Principles of SEM
1. Introduction
2. History
3. Instrument Overview
4. How does the SEM work?
   1. The electron gun
   2. Electron lenses and electron column
   3. Generating electrons and producing images
   4. Types of detectors
   5. Magnification
5. Electron Beam Interaction
   1. Origin of signal
   2. Material interaction volume
   3. Effect of spot size
   4. Aberrations
The scanning electron microscope (SEM) is a versatile tool for high resolution imaging and materials characterization. With this technique a high energy (kV) beam of electrons is focused to a fine point on a material and rastered/scanned along parallel lines. The interaction between the electron beam and the material generates multiple types of signals which are then collected and processed. These range from light and heat, to various types of electrons, and even X-rays characteristic to the elements comprising the material being analyzed. It is through this process that we can use the SEM to gain a better understanding of materials down to the micro and nanoscopic scale.
History

The history behind the scanning electron microscope (SEM) is one of extensive contributions and innovations from the past hundred years. Below are only part of the significant advances and events:

- Louis de Broglie, 1925: Corpuscule waves, or the material wave, tied a frequency and wavelength to the electron, leading to wave electron optics.
- Hans Busch, 1926: Laid the groundwork for electron optics through demonstrating that axially symmetric electric and magnetic fields could alter the trajectories of charged particles.
- Ernst Ruska and Max Knoll, 1931: Put the lens formula of Busch in to practice with the first transmission electron microscope (TEM).
- Max Knoll, 1935: Build the first scanning microscope, which was limited to a probe diameter of approximately 100 μm due to the lack of demagnifying lenses.
- Siemens, 1939: The first serially produced transmission electron microscope hits the market, with a resolution of 7nm at 70 kV.
- Jan Le Poole, 1941-1944: Construction of electron microscopes which would be the precursors to Phillips brand microscopes.
- Vladimir Zworykin, 1942: Built the first true SEM, demonstrated the use of secondary electrons in providing topographic contrast, and introduced the use of an electron multiplier tube.
- 1947: The first serially produced English electron microscope, the EM2, emerges under Associated Electrical Industries (AEI).
- Raymond Castaing, 1949-1951: Early development of the X-ray microanalyser.
- Tadahoshi Hibi, 1956: The pointed filament is introduced and applications explained.
- James Menter, 1956: Crystal lattice with dislocations is observed.
- Thomas Everhart and Richard Thornley, 1960: The Everhart-Thornley detector (ETD) is described.
- Cambridge Instrument Company, 1965: The first commercial SEM, the Stereoscan, is made available to the public.
- More to come - lots of history!

The following is a brief overview of the cross section of the SEM. Specific areas such as the electron gun and lenses will be discussed in the following sections in greater detail.

Many of the above features are hidden behind a shroud on the Apreo system. Additional items, such as the EDS and EBSD detectors not pictured above, are also present as seen in the image below.

How does the SEM work?

The modern SEM consists of two critical parts: the electron column and the operational console. The former consists of the electron gun, where the electron beam originates, in addition to a series of electron lenses which control the path and shape of said beam. The latter consists of a screen, either a cathode ray tube (CRT) or PC monitor depending on the age of the microscope, as well as operator controls in the form of physical knobs or software driven options for controlling the electron beam/column.

We can break the process of SEM imaging into a few key areas:

1. Generation of the electrons
2. Passage of the beam through the column
3. Collection of signal and image generation
In order to image materials we must first generate an electron beam. This is the purpose behind the electron gun; to provide a continuous beam of electrons which can be adjusted by the user. There are several common types of electron guns are seen in modern SEMs: tungsten "hairpin" thermionic emitter, the LaB thermionic emitter, as well as various types of field emission sources.
The oldest and most common type of electron gun is the tungsten electron gun, which consists of three main parts: a filament, a cap or Wehnelt, and an anode. The filament and anode act as negative and positive electrodes within the gun. Through resistive heating, up to and exceeding 2700 K due to the high work function of tungsten, the filament is heated to the point of thermionic emission.

