, which is indicated on the meter. The upper measuring range is limited by the glow discharge region (shown above). Unfortunately, the calibration of an ion gauge changes over time as the result of gradual contamination of the components.
6.3. Electron Sources

6.3.1. Introduction

Electron guns provide electrons for an electron beam by allowing them to escape from a cathode material. However, an electron must be supplied sufficient energy to kick it into a high energy state within the material and additional energy for it to escape the surface. Electrons with a component of velocity at right angles to the surface and kinetic energy at least equal to the work done in passing through the surface will be emitted. This total energy required to for a material to give up electrons is related to its work function, $E_w$. The work function of a material is given by

$$E = E_w + E_f$$

where $E$ is the total amount of energy needed to remove an electron to infinity from the lowest free energy state, $E_f$ is the highest free energy state of an electron in the material, and $E_w$ is the work function or work required to achieve the difference.

There are three main types of electron sources used in SEMs and microprobes: thermionic, field emission, and thermal-field.

6.3.2. Thermionic Sources

In thermionic sources, electrons are produced by heating a conductive material to the point where the outer orbital electrons gain sufficient energy to overcome the work function barrier and escape. There are two main types of thermionic sources: tungsten metal filaments and LaB$_6$ crystals. These two types of sources require vacuums of $\sim 10^{-5}$ and $\sim 10^{-7}$ torr, respectively.

Thermionic Emission. Thermionic Emission occurs when enough heat is supplied to the emitter so that electrons can overcome the work-function energy barrier, $E_w$, of the material and escape from the material. Electrons have a range of energies with the highest energy state called the Fermi Level, $E_f$. The energy required to place an electron in the
vacuum from the lowest energy state in the metal is $E$. When the emitter material is heated to a high temperature, a small fraction of the electrons at the Fermi level acquires enough energy to overcome $E_w$ and escape into the vacuum.

### 6.3.2.1. Tungsten Filament Cathode

The electron guns used in NAU's Cameca Camebax MBX microprobe and JEOL-6480LV scanning electron microscope employ tungsten filaments and have a triode (three part) configuration consisting of a cathode, Wehnelt cap and anode.

The tungsten filament is a thin (about 0.1 mm) wire bent into an "inverted V" to localize emission at the tip. This yields a coherent source of electrons emitted from a fairly small area; however, because the filament is bent in a single plane the geometry of the emission region is not perfectly circular. Tungsten is used because it withstands high temperatures without melting or evaporating. Unfortunately, the filament has a very high operating temperature (2700 K). Higher temperatures can deliver greater beam current, but the tradeoff is an exponentially decreasing lifetime due to thermal evaporation of the cathode material.
Electrons leave a heated filament with an average energy of:

\[ E = kT, \]

where \( k = \) Boltzmann's constant (8.617398 x 10^{-5} \text{ eV/K}), and \( T = \) filament temperature (K). At 2700 K, emitted electrons have energies of about 0.23 eV.

### 6.3.2.2. LaB\(_6\) Cathode

Another type of thermionic source uses LaB\(_6\) for the cathode material. LaB\(_6\) has \( E_w \) of \( \sim 2.5 \text{ eV} \), yielding higher currents at lower cathode temperatures than tungsten. Typically, LaB\(_6\) cathodes exhibit 10 times the brightness and more than 10 times the service life of tungsten cathodes. Additionally, the smaller, circular, emission region improves the final resolution of the SEM. However, LaB\(_6\) is reactive at the high temperatures needed for electron emission.
Three designs of LaB$_6$ cathodes. The Broers design uses a tungsten coil to heat the tip of the LaB$_6$ rod. A heat sink at the base helps cool the rod and decrease its overall reactivity. The Vogel design heats the LaB$_6$ rod by passing a current perpendicular to the length of the rod and uses graphite spacers between the rod and contacts to limit chemical reactivity between them. The Ferris design uses a short LaB$_6$ rod is supported by a ribbon or strip through which an electrical current is passed for heating. The ribbon is made of a material that does not react with LaB$_6$, such as graphite or tantalum. Image source and design explanations: http://www.sem.com/analytic/sem.htm.

6.3.2.3. Wehnelt

Electrons are emitted in all directions from the entire heated filament, so a way is needed to localize emission at one spot. This is accomplished by surrounding the filament with a negatively biased Wehnelt cap. The Wehnelt is biased -200 to -300 V with respect to the filament, producing a repulsive electrostatic field that condenses the cloud of primary electrons produced from the filament. Emission from the filament is localized at the tip above an aperture in the Wehnelt.

It is important that the filament is properly centered in relation to the opening of the Wehnelt cap and be the proper distance from the opening. Otherwise, an off-center beam that is either weak/condensed or bright/diffuse will be produced. The Wehnelt cap acts as a convergent electrostatic lens and serves to focus the cloud of electrons. The electrons converge at a point (10-100 μm in diameter) located between the base of the Wehnelt cap and the anode plate.
This point is called the "cross-over" and is the location of the effective electron source. The distance between the tip of the filament and the Wehnelt aperture is critical in determining the geometry of the lens. Movement of the filament tip is the major source of beam instability and even a displacement of 1° will produce a significant change.

The potential difference between the filament and Wehnelt is maintained using a bias resistor, which allows the gun to be self-regulating. Recall from high-school physics that \( V = I \times R \), where \( V = \text{voltage}, \ I = \text{current}, \) and \( R = \text{resistance} \). As the filament emits electrons, an emission current \( (I) \) flows from filament to Wehnelt. Any increase the emission current causes a larger voltage drop \( (V) \) across the bias resistor and a larger negative voltage is applied to the Wehnelt, reducing the current. As the emission increases, so does the voltage difference between Wehnelt and filament, causing the emission to plateau. Proper bias voltage also optimizes the electron beam brightness (current density per solid unit angle) providing the most focused electron beam.

![Schematic relationship between bias voltage, emission current, and beam brightness. Figure after Goldstein and Yakowitz, 1975, p. 25.](image)

The electron flux from a tungsten filament is minimal until a temperature of approximately 2500 K. Above 2500 K, the relationship predicts that the electron flux will increase essentially exponentially with increasing temperature, until the filament melts at about 3100 K. However, in practice, the electron emission reaches a plateau termed saturation due to the self-biasing effects of the Wehnelt cap. Emission currents, those that flow between the filament and anode, generally are on the order of 200 picoamperes \( (\text{pA}) \) \( (1 \text{ pA} = 10^{-12} \text{ A}) \), in SEMs. Proper saturation is achieved at the edge of the plateau; higher emission currents serve only to reduce filament life.
Sample current as a function of filament voltage in a self-biased gun. Emission and sample currents as a function of filament voltage in a self-biased gun (think of voltage as analogous to filament temperature). The operating voltage was 20 kV. Notice the false peak at about 3.4 volts caused by region of filament that reaches emission temperature before tip. After Figure 3.19 in Heinrich, 1981, p. 32.

The flux of electrons from the filament-Wehnelt configuration can be expressed by the ‘Richardson-Dushman’ equation (1923), which describes the current density emitted by a heated filament:

\[ J = A T^2 e^{-\frac{E_w}{kT}} , \]

where \( A \) is a material constant (60 amp cm\(^{-2}\) K\(^{-2}\) for tungsten), and \( E_w \) is the thermionic work function (about 4.5 eV for tungsten).

### 6.3.3. Field Emission Sources

In a field emission source, a large electrical field, \( 10^5 \) to \( 10^8 \) V/cm, is placed between cathode and anode. This field decreases the \( E_w \) of the cathode, a phenomenon called the "Schottky effect". At sufficiently high field strengths, \( E_w \) vanishes. The Schottky effect occurs even at room temperature and depends only very slightly on temperature, indicating that it is not a temperature activated process. Instead, it is a purely quantum mechanical effect called "tunneling". Field emission sources require vacuums of \( \sim 10^{-9} \) torr.
6.3.4. Thermal-field (TF) Source

Finally, in thermal-field source, the tungsten point in a field emission source is heated, incorporating both thermionic and field emissions; this is also referred to as a "Schottky cathode". A TF source requires a vacuum of \( \sim 10^{-8} \text{ torr} \).

Microprobes exclusively use thermionic electron guns. The electron source in an SEMs may use either type of thermionic cathode or utilize field emission. Field emission machines are markedly more expensive and are employed for ultrahigh resolution work. Most geological SEM applications, however, require relatively low magnifications and use thermionic guns.

6.3.5. Anode

The electrons emitted from the filament are drawn away from the cathode-Wehnelt assembly by the anode plate, which is a circular plate with a hole in its center. The anode is biased from +1 to +50 kV with respect to the filament (actually, the electron gun is held at a negative voltage relative to the anode, which is grounded). The voltage potential between the cathode and the anode plate accelerates the electrons down the column and is known as the accelerating voltage, which is usually given in terms of KV and abbreviated \( E_0 \). Together the Wehnelt cylinder and anode plate serve to condense and roughly focus the beam of primary electrons.

6.3.6. Electron Gun Alignment

Observation of the absorbed current (that remaining in the sample during electron beam bombardment) allows alignment of the electron gun. The electron gun is aligned by shifting the position of the filament assembly relative to the anode and the column beneath it, to maximize the absorbed current. In the Camebax microprobe this is accomplished by moving the electron gun using two alignment knobs. Alignment may have to be periodically checked.
The JEOL-6480LV SEM uses an electromagnetic alignment coil located beneath the anode to align the gun. The coil is software controlled.

6.4. Electron Column

6.4.1. Magnetic Lenses

The electron beam is divergent after passing through the anode plate and must be refocused. The simplest type of electron lenses are electrostatic lenses, which deflect beam electrons using electrically charged plates. While a charged particle is in an electric field, a force acts upon it. The faster the particle the smaller the accumulated impulse. Electrostatic lenses require a very clean and high vacuum environment to prevent arcing across plates.

At present, electrostatic lenses most commonly are used to deflect and focus ion beams in mass spectrometers. Microprobes and SEMs use magnetic lenses. Although electron lenses in principle behave the same as optical lenses, there are differences, and the quality of electron lenses is not nearly as good as optical lenses in terms of aberrations.
The first magnetic electron lenses were developed by M. Knoll and E. Ruska in Germany in 1932. The action of a magnetic lens is similar in principle to optical lenses, but electron lenses can only be made to converge, not diverge. They consist of two circularly symmetric iron pole-pieces with copper windings with a hole in center through which beam passes. The two pole pieces are separated by "air-gap" where focusing actually takes place. The magnetic flux diverges along the electron beam axis (figure below). Consequently, an off-axis electron is acted on by a magnetic force proportional to the cross product of the vectors $\mathbf{v}$ and $\mathbf{B}$:

$$\mathbf{F} = -e\mathbf{v} \times \mathbf{B},$$

where $\mathbf{v}$ = electron velocity and $\mathbf{B}$ = the magnetic field. This force causes the electron to move perpendicular to the axis of the lens. The resulting change in direction, in turn, yields a direction of force, deflecting the electron toward the axis. The combination of forces causes the electron to spiral around the beam axis.

Electron lenses are not as good as optical lenses in terms of defects of focus, called **aberrations**. Aberrations are of two types. Spherical aberrations, in which the outer zones of a lens focus more strongly than inner zones, are most important in magnetic lenses. The result is that electrons along beam axis are deflected less than electrons passing through beam periphery, yielding more than one focal point. Chromatic aberrations, in which electrons of slightly different energies are focused differently, are relatively minor because the electron gun produces electrons with essentially uniform velocities. Spherical aberrations are minimized by placing a *spray aperture* in front of the magnetic lens, confining electrons to the center. This results in greatly reduced, but still acceptable, beam currents.
6.4.2. Condenser Lens

The condenser lens controls the amount of current that passes down the rest of the column. This is accomplished by focusing the electron beam to variable degrees onto a lower aperture. The sharper the focus, the less of the beam intercepted by the aperture and the higher the current. The Cameca MBX microprobe has a second condenser lens that is used to provide a better focus for SEM work, but during normal operation, this lower condenser lens is "decoupled" (not used).

Condenser Lens. Schematic of a condenser lens, showing upper spray aperture and lower limiting aperture.
Controlling beam current with the condenser lens. (left) Change in probe current resulting from varying the condenser lens current. The points a, b and c correspond in both diagrams. The condenser lens should be operated to keep the focus point above the lower limiting aperture (currents along BC). Redrafted after Potts 1987.

Operation of the JEOL-6480LV condenser lens. The zoom condenser lens closely maintains beam focus at the sample, without needing operator intervention. The diagram does not depict the limiting aperture inside the condenser lens assembly.

The ultimate electron beam spot size depends on the beam current, which is controlled by the condenser lens, and the type of filament in use. A field emission gun provides the largest current with smallest beam diameter. However, these are expensive and tungsten filaments are most commonly used.
Microprobe electron beam currents are usually from about $10^{-9}$ A ($10^{-9}$ A = 1 nanoamp = 1 nA) to $10^{-7}$ A (100 nA). SEM electron beam currents of are usually on the order of 100 pA ($10^{-12}$ A = 1 picoamp = 1 pA). Microprobe beam currents are larger (resulting in a larger focus point) because there is no reason for beam diameter to be smaller than diameter of excited region (~2 µm). In contrast, SEM beam size is not limited by excitation area because SE are produced from just the upper ~100 pm of the sample. Scanning electron microscopes typically use beam diameters much less than 1 µm and use smaller final apertures to limit beam diameter. We tend to operate the JEOL-6480LV SEM as if it were a microprobe, using high sample currents.

The beam below the condenser lens is again divergent and must be refocused before striking the sample. Before this occurs, it is useful to be able to monitor and regulate the beam current. Emission from a tungsten filament can vary by up to 1% in 10 minutes; although this seems a small change, it will result in unacceptable analytical results. Thus, beam regulation is of critical importance in the electron microprobe. Such variations can be ignored in a scanning electron microscope, where the primary function is imaging; it takes no longer than about 2 minutes to acquire an image.

6.4.3. Beam Regulator

The beam regulator is located between the condenser and objective lenses in the electron microprobe column. The beam current may vary due to shifts in filament emission or changes in the current through the condenser lens. The regulator senses changes in the current and adjusts the condenser lens power accordingly.
Beam regulation in the MBX microprobe. Schematic of the microprobe electron column showing the electron beam in red with its width greatly expanded. The effects of the condenser lens (1) and objective or probe-forming lens (PFL) are shown schematically by divergent black arrows; the sample is located at (6). The beam regulation system, located below the cross-over from the condenser has two apertures: a limiting aperture (2) and lower aperture with a smaller diameter hole (3). The amount of current impinging on the lower aperture (3) is amplified (5) and used to control the condenser lens power supply (4). Increased current through the condenser causes the beam to widen and more of it is blocked by the limiting aperture. The lower aperture senses the small beam current and decreases the condenser lens current, allowing more beam to pass through it. If the beam current decreases and narrows, less is blocked by the limiting aperture, more strikes the lower aperture. The current through the condenser lens is increased, decreasing the beam current.

6.4.4. Beam Monitor

The beam can be measured by interposing a sensor along its path. In the microprobe, the beam is measured before every analysis. X-ray counts are normalized to the measured value as an additional control on drift in the beam current. This procedure requires that there be a linear correlation between beam current and X-rays produced from the sample; however, in practice this does not seem to be the case for large changes in current. For example, the beam normalized count rates on a standard vary significantly between a 10 nA and 25 nA beam current. Obviously, one would not expect such large changes to occur as a result of drift!

Beam monitoring is software controlled in the microprobe. The user must push a button to activate the SEM beam monitor. A Faraday cup is used to block and measure the electron beam current. This device consists of a cup of electron absorbing material, such as carbon, capped by a 25 to 100 µm aperture. The Faraday cup is swung into the beam current, completely blocking
it. The beam electrons enter the cup through the hole in the aperture and are trapped. Those backscattered hit the bottom of the aperture; similarly secondary and Auger electrons are unable to escape. The cup is grounded through a picoammeter, which measures the resulting current.

![Faraday cup diagram](image)

**Faraday cup.** Schematic diagram of a Faraday cup with the electron beam shown in blue.

The Cameca MBX microprobe does not use a true Faraday Cup, although later models (SX-50, SX-100) do. Instead, an aperture-capped cylinder of electron absorbing material, through which the beam always passes, is part of the electron column. The beam is measured by swinging a "trap door" to block the bottom of the cylinder, blocking the beam.

![SEM beam monitor](image)

**JEOL-6480LV SEM beam monitor.** The user presses a button to move the Faraday cup in and out of the beam.

6.4.5. **Objective Lens**

The objective, or "probe-forming", lens is located at the base of electron column just above sample. The beam is again divergent after passing through the apertures below the condenser lens and must be refocused. The objective lens focuses the electron beam onto the sample and controls final size and position. Hosted within it are **scanning coils** that allow the beam to be rastered across the sample surface, the astigmatism correction coils, the beam shift coils, and the visible light microscope optics.
6.4.5.1. Microprobe Minicoil Lens

The Cameca MBX microprobe has a complex "minicoil" objective lens. It is an asymmetrical magnet with less of iron casing located below its air gap to minimize the magnetic field at the sample. It is designed to accommodate the light optical system and scanning coils inside it and has apertures that allow X-rays from the sample to pass to the spectrometers at a 40° take-off angle. The apertures have electron traps, which use negative voltages to prevent backscattered electrons from entering X-ray spectrometers.

