Chapter 1
The Helium Ion Microscope

John Notte and Jason Huang

Abstract The key technologies that comprise the helium ion microscope are described in detail. Specific attention is given the cryogenic cooling system, the vacuum system, the gas delivery system, the ion-optical column, the detector and imaging system, and vibrational considerations.

1.1 Introduction

The family of charged particle microscopes (CPM) includes all instruments that rely upon a charged particle beam for precision imaging, or patterning of a sample. The family originated in the 1930s with the efforts led by Helmut Ruska, and independent work by Manfred von Ardenne, and now includes the SEM, TEM, STEM, FIB, and all of their variants. The CPM family has a long, rich history, and has played a distinguished role in enabling mankind to see matter at the smallest scale. One of the most recent entries into the family is the class of microscope that relies upon the gas field ion source (GFIS). The GFIS itself is an old concept, conceived some 50 years before its eventual commercialization [1] in 2007 in the form of the helium ion microscope (HIM). More broadly, this class of instrument is termed the gas field ion microscope (GFIM), which includes instruments with gas species other than helium (e.g. hydrogen, neon, nitrogen, or argon). The HIM instrument offers distinct imaging advantages such as high resolution [2], long depth of focus, surface sensitivity, high secondary electron yield [3], and several novel contrast mechanisms [4, 5]. The HIM has also been employed for non-imaging applications where it can be used to expose resist [6], directly pattern a substrate [7], drill nanopores [8], induce patterned chemical processes, implant stress [9], or perform material
analysis [10]. Variants of the HIM include the use of the GFIS with other gas species such as neon [11], which enables further nanofabrication [12] and analysis capabilities [13].

This chapter provides an overview of some of the specific technology sub-systems that are contained within the helium ion microscope (HIM). Special attention is given to the technologies that are of unique importance to this class of instrument. At the time of this writing, there is only one commercial provider of the GFIS microscope, but when possible, the technologies described herein are presented in a generic and even hypothetical way, as it might pertain to any such microscope. The topics which are addressed are the GFIS gun (Sect. 1.2), the cryogenic cooling system (Sect. 1.3) the vacuum system (Sect. 1.4), the gas delivery system (Sect. 1.5), the ion optical column (Sect. 1.6), an analysis of sample damage, (Sect. 1.7), the detectors and signal chain (Sect. 1.8), and vibrational considerations (Sect. 1.9).

1.2 The GFIS Gun

The term, gas field ion source gun or “GFIS gun” encompasses the hardware immediately surrounding the gas field ion source, and which enables its basic performance. The term “GFIS emitter”, or “GFIS source” is more specific and refers to just the emitter itself, and is addressed in detail in the subsequent chapter. The GFIS gun is distinctly different from other established technologies like the liquid metal ion source (LMIS) for gallium, the thermal field emitter (Schottky source) for electrons, or the cold field emitter (CFE) for electrons. Therefore, we begin this chapter with a review of the requirements for the GFIS gun, and a summary of its performance. The number of requirements, and the engineering challenges in simultaneously addressing them are the main reasons that the GFIS microscope did not become commercialized earlier. Indeed, it is quite likely that Mueller and his group at Penn State envisioned that their field ion microscope (FIM) might one day be the source which powers a powerful new ion beam microscope [14]. But the pathway from FIM to GFIS microscope was littered with a collection of dead ends, distractions, and failures lasting for 50 years. Even after the technology was demonstrated in the laboratory, there are additional hurdles to overcome before it could become a commercial product. A commercialized product has requirements of affordability, serviceability and easy operation in the hands of an operator whose interest lies in their sample—not in the microscope.