It is at this point that the electrons emitted from the source, held at a negative potential ranging from 0.1-30 kV (-100 to -30,000 V), are accelerated towards the anode which is held at a ground potential of 0 V. Between the two, the Wehnelt focuses electrons and restricts the amount of electron emission. It is important to note that the electric field is concentrated by the shape of the tip which is, in this case, rather blunt compared to other types of sources.
One of the primary benefits to the tungsten gun is a readily available and easily replaced filament. While they are stable emitters, tungsten sources have a shorter life span compared to other sources due to their operation at white-hot temperatures. The size of the generated beam at crossover (source size, $d$) is also larger compared to LaB filaments or field emitters, leading to a lack of resolution compared to these sources without sufficient demagnification of the beam to the final probe size.
LaB$_6$ Guns

LaB$_6$, or lanthanum hexaboride, electron guns are the next common type seen in electron microscopy. These guns offer an immediate benefit to their tungsten counterparts in that they are longer lived, offer increased brightness (current over area), and provide a smaller source size. This final point is achieved because of the size of the emitter, in this instance a small piece of single crystal LaB$_6$.

This crystal is further sharpened to a $\sim$1$\mu$m point for an exceedingly small point of emission when compared to the hairpin tungsten filament. This tiny area creates a condition that is easier for electrons to escape, as opposed to the much flatter surface of the hairpin filament, thus increasing electron emission for the same amount of heating temperature. Temperature also plays a role in keeping these emitters free of contamination, with 24-hour operation maintaining adequate temperature to remove any adsorbed gasses or contamination after filament exchange.

For the improvements the LaB gun bring there are also some draw backs compared to the hairpin filament. Most notable is the need for a higher quality vacuum near the gun; this requirement may call for differential pumping near the gun in the form of an added ion pump at the top of the electron column. Price point is also considerable as these sources are many times more expensive than their tungsten counterparts. Longevity will offset some of this cost, and thankfully commercial varieties are comparable in ease of replacement as tungsten filaments.
Field Emitters

Building on the themes of smaller areas of emission and lower temperatures needed to emit electrons, field emission electron guns are the next progression seen in modern SEMs. For these guns, the filament often consists of a wire (often tungsten) with a point of emission 100 nm or smaller in diameter. This allows for orders of magnitude higher brightness versus the thermionic emitter counterparts.

High field strength concentrated at the tips of these sources allows electrons to "tunnel" through the surface potential and escape into vacuum, a feat done at elevated temperatures in thermionic emitters to overcome the work function of the filament material. Three common types of field emitters exist, each with their own benefits and drawbacks.

1. Cold field emitters (CFE): Relying solely on the high field strength mentioned above to generate electrons this source is not cold, but rather, operates independent of temperature as opposed to thermionic emitters. A benefit of these sources is exceedingly high brightness due to the electron emission over a very small area, however there is a trade off as the source must be exceptionally clean. This requires the source to be "flashed", or heated to elevated temperatures above 2000 K, before operation. This process is typically repeated after hours of operation and leads to a dulling of the tip over time, yet even with this the longevity of these emitters are years longer than their LaB counterparts.

2. Thermal field emitters (TFE): Operate on the same principle (high field strength) as a CFE, but at elevated temperatures. This counteracts the need to flash the tip periodically and leads to a cleaner source while also reducing some of the dulling effects incurred in the process. The images below display a TFE, damaged and removed from the instrument (note the spherical melted end).

3. Schottky field emitters (SFE): Technically a thermionic emitter, the brightness and emission of the SFE is on par with the CFE and TFE. The tips of these sources are flat and coated with ZrO to lower the work function and promote electron escape. These guns are always on, producing a cleaner source like the TFE.
Field Emitters
Electron Lenses & Electron Column

After leaving the source, the electron beam passes through various electron lenses. These lenses play crucial roles in shaping and focusing the beam. These lenses are stacked atop one another and form what is known as the electron column.

What is an electron lens?