![MBX microprobe objective lens](image)

MBX microprobe objective lens. The light optics (4) and scanning coils (1) are located inside the minicoil probe-forming lens (2), which has a pole piece (7) made of one solid piece of metal. X-rays (3) are collimated by small apertures (6) in the lens. Electron traps (5) prevent backscattered electrons from entering the X-ray spectrometers.

6.4.5.2. SEM Objective Lens

The objective lens in the JEOL-6480LV SEM is much simpler than that used in the MBX microprobe, being a solid conical lens with the scanning coils housed inside it. There is no need for the objective lens to accommodate a light microscope system, allowing it to have a more pointed configuration. Consequently, there is no need for apertures for X-ray spectrometers to achieve comparable take-off angles.
6.4.6. Scanning Coils

Scanning (or "deflector") coils raster the beam across the sample for textural imaging. Beam deflection is accomplished by coils located within the objective lens. The coils consist of four radially oriented magnets that produce fields perpendicular to optical axis of the electron beam. The beam is rastered by varying the current through these magnets. Deflecting the beam off-axis introduces aberrations that limit the maximum deflection that can be used and still keep the sample in focus. Additionally, deviations from linearity in X and Y increase as amount of deflection increases.
6.4.7. Astigmators

Magnetic lenses do not have perfect symmetry and the electron beam may be elliptical, something termed **astigmatism** (stige means "mark" or "spot" in Latin). An astigmatic beam will collect signals from an elliptical region that will be rendered as a round spot on viewing screen or final image, producing a reduction in resolution. Astigmatism results from machining errors, asymmetry in magnetic windings, inhomogeneous magnetic fields in the iron, the buildup of contamination within electron column, and misalignment of the electron column. Only misalignment of the column is easily addressed because cleaning the column is a difficult and prolonged exercise.

**Effects of astigmatism.** If the effect of a magnetic lens is slightly tilted (or, equivalently, if the electron beam is tilted off axis), the focusing pattern becomes astigmatic. The beam first focuses along a horizontal segment (tangential focus), has a more or less circular shape at the focal plane, and focuses again along a vertical segment (sagittal focus) behind the focal plane. Image source: [http://www.optonlaser.com/pages_communes/glossaire/GlossaireLaser.htm](http://www.optonlaser.com/pages_communes/glossaire/GlossaireLaser.htm).

Both microprobes and scanning electron microscopes have "stigmators" housed with the scanning coils inside objective lens that compensate for imperfections in construction and those not corrected by alignment of column. These consist of eight weak electromagnetic lenses with alternating coils connected in series but wound in opposite directions to achieve different polarities. Stigmators apply a correcting magnetic field to produce symmetrical electron beam at the sample. The operator can change both the strength and orientation (angle) of the magnetic field produced by stigmators to control the final beam shape.

**Stigmators.** Schematic top view of the arrangement and polarities of the stigmator coils.
6.4.8. Column Alignment

The alignment of the microprobe column can be tested by changing the objective lens focus and examining the over- and under-focus relationships of the electron beam on a fluorescent sample. If the limiting aperture is not centered along the axis of the electron column, the image of the aperture hole will rotate as the setting of the objective lens is changed due to the helicoidal movement of electrons in the beam. The aperture should be centered using the limiting aperture knobs to eliminate this movement.

Alignment of the limiting aperture. A. Perfect alignment: the beam changes diameter symmetrically when changing objective lens focus. B. Poor alignment: the beam sweeps out an arc as the focus is changed; $r_1$ is the distance from the focused beam to the electron column axis. C. Fair alignment: this is the most common case, where alignment is slightly off. Note that rather than sweeping, the beam seems to expand and contract from one side as focus is changed; $r_2$ is the distance from the focused beam to the electron column axis. After Figure 3.25 from Heinrich, 1981, p. 42.

6.5. Sample Chamber

Geological microprobe samples are usually petrographic thin sections that have been ground flat and polished. In contrast, SEM samples are far more variable ranging from thin sections to actual large pieces of rock. The sample chambers reflect this difference. The Camebax microprobe has a sample chamber just large enough to permit movement of two entire thin sections under the electron beam (one usually reserved for standards materials). In contrast,
the JEOL-6480LV scanning electron microscope has a sample chamber that permits large and complex movements of the sample stage; for example, the SEM can accommodate samples up to 200 mm in diameter and ~80 mm high.

![JEOL-6480LV SEM sample chamber. The silver stage motor in the front is ~11 cm long.](image)

### 6.5.1. Low-Vacuum Mode

The JEOL SEM has the capability to also work with a relatively poor vacuum in the specimen chamber. Air (or another gas) is admitted into the sample chamber to neutralize the build-up of the negative electrostatic charge on the sample. In "low-vacuum (LV) mode," the sample chamber pressure, typically 20-30 Pa, is maintained at a specifiable value by a separate extra pump with a large fore-line trap. The electron column is separated from the chamber by a vacuum orifice (aperture) that permits the column and objective lens apertures to be maintained at a higher pressure. Unfortunately, secondary electrons are absorbed by the air in the sample chamber. Instead, the backscattered detectors are used to image the sample. LV mode allows the study of samples with poor or no electrical conductivity. Normally, these materials would be either coated with an electrically conductive coating, which takes time and hides the true surface, or studied at low voltages, which do not give backscatter or X-ray information ($E_0 < E_c$ for most elements of interest). Additionally, samples containing volatile substances (water) can often be studied directly.

### 6.5.2. Environmental SEM

An environmental SEM utilizes relatively low vacuum pressures (up to 50 torr ~ 6700 Pa) not only to neutralize charges, but to provide signal amplification. The electron column is gradually brought up to pressure by a series of pumps and apertures.
A positively charged detector electrode is placed at the base of the objective lens at the top of the sample chamber. Secondary electrons emitted from the sample are attracted to the positive electrode, undergoing acceleration. As they travel through the gaseous environment, collisions occur between electrons and gas particles, producing more electrons and ionization of the gas molecules and effectively amplifying the original secondary electron signal (electron cascade). This form of signal amplification is identical to that which occurs in a gas-flow X-ray detector. Positively charged gas ions are attracted to the negatively biased specimen and offset charging effects. The vapor pressures in environmental SEMs can high enough that liquids may be observed!

Dissolution and precipitation of halite observed in an ESEM. A series of electron micrographs showing the dissolution (a–c) and reprecipitation (d–f) of NaCl as water is condensed on the surface of the salt crystals and then evaporated. White bar at the lower center part of each micrograph = 20 μm. Image source: [http://aic.stanford.edu/jaic/articles/jaic33-02-008.html](http://aic.stanford.edu/jaic/articles/jaic33-02-008.html).

6.6. Microprobe Sample Stage

The sample chamber for a microprobe does not need to be large since all samples are flat. The only requirement is that the sample chamber have sufficient room for the motors and gearing that allow movement of the sample under the electron beam. Most microprobes can accommodate standard polished thin sections or 1" polished round sections and most permit examination of the sample with either reflected or transmitted light (see below). The number of sections and the maximum amount of movement depend upon the individual instrument.

The X, Y, and Z positioning of samples is accomplished by moving the stage with stepper motors. The MBX microprobe has a transmitted light stage installed that can be moved 20 mm in the X-direction (left/right) and 50 mm in the Y-direction (in/out). Movement in the Z-direction is limited to avoid striking the optical system. Each stepper-motor step represents 1 μm of stage travel. Consequently, stage motion is reproducible to about 1 μm. The worm gears on the stage drives and spectrometers (see below) are subject to mechanical "slop", even with very high tolerance machining. As a result, peak positions appear to shift when approached from different directions (up or down the drive screw) and stage positions are not reproducible. The MBX microprobe analytical programs use a "backlash" to compensate for this slop in which all peaks and stage positions bypassed by a certain number of steps and always approached from the same direction. The newest Cameca microprobes (SX-50, SX-100) use optically encoded light bars to control stage (and spectrometer) movement. The direction of movement does not matter since the positions are located using absolution positions on the encoders. The
stage and spectrometers are moved using DC motors rather than AC stepper motors, and movement is very rapid and precise.

![MBX sample stage](image)

**MBX sample stage.** (left) The drive motors are located on the left of the stage assembly. Note the forked sample holder that can accommodate two samples simultaneously. The technician is holding the transmitted light assembly prior to installation. (right) Sample holder with a sample installed in right-hand fork. The transmitted light assembly inside the column is positioned as indicated.

In microprobes, sample changes are accomplished using an airlock to avoid having to vent the entire column. The airlock cover plate is removed, and a sample change unit attached. This changer is evacuated to a pressure of about 2 torr, whereupon the airlock door can be opened, and the sample inserted. In the MBX microprobe the sample slides into a narrow forked holder that allows a transmitted light source to be positioned beneath them.

### 6.7. Light Optics

The MBX microprobe is equipped with a light microscope that allows simultaneous examination of electron beam and specimen during operation. Beam position, size and shape are observed using a fluorescent sample such as willemite, or benitoite. This allows the operator to locate the beam position prior to analysis of minerals that do not fluoresce. The microscope visible light optical system is housed within the objective lens. The optical system consists of a series of mirrors with holes in them through which electron beam passes. The resolution of the microscope is 0.55 µm (the smallest tick marks on optical reticule are 1 µm apart); the field of view is 500 µm wide. The light optics have a small depth of focus (~0.9 µm), insuring good focus for analysis.
Samples can be observed in both reflected and transmitted light. The transmitted light assembly consists of a right angle mirror positioned beneath the optical microscope and a horizontally oriented light source located outside of the evacuated electron column. There is only one magnification (400x) available for observation on the MBX probe; this magnification is higher than that used during standard petrographic examination, making it sometimes difficult to locate points of interest for analysis. Other microprobes have additional magnifications available. Regardless, all samples should be thoroughly characterized petrographically and grains of interest be carefully located before microprobe analysis. One popular method is to use thin lines of India ink leading to or circling grains of interest. Additionally, photographs taken in transmitted or (better) reflected light serve as useful maps of a sample.
Chapter 7 – Wavelength Dispersive Spectrometry (WDS)

7.1. Introduction

Wavelength spectrometers are used to select the X-ray of interest for analysis. This selection is made by Rayleigh scattering of the X-rays from a systematic crystal located between the sample and X-ray detector. By changing the angle of incidence, a crystal can be made to constructively diffract X-rays of different wavelengths. The diffracted X-rays are counted using X-ray detectors that must be moved to accommodate the changing incident angles on the crystal.

![Schematic diagram of a wavelength dispersive spectrometer.](image)

7.2. Bragg’s Law

Scattering and diffraction of a coherent beam of X-rays can occur from parallel planes of atoms spaced distance, $d$, apart. For a given angle, $\theta$, in order for constructive interference to occur, the extra distance traveled by a wave diffracting off the lower layer (CBD in the figure below) must be an integer number of wavelengths.

Since

$$CB = BD = d \sin \theta,$$

the conditions necessary for constructive diffraction of X-rays (and light) by a crystal are
\[ n\lambda = 2d \sin \theta \]

where \( n = \) integer (1, 2, 3 ...). X-rays with other wavelengths will produce destructive interference. Note, however, that the wavelengths 1\(\lambda = 2(\lambda/2) = 3(\lambda/3) \ldots \) Thus, specific shorter wavelengths will also experience constructive interference. These are called higher-order wavelengths (abbreviated 2°, 3°, etc.) because their diffraction occurs when \( n > 1; \) the value of \( n \) specifies the "order" of the diffraction. Higher order diffracted wavelengths generally have lower intensities than those produced by first order (1°) diffractions.

![Wavelength dispersive spectrometer scan of an augite pyroxene using a TAP crystal. Note the high-order peaks at high sin\(\theta\) values.](image)

### 7.3. Spectrometer Geometry

The WDS spectrometers on the microprobe and JEOL-6480LV SEM use what are termed linear-focusing drives. For efficient X-ray focusing, the sample, crystal, and detector must lie on a circle called the Rowland circle and remain on it for all wavelengths of interest. Since the sample is fixed in place, the crystal and detector must both move to remain on the Rowland circle. This arrangement is mechanically complex, since to maintain the appropriate geometry the crystal must rotate as source-crystal distance changes. However, because the Rowland circle radius and the take-off angle do not change, the mechanical problems are not insurmountable. Precision internal gearing controls the movement of the detector along a complex path and rotation of the detector. Crystal and detector movement are driven by a worm gear connected to a stepping motor. Each step corresponds to 10000 \( \times \sin\theta \) on the MBX microprobe. Thus, a position where \( \sin\theta = 0.45 \) corresponds to a motor value of 45000.
Spectrometer geometry. The geometry of a linear, fully focusing X-ray spectrometer. This type of spectrometer is used in most commercial wavelength-dispersive electron microprobes (after Williams 1987).

A larger diameter Rowland circle is more efficient than a small diameter circle in providing peak separation at given angle. The diameter of the Rowland circle of all Cameca microprobes is 16 cm, whereas the JEOL 733 Superprobe has a 14 cm Rowland circle. Thus, the Cameca microprobes provide better peak resolution. However, larger Rowland circles require larger spectrometer housings with the attendant vacuum and positioning problems. A very large Rowland circle could provide better peak resolution, but it is not practical.

MBX Microprobe Spectrometer. (left) The armature and gearing that move the detector are visible in the center of this photograph of a Cameca SX-50 vertical spectrometer; (upper right) Detail of the crystal holder and flipping motor; (lower right) X-rays from the sample come through the X-ray window (circular hole into the column) and are diffracted by the analyzing crystal into the detector. The high voltage feed and gas-flow lines trail off from the detector towards the lower right corner.
Spectrometers may be mounted vertically or inclined relative to the top surface of the sample. The NAU microprobe has one inclined spectrometer (#1) and two vertical spectrometers (#2, #3). An inclined Oxford spectrometer is attached to the JEOL-6480LV SEM.

Inclined spectrometers are less sensitive to changes in sample focus and topography, but, because of their orientation, only a few can be mounted together around the electron column. A small variation in sample height simply displaces the X-rays to a different portion of the analyzing crystal.
Focusing in an inclined spectrometer. The effect of a small variation in sample height on the focusing of the crystal is very small in an inclined spectrometer, as is the corresponding variation of the Bragg angle. For clarity, only the front part of the Rowland circle is shown (after Maurice et al. 1979).

Vertical spectrometers are more sensitive to changes in focus, but many can be mounted together. For example, the old Smithsonian microprobe had 12 vertical spectrometers!

Focusing in a vertical spectrometer. The effect of a small variation in sample height on the focusing of the crystal produces a correspondingly large variation of the Bragg angle (after Maurice et al. 1979).

7.5. Take-off Angle

The angle between the surface of the sample and the detecting crystal is termed the "take-off" angle, $\phi$. High take-off angles decrease absorption of the X-rays and other effects produced in the sample and yield in higher count rates. Low take-off angles increase absorption because the X-rays must traverse more material. Thus, a low $\phi$ take-off angle is more sensitive to topographic effects. For this reason, secondary electron counters are located at low angles to enhance the topographic resolution of secondary electron images.
Schematic illustration of the take-off angle. For a given angle of electron incidence, the length of the absorption path is directly proportional to the cosecant of the take-off angle, \( \phi \).

Absorption increases rapidly for take-off angles of under 25° but is relatively constant for angles greater than 40°. All Cameca and JEOL microprobes have take-off angles of 40°, old ARL machines have angles of 52.5°, and MAC and ETEC machines angles of 38.5°. Count rates on ARL microprobes are about 30% higher than on MAC microprobes simply due to the higher take-off angle.

### 7.6. Analyzing Crystals

#### 7.6.1. Introduction

Space constraints limit the range of Bragg angles possible in a wavelength dispersive spectrometer. The incident angle, \( \theta \) ranges from 12.5° to 56° in the MBX microprobe; the range is slightly larger in a JEOL 733 microprobe (12.5° to 65°). The range for the Oxford wavelength dispersive spectrometer attached to the JEOL-6480LV SEM is 16.5° to 67.5°. Thus, crystals with different d-spacings are necessary to cover the entire range of X-ray wavelengths of interest (~1 to 20 Å). Many microprobes have multiple crystals in each spectrometer to increase analytical flexibility; the NAU microprobe has pairs of crystal that can be flipped. The most commonly used crystals are:

- Lithium fluoride 200 (LIF)
- Pentaerythritol 002 (PET)
- Thallium acid pthalate 1011 (TAP)
- Lead sterate or Lead octodecamoate (ODPB)

LIF is an ionic solid, PET and TAP are organic crystals, and ODPB is a pseudo-crystal, in which Pb atoms are interlayered with a fatty-acid salt to produce regularly spaced planes equivalent to interplanar crystal spacing; it is periodic in only one direction. There are now many types of large d-spacing synthetic "crystals." For example, the JEOL-6480LV has an LSM-060 crystal, which is made of W and Si, and has a d-spacing of ~61Å. The ranges for each of the crystals are shown below.
A given X-ray line can be diffracted using any crystal, but spectrometer mechanical limits will restrict which crystals can actually be used. For example, Fe-Kα with a wavelength of 1.937 Å, is located at θ of 28.75° on LIF, 12.8° on PET, and 4.3° on TAP. Fe-Kα radiation is very near or beyond the mechanical limits of the MBX spectrometers for PET and TAP, necessitating use of a LIF crystal. Recall that although a spectrometer may be able to reach a certain elemental peak, insufficient overvoltage may make the peak unusable.