GFIS Gun Performance

While a detailed description of the GFIS emitter is addressed in Chap. 2, it is necessary here to give an overview of the overall gun performance before we discuss the technology sub-systems which enable it. Table 1.1 summarizes the key GFIS emitter and gun characteristics under typical operating conditions.
### Table 1.1 Typical performance metrics for a GFIS gun

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Units</th>
<th>Value</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating extraction voltage</td>
<td>kV</td>
<td>25–35</td>
<td>This is typical for helium operation, but depends on the shape of the emitter</td>
</tr>
<tr>
<td>Operating temperature</td>
<td>Kelvins</td>
<td>60–90</td>
<td>Colder is better, but with limits (see Sect. 1.3)</td>
</tr>
<tr>
<td>Base pressure</td>
<td>Torr</td>
<td>&lt;4 × 10^{-10}</td>
<td>(See Sect. 1.4)</td>
</tr>
<tr>
<td>Operating gas pressure</td>
<td>Torr</td>
<td>1 × 10^{-7}–5 × 10^{-6}</td>
<td>This is the partial pressure of helium or neon indicated on the gun gauge, not the actual pressure at the emitter. (See Sect. 1.5)</td>
</tr>
<tr>
<td>Brightness</td>
<td>A cm^{-2} sr^{-1}</td>
<td>5 × 10^9</td>
<td>This is under optimized conditions</td>
</tr>
<tr>
<td>Reduced brightness</td>
<td>A m^{-2} sr^{-1} V^{-1}</td>
<td>1 × 10^9</td>
<td>This is under optimal conditions, with an extraction voltage near 30 keV</td>
</tr>
<tr>
<td>Pressure normalized brightness</td>
<td>A cm^{-2} sr^{-1} Torr^{-1}</td>
<td>2.5 × 10^{12}</td>
<td>The reference pressure is the species-corrected pressure in the inner gun</td>
</tr>
<tr>
<td>Energy spread</td>
<td>eV</td>
<td>~1</td>
<td>This is measured as a FWHM</td>
</tr>
<tr>
<td>deBroglie wavelength</td>
<td>pm</td>
<td>~0.080</td>
<td>This is for helium at 30 keV</td>
</tr>
<tr>
<td>Virtual source size (diameter)</td>
<td>nm</td>
<td>&lt;0.25</td>
<td>This is inferred from the atomic spacing of emitter atoms (expressed as a FWHM) and represents an upper bound</td>
</tr>
<tr>
<td>Total emitted current</td>
<td>pA</td>
<td>~150</td>
<td>This is the total emission from three atoms of a trimer under typical conditions, (gun gauge pressure indicates 2 × 10^{-6} Torr)</td>
</tr>
<tr>
<td>Angular current density</td>
<td>A/sr</td>
<td>1 × 10^{-6}</td>
<td>This is the peak angular current density after extraction under typical conditions (gun gauge pressure indicates 2 × 10^{-6} Torr)</td>
</tr>
<tr>
<td>Angular emission</td>
<td>Degrees</td>
<td>0.5°</td>
<td>This is typical for each atomic emission site after extraction, expressed as FWHM measurement</td>
</tr>
<tr>
<td>Emission current stability</td>
<td>%/h</td>
<td>&lt;1</td>
<td>This value is typical for helium. The value for neon is considerably greater</td>
</tr>
<tr>
<td>Trimer lifetime</td>
<td>µTorr × h</td>
<td>~80</td>
<td>This value is typical for helium. The value is considerably less for neon</td>
</tr>
</tbody>
</table>
GFIS Gun Requirements
The GFIS gun has several prerequisites that must be met before it can be operated—just as the LMIS, Schottky field emitter, or the CFE have their own requirements. For the GFIS, the requirements are relatively simply stated: (1) An emitter in the shape of a sharpened wire should be biased with a positive voltage respect to an adjacent electrode. The details of the emitter, and its composition, and shaping will be discussed in detail in the subsequent chapter. (2) Cryogenic cooling should be provided to cool the GFIS gun to a working temperature of no more than 90 Ks. The reasons for the cryogenic cooling and the practical issues of providing such cooling are the focus of Sect. 1.3. (3) The GFIS gun must be housed in a UHV vacuum vessel to provide a base pressure in the range of $10^{-9}$ Torr. An analysis of the vacuum requirements and some practical examples are discussed in Sect. 1.4. (4) The GFIS gun must allow the introduction of helium (or other chosen gas) at a sufficient purity and flow rate. The details of the gas delivery system are provided in more detail in Sect. 1.5.