Unlike a physical lens that is shaped to focus a beam of light, an electron lens relies on electromagnetism to act on an electron beam. As seen below, these lenses consist of coiled wire housed inside an iron casing, which surrounds the narrow path the electrons travel down from the gun. When powered with a current (I) through a given number of turns of the wire (N) a magnetic field is created which extends across the electron column through openings called polepieces. This projected field can then alter the trajectory of electrons in the column and push them towards the optic axis, or center of the column. The distance between the initiation of this push and the electron crossing the optic axis is referred to as the focal length (f).

Electron Lenses & Electron Column

The Condenser Lens
Critical to the formation of the final electron probe size is the process of demagnification of the electron beam as it passes through the column. This process restricts the diameter of the beam at crossover after leaving the electron gun \( (d) \) in progressively smaller increments.

In the example below we see the effects of low current (left) versus high current (right) in the condenser lens in a multi-lens system, and the effects this has on various beam diameters. The \( d \) diameter (crossover below the gun) is further decreased after each lens to \( d \) and \( d \). Also effected are the focal length \( (f) \) and angular spread of electrons \( (\alpha) \). These changes alter the total number of electrons landing on a sample, and thus the current and size of the probe or "spot size".

Electron Lenses & Electron Column

The Objective Lens
Below the condenser lens is situated the objective lens, an electron lens with the primary task of focusing the electron lens on to the sample surface. This is the lowest lens in the electron column and can be seen in the chamber scope of many instruments.

Scan Coils
While the objective lens focus the beam to the sample surface, the scan coils (also referred to as deflection coils) are responsible for moving the beam across the surface of a material.
Generating Electrons & Producing Images

Producing images in the SEM combines multiple steps and is a complex process. This section will take a brief overview of the generation of electrons and the different styles of detectors used to collect them.

**Incident Beam Scanning**

The first step in the process involves the electron beam, often referred to as the incident beam, which has just traveled through the electron column and will be focused on the material surface. The interaction of incident beam and material produces various types of electrons (SE, BSE) as well as characteristic X-rays. These electrons are collected on a per pixel basis in X–Y coordinate space before the beam is shifted or translated to the next X–Y point. The incident beam is translated laterally, or "rastered", across the field of view and upon reaching the end of one row (1) is moved to the beginning the next (2). The intensity for each point is recorded; this is what forms the basis of the image we see on screen. There is no true image (e.g., captured by film inside the chamber) for the SEM.
At each point the incident beam interacts with the material, multiple types of signal are generated (see figure below). For the purposes of this section the most important are: 1) **Secondary electrons** (SE), low energy (50 eV or less) valence electrons generated through inelastic scattering (loss of energy from the incident beam). These electrons have a very small mean free path $\lambda$ (average distance between collisions) and are limited to generation within ~$5\lambda$ of the material surface. 2) **Backscattered electrons** (BSE), electrons with a much higher energy on par with the incident beam. These electrons generally elastically scatter through the material, in the process generating SE, before emerging from the surface in relatively straight trajectories. The proportion of electrons scattered back is tied to mean atomic number, with higher atomic numbers causing more electrons to be backscattered.
Types of Detectors

**Everhart-Thornley Detectors (ETD)**

The most common detector seen in the SEM is the Everhart-Thornley Detector (often seen as E-T or ETD). The collection of secondary electrons with this detector leads information about sample topography. This detector collects a combination of secondary as well as backscattered electrons with a wide solid angle of collection, meaning signal can be collected in a variety of angles between sample and detector. The detector is based on the premise that an energetic electron (~10 kV) strikes a scintillator (displays light/lumination in excitation by ionizing radiation), and this light passes through a light pipe to a photomultiplier tube. In this unit light is converted back into electrons, which are then amplified through acceleration and deflection between multiple electrodes.