<table>
<thead>
<tr>
<th>Crystal</th>
<th>Cameca MBX</th>
<th>JEOL SEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>2d (Å)</td>
<td>Energy Range (Å)</td>
</tr>
<tr>
<td>TAP</td>
<td>25.75</td>
<td>5.573 to 23.337</td>
</tr>
<tr>
<td>PET</td>
<td>8.742</td>
<td>1.892 to 7.923</td>
</tr>
<tr>
<td>LIF</td>
<td>4.0267</td>
<td>0.872 to 3.649</td>
</tr>
<tr>
<td>ODPB</td>
<td>100</td>
<td>21.644 to 90.631</td>
</tr>
<tr>
<td>LSM-060</td>
<td>61</td>
<td>N/A</td>
</tr>
</tbody>
</table>

7.6.2. Crystal Characteristics

Not every regular crystal makes a good analyzing crystal.

- Analyzing crystals must be chemically stable; for example, although NaCl has a useful 2d-spacing, it dissolves in humid air.
- A crystal must also not be too perfect. If a spectral peak is too sharp, its position is difficult to reproduce consistently.
- Low X-ray absorption is also critical. An analyzing crystal is constantly bombarded with X-rays; while much radiation is diffracted, some is absorbed and re-emitted as secondary fluorescence radiation, which may reach the detector.
- Analytical crystals must effectively separate adjacent spectral lines, something called dispersion efficiency. Dispersion efficiency may be expressed as

\[ \frac{d\theta}{d\lambda} = \frac{n}{2d \cos \theta} \]

Note that higher order reflections are better dispersed and that the 2d-spacing of the crystal is an important factor. Where high resolution is desired, it is best to avoid low θ and large 2d-spacings even at the loss of peak intensity. For example, an LIF crystal will separate the peaks for Ti-Kα (4508 eV) and Ba-Lα (4466 eV) better than a PET crystal.
• Such conditions pertain when analyzing trace amounts of V in Ti-bearing materials and the analysis of rare earth elements (REE).

• Crystals also must have high reflection efficiency. Reflection efficiency depends on the number of electrons in orbitals of atoms in the crystal, the angle of diffraction and the wavelength of the X-rays.

• Finally, crystals should not be sensitive to temperature fluctuations, having a low coefficient of thermal expansion. The effect of thermal expansion is most pronounced at higher diffraction angles. Of the commonly used crystals, PET is most sensitive to changes in temperature, but the effect is not a problem if temperature is adequately regulated.

![Graph showing the effect of temperature on analyzing crystals](image)

**Effect of temperature on analyzing crystals.** The effect of thermal expansion is most pronounced at high diffraction angles with PET showing the greatest changes (after Jenkins & de Vries, 1967).

### 7.6.3. Crystal Design

The efficiency of an X-ray spectrometer, in part, depends on the crystal design and crystals are bent and/or ground to improve their efficiency. This can be done in two ways:

- **Johann optics**, where the dispersing crystal is bent to a radius equal to the diameter of the Rowland circle, R. This results in some broadening and asymmetry of the focusing 'point'.

- **Johanson optics**, where the crystal is bent as in Johann optics but also ground to a radius of R/2, resulting in focusing over the entire range of angles.
The essential geometry of fully focusing Johanson (left) and semi-focusing Johann (right) X-ray spectrometer optics. These diagrams are not drawn to scale and exaggerate the loss of resolution in a semi-focusing spectrometer. The effect is much less pronounced than the diagrams indicate; semi-focusing spectrometers are used in microprobes (after Williams 1987).

7.7. X-ray Detectors

7.7.1. Introduction

Detectors convert X-rays photons through ionization into voltage pulses which can be counted. There are three types of proportional detectors used in microprobes to detect X-rays: gas-flow, sealed-gas and semiconductor. In a "proportional" counter the size of pulse is proportional to the energy of the X-ray that produced it. However, the first two types have poor energy discrimination and must be used with an analyzing crystal to preselect the X-ray energy. Semiconductor detectors have excellent energy discrimination and are used without analyzing crystals. They will be discussed in the Energy Dispersive Spectrometry section.

7.7.2. Gas-flow and Sealed Detectors

Gas-flow and sealed-gas detectors have identical designs except that in the first case the detector gas is constantly changed and replenished while in the latter the gas is static. Both detectors consist of a cathode tube with a thin (20-100 mm thick) W-wire anode running through the center. A 1-2 kV voltage is applied between wire and tube. Sealed-gas detectors were first developed to a high degree of reliability by H. Geiger and W. Muller in 1928.
The detector tube has a window through which X-rays from the analyzing crystal enter the detector. The window is usually made of either Mylar (polyethylene terephthalate) or polypropylene window. Unfortunately, whatever the thickness of the window, it will still absorb a significant portion of the X-rays from the analyzing crystal. For example, a 5.5 mm Mylar window absorbs ~50% Al-Kα, ~70% Mg-Kα, ~85% Na-Kα and ~98% F-Kα X-rays. Thinner polypropylene windows absorb about 60% less than Mylar windows and are routinely used for light element detectors. The NAU microprobe uses a 1 mm polypropylene window for the light element spectrometer (TAP crystal) and 6 mm Mylar windows for more energetic X-rays (PET and LIF).

Detectors usually have a second window opposite the one that allows the X-rays from the crystal to enter. These allow high-energy X-rays to exit, preventing them producing secondary radiation from the detector itself. All MBX detectors have thicker (25 mm) Be windows in their backs, as does the light element detector of the Oxford spectrometer on the JEOL-6480LV microscope.

In a gas-flow detector, a mixture of 90% argon and 10% methane, called P-10, continuously flows through the detector. The detector tube is not just sealed with P-10 gas inside, because some gas escapes through the thin windows and must be constantly replenished. The detector gas must be ultra-pure to avoid reaction with the wire and subsequent loss of detector sensitivity. Electronegative impurities like O²⁻ and CO⁻ are especially bad. The efficiency of a gas-flow detector can be improved by increasing the pressure and most probes have high and low pressure detectors.

Sealed detectors filled with Xe or Kr gas are used for higher energy X-rays since these gases are ionized more efficiently than Ar by shorter wavelengths. Since they are used for more energetic
X-rays, the detector windows can be thicker to prevent loss of the gas by diffusion. A window ~25 mm thick made of Be or Al is commonly used because the high-energy X-rays gradually destroy Mylar or polypropylene windows.

### 7.7.3. Detection Process

In gas-flow and sealed detectors, the X-rays ionize an inert detector gas, ejecting an outer shell electron to produce an electron-ion pair: The first ionization potentials of the inert gases are small (less than 25 eV); however, the effective ionization potential required to produce an electron-ion pair is somewhat higher due to competing processes which absorb incident photon energy without causing ionization.

<table>
<thead>
<tr>
<th></th>
<th>1st Ionization Potential (eV)</th>
<th>Effective Ionization Potential (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>24.5</td>
<td>27.8 eV</td>
</tr>
<tr>
<td>Ar</td>
<td>15.7</td>
<td>26.4</td>
</tr>
<tr>
<td>Xe</td>
<td>12.1</td>
<td>20.8</td>
</tr>
</tbody>
</table>

The average number of electron-ion pairs, $n$, produced by an X-ray is

$$n = \frac{E}{e_i}$$

where $E$ = incident X-ray energy (eV) and $e_i$ = effective ionization potential of the detector gas (eV). Consider a Cu-Kα X-ray, which has energy of 8.04 kV. With Ar detector gas, $n = 8040/26.4 = 304$ primary electron-ion pairs. This number is too small to detect by itself but placing a potential across the gas from wire to tube wall will produce amplification. The electrons produced by the incoming X-ray are accelerated towards the anode wire by the detector voltage and can in turn ionize other Ar atoms producing another electron-ion pair and so on. The resulting **electron avalanche** is still proportional to the initial signal.
**Principle of detection of an X-ray photon.** Incident X-rays ionize the Ar detector gas (purple spheres) losing an average of 26.4 eV, then continue to ionize other Ar atoms. The resulting secondary electrons are accelerated toward the detector wire, gaining sufficient energy to ionize other Ar atoms, producing an electron avalanche. The Ar ions are neutralized by electrons donated by methane molecules in the detector gas mix.

The chain of ionizations causes a momentary voltage across the detector producing a pulse. The amount of amplification produced by the gas depends on the amount of voltage applied to the detector. At very low voltages in the region of "undersaturation", the detector potential difference is too small to prevent recombination of electron-ion pairs formed by incident X-rays before they reach the collecting wire. At slightly higher voltages in the ionization chamber region, the potential is just sufficient to counter recombination so that the number of electron-ion pairs produced by X-rays equals the number reaching the anode wire; the resulting amplification (gain) is 1. Further increases in the detector voltage produce the avalanche effect and significant amplifications. At voltages in the proportional counter region, the pulse height is proportional to the energy of the incident X-ray. Too high a voltage drives the detector out of the proportional region and into the Geiger region.

**Effect of detector voltage.** Changing the applied anode voltage affects (a) the gas amplification factor and (b) the observed count rate for a gas proportional counter. Note the rapid increase in observed count rate at the threshold of Geiger breakdown (after Potts 1987).

Gain is defined as

\[
G = \frac{N}{n}
\]

where \(N\) = number of electrons reaching the anode wire, \(n\) = number of electrons produced by X-ray ionization. Detector gains are typically on the order of \(10^4\) to \(10^5\). With a gain of \(10^5\), the 304 electron-ion pairs formed by a Cu-K\(\alpha\) X-ray yield \(3.04 \times 10^6\) electrons at the anode wire. The size of the resulting pulse can be calculated from

\[
V = \frac{Nc_e}{C}
\]

where \(c_e\) = charge on the electron \((1.6022 \times 10^{-19}\) coulomb\), \(C\) = capacitance of the detector (typically \(10^{-10}\) farad). Thus, the Cu-K\(\alpha\) photon will generate a voltage, \(V\), of
or 4.87 mV. Recall that we've assumed a gain of 10,000! The resulting pulse is still very small and needs further electronic amplification by a preamplifier located near the detector on the spectrometer housing.

### 7.7.4. Escape Peaks

If the energy of the incoming X-rays is greater than the absorption edge of the detector gas, it can produce characteristic X-rays from the gas and produce what is termed an escape peak. The name derives from the fact that some energy of the incoming X-ray is escaping as characteristic X-rays from the detector gas.

For example, $E_c$ for Ar is 3.2 kV and any X-rays with higher energy can excite Ar-Kα X-rays ($E = 2.95$ kV). The production of characteristic X-rays from the gas decreases the apparent energy of the incident X-ray and yields a separate peak offset towards lower energy by 2.95 kV. The size of the escape peak is a function of fluorescent yield, $\omega$, and is generally smaller than the characteristic peak because relatively few X-rays excite the gas compared with the number detected; however, in some cases its size can be quite large.

![Escape Peaks Diagram](image)

Examples of escape peaks. (A) Iron escape peak in argon ($\omega_x = 0.12$); (B) Molybdenum escape peak in krypton ($\omega_x = 0.65$) (after Maurice et al. 1979).

Escape peaks can be a problem if they interfere with peaks of elements of interest.

### 7.7.5. Dead Time

All detectors suffer from dead time due to the physical constraints on X-ray detection. The time between X-ray entry and electron-ion pair avalanche production is on the order of $3 \times 10^{-7}$ seconds. The avalanche produces a sharp voltage drop that requires $\sim 10^{-4}$ seconds to decay back to zero. The rate-limiting process in this decay is the slow movement of $\text{Ar}^+$ ions to the cathode tube, which is about $10^3$ to $10^4$ times slower than electron movement. This slow neutralization of the $\text{Ar}^+$ ions prevents the production of another electron-ion avalanche and
produces what is called a dead time, an interval during which the detector is inoperative and cannot detect incoming X-rays.

To counter the effect of the slow movement of Ar⁺ ions, methane is added to the detector gas ("P-10" indicates 10% methane). Methane is an electron donor and provides "quench" electrons to the Ar⁺ ions, allowing the detector to rapidly reset, decreasing dead time, because neutralization of the Ar⁺ ions is not dependent on them actually reaching the cathode tube. In addition, the electronics are often set to "clip" the decay pulse after about 10⁻⁶ seconds. Dead times are usually about 1 to 2 µsec.

Effect of dead time. Plot of output count rate as a function of input count rate, for four dead time constants.

All X-ray count rates are corrected for the effect of dead time by the analysis software. Such dead time corrections become very important at count rates exceeding 3000 cps. The detector dead time correction is

\[ I_t = \frac{I_m}{1 - I_m t_d}, \]

where \( I_t \) = true counts, \( I_m \) = measured counts, and \( t_d \) = dead time. For example: if \( I_m = 10^5 \) and \( t_d = 2 \times 10^{-6} \) seconds, then \( I_t = 1.25 \times 10^5 \) (25% higher than observed). Dead time constants can be determined by accumulating a series of intensities, \( I \), at different beam currents, \( i \). For any current, \( i \), the measured intensity is proportional to the current:

\[ I_m = Ki. \]

Thus, we can rewrite the dead time correction equation:

\[ \frac{I_m}{i} = K(1 - I_m t_d). \]

Plotting \( \frac{I_m}{i} \) against \( I_m \) will yield a straight line; the value \( K \) can be determined by extrapolating the line to \( I_m = 0 \). Once \( K \) is known \( t_d \) can be obtained from any point on the line by

\[ t_d = \frac{1 - (\frac{A}{K})}{I_m}, \]
where $A = (I_m/i)$. Values calculated from higher values of $I_m$ will generally be more accurate because of better counting statistics.

### 7.7.6. Resolution

The statistical process of pulse formation yields a Poisson distribution of voltages coming from the detector.

![Detector Resolution](image)

**Detector Resolution.** Peak shape and energy resolution of the detected X-ray peak (A) and of the proportional counter (B). Resolution of the X-ray peak includes the additional effects produced by the analyzing crystal, whereas resolution of the proportional counter reflects the effects of the detector electronics.

At sufficiently high count rates a Poisson distribution approaches a normal distribution, and the relative standard deviation may be expressed as

$$\frac{K}{\sqrt{E}},$$

where $K = \text{constant (close to 16 for Ar-filled detectors)}$, and $E = \text{energy of incident X-rays (kV)}$. The pulse height distribution is wider for longer wavelengths (lower $E$). For example, the standard deviation is 13.1% for Al-Kα and 6.3% for Fe-Kα.

A more refined way of calculating theoretical resolution is expressed by:

$$E_{FWHM} = 2.355 \sqrt{E e_i F}$$

where $E_{FWHM} = \text{full width at half maximum of the energy distribution (eV)}$, $E = \text{energy of incident X-rays (eV)}$, $e_i = \text{effective ionization potential (eV)}$, and $F = \text{Fano factor that depends on gas type}$. The Fano factor is required because it has been demonstrated that the resulting variance is only a fraction of what would be expected. This factor ranges from 0.5 to 0.22 for Ar/methane mixtures. The 2.355 factor in the formula relates 1 standard deviation (1σ) to the FWHM of the pulse-height distribution. For example, with $F = 0.22$, the Al-Kα energy distribution is 218 eV (14.7%) at FWHM.

It is usually better to not rely on theoretical expressions and simply measure the detector resolution, expressing the result as a percentage of the X-ray energy. The observed resolution is defined as the full width in energy units of the peak at half-maximum, $E_{FWHM}$, divided by the energy of the X-ray, $E$: 
\[ R(\%) = 100 \left( \frac{\Delta E_{FWHM}}{E} \right). \]

Once the resolution of a detector has been established for given energy the relative resolutions for other energies may be calculated using

\[ R_2 = R_1 \sqrt{\frac{E_1}{E_2}}, \]

where \( R_1 \) = observed resolution at \( E_1 \), \( E_2 \) = energy of interest, and \( R_2 \) = resolution at energy of interest.

### 7.8. Single Channel Analyzer

#### 7.8.1. Introduction

Recall that the amplitude of the pulse from gas-flows detectors is proportional to the energy of the incident X-ray photon. This permits the electronic circuitry to select for a specific X-ray photon by excluding anything but the desired amplitude. A single-channel analyzer (SCA) selects pulses of interest and outputs pulses suitable for the counting electronics.

The SCA can be set to reject pulses of lower energy than a certain threshold (baseline) and higher than the baseline plus a "window." Sometimes, these discrimination values are expressed as "lower" and "upper" energy limits, \( E_L \) and \( E_U \), rather than as a baseline and window. Selection of desired pulse energy is termed pulse-height analysis, and SCAs are commonly called "pulse height analyzers" (PHAs).

#### 7.8.2. Differential and Integral Mode

When both baseline and window are in use, the PHA is described as being in “differential” mode. The baseline is usually used to eliminate low-energy electronic noise. The upper limit (window) may be used to eliminate the interference from higher order peaks of other
undesired elements which also satisfy Bragg’s Law. In addition to the desired wavelength, an analyzing crystal will also diffract $2(\lambda/2)$, $3(\lambda/3)$, etc. However, only the desired wavelength has energy of $E$; the $2^\circ$ peak has energy of $2E$, the $3^\circ$ peak has energy of $3E$, etc. Setting the window (upper limit) to screen out these higher energies ensures that only photons with the correct intensities are counted.

As an alternative, the PHA can be set in “integral” mode with the window wide open to accept all pulses greater than a given baseline. The use of integral mode is recommended in most cases to limit the problems associated with a narrow window caused by drift in the detector electronics, and changes in P-10 gas pressure and room temperature. The significance of the high-order interference depends on how much of the interfering element is in the sample and the effective intensity of its X-ray line. For the elements most abundant in geological samples, high-order interferences are rare, because elements with X-ray lines twice (or three times, etc.) the energy of the element of interest are rare. Obviously, in trace element work, where the peaks of interest are very small, very careful evaluation of interferences is required.