1.3 Cryogenic Cooling

Unlike conventional sources of ions or electrons, the GFIS ion source must be cooled to a steady temperature between 60 and 90 K. Experiments have indicated that emitted current increases markedly with decreased temperature (about 5% per Kelvin). The cold temperatures are thought to improve the collection and surface accommodation [15] of the provided imaging gas. Continued benefits, however, are not realized with further reduced temperature as the emission tends to become erratic. It is suspected that the imaging gas becomes less mobile at the lowest temperatures and cannot be surface transported from the shank and periphery of the emitter to the ionization sites. The low temperature also helps to assure the mechanical integrity of the emitter in the presence of the large electric fields. Thermal motions of the emitter, which could impact virtual source size or energy spread, are also reduced at the cryogenic temperatures. And lastly, the low temperature provides some level of cryo-pumping and cryo-trapping to reduce many of the otherwise corrosive impurities from reaching and damaging the emitter. To achieve these requirements there are entire families of refrigeration technologies capable of achieving these temperatures (e.g. Joule-Thompson, Gifford-McMahon, Pulse Tube, etc.). However, many of these technologies cannot be scaled to provide the cooling power (>2 W) that is required for a practical GFIS gun with the unavoidable heat loads it must contend with. It should be recognized that most of the refrigeration techniques introduce vibrations which can adversely impact the utility of the produced beam. A pulse tube cooler operated with “active” and “quiet” modes has been demonstrated successfully on several of the first commercial HIM systems. However, a standard cryogen in the liquid or solid phase is the most preferred solution because of the incredible cooling power, simplicity, and the lack of any moving parts near to the GFIS gun. When operating over a phase change
The cryogens maintain a fixed temperature as they absorb heat from the system, eliminating thermal drift effects. Both liquid neon and nitrogen have been considered, with liquid nitrogen being more readily available.

The current commercial product, the ORION NanoFab, provides the necessary cryogenic cooling through an integrated dewar (Fig. 1.2) that is thermally connected to the internal GFIS gun through a cold finger. The cryogen is surrounded by an evacuated “jacket” region to provide the necessary thermal insulation. According to a prescribed schedule, the contents of the integrated dewar is filled with liquid nitrogen via a supply valve until it reaches its capacity. As soon as the fill process is completed, the supply valve is closed, and the vacuum valve is opened to draw a vacuum over the contents of the dewar. This causes the liquid nitrogen to change phase to solid nitrogen (removing about 25% of the product in the process) and be further cooled to about 60 K. This temperature is maintained, given the total heat load, for several hours. Thereafter, as per the software scheduled fill time, the supply valve opens again to fill the integrated dewar, now almost empty, with liquid nitrogen from the facility provided LN2 tank. The low temperature achieved with solid nitrogen compared to liquid nitrogen helps to make up for the unavoidable temperature drop along the thermal connections from the integrated dewar to the GFIS gun. The temperature as measured at the GFIS gun is typically 75–80 K. One of the advantages of the solid nitrogen cryogen is the elimination of the boiling, and the vibrations it invariably introduces into the GFIS emitter. As shown in Fig. 1.1, as the heat load is delivered to the dewar, the cryogen will change phases. If
operated at atmospheric pressure, the heat load drives the nitrogen from point A to A’, thereby involving significant boiling and achieving a temperature of only 78 K at the bottom of the dewar. When operated at a vacuum level sustained at about 3 Torr, the same heat load drives the nitrogen from point B to B’, which involves only sublimation, a process which transfers no vibration to the GFIS gun.