There are two important distinctions between electrons to be made here: 1) secondary electrons are much lower energy (50 eV or less), and 2) backscattered electrons are much higher energy, close to that of the incident beam. Because of this SE not only need to be pulled towards the ETD with a positive bias voltage, but also need to be accelerated to the point they will generate light upon striking the scintillator. This is accomplished by using a strong bias potential, ensuring both types of electrons are detected efficiently.
Types of Detectors
Types of Detectors

Backscattered Electron Detectors

In contrast to low energy SE which need to be pulled to a detector, BSE detectors can take advantage of the higher energy, relatively straight trajectory of the BSE, and especially close proximity to samples to increase the solid angle of collection relative to ETD detectors. In fact, these detectors are only sensitive to higher energy BSE and do not see signal from lower energy SE. Dedicated solid state BSE detectors are often mounted, or inserted, just below the objective lens and often have an annular (ring) shape further divided into various sections or arrays. The example below is an annular detector with two sections, A and B.

BSE scattered by the incident beam travel up to the detector and strike the detector, a semiconductor, and in doing so excite electrons to cross the band gap and move from the valence band to the conduction band. This action produces an absence, or hole, due to the electrons change in band; this action also produces a measurable current.

Materials such as silicon are used in the construction of these detectors due to their nature of producing many band gap crossings when struck by the high energy of the BSE.
The magnification displayed on screen is tied directly to the length of the line scanned or “rastered” by the electron beam, a process seen in section 4.3. Specifically, magnification is determined by the following equation:

\[ M = \frac{L}{l} \]

Here, magnification (M) is a function of the ratio between the length of the rastered line on the material (l) and the length of the line displayed on screen (L). Because the size of the monitor is fixed, changing the length of the rastered line (or the area scanned) changes the magnification of the image produced. With smaller and smaller scan areas the magnification increases, while larger scan areas lead to a low magnification and large field of view.
Electron Beam Interaction

Knowledge regarding the interaction of the electron beam with a material is fundamental for proper operation of the SEM. In the following sections topics such as where signal originates in a material, interaction volumes, effects of spot size/probe current, and more will be discussed.

Origin of Signal

When the incident electron beam lands on a material it interacts with a volume of that material and produces multiple types of signal. These include:

- Visible light: luminescence

- Electrons:
  1. Backscattered electrons (BSE) – High energy incident beam electrons almost entirely inelastically scattered (no loss of incident beam energy) through a material which eventually scatter out from the surface of the material. Interaction with the material will also generate secondary electrons.
  2. Secondary electrons (SE) – Low energy (<50 eV) valence electrons ejected from inelastic scattering (loss of energy from the incident beam). These electrons have a very limited mean free path \( \lambda \) (average distance between collisions) and tend to originate no deeper than 5\( \lambda \) within a material. Conductive materials tend to have a lower \( \lambda \), while insulators have a higher \( \lambda \).

- Characteristic X-rays: X-rays, described below, are also generated via inelastic collisions.
Electron Beam Interaction

An approximation of origins of different types of signal from the interaction volume between material and incident beam.
Electron Beam Interaction

Energy distribution of electrons from a target struck by the incident beam. Note the low energy range of secondary electron emission (I) compared to BSE and auger emissions (II) and BSE emissions (III) as E approaches E or the incident beam energy.
Electron Beam Interaction

**Characteristic X-rays**

Electrons from the incident beam will experience inelastic collisions with inner shell electrons of atoms within a material, causing an ejection of a shell electron. This inner shell ionization leaves the atom in an excited or elevated energy state (in the example below this is due to a K-shell electron ejection). To return to a lower state an electron from a higher shell must fill this new vacancy, in this case from the L-Shell. The energy difference between the K-Shell and L-Shell electrons during this process is called the transition energy.

This process happens in two ways:

1. **Auger emission**: The transition energy is transmitted to another L-Shell electron. This electron is then ejected with a unique kinetic energy depending on the material or phase it originated from.

2. **X-ray emission**: A X-ray is generated with an energy (eV) equal to the transition energy between the K and L-Shell electrons involved. X-rays generated in this fashion have a very narrowly generated energy range that is "characteristic" of the atom and shells involved (e.g. K-L, K-M). For example, below is an oxygen atom. The characteristic X-ray energy would be the difference between the K (532 eV) and the L (7.1 eV), or 524.9 eV.
Electron Beam Interaction

K-Shell

L-Shell

e\text{ beam}

Auger emission

X-ray emission
Material Interaction Volume

The accelerating voltage, the voltage at which electrons are accelerated down the electron column, has a direct impact on the interaction volume between beam and material. As a general rule the size of interaction increases with increased voltage. Furthermore, the mean atomic number of the material interacting with the beam will also affect the interaction volume, as seen below. The majority of electron movement within the material forms an approximate “tear drop” which trends closer to a hemisphere as the mean atomic number increases.
Where is your signal coming from?