One important example of interference is that of Ca-Kβ1 ($2^\circ$) on P-Kα1, which affects analysis of the mineral apatite. Fortunately, the effect is essentially uniform in all apatite and can be ignored if apatite unknowns are analyzed using apatite standards. Another important case is that of analyzing Hf-Lα1 in zircon, owing to an interfering Zr-Kα1 ($2^\circ$) peak; here, setting an upper limit is essential.

In summary, the operator user must consider three possibilities:

1. If the X-ray of interest has energy exceeding $E_c$ of Ar (2.96 kV), escape peaks can be a problem, but high-order diffractions are not. The baseline may be used to eliminate the escape peak and the PHA run in integral mode.
2. If the X-ray energy is less than $E_c$ of Ar and greater than about 30 times the ionization potential of Ar ($30 \times 26.4 \text{ eV} = 0.8 \text{ kV}$), there are neither escape peaks nor higher-order diffractions. Again, integral mode is appropriate.

3. If the X-ray energy is less than 0.8 kV, there are likely to be wavelength shifts and higher-order peaks can cause significant interferences. Differential mode, with carefully selected baselines and windows is appropriate.
Chapter 8 – Energy Dispersive Spectrometry (EDS)

8.1. Introduction

The semiconductor detectors used in energy dispersive spectrometry (EDS) are proportional detectors like the gas-flow detectors used in WDS analysis. However, in contrast to the gas detectors, a semiconductor has good energy discrimination by itself and can be used without preselecting energies using analyzing crystals. It might be worth reviewing the discussion of semiconductors in the Cathodoluminescence section.

Semiconductors detectors are made of single crystals of either Si or Ge. If the crystal structure were perfect, there would be no local abundances or shortages of electrons; however, all crystals have imperfections (lattice defects, impurities, etc.) that result in electron-deficient areas (termed "holes") and extra free electrons within the crystal lattice. These holes and free electrons will act as charge carriers when an electrical field is applied across the crystal. Thus, a pure crystal, with fewer holes and free electrons, allows less current to pass than an impure crystal.

X-ray analyzing instruments use Si semiconductor detectors; Ge detectors are generally used for γ-ray counting (e.g., in instrumental neutron activation analysis, INAA). Pure Si is an intrinsic semiconductor, providing an excellent material for a detector; however, even purest available Si contains some residual impurities such as B of Al, causing it to become a conductor by creating holes in the valence band of Si (B has fewer valence electrons than Si). These holes become charge carriers, allowing current to flow; the result is called a p-type semiconductor. Addition of Li or P to the Si crystal adds electrons to conduction band (more valence electrons than Si) and swamp out the effects of B impurities; this is termed an n-type semiconductor. Addition of trace impure such as Li to a crystal is called "doping." P-type and n-type semiconductors are also termed "extrinsic" semiconductors, because their semiconductive properties result from the introduction of impurities into the lattice.

![Types of semiconductors](http://britneyspears.ac/physics/basics/basics.htm)

Different types of semiconductors can be combined to yield useful electronic devices. For example, a p-n junction (also called a diode), will only allow current to flow in one direction. Placing a voltage across a diode causes the electrons in the n-type zone to move across the contact and towards the positive terminal and the holes in the p-type zone to move the
opposite way across the contact towards the negative terminal. When the electrons and holes reach each other, they recombine causing a current flow to through the diode. Although the electrons and holes flow in opposite directions, the current only flows one way because the charge carriers have opposite polarities.

**PN junction (diode).** A forward-bias (voltage) situation applies when the p-type silicon is connected to the positive terminal of a battery and the n-type silicon is connected to the negative terminal; a current will flow. With a reverse bias (p-type silicon is connected to the negative terminal of a battery, etc.), no current will flow. Image modified from images at: http://www.mpoweruk.com/semis.htm.

### 8.2. Detector Crystals

Detector crystals are doped with Li (an electron donor) in a process called "drifting." Li drifting makes Si a better semiconductor by swamping out the effects of impurities. Li-drifted crystals are called "silly" for Si(Li) and "jelly" for Ge(Li). A detector crystal consists of 2-5 mm thick Si crystal, with gold contacts on its ends. A bias is applied across the crystal causing a current to flow. The Si crystal consists of a "Li-drifted" intrinsic region facing the specimen (p-type) and an adjacent Li-free region (intrinsic). The front contact, Li-free region and Li-drifted intrinsic region form a p-i-n junction. The crystal is maintained at low temperatures to prevent diffusion of Li from the intrinsic region to the Li-free region. In general, the diffusion of Li is only a problem when there is a voltage across the detector.
Cross section of a typical lithium-drifted silicon detector. X-rays create electron-hole pairs in the intrinsic region of the semiconductor; these charge carriers then migrate to the electrodes under the influence of an applied bias voltage (After KVex Corporation 1983).

Immediately beneath the gold surface facing the sample is a "dead-zone" where the Li drifting is inadequate to deal with impurities. In this zone there is a large excess of holes, that trap charges produced by X-ray interaction in the crystal and produce a low energy "tail" on the side of the peak.

### 8.3. Detection Process

When an X-ray hits the detector crystal, it is absorbed and produces a high-energy photoelectron. The photoelectron knocks a valence band electron into the conduction band producing an electron-hole pair. This interaction uses up some of the energy of the photoelectron, and the slightly less energetic photoelectron continues to produce more electron-holes pairs until its energy is dissipated. On average for a Si crystal, 3.8 to 3.9 eV are dissipated per electron-hole pair created. Thus, X-rays with energy more than 1 kV make many holes. The bias placed across the detector crystal causes the electron-hole pairs to migrate and increases the conductivity of the crystal. The more holes created, the greater the increase in conductivity and the associated drop in resistance. The total charge conducted is directly proportional to the energy of the absorbed X-ray; unlike a gas-flow detector there is no internal gain. Baseline conductivity due to thermal excitation of electrons in the detector crystal produces a leakage current and increased signal noise. Consequently, the crystal is sheathed in a liquid-N₂-filled cryostat producing a temperature of about 77 K (-209 °C) to decrease thermal noise.

The X-ray detection process in the Si(Li) detector. The energy of incident X-ray is absorbed by production of Auger electrons and by electron-hole pairs. Incident X-rays may cause ionization in the Si of the detector. Escape peaks in a Si(Li) detector occur 1.84 kV lower in energy ($E_{e}$ for Si-Kα) than the incident X-rays and will be present for all parent peaks above 1.84 kV. However, the magnitudes of the escape peaks are usually just a few percent of the parent peak (after Goldstein et al. 1981).
8.4. Resolution

Although the number of electron-hole pairs created is a function of the energy of the X-ray, it is a statistical process and there is a normal distribution around the actual X-ray energy. Detection is a statistical process and resolution is treated as for gas-flow detectors. Fano factor for Si(Li) semiconductor detector ranges from 0.1 to 0.13. Resolution depends not only on the detector but also the electronics. The standard industry EDS resolution test is on Mn-Kα radiation (5.9 kV) at 1000 cps with an 8 msec time constant; resolutions are ~150 eV. WDS systems can handle a much higher count rates without loss of resolution than EDS systems. For WDS the limit is about 50000 cps, whereas EDS functions best at 2000 cps.

8.5. Signal Amplification and Processing

The pulses produced by the Si(Li) crystal require amplification before they can be shaped, analyzed and measured. The preamplifier is located physically near the detector crystal to minimize stray electronic noise. It incorporates a field-effect transistor (FET) that is used to reset the pulse amplification circuitry. Both the preamplifier and FET are cooled with liquid N₂ to decrease electronic noise. The resulting signal is sent to a main amplifier, which provides linear, low noise amplification of the preamplifier signal. Most importantly, the main amplified must be able to recover quickly after processing one pulse to be ready for the next, because pulses are produced rapidly.

At high count rates, a second pulse may reach the amplifier before the first has been fully processed and pulse pileup results. This results in a total pulse that is a combination of two pulses and, thus, is meaningless. Pulse pileup rejection circuitry, which assures that processing of one pulse is finished before accepting another, is used to avoid this. Unfortunately, the pulse pileup circuitry significantly increases the dead time.

**Pulse Pileup.** Failure to discriminate pulses 1 and 2 leads to an anomalously large pulse (1+2) being recorded (after KVex Corporation 1983).

A multichannel analyzer (MCA) processes the pulses from the main amplifier. Recall that the amplitude of the pulse is proportional to the energy of the incident X-ray photon. The MCA contains an analog-to-digital converter (ADC) and as each pulse is received it is converted into a
number. Large pulses are converted into larger numbers. The resulting ADC output is used as the address of a memory location and one is added to the value at that location. In effect, the memory acts as a set of independent counters, each covering a narrow range of energy. Counts are accumulated into these "bins" or "channels" for a preset "live" time (counting time corrected for dead time) and the results can then be displayed.

The EDS system on the Cameca MBX microprobe consists of a Quantum detector system connected to a 4Pi MCA board, which is located in the G3 computer. The MCA board and has a 1024-channel memory and can accumulate up to 16,777,215 (2^24-1) counts in each channel. Each channel is approximately 10 eV wide. The system can be calibrated using known energies to allow the use of a set of KLM marker lines. The INCA system on the JEOL-6480LV is similar.

8.6. Spectral Fitting

The resulting energy spectra can be used qualitatively or qualitatively. Commonly available software allows the analyst to rapidly identify the X-ray lines present in an accumulated spectrum. In general, elements with concentrations of about 0.5 wt. % produce distinct identifiable EDS peaks. Peak heights can be used as a rough proxy for the element abundances, although the effect of variable fluorescent yields makes this problematic. For example, the heights of the Si-Kα and Fe-Kα peaks in a fayalite spectrum have similar heights, but the concentration of Fe is about three times that of Si in this mineral. However, for a given spectra line, variation in peak height have an approximately linear correlation with concentration. Thus, if the Fe-Kα peak is twice as high in a mineral when compared with another mineral, the concentration of Fe is also about twice as high.

Quantification of the spectrum requires a procedure called "peak fitting," in which reference peaks are used in variable proportions to recreate the spectrum. The reference peaks are acquired on reference materials using identical operating conditions. Peaks sizes are adjusted to minimize the least-squares deviation of the modeled spectrum from the actual one. This process is sometimes termed "deconvolution."
**Peak fitting (deconvolution).** Acquired rare-earth element spectrum compared with the modeled spectra of all elements determined by deconvolution. Note how each element has numerous L peaks that overlap. Image source: [http://microanalyst.mikroanalytik.de/info5.phtml](http://microanalyst.mikroanalytik.de/info5.phtml).

The spectra used in the fit and the spectrum of the unknown should be collected long enough to eliminate the effects of noise; this is especially critical for the quantification of minor elements. Acquisition times of ~200 seconds yield generally excellent spectra. Perhaps the biggest problem with quantitative analysis is the presence of unresolved peak overlaps. Although many potential overlaps are trivial in geological materials, a few make absolute identification and quantification impossible. For example, there is an almost exact overlap between Pb-Mα and S-Kα, making analysis of Pb in sulfides impossible. Other geologically significant overlaps include:

- F-Kα and Fe-Lα
- Na-Kα and Zn-Lα
- Si-Kα and Sr-Lα
- P-Kα and Zr-Lα
- Cr-Kβ and Mn-Kα
- Ti-Kβ and V-Kα
- Fe-Kβ and Co-Kα
Chapter 9 – Detection of Non-X-ray Signals

9.1. Backscattered Electrons

9.1.1. Detection Process

Backscattered electrons are detected using PN junctions positioned at the base of the objective lens at a high take-off angle. Current flows when backscattered electrons strike the semiconductor chips imparting energy. The size of the current (signal) is proportional to the number of electrons hitting the PN junctions. The BSE detectors are sensitive to light and cannot be used when the sample illumination is on.

The BSE detector system for the MBX microprobe consists of four semiconductors (PN junctions) located at the bottom of the column at a 40° take-off angle. The JEOL-6480LV SEM has two semicircular plates at the base of the objective lens and one offset plate.

The spatial resolution of BSE images is poor (usually ~1 μm; at best ~0.1 μm) because BSE are produced from the entire upper half of the interaction volume; however, BSE provide valuable
information because of their sensitivity to atomic number variations. BSE images are displayed on the MAC 3 computer and acquired using the NIH imaging software.

Example MBX image. Backscattered electron image of exsolution of orthopyroxene lamellae in large iron-rich meteoritic pigeonitic pyroxene.

Example JEOL-6480LV image. Backscattered electron image of a chondrule from an ordinary chondrite meteorite with large, zoned olivine crystals and glassy matrix with tiny quench crystals. The strong iron enrichment in the rims of the olivine is easily observed in the backscattered image.

9.1.2. Signal Types

The BSE signal reflects two parameters: specimen composition (average Z) and specimen topography. The JEOL-6480LV uses paired detectors to combine signals to emphasize these two signal components. In TOPO mode the signals are processed to emphasize topography (TOPO),
in COMPO mode to emphasize composition. The SEM also has a shadow mode, which further enhances topographic effects. The quadripole BSE detectors of the MBX microprobe were designed to allow the instrument to act as a scanning electron microscope; the signals are simply summed when used as a microprobe.


9.2. Secondary Electrons

9.2.1. Detection Process

Secondary electrons emitted from the specimen are detected using a scintillator-photomultiplier, "Everhart-Thornley," detector. Low energy secondary electrons are emitted from the sample in all directions. They are gathered by a charged collector grid (or cage), which can be biased from -50 to +300 V. This draws the secondary electrons towards the scintillator. The scintillator is a thin plastic disk coated with a short-persistence phosphor that is highly efficient at converting the energy contained in the electrons into ultraviolet light photons (4000 Å). The response time of the phosphor is fast and permits high resolution scanning. The outer layer of the scintillator is coated with a thin layer [10-50 nm] of aluminum, positively biased at approximately 10 kV, which accelerates the electrons to the scintillator surface. The charged collector grid, in addition to collecting secondary electrons from the sample, helps to alleviate some of the negative effects of the scintillator aluminum layer bias, which can actually distort the incident beam.
Secondary Electron Detection. Low-energy secondary electrons (trajectories shown by dashed lines) are collected by applying a suitable bias to the Faraday grid. These electrons are further accelerated in order to give them sufficient energy to scintillate the phosphor (after Potts 1987).

The aluminum layer also acts as a mirror to reflect the photons produced in the phosphor layer down the light pipe, which consists of a Plexiglas or polished quartz pipe, and out through the specimen chamber wall. The photoelectrons strike a photocathode that converts the UV photons back into electrons. The electron multiplier consists of a series of dynodes, each held at a more positive voltage than the previous one. The tube interior is a vacuum. Electrons are accelerated towards the first dynode, strike it, and emit more low energy electrons. These in turn are accelerated towards the second dynode, and so forth. The dynodes are coated with a material with high SE yield so at least two (sometimes as many as ten) secondary electrons are emitted for each photoelectron. The result is a cascade of electrons that eventually strike the anode; a single photon produces \( \sim 10^6 \) final electrons.

The number of cascade electrons produced by the PMT depends on the voltage applied across the cathode and anode of the PMT, in a manner analogous to how a gas-flow X-ray detector works. One can increase the gain by increasing the voltage to the PMT (this is essentially what is accomplished when adjusting the contrast). The amplified electrical signal is sent to further electrical amplifiers, which increase the electrical signal thus increasing brightness.


The topographical aspects of a secondary electron image depend on how many of electrons actually reach the detector. When the incident electron beam intersects the edges of topographically high portions of a sample at lower angles, it puts more energy into the volume of secondary electron production. Thus, high points produce more secondary electrons, generating a larger signal. Faces oriented towards the detector also generate more secondary
electrons. Secondary electrons that are prevented from reaching the detector do not contribute to the final image and these areas will appear as shadows or darker in contrast than those regions that have a clear electron path to the detector. It should be noted that those backscattered electrons directed at the scintillator will also contribute to the signal that reaches the scintillator and form part of the secondary electron image.

9.3. Cathodoluminescence

9.3.1. Detection Process

A photomultiplier tube is also used to detect the cathodoluminescence signal, but rather than converting electrons into photoelectrons before amplification, photons are amplified directly. The CL signal contains two types of information: intensity and wavelength; however, most CL detectors respond only to the former, combining photons of all wavelengths to generate an intensity signal.

Information in CL Signal. (left) The actual CL signal produced from a sample includes both color and intensity information. Colors reflect the nature of the activating elements. (right) The same image processed to remove the color information. Although different phases can be distinguished by intensity variations, but the activator information is gone. Original image source: http://serc.carleton.edu/images/research_education/geochemsheets/cathodoluminescence_2.jpg.

Some CL detectors can also access the wavelength information. This is accomplished by using a diffraction grating (monochromator) to select a particular wavelength in a manner directly analogous to that of using X-ray analyzing crystals, separating the wavelength of interest from other wavelengths emitted from the sample. A grating is a reflective surface, scored either mechanically or holographically with parallel grooves.
Design of a typical monochromator. A monochromator consists of the diffraction grating (dispersing element), slits, and spherical mirrors. Note how the blue light is diffracted into the exit slits, whereas other wavelengths (such as red) are not. Image source: http://weather.nmsu.edu/Teaching_Material/SOIL698/Student_Reports/Spectroscopy/monochromator_files/image002.gif.