1.4 Vacuum System

Adopting the terminology from the FIM literature, the term “base vacuum” refers to the pressure in the absence of the imaging gases (helium, neon, or argon for example). The various constituents of the base vacuum present a risk of adsorption near the emission sites as adatoms, which would destabilize the emission process [16]. The base vacuum levels required for the successful GFIS emitter cannot be underestimated, given that the selected emission site corresponds to a single atom which must be free from adsorbates for about one week at a time! A simple calculation of the monolayer formation time and the condition of less than 10% coverage, suggests a vacuum requirement on the order of $10^{-13}$ Torr is required in the region of the emitter. There are technologies and techniques that make this otherwise impossible task more tractable. Most importantly, the modern turbo-molecular pump provides extremely high pumping speeds for almost all gas species. All of the commercial GFIS microscopes incorporate a combination of turbo pumps with a total pumping speed of nearly 1000 liter per second! Multiple pumps are required for the multiple chambers separated by differential pumping apertures, since the emitter must be kept under extreme high vacuum (XHV) conditions while the sample chamber must accept the frequent loading and unloading
Fig. 1.3 Schematic of (a) the main system and (b) the gas delivery system of the HIM as presented in the vacuum user interface.
of samples (Fig. 1.3). Additionally, the previously mentioned cryogenic surfaces surrounding the gun can provide an important additional degree of cryo-pumping and cryo-trapping, but this is only for select species. And lastly, there is a unique property of the high electric field surrounding the emitter which impacts its local vacuum. The high field will tend to polarize any neutral gas atoms in the area, and when subjected to the strong gradient they are drawn in towards the emitter with a force that is proportional to the electric field squared [17]. The undesired gas atoms are ionized as they approach close to the emitter where they experience an electric field surpassing their field ionization limit which is about 2 V/Å. Thereafter, the positive ions are accelerated away from the emitter, leaving a small “bubble” of extreme high vacuum bounded by the contour of critical electric field magnitude. Operating the emitter with a larger extraction field increases the volume of the bubble. In contrast, the helium or neon atoms are typically not ionized until they reaches the surface of the emitter.

1.5 Gas Delivery System

When the imaging gas (neon for example) is turned on, the imaging gas atoms need to be delivered to the emitter with a minimum of contaminants to preserve the trimer and its immediate neighborhood. As described earlier a single adatom adsorbed near the selected atom can result in erratic emission or even cause the selected atom to be lost. Typically, the flux of delivered imaging gas is roughly, \( Q = 4 \times 10^{-4} \text{ Torr liter/s, or } 10^{16} \text{ atoms/s, or } 0.03 \text{ sccm} \). A high quality, “research grade” neon might have contaminants at the 5 ppm level, implying that there may be as many as \( 5 \times 10^{10} \) impurity atoms/s delivered to the gun. It is for this reason that HIM instruments incorporate an active chemical purifier integrated into the gas delivery system. The purifier is initially activated as a part of the gas box baking routine, and reaches a peak temperature of about 400 °C. Thereafter, when operating at room temperature, the purifier helps to diminish the abundance of a select subset of the impurities. The purity of the delivered gas is also improved upon its passage through a cryogenically cooled delivery tube to the gun—allowing select gases to be cryo-pumped or cryo-trapped. The rate of gas delivery is controlled by a high precision UHV leak valve which relies upon the adjustable separation between a sapphire surface and an opposing copper surface. The resulting leak rate can be adjusted manually, but more commonly a motorized actuator and regulation software work together to keep the delivered gas pressure constant to within a few percent. Needless to say, the stringent requirements for gas purity require that the entire manifold from gas cylinders to the GFIS gun be constructed of the highest quality vacuum materials, and be thoroughly baked to maintain surface cleanliness.
1.6 The Ion Optical Column

The primary functions of the optical column in the HIM are: (1) To image the gas field ion source onto the specimen to form a finely focused probe, with a specified probe current; and (2) to scan the probe on the specimen plane to perform various applications, such as imaging or patterning. The first function is often called the DC property of the column, while the second function called the AC property. To accomplish these functions, a typical HIM column can be implemented with a design depicted in Fig. 1.4.

In the column shown in Fig. 1.4, there are two lenses after the extraction optics. The extraction optics extracts the ion beam out of the tip. The first lens (typically called the condenser lens) images the source to a cross-over inside the column. The second lens (typically called the objective lens or final lens) images the cross-over onto the specimen plane. A two-lens system, combined with a variable beam limiting aperture, allows independent control of the column magnification and probe current.
One can estimate the diameter of the probe formed with certain column setting, using the widely known method developed by Barth and Kruit [18]. This method calculates the diameter of the probe containing 50% of the probe current. The calculated diameter using this method is often cited as the $d_{50}$ of the probe. Following this approach, the overall probe size $d_p$ consists of four components:

- $d_I$: the size of the source image
- $d_A$: the contribution due to diffraction
- $d_S$: the contribution due to spherical aberration
- $d_C$: the contribution due to chromatic aberration

For Helium ions, these components can be calculated using the following formulae:

\[ d_I = \sqrt{\frac{4I_p}{B_r V_p^2 \alpha_i^2}} \]  \hspace{1cm} (1.1)

where $I_p$ is the specified probe current, $B_r$ is the reduced brightness of the source, $V_p$ is the accelerating voltage of the beam and $\alpha_i$ is the half angle of the beam on the image side.

\[ d_A = \frac{7.78 \times 10^{-12}}{\sqrt{V_p \alpha_i}} \]  \hspace{1cm} (1.2)

\[ d_S = 0.18 C_s \alpha_i^3 \]  \hspace{1cm} (1.3)

where $C_s$ is the spherical aberration coefficient of the column at a given column magnification.

Fig. 1.5 The $d_{50}$ probe size vs image-side half angle for a 30 keV landing energy helium beam. The two graphs are for optimal conditions a for a 0.5 pA helium ion beam, and b for a 20 pA beam
where $C_c$ is the chromatic aberration coefficient of the column at a given column magnification, and $\Delta V$ is the energy spread of the ions emitted from the gas field ion source.

The four components are then combined together to form the overall $d_{50}$ probe size $d_p$:

\[
d_p = \sqrt{\left( (d_A^4 + d_s^4)^{1/4} + d_{1.3}^{1.3} \right)^{1/2} + d_C^{1/2}}
\]

From (1.5) we can calculate the theoretical probe size (or resolution) at a given probe current, and at a given column magnification.

From (1.1) through (1.5) it is clear that the source properties have a very high impact on the probe size. The gas field ion source is unique compared to other sources in that it has a very high brightness, an atomic-size virtual source size, low energy spread, and an extremely short de Broglie wavelength. Specifically, the helium ion source is characterized by the optical properties presented Table 1.1.

For a given current, $I_p$ (which relates to the source-side half angle by the angular intensity), one can calculate the $d_{50}$ probe size as a function of the image-side half angle, $\alpha_i$. Figure 1.5 shows some typical $d_{50}$ probe sizes versus $\alpha_i$ plots for helium beams with a 30 keV landing energy. As we can see in these two cases (either 0.5 pA or 20 pA), the probe size of the helium ion beam is primarily limited by source brightness and chromatic aberration.

All four components are also plotted along with the total probe size, and from the plot we can see the optimal probe size for the 30 kV 0.5 pA probe is about 0.35 nm, achieved at an $\alpha_i$ of 0.35 mrad. The corresponding column magnification for this half angle can be calculated using the following equation:

\[
M = \frac{\alpha_o}{\alpha_i} \sqrt{\frac{V_{ext}}{V_p}}
\]

where $\alpha_o$ is the source-side half angle, $V_{ext}$ is the extraction voltage, and $V_p$ is the landing energy of the probe. For the 30 kV 0.5 pA probe considered here, the optimal probe size is achieved with the column magnification of 0.72.

As we can see in these two cases (either 0.5 pA or 20 pA), the probe size of the helium ion beam is primarily limited by source brightness and chromatic aberration.

One important implication from the $d_{50}$ analysis above is that the optimal probe size for helium beam is achieved at a relatively high column magnification (close to 1), as compared to other commonly used charged particle sources (Schottky field emitter has an average source size of 25 nm, Ga source has a source size from 35 to 50 nm; they all operate with a much lower column magnification). If there is any mechanical vibration on the source, the magnitude of the vibration will be imaged
onto the specimen plane by the column magnification factor. A higher optimal column magnification means higher sensitivity to source vibration. Great care and engineering effort are therefore expended to minimize the source vibrations within the HIM (see Sect. 1.9).