A very important consideration when choosing a voltage is interaction volume; with higher voltages increasing signal from depth, information near the surface may be degraded or missed.

Both images are taken from tin spheres on a carbon backing material. Note how much greater the surface detail is at 2 kV (left) than 30 kV (right), where many small features appear washed out.
Material Interaction Volume

Effects from voltage

The simulations below demonstrate the changes in interaction volume within a material, in this case iron, at 10, 20, and 30 kV. While images have been enlarged to enhance this effect, notice the areas marked in μm in the upper right.
Material Interaction Volume

Effects from atomic number

In addition to the voltage of the incident beam, the mean atomic number of the material being analyzed must also be considered. In the simulations below, the electron paths generated by a 20 kV beam interact with carbon (left), iron (center), and uranium (right) are displayed. As with the previous simulations, note the areas indicated in the upper right and how they shrink as atomic number increases.
Effect of Spot Size

A key factor in resolving features in a material is the spot size or probe current. The two are tied together, as seen in the figure below depicting the incident electron beam. As the final diameter of the electron beam \((d)\) is increased or decreased, the number of probe current \((i)\) or number of electrons striking the sample surfaces increases or decreases as well. Thus, a larger spot size translates to a higher probe current and vice versa. It should be noted that there is also an association with accelerating voltage, with higher voltages leading to a smaller spot size.

1: Accelerating voltage \(V_0\) (kV)
2: Probe current \(i_p\)
3: Convergence angle \(\alpha_p\); \(1/2\) the angle of the electron beam cone
4: Probe diameter/spot size \(d_p\)
Effect of Spot Size

For higher resolution images it is recommended to work with a smaller spot size in order to "resolve" between increasingly small features (above figure). However, users should not be focused solely on "high resolution". Larger spot sizes can be used to ensure lower magnification images are crisp, and a range of magnifications should be observed to truly characterize a material rather than solely striving for the highest magnifications possible. Furthermore, where techniques such as EDS or EBSD are concerned, the increased signal generated from a larger spot and higher probe current can be beneficial for the purposes of characterization.
Aberrations

Any lens, either a physical lens used to bend light or an electromagnetic lens used to move electrons, is prone to aberrations (defects) which affect the final image. Fortunately, in the case of the electron microscope these defects can be minimized by the user. The final spot size \( d \) of the image is a product of the root sum of the squares of each of these defects, which by nature cause electrons of the primary beam to land in a disc rather than a focused point.

**Astigmatism**

Perhaps one of the most noticeable aberrations, astigmatism is often seen as stretching of features in perpendicular directions relative to a point of focus. So common is this aberration that knobs on the MUI directly control a set of opposing magnetic poles designed to force the beam into a circular shape, removing the stretched features from an image. Due to an array of potential causes, from dirty apertures to errors in lens construction and more, electrons emerging from a point \( P \) can form two different foci at right angles to one another as seen in figure A. This can lead to a stretching or smearing of an image.

**Aperture Aberration**

Looking at figure B, primary beam electrons can diffract at the edges of small apertures, resulting in a "Airy disc" \( d \) intensity distribution upon landing.

**Spherical Aberration**

In figure C, the trajectories of electrons traveling farther from the optic axis \( B \) are changed more significantly than those near the optic axis \( A \). This effect causes electrons to land in a disc, rather than a point. The region at which these two trajectories are the closest is the "disc of minimum confusion", \( d \).

**Chromatic Aberration**

Seen in figure D, electrons emerging from a point \( P \) which have different energies \( E \) vs \( E - \Delta E \) will be focused to different points on a sample after passing through the lens.
Aberrations