The resulting intensities are different wavelengths of light yield a spectrum that can be deconvolved into individual peaks using a computer.

Example of multiple Gaussian curves fit to CL spectrum. Note how the bumps in the spectrum result from the addition of several peaks with variable intensities and widths. Image source: http://www.gatan.com/images/monocl3_gaussiancurve.gif.

The Gatan CL unit attached to the JEOL SEM in the Geology Department detects only intensity variations. It must be inserted manually into the sample chamber after carefully checking that the stage and sample are not in the way. Once inserted the sample is brought as close as possible to the light pipe and tilted to face it by approximately 20° to direct more photons into the light pipe. In addition, there is a paraboloidal collecting mirror attached to the end of the pipe to capture even more phonons. The resulting phonon signal is amplified using a photomultiplier tube controlled by the CL controller box. Increasing voltage on the tube increases the overall intensity of the resulting image and contrast enhances variations in signal intensity.
Gatan MiniCL Detector. (left) Gatan MiniCL detector showing end of light pipe and paraboloidal collecting mirror. (right) View outside of the sample chamber. The detector electronics are housed in the cylinder with the logo.

9.3.2. Application

Very few minerals have intrinsic CL; scheelite (CaWO$_4$) is one of them. In contrast, minerals with extrinsic CL are fairly common, among them quartz, feldspar, and carbonates. The impurities that cause luminescence are called **activators** and include Mn and the rare earth elements (REE). Other elements, called **quenchers**, can absorb CL before it escapes a material. Iron is perhaps the most important quencher.

Activators and quenchers have been studied extensively in feldspar, carbonates, and zircon. Activators in carbonates include Mn$^{2+}$, Sm$^{3+}$, Tb$^{3+}$, Eu$^{2+}$ and Eu$^{3+}$ at concentrations of more than 10-20 parts per million (ppm). The transition elements, Fe$^{2+}$, Fe$^{3+}$, Ni$^{2+}$ and Co$^{2+}$, act as quenchers at concentrations of over 30-35 ppm. Much of the documented variation in cathodoluminescence in calcite involves Mn$^{2+}$ and Fe$^{2+}$ with the latter being the quencher. Documented activators in feldspar include Ce$^{3+}$ (bluish-green, peak at 490 nm), Ti$^{3+}$ (blue, peak at 460 nm), Eu$^{2+}$ (blue, peak at 420 nm), Mn$^{2+}$ (greenish-yellow, peak at 540-560 nm), and Fe$^{3+}$ (red/infrared, peak at 68-780 nm); all must be present in low concentrations. Blue CL emission has also been linked to Al-O-Al defects in the feldspar structure. Zircon, which commonly displays complex CL zoning, incorporates U, Th, heavy REE, Y, P and Hf into its structure. Studies of synthetic crystals have demonstrated a correlation between REE content (especially Dy$^{3+}$) and CL. However, in zircons from the Carpathian Mountains, CL brightness does not correlate with concentrations of REE, Y, or Th; the authors propose U is probably responsible for quenching or suppressing the CL signals.
Cathodoluminescence in calcite, dolomite, and pyroxene. (left) Image in crossed-polarized light XPL; (right) Cathodoluminescent image. Note the distinct color difference in CL between dolomite (dark red) and calcite (orange, yellow). Images source: http://www.lumic.de/gallery/index.html.

Cathodoluminescence in quartz. This image shows total polychromatic CL intensity as gray scale intensity. The quartz shows both sector and concentric zoning in hydrothermal vein quartz, possibly due to Al3+ variation. Image source: http://www.see.leeds.ac.uk/research/igt/people/lloyd/eg_cl.htm.

Cathodoluminescence in zircon. This image shows polychromatic CL intensity as gray scale intensity. The zircon shows complex zoning. Image source: http://cda.morris.umn.edu/~jonesjv/mineralogy/zircon.jpg.

Cathodoluminescence can be used for determining modal abundances in cases where minerals display different luminescence (as in the calcite-dolomite sample above; however, CL is not always present or characteristic. CL also has application in discriminating synthetic from natural
gems. It is particularly valuable in documenting the distribution of trace elements in minerals such as carbonates and zircon, which reflect chemical changes during overgrowth and dissolution. Spectral analysis of CL allows the identification of specific trace elements. For example, the CL zoning in zircon reflects variations in Dy concentration.
Chapter 10 – Statistics

10.1 Populations & Samples

Before proceeding with a discussion of basic statistics, two terms must be defined. A population consists of a complete set (either finite or infinite) of measurements of the variable of interest. For example, measurements of the heights of all the students in a class, or their grades, would constitute a population of measurements. In contrast, a sample consists of a set of random measurements representative of the population. Each observation must have an equal chance of being chosen. Heights of randomly select people in a busy airport terminal would constitute a sample of the population. The number of X-rays reaching a detector in a second similarly would represent a sample, because not every X-ray produced from the material is counted. The distinction between sample and population vanishes as the number of measurements grows large.

10.2. Sample Distributions

10.2.1. Poisson Distribution

The Poisson distribution applies to a wide range of phenomena in the sciences. It describes the probabilities inherent when an event occurs with a constant probability per unit time. Geological examples include radioactivity (number of unstable nuclei that decay within a given period) and X-ray production (number of X-rays counted in each period).

In 1831, Simon Denis Poisson (1781-1840) determined the expected distribution of the number detected in various statistically independent time intervals. This distribution assumes that:

1. The events occur independently (their occurrence does not depend upon previous or following events,
2. The probability of an event occurring is proportional to the length of time since the last event, and
3. Two events cannot occur simultaneously.

The resulting probability, $P$, of a event during a time interval is

$$P(a) = e^{-m} \frac{m^a}{a!},$$

where $m = $ average number of random occurrences per interval. As the average number of random occurrences per interval, $am$, increases, the probability distribution moves to the right and broadens. For each $m$ value, sum of all probabilities = 1.0. It should be emphasized that the Poisson distribution is not a continuous function, but rather gives the probabilities for discrete values of $a$. 
Poisson distribution. Notice how the distribution for \( m = 10 \) resembles a "bell-shaped-curve." The connecting lines are for convenience, probabilities only apply at the integer intervals shown by symbols.

10.2.2. Gaussian Distribution

For larger values of \( m \), usually 30 or more, the Poisson distribution approaches the normal or Gaussian distribution.

Probability distributions. The Poisson distribution for \( m = 10 \) is almost identical to the Gaussian distribution. Note that the curve width at half-maximum height is the square-root of \( m \).

For populations of measurements that have a Gaussian distribution, we expect that about 68.3% of the data will be within 1 standard deviation of the mean (i.e., in the range \( \bar{x} \pm s \)). A more complete description of the distribution of data in a Gaussian distribution is given in the table.
Thus, we should expect that 95% of the data would be within 1.96 standard deviations of the mean (i.e., in the range $X_{\text{avg}} \pm 1.96 \sigma$). This is called a 95% confidence interval for the sample.

Gaussian distribution. From Figure 3.13 in Till, 1974, *Statistical Methods for the Earth Scientist: an introduction.*

### 10.3. Central Tendency

When successive measurements of the same quantity are repeated, a distribution of values is obtained. There are several measures of "central tendency," including the median, mode and mean. We will use the **mean**, which for a sample is

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

where $x_i =$ value of the $i^{\text{th}}$ measurement and $n =$ total number of measurements. The formula here uses "sigma" notation to indicate that a series of numbers are added together:
$$\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots x_n.$$ 

Note that for a population, the symbol \( \mu \) is used to denote the mean.

### 10.4. Variability

We also wish to characterize populations of measurements in terms of their variability. Measures of variability include range, standard deviation, the coefficient of variation, variance, and standard error. The most useful for our work is the **standard deviation**, which for a sample is:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}.$$ 

Note that because we are dealing with a sample of the total population (the usual situation), the weighting factor is \((n-1)\). If we had data for the entire population, this factor would be \(n\):

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

Note that for a population, the symbol \( \sigma \) is used to denote the standard deviation.

Results are usually reported as **mean ± standard deviation**. In other words, we are saying that, assuming a Gaussian distribution of values, there is a probability of 68% that the measured value is within the range defined by the standard deviation. The standard deviation is related to the other descriptions of variation. The **variance** is \( \sigma^2 \) and the **coefficient of variation**, \( \varepsilon \), is:

$$\varepsilon = \frac{s}{\bar{x}}.$$ 

The coefficient of variation is also called the **fractional standard deviation, relative error** or **relative standard deviation** and is often expressed as a percentage.

### 10.5. Accuracy & Precision

#### 10.5.1. Accuracy

Accuracy is defined as how close a measured value is to the "true" value, consequently, it has an absolute description. It is very difficult to establish "true" values. Determining the accuracy of a measurement requires calibration of the analytical method with known standard material. Accuracy is affected if the compositions of the reference standards are not well known. The geostandards community constantly works to provide well-characterized reference materials. In addition to having well known compositions, microprobe standards must also be homogeneous, having the same composition throughout. There are a large number of minerals, glasses, and metal available as microprobe standards.
10.5.2. Precision

Precision refers to how well a given measurement or result can be reproduced and has a relative description. Usually this takes the form of a standard deviation around a mean value. Values can be very precisely determined and still be very inaccurate; conversely, imprecise analyses may average to very accurate values (albeit with a large standard deviation). Precision effectively is limited by counting statistics when dealing with X-ray analysis as described below.

10.5.3. Factors Affecting Accuracy and Precision

Many factors introduce errors to affect the accuracy and precision of analyses obtained by electron microprobe analyses. It is useful to subdivide these errors into two categories: systematic and random. Systematic errors introduce a constant bias into the results. Unlike random errors, which can be reduced by repeated measurements, systematic errors cannot be detected by statistical means. Their general effect is to shift the measured quantity away from the "true" value. Factors that can affect probe accuracy and precision include the following (S = systematic error, A = accuracy, R = random error, P = precision, → = affects):

- The random nature of X-ray generation and emission (R → P)
- Long-term Instrumental drift (S → A, P)
- Short-term filament instability (R → P)
- Specimen surface irregularities (R → P)
- Focusing inconsistency (R or S → A, P)
- Interaction volume intersecting two phases or secondary fluorescence from phases below surface (R → P)
- Sample damage such as sodium- or volatile-loss (S → A, P)
- Incorrect standard composition values (S → A)
- Errors in matrix factors (S → A)
- Incorrect system parameters such as take-off angle (S → A)
- Sample tilt (S → A)
- Variations in C-coat thickness relative to standard materials (S → A)
- Sample charging (R → P)
- Incorrectly located background measurements (S → A)
- Peak wavelength shifts (S → A)
As can be seen in general systematic error affect accuracy, whereas random errors affect precision.

10.6. Propagation of Errors

10.6.1. Methodology

It is occasionally necessary to propagate errors to determine the total error that results from performing mathematical operations using several numbers that have associated errors. The discussion below is purely practical; a very complete discussion of error propagation is located at: http://www.itl.nist.gov/div898/handbook/mpc/section5/mpc55.htm. At high count rates, X-ray production follows a Gaussian distribution. The combination of two Gaussian distributions, with standard deviations of $s_1$ and $s_2$, is another Gaussian distribution with a standard deviation, $s$:

$$s^2 = s_1^2 + s_2^2.$$

This extremely useful property allows us to "propagate" the errors on individual measurements to determine a total error. The appropriate formulae are given in the table.

<table>
<thead>
<tr>
<th>Function, $Z(A, B)$</th>
<th>Error Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z = A + B$</td>
<td>$s_Z = \sqrt{s_A^2 + s_B^2}$</td>
</tr>
<tr>
<td>$Z = A - B$</td>
<td></td>
</tr>
<tr>
<td>$Z = A \times B$</td>
<td>$\epsilon_Z = \sqrt{\epsilon_A^2 + \epsilon_B^2}$</td>
</tr>
<tr>
<td>$Z = A / B$</td>
<td>$s_Z = \frac{s_A}{A^2} = Z^2 s_A$</td>
</tr>
<tr>
<td>$Z = 1/A$ (assuming no uncertainty in B)</td>
<td>$\epsilon_Z = B s_A$</td>
</tr>
<tr>
<td>$Z = \exp(A) = e^A$</td>
<td>$s_Z = e^A s_A = Z s_A$</td>
</tr>
<tr>
<td>$Z = \ln(A)$</td>
<td>$s_Z = \frac{s_A}{A}$</td>
</tr>
<tr>
<td>$Z = 10^A$</td>
<td>$s_Z = s_A 10^A \ln 10 = Z s_A \ln 10$</td>
</tr>
<tr>
<td>$Z = \log(A)$</td>
<td>$s_Z = \frac{s_A}{A \ln 10}$</td>
</tr>
</tbody>
</table>

10.6.2. Example

It is often necessary to combine error expressions to calculate the total error. Consider the following expression with the errors given as $1s$ values:

$$[(56 \pm 7) + (23 \pm 5)] \times (12 \pm 3).$$

First, we can ignore the errors and determine the result:

$$[56 + 23] \times 12 = 948.$$
Now let's find the error on this value. First combine the errors in the addition portion of the expression:

\[ s = \sqrt{7^2 + 5^2} = 8.6. \]

Next, we must combine this error on the numerator with the error associated with the denominator. We must combine coefficients of variation because multiplication is involved:

\[ \varepsilon = \sqrt{\left(\frac{8.6}{79}\right)^2 + \left(\frac{3}{12}\right)^2} = \sqrt{0.1089^2 + 0.25^2} = \sqrt{0.0744} = 0.2727. \]

This is the coefficient of variation of the result (948); we must convert it back into a standard deviation:

\[ s = 948 \times 0.2727 = 258.5 \]

So, the final answer is 948 ± 258.5 (1s). Note that although the absolute value of the error of the numerator is larger (8.6 vs. 3), most of the error on the result (>90%) comes from the large error on the denominator (25%).

### 10.7. Significant Digits

Numbers (analytical results, ratios, etc.) should only be reported to the proper number of significant digits, because the number of significant digits and associated error are indications of the precision of the analytical results. Correct handling of significant digits (and error) and retention of the available precision requires an understanding of the propagation of significance in calculations.

Generally, if not specified, the precision may be assumed to be ±1 in the last reported digit, which is termed the least significant digit. However, some values effectively have an infinite number of significant digits. For instance, 1 inch is defined as exactly 2.54 centimeters (2.54 with an infinite number of zeros following) and each value is infinitely precise for purposes of conversion. In addition, for practical purposes, many constants (speed of light, Planck's constant, etc.) are comparably precise and do not limit the precision of the results of calculations involving them.

The number of significant figures is defined as the quantity of digits in the number excluding leading or trailing zeros. For example, 3.142 has 4 significant figures; 23,459,000 has 5 significant figures; 0.31910 has 4 significant figures (the last zero does not count); and 0.0004086 has 4 significant figures (the zero between 4 and 8 is not a leading or trailing zero and so is counted). Trailing zeros are a main source of confusion, but use of scientific notation allows the writer to indicate the precision by only showing significant figures. Consider the number 2000 (which when written this way has 1 significant digit). The best way to indicate the number of significant digits is to use scientific notation:

\[ 2 \times 10^3, \text{ 1 significant digit} \]
\[ 2.0 \times 10^3, \text{ 2 significant digits} \]
\[ 2.00 \times 10^3, \text{ 3 significant digits} \]
One should retain all digits when performing calculations and when finished, round the result to the appropriate number of significant digits. For addition and subtraction, the result should have the same number of significant digits as the least precise number in the calculation. For example,

$$14.72 + 1.4331 + 0.00235 = 16.16.$$ 

In contrast, theoretically the only way to determine the correct number of significant digits for the results of calculations involving multiplication and division is to propagate significance as one would propagate error. Thus, the precision of the result cannot be better than the square root of the sum of the squares of the relative errors. For example, a measurement of 52.3 has an implied error of ±0.1, corresponding to a relative error of 0.0019. If we wished to square this value, the relative error of the result is

$$\epsilon = 0.00268.$$ 

Now, $52.3^2 = 2735.29$, so the relative error corresponds to an absolute error of $2735.29 \times 0.00268 = 7.3$. The limit on precision is thus 7.3 (rounding to 1 in the tens place), and the result should be presented as $2.74 \times 10^3$. In practice this procedure is cumbersome, and usually, unnecessary. Note that the result has the same number of significant digits as the two numbers, which were multiplied.

Generally, one can simply follow the rule of rounding the result to the same number of significant figures as the least precise quantity used in the calculation. As examples, the limiting factors for some common calculations are given below:

<table>
<thead>
<tr>
<th>Calculation</th>
<th>Limiting Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Weight</td>
<td>smallest number of significant digits in atomic weights used</td>
</tr>
<tr>
<td>Weight Fraction</td>
<td>and smallest number of significant digits in the conversion factor</td>
</tr>
<tr>
<td>Energy from Wavelength</td>
<td>or known quantity</td>
</tr>
<tr>
<td>Wavelength from Energy</td>
<td></td>
</tr>
<tr>
<td>Spectrometer Position</td>
<td>smallest number of significant digits in the 2d values or spectrometer position</td>
</tr>
<tr>
<td>Mass Absorption Coefficient</td>
<td>smallest number of significant digits in the elemental absorption coefficients</td>
</tr>
</tbody>
</table>

Finally, when it is necessary to reduce the number of digits in a result this should be accomplished by rounding. If the number after the last significant digit is greater than 5, one should round the final digit up; if less, round down. If the digit is exactly 5, round up if the digit preceding it is odd (and down if it is even) to average out the effects of rounding.