In summary, a working optical column for GFIS needs to be able to provide a focusing property with a column magnification factor in the neighborhood of 1 (typical range of 0.3–2) in order to cover the optimal focusing for a range of probe currents. The column should be relatively immune to mechanical vibration, and the entire column resides in high to ultra high vacuum. In actual implementation, electrostatic lenses are often the lenses of choice because of the weak effect magnetic fields on ions in the energy range in question. In addition to the DC focusing property, the column needs to be able to provide AC forces to deflect the beam to different locations on the specimen plane.

**Beam Extraction**

In order to ionize the gas atoms and extract them away from the emitter, the gas field ion source is placed at a close distance in front of an extraction electrode (often called the extractor). The extractor is biased to a negative potential with respect to the source in order to generate the electric field needed for ion emission (Fig. 1.6). For helium ions, the field at the apex of the source needs to be at least 4.4 V/Å, and for neon ions, the field is 3.3 V/Å.

The extractor also generates the field needed to select a certain atomic configuration at the source apex for emission site by field evaporating the substrate atoms. This field is typically higher than the extraction field for the imaging gases. After the preferred atomic configuration is achieved, the extraction potential is adjusted to maximize the field only at the selected atomic sites (such as the trimer), and to minimize emission from elsewhere for optimal source brightness.

In implementing the high voltage for the extraction, there are several different topologies available. A commonly used one is floating the extraction power supply...
on the accelerating voltage power supply for the source. Alternatively, a ground-based extraction supply could also be used. In either case, it is important to prevent the source arcing to any neighboring ground potential in the GFIS gun, because that would momentarily turn on the source into an electron emitter (rather than ion emitter). A safety interlock that locks the extractor potential to remain negative with respect to ground would effectively prevent electron emission due to arcing.

The extractor electrode must be fabricated with mirror finish, with materials that are compatible with high voltage and ultra-high vacuum. Materials with low ion sputtering yield are preferred. The angle that the extractor aperture subtends with respect to the source should be adjusted based on practical operating voltage, effective gas pressure at the apex of the source, distance-pressure-product between the apex and the exit of the aperture, and the exposure to sputtering ions and neutrals.

**Beam Alignment**

In principle, to achieve the best optical probe size, the ion beam should travel along the optical axis of the column. In reality, the beam emitted from the source may not be aligned with the axis of the condenser lens, and the axis of the condenser lens

---

**Fig. 1.7** Example of electronic beam alignment with deflectors. The optical axis of the objective lens is shown as a *dotted line*, while the beam is shown as a *solid green line*.
may not overlap with the axis of the objective lens due to mechanical tolerances. Therefore, mechanisms are provided within the column to align the beam with the optical axes (Fig. 1.7).

The first alignment typically is to align the beam emerging from the source to the optical axis of the condenser lens. This can be done mechanically with tilt and shift mechanisms, or electronically with beam deflectors, or a combination of both. The alignment criteria of this step would be to minimize image movement while varying the strength of the condenser. The alignment position varies slightly from one atomic configuration of the source to another; therefore, realignment is needed after a new trimer formation or emission site is selected.

The second alignment is to align the beam emerging from the condenser to the axis of the objective lens. A double deflection system above the aperture plane achieves this goal. The upper and lower deflector sets (each with X and Y components) allow the outgoing beam to be correctly aligned to the axis of the objective lens. For convenience, the controls are mapped to provide virtual pivot points which are not necessarily at the planes of the deflectors. This allows the operator to first optimize for illumination, and then for optimize for minimal beam motion with objective lens modulation (“wobble”).

**Aperture Mechanism**
The aperture radius defines the maximum ray angle for a given ray path. Combined with the position of the cross-over point, it defines the probe size and probe current. Therefore, when deciding where to put the aperture and what size should be used, there are two general considerations. First, the aperture plane should be positioned as far away from the condenser lens as possible for maximum probe current adjustment range (while the condenser is positioned as close to the source as possible). There are practical constraints that would not allow positioning the aperture plane immediately before the objective lens. If a double condenser system is used, then the aperture plane should be positioned in between the two condenser lenses, and far away from the first condenser. Secondly, the size of the aperture used for a particular probe current should be optimized. This is done by first computing the optimal ray path for that probe current, and from the position of the aperture plane we can calculate the diameter of the beam on that plane. Such diameter becomes the optimal aperture size for that probe current.