### 10.8. Statistics of X-ray Counting

#### 10.8.1. Error on Count Ratio

The production of X-rays is a Poisson process and can be analyzed using statistical methods. As noted above, at sufficient high values of $m$ (count rates), the Poisson distribution is identical to the normal distribution, so
\[ \sigma_{\text{theory}} = \sqrt{C} \]

Where \( \sigma_{\text{theory}} \) = theoretical population standard deviation and \( C \) = total number of counts. In quantitative analysis, one determines the total number of counts on the peak of interest and corrects them for the background of the continuum. Corrected peak counts on the unknown are divided by corrected counts on a known standard material to determine the "K" ratio, which is used as input for the data reduction routine. The combined uncertainty on the K ratio due solely to X-ray counting statistics is

\[
\varepsilon = \left( \frac{(R_p + R_b)}{(R_p - R_b)^2} \right)^{\text{unknown}} + \left( \frac{(R_p + R_b)}{(R_p - R_b)^2} \right)^{\text{standard}}
\]

where \( R_b \) = background rate (cps), \( R_p \) = peak rate (cps), \( T_b \) = background counting time, and \( T_p \) = peak counting time. Derivation of the equation is left as an exercise to the student.

Consider a case where the standard peak is 100 cps, the backgrounds on standard and unknown are 2 cps, the unknown peak is 10 cps, and the count times are the same (\( T_b = T_p = 15 \) sec). The error on the ratio is 11.5%. If we count twice as long, (\( T_b = T_p = 30 \) sec), \( \varepsilon \) decreases to 8.2%. Most of the error comes from the relatively low number of counts on the unknown; if they were doubled (\( R_b = 50 \) cps), \( \varepsilon \) would only be 3.3%. Note also that decreasing \( T_b \) for the standard to 1 second (assuming all other initial assumed values above) would only increase the total error on the K ratio to 11.6%. This is because at high values of \( R_p \), the background becomes relatively unimportant.

10.8.2. Optimum Counting Times

To minimize the uncertainties due to the counting statistics, the times spent measuring the peak and the associated background should satisfy following relationship:

\[
\frac{T_p}{T_b} = \sqrt{\frac{R_p}{R_b}}
\]

10.8.3. Detection Limit

If peak counts are much greater than background counts, the background measurement hardly matters. By choosing standard materials with high concentrations of the element of interest, it is not necessary to count its background very long. However, if peak count rates are similar to background count rates (as is the case with trace elements), it is essential to spend as much time establishing the background rate as counting on the peak.

10.8.3.1. Derivation

When considering trace elements, we need to determine the "detection limit." In essence, we are confronted with the signal to noise problem we've encountered before. Detection limits are a function of the background counts in unknown (controlling what is considered as detected)
and the peak counts in the standard (allowing conversion of counts into concentration). Recall that the corrected peak counts, $C_{corr}$, and standard deviations associated with it are:

$$C_{corr} = C_p - C_b$$
$$\sigma_p = \sqrt{C_p}$$
$$\sigma_b = \sqrt{C_b}$$

Using these relationships, the combined standard deviation associated with $C_{corr}$ is:

$$\sigma_{corr} = \sqrt{\sigma_p^2 + \sigma_b^2} = \sqrt{C_p + C_b}.$$  

For detection, the corrected X-ray peak height must be greater than error associated with it by some factor, $z$:

$$\left( C_p - C_b \right) > z \sigma_{corr}.$$  

If the population standard deviation $\sigma$ is known, confidence limits about a single result may be calculated. The coefficient $z$ is the limiting value of the t-distribution function $(t_{p,\infty})$ for $\nu = \infty$ confidence level $1-\alpha$:

<table>
<thead>
<tr>
<th>Confidence Interval</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.9% ($\alpha = 0.001$)</td>
<td>3.090</td>
</tr>
<tr>
<td>99% ($\alpha = 0.01$)</td>
<td>2.326</td>
</tr>
<tr>
<td>95% ($\alpha = 0.05$)</td>
<td>1.645</td>
</tr>
<tr>
<td>90% ($\alpha = 0.1$)</td>
<td>1.282</td>
</tr>
</tbody>
</table>

For example, for a 99% confidence interval ($\alpha = 0.01$), the value of $z = 2.326$. The user should select a confidence interval that meets the situation. In electron microprobe analysis, the 99% confidence is more than sufficient. Near the detection limit, $C_p \approx C_b$, so we may specify the detection limit (D.L.) as

$$D.L. = z\sqrt{2C_b}.$$  

We can refine this expression by recognizing that counting for a longer time improves our counting statistics. Substituting

$$C_b = R_b T_b$$

into the equation, yields

$$D.L. = z\sqrt{2R_b T_b}.$$  

So far, we have defined the detection limit in terms of counts (or time and count rate), but we are interested in absolute concentrations. At such small concentrations, it is not necessary to make matrix corrections, so we may use a simple conversion factor, $K$, between counts per second and concentration. The factor $K$ has units of counts per time per concentration (cps/wt. % element or cps/wt. % oxide). However, we are working with absolute numbers of counts so
the K value (measured on the standard) must in turn be converted into counts per wt. % by multiplying by time. Rewriting the detection limit expression to include K yields:

\[
D.L. = \frac{\sqrt{2R_b T_b}}{KT_b} = \frac{z\sqrt{2R_b}}{K\sqrt{T_b}}.
\]

All values except time, \(T_b\), are constants (with \(z\) selected by the analyst). The detection limit decreases as the square-root of time; thus, in order to halve it, \(T_b\) must be increased by a factor of 4!

### 10.8.3.2. Example

As an example, consider the detection of Ni in olivine. The peak count rate on a standard that contains 30 wt. % Ni is 1460 cps. The background rate on the olivine is 3 cps. The value of K is 48.67 cps/wt. % Ni. If we assume a count time of 30 seconds and a confidence level of 95% (\(z = 1.645\)), the detection limit is

\[
D.L. = \frac{1.645\sqrt{2\times3}}{48.67\sqrt{30}} = 0.015 \text{ wt. % Ni}.
\]

To convert this number into parts per million (ppm), recall that 100 wt. % = 1,000,000 ppm; thus, 1 wt. % = 10,000 ppm. Thus, 0.015 wt. % Ni = 150 ppm Ni.

### 10.8.3.3. Caveats

Measuring trace elements requires careful determination of the background count rate on the unknown. It is usually wisest to slowly scan the spectrometer over the range of interest to characterize the shape and slope of the continuum. Analysts should be wary of non-linear backgrounds, peak interference(s) at the background locations, and for the presence of absorption edges, which put a step in the background.

**Background complications.** Correct determination of the background when measuring the concentrations of trace elements is critical. Usually, it is assumed that the true background can be determined by measuring the background...
away from the peak and interpolating (case a). However, potential complications arise if there are interferences from high order peaks (case b), nearby absorption edges (case c), or non-linear backgrounds (case d).

One particularly good way to determine the background count rate is to use a standard that does not contain the element of interest and actually count X-rays at the peak position. Detection limits may be improved by increasing counting times and beam current to yield higher total counts. In addition, changing the accelerating voltage to optimize the overvoltage for the line of interest will improve detection limits.

10.8.4. Sample Homogeneity

The precision due to counting statistics is the same for one long count as for several shorter ones done for the same total time. The latter method is useful for determining whether counting statistics are the limit to attainable precision. If counting statistics is the principal source of X-ray count variation for a homogeneous sample, then comparisons of observed and calculated parameters can be used to evaluate sample homogeneity. Consider analyses of a number of different grains (or different points on one grain). Are the observed variations the result of counting statistics or do they demonstrate true compositional variations? Define "homogeneity index," H.L., as the ratio of the standard deviations (in microprobe analysis this ratio is also called the "sigma ratio"):

\[ H.L. = \frac{\text{Observed}}{\text{Theory}}. \]

We can use a statistical procedure called the F-test, which assumes that all data were sampled on a random basis and are normally distributed, to test for homogeneity. We will use the F-test to evaluate the hypothesis that there is no difference between the two variances, \( \sigma_1^2 \) and \( \sigma_2^2 \) (as is the case for homogeneity); however, because we are not stating which variance is larger, we must use a "two-tailed" test and double the probabilities below (5% becomes 10%, etc.). Because F-values are greater than one, we must place the larger value in the numerator.

The F ratio compares variance rather than standard deviations, so we square the homogeneity index to get the F ratio. Recall that in a Gaussian distribution, \( \sigma_{\text{Theory}} \) is the square root of the counts; thus

\[ F = \frac{s^2}{c}. \]

If \( F = 1 \), the observed variance is same as the variation expected just from X-ray counting statistics and we can argue the analyses are of a homogeneous sample. However, the F test allows us to assign probabilities to the likelihood of homogeneity. Tables of F-ratios and probabilities of equal variance (homogeneity) are available in many sources, but it is simplest to use the FDIST function available in EXCEL to determine the probability of homogeneity. This function requires the parameters of:

- F-ratio
- Number of degrees of freedom for the numerator of the F-ratio (\( n - 1 \)), and
- Number of degrees of freedom of the denominator (taken to be very large or infinite; enter "1E7" for \( 1 \times 10^7 \)).
As an example, consider a set of three X-ray counts on a standard, with a sigma ratio (H.I.) of 1.1. The FDIST function tells us that there is a 29.8% probability of homogeneity. The FDIST function number is “single-sided” and must be doubled to reflect the total probability, in this case to 59.6%. Other example values are given below.

<table>
<thead>
<tr>
<th>H.I.</th>
<th>F</th>
<th>Degrees of Freedom (n-1)</th>
<th>Probability of Homogeneity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>1.21</td>
<td>2</td>
<td>59.6</td>
</tr>
<tr>
<td>1.1</td>
<td>1.21</td>
<td>3</td>
<td>60.9</td>
</tr>
<tr>
<td>1.5</td>
<td>2.25</td>
<td>2</td>
<td>21.1</td>
</tr>
<tr>
<td>1.5</td>
<td>2.25</td>
<td>3</td>
<td>16.1</td>
</tr>
<tr>
<td>1.5</td>
<td>2.25</td>
<td>4</td>
<td>12.2</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3.7</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>2</td>
<td>0.02</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>3</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Rigorous statistical tests of homogeneity have not been applied much in the geological literature, but H.I. is often reported. A value of for H.I. greater than 3 has been cited as proof of sample inhomogeneity. However, as shown, if counting statistics were the only cause of variation, then much smaller values of H.I. effectively would demonstrate inhomogeneity.

**Sample Homogeneity.** Note that as the degrees of freedom (number of data points) increase, homogeneous samples are permitted to display smaller amounts of variability as indicated by H.I.

Since counting statistics are only one source of X-ray count variability, the F test must be applied with care. Often, correlated variations (e.g., Ca decreases when Na increases) can reveal inhomogeneity.
Chapter 11 – Quantification

11.1. Introduction

The X-ray signals measured from an unknown must be converted into concentration data. Element concentration is not directly proportional to the corrected intensity ratio of an element’s X-ray line because the ratio is affected by X-ray absorption, secondary fluorescence, electron backscattering, and the electron stopping power of the sample. There are several methods of taking these factors into account and determining concentrations from X-ray counts:

- Calibration curves, which use an empirically determined relationship between counts and concentration;
- Empirically based matrix corrections (Bence-Albee); and
- ZAF (or PAP), which uses fundamental factors to correct for the effects of atomic number, absorption and fluorescence.

These methods are summarized below. Those wishing more detail should examine one of the references.

11.2. Calibration Curves

Calibration curves may be constructed to relate X-ray counts and element concentration. However, these curves require a large number of well-characterized standards with compositions that bracket the unknowns. In wet chemical analysis, such as atomic absorption spectroscopy (AAS), it is relatively easy to make solutions with the appropriate concentrations. In contrast, production of standards for microprobe analysis is far more difficult, requiring apparatus to produce homogeneous glasses. In both cases, however, the data used to construct a curve must be taken at identical operating conditions (take-off angle, accelerating voltage and beam current) as will be used during analysis.

Calibration curve for carbon in nickel steel. Redrafted from Figure 8.26 in Goldstein et al., 1981, Scanning Electron Microscopy and X-ray Microanalysis.
For geological applications, calibration curves must be constructed for each mineral group; however, because the unknowns approximate the standards in compositions, matrix corrections are unnecessary. The requirement of a great number of standards is especially difficult to satisfy especially with the accuracy required. In addition, this technique does not allow confident analysis of a truly unknown material.

Calibration curve of Mg-Kα counts per second vs. weight percent MgO for chemically analyzed biotite, chlorite, staurolite, chloritoid, and garnet. The scatter in the data results from absorption (or fluorescence) due to other oxides other than MgO in the samples. Redrafted after Bence and Albee, 1968.

11.3. Empirical Matrix Corrections

11.3.1. Introduction

Empirical correction schemes are based upon empirical determinations of inter-element effects, which are then applied to unknown materials. These empirical factors must be determined for all operating conditions of interest (take-off angle and \( E_0 \)). As with the calibration curves the standards must be suitable for microprobe analysis, but many fewer are required.

In 1963-4, Ziebold and Ogilvie developed corrections for some binary metal alloys using an equation of the form

\[
\frac{(1-K_A)}{K_A} = \alpha_{AB} \frac{(1-C_A)}{C_A},
\]

where \( C_A \) = concentration (fraction) of element A in the mixture AB, \( \alpha_{AB} \) = factor describing matrix effects, and \( K_A \) = background-corrected intensity ratio of A counts in mixture AB relative to those in pure element A. This equation can be rearranged as:

\[
\frac{C_A}{K_A} = \alpha_{AB} + (1 - \alpha_{AB})C_A.
\]

A plot of \( C_A/K_A \) against \( C_A \) yields a straight line with slope \( (1 - \alpha_{AB}) \); thus, X-ray counts collected on a set of variable mixtures allows one to determine \( \alpha_{AB} \). Although this hyperbolic relationship
between $C_A$ and $K_A$ was demonstrated for several alloy and oxide systems, it was difficult to find appropriate intermediate compositions for many binary systems. Ziebold and Ogilvie showed that $\alpha$ corrections also could be developed for some ternary metal alloys. It should be emphasized that the resulting $\alpha$ factors are only appropriate for a particular accelerating voltage and take-off angle.

![Graph showing intensity of Fe-Kα in different binary mixtures.](image)

**Intensity of Fe-Kα x-rays in Different Binary Mixtures.** The intensity of Fe-Kα is affected by the nature of the coexisting element. The presence of Ni causes enhanced Fe-Kα x-ray intensity due to fluorescence by Ni-Kα. Conversely, Cr absorbs Fe-Kα x-rays and reduced their total intensity. The presence of Mn in the mixture has almost no effect. The nature and extent of binary interactions can be determined by measuring x-ray intensity for a range of binary mixtures.

### 11.3.2. Bence-Albee Corrections

The most commonly used empirical method of matrix correction in the 1970s was developed by T. Bence and A. Albee; it is still used. The Bence-Albee correction method utilizes 'alpha factors' ($\alpha$) to describe the effects of atomic number, absorption, and fluorescence. The $\alpha$ factors are established for binary systems as shown in the above figure. For cases where the presence of a second element causes the absorption of X-rays from the element in interest, the factor $\alpha > 1$. Where fluorescence of the second element by the x-rays from the element of interest occurs, the factor $\alpha < 1$. The ideal case, where there are no inter-element effects, has $\alpha = 1$. 
Form of the calibration curves for C vs. K and different values of \( \alpha \). Mixtures exhibiting absorption will have upward concave curves with \( \alpha \) factors > 1. Those where fluorescence dominates will have concave downward curves with \( \alpha \) factors < 1 (redrafted after Bence and Albee 1968).

Quantifying the shape of the curve is difficult, but a linear relation exists between \( \frac{C_A}{K_A} \) and \( C_A \) and the value of \( \alpha \) can be determined by finding the intercept (where \( C_A \) goes to zero).

Plot of \( \frac{C_A}{K_A} \) vs. \( C_A \) for the curves shown above. A linear relationship results when the ratio of \( C_A \) to \( K_A \) is plotted against \( C_A \). (redrafted after Bence and Albee 1968).

For the binary system AB, we may write:

\[
C^A_{AB} = \alpha^A_{AB}K^A_{AB}.
\]

There are theoretical grounds for the approximate validity of this correction, but it is unreasonable to expect such a simple equation to be exact in light of the complexity of electron scattering processes. Bence and Albee (1968) applied this correction scheme to silicates and oxides (in which absorption effects dominate and fluorescence effects are generally minor).
using oxide rather than element concentrations. For a multicomponent system, the concentration, \( C \), of element A in the unknown can be calculated as

\[
C^A_{\text{unk}} = \beta^A_{\text{unk}} K^A_{\text{unk}}
\]

where \( K^A_{\text{unk}} \) = background-corrected intensity ratio of the counts on the unknown relative to the counts on a pure element standard, and

\[
\beta^A_{\text{unk}} = \left( \frac{K^A_{\text{unk}} a^A_{\text{unk}} + K^B_{\text{unk}} a^B_{\text{unk}}}{K^A_{\text{unk}} + K^B_{\text{unk}}} \right).
\]

### 11.3.3. Example Bence-Albee Correction

Unknown concentrations are determined using an iterative process. Initially, \( K \) values are inserted into the equation for \( \beta \) in the place of concentrations to find starting concentrations. These new concentrations are reinserted to find new \( \beta \) values and the iteration is repeated until convergence is achieved. Two or three iterations usually result in values very close to the theoretical composition.

As an example, we will calculate the concentration of an "unknown" synthetic grossularite garnet, \( \text{Ca}_3\text{Al}_2\text{Si}_3\text{O}_{12} \). The standards that we will use are:

<table>
<thead>
<tr>
<th>Standard</th>
<th>Oxide Weight Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \text{Al}_2\text{O}_3 )</td>
</tr>
<tr>
<td>Corundum</td>
<td>1.000</td>
</tr>
<tr>
<td>Quartz</td>
<td>-</td>
</tr>
<tr>
<td>Wollastonite</td>
<td>-</td>
</tr>
</tbody>
</table>

The appropriate \( \alpha \) values from Albee and Ray (1970) are:

<table>
<thead>
<tr>
<th>X-ray</th>
<th>( \alpha ) Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \text{Al}_2\text{O}_3 )</td>
</tr>
<tr>
<td>Al-K( \alpha )</td>
<td>1.00</td>
</tr>
<tr>
<td>Si-K( \alpha )</td>
<td>1.54</td>
</tr>
<tr>
<td>Ca-K( \alpha )</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Note that the \( \alpha \) values all exceed 1.00, indicating that absorption dominates. Calculating the \( \beta \)s for the standards yields:

<table>
<thead>
<tr>
<th>Standard</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corundum</td>
<td>( \beta\text{-Al} = 1.00 \times 1.000 = 1.000 )</td>
</tr>
<tr>
<td>Quartz</td>
<td>( \beta\text{-Si} = 1.00 \times 1.000 = 1.000 )</td>
</tr>
<tr>
<td>Wollastonite</td>
<td>( \beta\text{-Ca} = 1.10 \times 0.517 + 1.00 \times 0.483 = 1.052 )</td>
</tr>
</tbody>
</table>

Measured x-ray counts corrected for background and dead time are:

<table>
<thead>
<tr>
<th></th>
<th>Unknown</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-K( \alpha )</td>
<td>29146</td>
<td>140126</td>
</tr>
<tr>
<td>Si-K( \alpha )</td>
<td>19764</td>
<td>56471</td>
</tr>
<tr>
<td>Ca-K( \alpha )</td>
<td>26971</td>
<td>35045</td>
</tr>
</tbody>
</table>
First, we make an initial approximation of the composition of the unknown by calculating initial $K$ factors:

$$K = \frac{w_{unk}}{w_{std} \beta_{std}}.$$ 

So,

- $K_{\text{Al}_2\text{O}_3} = \frac{29146/140126 \times (1.000 / 1.000)}{0.2080} = 0.2080$
- $K_{\text{SiO}_2} = \frac{19764/56471 \times (1.000 / 1.000)}{0.3500} = 0.3500$
- $K_{\text{CaO}} = \frac{26971/35045 \times (0.483 / 1.052)}{0.3533} = 0.3533$

These $K$-factors total to 0.9113. Next, we make an approximation of the $\beta$ values using these $K$ values and the $\alpha$ factors above. For this initial approximation, the $K$ values are used as estimates of the concentrations:

- $\beta_{\text{Al}_2\text{O}_3} = \frac{1.00 \times 0.2080 + 1.05 \times 0.3500 + 1.18 \times 0.3533}{0.9113} = 1.089$
- $\beta_{\text{SiO}_2} = \frac{1.54 \times 0.2080 + 1.00 \times 0.3500 + 1.07 \times 0.3533}{0.9113} = 1.150$
- $\beta_{\text{CaO}} = \frac{1.08 \times 0.2080 + 1.10 \times 0.3500 + 1.00 \times 0.3533}{0.9113} = 1.057$

Now, we can use these $\beta$s and the $K$ values to make a better approximation of the actual concentrations:

- $\text{Al}_2\text{O}_3 = 1.089 \times 0.2080 = 0.2265$
- $\text{SiO}_2 = 1.150 \times 0.3500 = 0.4026$
- $\text{CaO} = 1.057 \times 0.3533 = 0.3733$

The new total is 1.0025. This can be improved to some degree by iteration. We refine the estimates of the $\beta$ values using the concentrations just calculated:

- $\beta_{\text{Al}_2\text{O}_3} = \frac{1.00 \times 0.2265 + 1.05 \times 0.4026 + 1.18 \times 0.3733}{1.0025} = 1.087$
- $\beta_{\text{SiO}_2} = \frac{1.54 \times 0.2265 + 1.00 \times 0.4026 + 1.07 \times 0.3733}{1.0025} = 1.148$
- $\beta_{\text{CaO}} = \frac{1.08 \times 0.2265 + 1.10 \times 0.4026 + 1.00 \times 0.3733}{1.0025} = 1.058$

None of the $\beta$ values changes by much! Using the refined $\beta$s to get new concentrations:

- $\text{Al}_2\text{O}_3 = 1.087 \times 0.2080 = 0.2261$
- $\text{SiO}_2 = 1.148 \times 0.3500 = 0.4018$
- $\text{CaO} = 1.058 \times 0.3533 = 0.3738$

Our new total is 1.0018. Continued iteration might improve these numbers a bit, but let’s compare them with the ideal composition of grossularite:

<table>
<thead>
<tr>
<th></th>
<th>Actual</th>
<th>Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Al}_2\text{O}_3$</td>
<td>0.2263</td>
<td>0.2261</td>
</tr>
<tr>
<td>$\text{SiO}_2$</td>
<td>0.4002</td>
<td>0.4018</td>
</tr>
<tr>
<td>$\text{CaO}$</td>
<td>0.3735</td>
<td>0.3738</td>
</tr>
</tbody>
</table>

Not bad...

Standards with compositions similar to the unknowns are often used with the Bence-Albee method to minimize the differences in absorption between the standards and the unknowns. (In geological materials, the effect of fluorescence is very small and only absorption is
significant.) Use of the Bence-Albee reduction scheme was very common in the 1960-70s, because the calculations required very little computing power and memory, but now that very powerful computers are available (and inexpensive), this is no longer an advantage. Today, the more complicated ZAF correction scheme is usually employed.

11.4. ZAF & PAP Corrections

11.4.1 Introduction

The ZAF correction scheme is based on first principles and provides data reduction for all operating conditions. If there were no inter-element (matrix) effects, the corrected intensity ratio could be converted to concentration using the formula called Castaing’s 1st Approximation,

\[ C_{\text{unk}}^A = C_{\text{std}}^A \left( \frac{I_{\text{unk}}^A}{I_{\text{std}}^A} \right), \]

where \( C_{\text{unk}}^A \) = concentration of A in the unknown, \( C_{\text{std}}^A \) = concentration of A in the standard, \( I_{\text{unk}}^A \) = background corrected intensity of A x-rays in the unknown, and \( I_{\text{std}}^A \) = background corrected intensity of A x-rays in the standard. This is equivalent to the case of ideality in Bence-Albee (\( \alpha = 1 \)). However, as we noted, matrix effects due to absorption and fluorescence of X-rays within the sample and atomic number effects are significant. The matrix corrections to be applied to the estimated initial composition from this approximation can be expressed in the form:

\[ C^A = C_{\text{unk}}^A \left( \frac{F_{\text{unk}}^A}{F_{\text{std}}^A} \right) \]

where \( F_{\text{unk}}^A \) and \( F_{\text{std}}^A \) are matrix or ZAF factors for unknown and standard respectively. The derivation of these factors is described below. Since the factors are dependent on composition, which is initially unknown, ZAF works in an iterative fashion similar to that of Bence-Albee. The intensity data are used with Castaing’s 1st Approximation to get initial elemental concentrations that are refined in subsequent iterations until they converge. ZAF is not very good for elements with X-ray energies less than 1 kV because of a lack of knowledge of the factors discussed below. For these elements it is best to use a standard of similar composition to minimize matrix effects. For example, when analyzing F in apatite, use a fluorapatite standard rather than a fluorphlogopite standard.

11.4.2. Atomic Number Correction (Z)

The atomic number effect controls the amount of incident energy from the electron beam that is actually put into the sample. This effect consists of two components: backscattering and electron-stopping power (or retardation). Both are a function of average Z and, to a lesser degree, the accelerating voltage.

Backscattered electrons leave the sample carrying energy without producing X-rays. The fraction of electrons backscattered from the sample, \( n_b \), ranges from about 0.12 for Al to 0.55
for U. At lower Z, more electrons stay within the sample to produce X-rays. The backscatter correction factor, $F_b$, reflects the X-ray intensity lost due to backscattering and is expressed as a fraction ($r$) relative to the intensity that would be produced with no backscattering:

$$F_b = \frac{1}{1-r}.$$  

This fraction in part depends on the overvoltage, but not to a significant degree. The factor rarely exceeds 1.2 for Z less than 30.

The stopping-power correction, $S$, relates the amount of incident energy to the amount absorbed by atoms of specific Z. The electron-capture cross-section decreases with increasing Z. Thus, light atoms have a higher ratio of atomic number to atomic weight ($Z/A$) and interact with disproportionately larger numbers of incident electrons. The characteristic intensity per unit concentration of a heavy element is less when combined with a light element than in a sample of the pure element. $F_s$ for a given element is:

$$F_s = \frac{S_{\text{sample}}}{S_{\text{pure element}}}.$$  

$S$ may be calculated using

$$S = k \left( \frac{Z}{A} \right) \ln \left( \frac{E}{J} \right),$$  

where $A$ = atomic weight, $Z$ = atomic number, $J$ = mean ionization potential, $k$ = constant, and

$$E = \frac{(E_0 + E_c)}{2}.$$  

where $E_0$ = accelerating voltage and $E_c$ = critical excitation potential. As a simplification, the mean ionization potential may be approximated by $11.5 Z$.

Values for $S$ must be calculated for each element separately because each has a different $E_c$. The average value of $S$ for the unknown is calculated in a manner analogous to the mass absorption coefficient, using the mass concentration fractions, $C_i$, of the elements:

$$S = \sum_{i=1}^{n} C_i S_i.$$  

The net Z correction factor for both effects is: $F_b \times F_s$. The effects of backscattering and stopping power are opposing and can be canceling. For example, in the analysis of Fe₂O₃, with pure iron as a standard, the stopping power of oxygen reduces the intensity of Fe-Kα from the oxide relative to that from the pure standard. However, the lower mean atomic number of Fe₂O₃ (15.2) compared with pure iron (Z = 26) results in a smaller backscattering loss for the oxide when compared with the metal. The net effect is that the intensity of Fe-Kα is 7.2% lower (at 15 kV) than calculated on a basis of the relative concentrations in the two materials.

**11.4.3. Absorption Correction (A)**

Since X-rays are generated below the surface of the sample, the emergent radiation undergoes absorption prior to detection. The absorption correction is a function of the take-off angle (length of path traversed by the X-rays), the distribution of X-ray generation, the wavelength of
the emergent X-ray and the elements present. As the take-off angle increases, the intensity of characteristic radiation decreases due to an increase in path length. Less energetic X-rays are more easily absorbed. Absorption can also be strongly affected by surface irregularities -- a good sample polish is thus critical. Most X-rays are generated at relatively shallow depths within the excitation volume and relatively close to the beam axis because it is in this region that electron energies are greatly attenuated by ionization or electron scattering. Several models have been used to describe the depth distribution of X-ray generation termed $\phi(p,Z)$, "phi-rho-zee".

![Diagram](image)

**The forms of the depth and lateral generated intensity functions $\phi(p,Z)$ and $\psi(y)$**. The vertical dimension in this figure is depth beneath the sample surface, expressed as mass thickness, $\rho z$. The horizontal dimension is lateral distance from the axis of the electron beam in arbitrary units (after Williams 1987).

Approximations are commonly used to represent the emission volume because of its complex shape. The simplest is that of Bishop, which assumed a puck-like body of uniform thickness. Most improvements in quantification have resulted from improvements of the $\phi(p,z)$ relationship (for example that of Philibert).

![Diagram](image)

**The Bishop rectangular approximation of $\phi(p,Z)$, compared to experimental data and Philibert's analytical approximation**. The area of the Bishop rectangle is assumed to approximate to that under the $\phi(p,Z)$ curve, and the continuous variation in $\phi(p,Z)$ is replaced by the concept of a 'mean mass depth' at half the depth of the Bishop
rectangle. This greatly simplifies the calculations required to correct for the effects of variable compositions on the depth generation of analytical X-rays (after Williams 1987).

Once the shape of the emission volume is described, absorption can be quantified using Lambert's Law. The result is an expression that accounts for the effects of take-off angle, accelerating voltage, composition, etc. One such formulation is the Philibert-Duncumb-Heinrick equation. Using this expression, the absorption factor for samples examined by microprobe analysis, $F_a$, is expressed:

$$F = \left(1 - \frac{\chi}{\sigma}\right)\left(1 + \frac{h}{(1+h) \left[\frac{\chi}{\sigma}\right]}\right),$$

where

$$\chi = \mu \csc \varphi,$$

$$h = 1.2 \frac{A}{Z^2},$$

$$\sigma = \frac{4.5 \times 10^5}{(E_0^{1.65} - E_c^{1.65})},$$

and $Z = \text{atomic number}$, $A = \text{atomic weight}$, $m = \text{mass absorption coefficient}$, $\varphi = \text{take-off angle}$, $E_o = \text{accelerating voltage}$, and $E_c = \text{critical ionization potential}$. For compounds, a mean value of $h$ is used:

$$\bar{h} = \sum_{i=1}^{n} C_i h_i,$$

where $c_i = \text{mass concentration of element } i$. The $\sigma$ factor above accounts for the voltage dependence of absorption of primary electrons. Errors in the calculation of $F_a$ can be decreased by using a high take-off angle to minimize $\chi$ and low overvoltages to maximize $\sigma$.

### 11.4.4. Fluorescence Correction (F)

The correction for fluorescence is a function of the elements present, their concentrations, their values of $E_c$ and mass absorption coefficients, and the take-off angle, $\varphi$. The most important factors are the concentrations of fluorescing elements and elements fluoresced. In general, the correction for fluorescence is the least important factor in the ZAF correction. The fluorescent yield decreases with decreasing $Z$ and is not important for the light elements, which dominate geological samples. The factor, $F_f$, can be expressed as:

$$F_f = \frac{1}{(1+0.5C_i \omega_i)},$$

where $C_i = \text{concentration of the fluorescing element}$, and $\omega_i = \text{the fluorescent yield of that element}$. The fluorescent yield increases rapidly with increasing atomic number and $F_f$ is negligible for K-lines of elements below atomic number 20. In silicates and oxides, absorption dominates, and fluorescent enhancement is rarely greater than a few percent.

### 11.5. PAP Model
PAP is a general model for calculating X-ray intensities in contrast to ZAF, which is conceived as a matrix correction procedure. PAP can be used to calculate X-ray intensities for a wide range of energies (100 eV to greater than 10 kV) and accelerating voltages (1-40 kV). The improvement over ZAF results from using better expressions for the Z effects (backscattering factor, retardation of the electrons, and effective ionization cross-section) and a better determination of the absorption effects. The latter is the most significant of these improvements and uses a modified version of the $\psi(\rho,Z)$ polynomial used in standard ZAF correction scheme to better fit the experimentally determined X-ray distribution. Although PAP corrections take slightly longer to calculate than ZAF corrections, the improvement is well worth it.

![Graph](image)

Depth distribution of Mg-Kα in aluminum at 25 kV. Comparison of theoretical functions with experimental results of Castaing and Henoc (after Pouchou & Pichoir, 1984).

Chapter 12 – Mineral Recalculation

12.1. Methodology

The calculation of mineral chemistry from probe analyses entails the conversion of oxide data to moles of cations and oxygen and then renormalization on the basis of a given number of cations or oxygens. In analyses by electron microprobe the oxidation state of Fe is assumed to be 2+ (FeO).

To determine the mole of cations present, one divides the measured oxide weight percents by the cation-equivalent molecular weights given in the table below. These numbers take into account the fact that some oxides contain two or more cations. Moles of oxygen are calculated assuming that all cations are bonded to oxygen. Thus, we may simply multiply the moles of cations by the oxygen factor to get moles of oxygen.
The results of the molar calculations must next be recalculated. This is usually done by assuming a known number of oxygen atoms in the mineral formula. One can make an attempt at determining the oxidation state of Fe, by using cation formulation, however, the validity of the results depend upon the overall quality of the analysis. Minor errors in the determination of SiO$_2$, for example, can produce large changes in the calculated Fe$_2$O$_3$. This will be covered in more detail below. Appropriate normalization numbers using oxygen (O) and/or cations for some minerals are:

<table>
<thead>
<tr>
<th>Mineral</th>
<th>Normalization</th>
<th>Mineral</th>
<th>Normalization</th>
<th>Mineral</th>
<th>Normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphibole</td>
<td>$O = 23$</td>
<td>Andalusite</td>
<td>$O = 20$</td>
<td>Apatite</td>
<td>$O = 25$</td>
</tr>
<tr>
<td>Beryl</td>
<td>$O = 36$</td>
<td>Biotite</td>
<td>$O = 11$</td>
<td>Carbonates</td>
<td>$O = 6$</td>
</tr>
<tr>
<td>Chlorite</td>
<td>$O = 14$</td>
<td>Chloritoid</td>
<td>$O = 14$</td>
<td>Cordierite</td>
<td>$O = 18$</td>
</tr>
<tr>
<td>Epidote</td>
<td>$O = 12.5$</td>
<td>Feldspar</td>
<td>$O = 8$</td>
<td>Garnet</td>
<td>$O = 12$</td>
</tr>
<tr>
<td>Kyanite</td>
<td>$O = 5$</td>
<td>Muscovite</td>
<td>$O = 11$</td>
<td>Mullite</td>
<td>$\Sigma$cation = 6</td>
</tr>
<tr>
<td>Olivine</td>
<td>$O = 4$</td>
<td>Pyroxene</td>
<td>$O = 6$, or $\Sigma$cation = 4</td>
<td>Sillimanite</td>
<td>$O = 5$</td>
</tr>
<tr>
<td>Sphene</td>
<td>Si = 4</td>
<td>Spinel</td>
<td>$O = 4$ or $\Sigma$cation = 3</td>
<td>Staurolite</td>
<td>$O = 48$</td>
</tr>
<tr>
<td>Topaz</td>
<td>$O = 20$</td>
<td>Vesuvianite</td>
<td>$\Sigma$cation = 50</td>
<td>Wollastonite</td>
<td>$O = 18$</td>
</tr>
<tr>
<td>Zircon</td>
<td>$O = 16$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For example, recalculation of feldspar analyses is done on an 8-oxygen basis; thus, the conversion factor is 8 divided by the total moles of oxygen. All the cation (or oxygen) values are multiplied by this conversion factor to yield normalized molar amounts. In addition, if the concentrations of F and/or Cl have been determined, their abundances are divided by atomic weight (18.9984 and 35.453, respectively) and multiplied by the same conversion factor.

Using the normalized cation amounts, we finally may fill the cation sites in the minerals. Cations are assigned in the order shown in the table below for the structural formulae. For example, for amphibole, one first fills the T site, then the M123 site, then the M4 site, and finally, the A site. Site assignments are a subject of research and debate, and you may wish to use a different site filling scheme.
For analyses with F and/or Cl, we subtract the F and Cl values from the number for the OH-site to determine the OH content of the site. We can calculate the H2O content of a mineral by multiplying the OH value by 9.0076 to get moles of H in the OH-site and dividing by the conversion factor. This number is added to analytical original total. The oxygen equivalents of F and Cl are calculated by multiplying moles of F by 0.4211 and moles of Cl by 0.2256, and the oxygen-equivalents for the F and Cl are subtracted from the analytical total.

### 12.2. Oxygen Normalization Examples

#### 12.2.1. Phlogopite mica

As an example of the recalculation procedure, consider the following microprobe data for a phlogopite (biotite) from the volcanic rocks of the Hickey Formation of central Arizona. Note...
that more digits than are significant are retained throughout the calculations and rounding only occurs at the end.

<table>
<thead>
<tr>
<th>Oxide</th>
<th>Oxide Wt. %</th>
<th>Single Cation Mol. Wt.</th>
<th>Cation Moles</th>
<th>Oxygen Factor</th>
<th>Oxygen Moles</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO₂</td>
<td>37.7</td>
<td>60.0848</td>
<td>0.6275</td>
<td>2</td>
<td>1.2549</td>
</tr>
<tr>
<td>TiO₂</td>
<td>2.26</td>
<td>79.8988</td>
<td>0.0283</td>
<td>2</td>
<td>0.0566</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>14.45</td>
<td>50.9806</td>
<td>0.2834</td>
<td>1.5</td>
<td>0.4252</td>
</tr>
<tr>
<td>Cr₂O₃</td>
<td>0.60</td>
<td>75.9951</td>
<td>0.0079</td>
<td>1.5</td>
<td>0.0118</td>
</tr>
<tr>
<td>FeO</td>
<td>10.1</td>
<td>71.8464</td>
<td>0.1406</td>
<td>1</td>
<td>0.1406</td>
</tr>
<tr>
<td>MnO</td>
<td>0.08</td>
<td>70.9374</td>
<td>0.0011</td>
<td>1</td>
<td>0.0011</td>
</tr>
<tr>
<td>MgO</td>
<td>20.4</td>
<td>40.3044</td>
<td>0.5061</td>
<td>1</td>
<td>0.5061</td>
</tr>
<tr>
<td>BaO</td>
<td>0.50</td>
<td>153.3394</td>
<td>0.0033</td>
<td>1</td>
<td>0.0033</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.71</td>
<td>30.9895</td>
<td>0.0229</td>
<td>0.5</td>
<td>0.0115</td>
</tr>
<tr>
<td>K₂O</td>
<td>9.16</td>
<td>47.0977</td>
<td>0.1945</td>
<td>0.5</td>
<td>0.0972</td>
</tr>
<tr>
<td>F</td>
<td>3.02</td>
<td>47.0977</td>
<td>0.1945</td>
<td>0.5</td>
<td>0.0972</td>
</tr>
<tr>
<td>Totals</td>
<td>98.9</td>
<td></td>
<td></td>
<td></td>
<td>2.5083</td>
</tr>
</tbody>
</table>

We want to recalculate atom concentrations on an 11-oxygen basis (10 O²⁻ + 2 OH⁻ represent the same total negative charge as 11 O²⁻); the factor needed to convert 2.5083 into 11 is:

\[ \text{C.F.} = \frac{11}{2.5083} = 4.3854. \]

Next, we multiply the cation values by this conversion factor:

<table>
<thead>
<tr>
<th>Cation</th>
<th>Normalized Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>2.752</td>
</tr>
<tr>
<td>Ti</td>
<td>0.124</td>
</tr>
<tr>
<td>Al</td>
<td>1.243</td>
</tr>
<tr>
<td>Cr</td>
<td>0.035</td>
</tr>
<tr>
<td>Fe</td>
<td>0.617</td>
</tr>
<tr>
<td>Mn</td>
<td>0.005</td>
</tr>
<tr>
<td>Mg</td>
<td>2.220</td>
</tr>
<tr>
<td>Ba</td>
<td>0.014</td>
</tr>
<tr>
<td>Na</td>
<td>0.100</td>
</tr>
<tr>
<td>K</td>
<td>0.853</td>
</tr>
</tbody>
</table>

Now we can fill the cation sites. First, we fill the T-site with Si and enough Al to bring its sum to 4.00. If the amount of Si exceeds 4.00 or if Si and Al are insufficient to fill the T site (within a 2% error), the analysis is bad. We get 3.994.

\[ (\text{Si}_{2.752}\text{Al}_{1.243})^T \]

Next, we fill the M-site with the remaining Al, Ti and Fe and then add Cr, Mg, and Mn. This should bring the total to 3.00. The acceptable range is about 2.94 to 3.06 (±2% error); the data sum to 2.965.
Finally, we fill the A-site with Na, K, Ca and Ba. The A-site should total to near 1.00; the data yield 0.968.

\[(\text{Na}_{0.100}\text{K}_{0.853}\text{Ba}_{0.014})^\text{A}\]

We can now calculate the occupancy of the (OH) site. First, we divide F by its atomic weight (18.9984) and multiply by the conversion factor. Subtracting the result from 2.00 determines the OH occupancy in the OH-site:

\[
F = \frac{3.02}{18.9984} \times 4.3854 = 0.697
\]

\[
OH = 2.000 - 0.697 = 1.303
\]

So, the site occupancy is

\[(F_{0.697}\text{Cl}_{0.000}\text{OH}_{1.303})^\text{OH} \]

We also can back-calculate the H2O content of the mica by multiplying the OH value by 9.0076 to get moles of H and dividing by the conversion factor:

\[
H_2O = \frac{1.303 \times 9.0076}{4.3854} = 2.68 \text{ wt. %}
\]

Finally, let’s add this to the total above and subtract the oxygen-equivalents for the F:

\[
(\text{OH} = F) = 3.02 \times 0.4211 = 1.27
\]

\[
\text{Total} = 98.9 + 2.68 - 1.27 = 100.3
\]

The site occupancies should be reported to no more than 3 significant digits:

\[(\text{Na}_{0.10}\text{K}_{0.85}\text{Ba}_{0.01})^\text{A}(\text{Ti}_{0.12}\text{Fe}_{0.62}\text{Cr}_{0.03}\text{Mg}_{2.22}\text{Mn}_{0.01})^\text{M}(\text{Si}_{2.75}\text{Al}_{1.24})^\text{T}(\text{F}_{0.70}\text{OH}_{1.30})^\text{OH} \]

### 12.2.2. Alkali Feldspar

As another example, consider an alkali feldspar (#1 in Table 37, The Rock-Forming Minerals, 2nd ed., Deer, Howie & Zussman, 1992). This is a wet-chemical analysis, and we will ignore the water, because it reflects the fact that the analysis was done on a mineral separate. Additionally, we will ignore the elements not normally found in a feldspar structure (TiO2, MgO) because these probably result from tiny inclusions in the feldspar grains analyzed. Here is a summary table for the calculation:

<table>
<thead>
<tr>
<th>Oxide</th>
<th>Oxide Wt. %</th>
<th>Single Cation Mol. Wt.</th>
<th>Cation Moles</th>
<th>Oxygen Factor</th>
<th>Oxygen Moles</th>
<th>Normalized Cations</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO2</td>
<td>65.76</td>
<td>60.0848</td>
<td>1.0945</td>
<td>2</td>
<td>2.1889</td>
<td>2.934</td>
</tr>
<tr>
<td>Al2O3</td>
<td>20.23</td>
<td>50.9806</td>
<td>0.3968</td>
<td>1.5</td>
<td>0.5952</td>
<td>1.064</td>
</tr>
<tr>
<td>Fe2O3</td>
<td>1.19</td>
<td>79.8441</td>
<td>0.0023</td>
<td>1.5</td>
<td>0.0034</td>
<td>0.006</td>
</tr>
<tr>
<td>CaO</td>
<td>1.19</td>
<td>56.0774</td>
<td>0.0212</td>
<td>1</td>
<td>0.0212</td>
<td>0.057</td>
</tr>
<tr>
<td>BaO</td>
<td>0.63</td>
<td>153.3264</td>
<td>0.0041</td>
<td>1</td>
<td>0.0041</td>
<td>0.011</td>
</tr>
<tr>
<td>Na2O</td>
<td>8.44</td>
<td>30.9895</td>
<td>0.2724</td>
<td>0.5</td>
<td>0.1362</td>
<td>0.730</td>
</tr>
<tr>
<td>K2O</td>
<td>3.29</td>
<td>47.0977</td>
<td>0.0699</td>
<td>0.5</td>
<td>0.0349</td>
<td>0.187</td>
</tr>
<tr>
<td>Totals</td>
<td>99.72</td>
<td></td>
<td></td>
<td>2.9840</td>
<td>4.990</td>
<td></td>
</tr>
</tbody>
</table>
The correction factor is: 8 / 2.9840 = 2.6810. The corrected cations are given in the last column. The resulting formula is

\[(Na_{0.73}K_{0.19}Ca_{0.06}Ba_{0.01})^4(Si_{2.93}Al_{1.06}Fe_{0.01})^8O_8.\]

There are several internal checks we can make:

1. The A site should total 1.00 ... it's 0.99
2. The T site should total 4.00 ... it's 4.00
3. Because of charge balance constraints, Al should equal (1.0 + Ba + Ca) = 1.07 ... it's 1.06
4. For the same reason, Si should equal (2.0 + Na + K) = 2.92 ... it's 2.92.

We can conclude that this is a very good feldspar analysis (why would I pick a bad one?).

12.3. Cation Normalization Example

12.3.1. Augite Pyroxene

In some cases, it may be possible to use cation normalization to determine the valence state of Fe; however, this will only work if the analysis is very good. As an example, we will use an augite microprobe standard, a material that was analyzed very carefully by wet chemical methods. We will combine FeO and Fe₂O₃ from the wet chemical analysis (2.66 and 5.70 wt. %, respectively) as would be the case with a microprobe analysis into total FeO. Using this analysis, let's see if charge balance constraints and cation normalization will yield the same Fe²⁺/Fe³⁺ ratio as the wet chemical analysis. Our "microprobe" analysis of this material is given in the table below.

<table>
<thead>
<tr>
<th>Oxide</th>
<th>Oxide Wt. %</th>
<th>Single Cation Mol. Wt.</th>
<th>Cation Moles</th>
<th>Oxygen Factor</th>
<th>Oxygen Moles</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO₂</td>
<td>52.7</td>
<td>60.0848</td>
<td>0.8772</td>
<td>2</td>
<td>1.7543</td>
</tr>
<tr>
<td>TiO₂</td>
<td>0.34</td>
<td>79.8658</td>
<td>0.0043</td>
<td>2</td>
<td>0.0085</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>1.84</td>
<td>50.9806</td>
<td>0.0361</td>
<td>1.5</td>
<td>0.0541</td>
</tr>
<tr>
<td>FeO</td>
<td>7.33</td>
<td>71.8444</td>
<td>0.1020</td>
<td>1</td>
<td>0.1020</td>
</tr>
<tr>
<td>MnO</td>
<td>0.16</td>
<td>70.9375</td>
<td>0.0023</td>
<td>1</td>
<td>0.0023</td>
</tr>
<tr>
<td>MgO</td>
<td>15.15</td>
<td>40.3044</td>
<td>0.3759</td>
<td>1</td>
<td>0.3759</td>
</tr>
<tr>
<td>CaO</td>
<td>21.58</td>
<td>56.0774</td>
<td>0.3848</td>
<td>1</td>
<td>0.3848</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.49</td>
<td>30.9895</td>
<td>0.0158</td>
<td>0.5</td>
<td>0.0079</td>
</tr>
<tr>
<td>Totals</td>
<td>99.59</td>
<td>1.7982</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this case, we'll normalize the cations to a total of 4.00. Our correction factor is 4 / 1.7982 = 2.2244. We will multiply the cation values and the oxygen values by this factor:

<table>
<thead>
<tr>
<th>Cation</th>
<th>Normalized Cations</th>
<th>Normalized Oxygens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>1.9510</td>
<td>3.9021</td>
</tr>
<tr>
<td>Ti</td>
<td>0.0009</td>
<td>0.0189</td>
</tr>
<tr>
<td>Al</td>
<td>0.0803</td>
<td>0.1204</td>
</tr>
<tr>
<td>Fe</td>
<td>0.2269</td>
<td>0.2269</td>
</tr>
<tr>
<td>Mn</td>
<td>0.0050</td>
<td>0.0050</td>
</tr>
</tbody>
</table>
Notice that the oxygen total is not 6 as would be expected from the pyroxene formula. Some oxygen is missing, indicating that some of the Fe is in the 3+ state. We can estimate how much by converting Fe$^{2+}$ into Fe$^{3+}$:

$$2\text{FeO} + \text{O} = \text{Fe}_2\text{O}_3$$

Notice that for every two Fe$^{2+}$ cations we convert into Fe$^{3+}$ we get to add one oxygen. We need to add 0.0170 more oxygens (6 - 5.9830) to bring the total to six; thus, we need to convert twice as much Fe$^{2+}$ into Fe$^{3+}$. This results in the following changes:

$$\text{Fe}^{3+} = 2 \times 0.0170 = 0.0340$$
$$\text{Fe}^{2+} = 0.2269 - 0.0340 = 0.1929$$

The resulting ratio Fe$^{2+}$/($\text{Fe}^{2+} + \text{Fe}^{3+}$) = 0.8502, so this fraction of the total iron should be expressed as FeO:

$$\text{FeO} = 0.8502 \times 7.33 = 6.23 \text{ wt. %}.$$ 

The remainder of the iron will be Fe$_2$O$_3$. We convert the remaining FeO into Fe$_2$O$_3$:

$$\text{Fe}_2\text{O}_3 = 1.1113 \times (7.33 - 6.23) = 1.22 \text{ wt. %}.$$ 

These two numbers are close to the analytical values: 6.23 % vs. 5.42 % for FeO and 1.22 % vs. 2.12 % for Fe$_2$O$_3$. The reason that this result is not closer probably reflects the fact that one or more of the other elements is not determined with sufficient accuracy. For example, if CaO were 22.0 wt. %, the results would be FeO = 5.70 % and Fe$_2$O$_3$ = 1.81 %. It is most likely that SiO$_2$ is the culprit because it constitutes over 50% of the mineral by weight; thus, small errors in its determined composition can make substantial differences. Conclusion? It is probably not wise to trust individual microprobe analyses to determine the abundances of the two species of iron! Averages of a large number of points are probably better, because random errors tend to cancel out. Similar procedures can be used to recalculate spinel analyzes (for example, magnetite).
13. References


Morgan GB & London D, 1996, Optimizing the electron microprobe of hydrous alkali aluminosilicate glasses, American Mineralogist, 81, 1176-1185.


Raudsepp M, 1995, Recent advances in the electron-probe micro-analysis of minerals for the light elements. Canadian Mineralogist, 33, 203-218


Smith JV, 1966, X-ray emission microanalysis of rock-forming minerals VI. Clinopyroxenes near the diopside-hedenbergite join, J. Geol., 74, 463-477.


Stormer JC Jr., & Pierson ML, 1995, Variation of F and Cl X-ray intensity and electron microprobe analysis of apatite: An addendum, Unpublished manuscript (3 pages)


Willich P, 1992, EPMA of submicron coatings containing ultralight elements, in Proceeding of 50th Annual Meeting of Electron Microscopy Society of America (Bailey, Bentley and Small, eds.), p 1628-29 (H2 Jet)


Ziebold TO, 1967, Precision and sensitivity in electron microprobe analysis: *Analytical Chemistry* 39: 858.

Ziebold TO & Ogilvie RE, 1963, Quantative analysis with the electron microanalyzer9: *Analytical Chemistry* 35: 621-627.