In terms of aperture materials, for the helium beam it is found that the gold thin film type apertures helps improve resolution because of reduced beam-aperture scattering and reduced contamination. The lifetime of this type of thin aperture is however shorter because helium sputters away the aperture over time. For the neon beam, the preferred aperture material is molybdenum for longer lifetime.

**Deflection System**
The column components mentioned above, extractor, lenses, beam alignment and aperture, all strive to provide a finely focused probe at a given current on the specimen plane. In order to acquire a 2-dimensional image, this probe needs to be
deflected to an array of locations on the specimen plane corresponding to the pixels in the final image. This task is accomplished with the deflection system.

There are many ways to deflect an ion beam. For the same reason as in the focusing system, an electrostatic field is a convenient way to implement such deflection. Choices of the deflection could be single deflection, double deflection, or multiple-levels of deflection serving different field of views. The position of the deflection system could also be quite flexible. It could be located before the objective lens (pre-lens deflection) or after the lens (post-lens deflection). Pre-lens deflection has a smaller field of view, but allows a shorter working distance between the lens and the specimen. Post-lens deflection on the other hand can give higher field of view at the expense of working distance.

A pre-lens, double deflection system is shown in Fig. 1.8 as an example. The incoming beam is first deflected by an angle $\alpha$ by the upper deflector, followed by a second deflection to the opposite direction by an angle $\beta$ by the lower deflector. In the passage of each deflector, the following equation applies:

$$\text{deflection angle} = \frac{qVL}{2ED}$$

(1.7)

where the deflection angle is $\alpha$ or $\beta$ corresponding to the upper or lower deflection plates, and $q$ is the charge, $V$ is the voltage across the deflector, $E$ is the energy of the particle, $L$ is the length of the deflector and $D$ is the distance between the opposite plates. By applying the appropriate voltages on the double deflection system, the beam should cross the optic axis again after two deflections at an angle of $\beta - \alpha$. The field of view therefore is:

$$FOV = 2(\beta - \alpha)P$$

(1.8)

where $P$ is the distance between the pivot point and the specimen. When the microscope changes its magnification, what is changing is the amplitude of the voltage ramps applied to the deflectors, which causes the range of deflection angles to change.

Varying the voltage ratio on the upper and lower deflectors moves the pivot point up and down along the optic axis. There are applications requiring putting the pivot points at various of locations, but typically this pivot point should be located...
at the distortion-free point of the objective lens to minimize the beam blur due to
deflection aberrations. By having the pivot point at the distortion-free point also
help reduce “barrel” distortion (a result of the pivot point located in front of the
distortion-free point) and “pincushion” distortion (a result of the pivot point
located behind the distortion-free point). A side effect of having a pivot point not at
the distortion-free point of the lens is, as the focus value varies, the ray is refracted
in the lens and results in different field of views.

From a practical engineering point of view, the design of the deflection system
requires a complex balance of several often conflicting factors, such as deflection
sensitivity, bandwidth, and noise of the driving electronics. For example, for fast
patterning applications it is desirable to increase the bandwidth, but that would
result in higher noise which would broaden the probe size.

In addition to the dynamic scanning signals discussed above, DC voltages are
superimposed on the deflectors to perform static corrections, such as image shift, or
astigmatism correction. Image shift or “pan” is the same as scanning, but static.
Astigmatism correction is different. Instead of using a dipole field as in the scanning
deflection, a stigmator uses a quadrupole field to correct for typical asymmetric
beam profiles. An example of quadrupole field is shown in Fig. 1.9. Four electrodes
are needed to generate such field. By having two quadrupole fields orientated at an
angle (typically 45°), any arbitrary first order astigmatism can be corrected. The
sensitivity of the stigmator is proportional to the diameter of the beam at the plane
of the stigmator. Therefore for a divergent beam entering the objective lens, the
stigmator ideally should be placed close to the final lens. In practice, the stigmator
is implemented using the deflection system already in place. When multiple stages
of deflection are used, the stigmator should be implemented on the first stage of the
deflection, where the beam is closest to the axis.

Dynamic deflection is also used to view the GFIS emission pattern itself, rather
than the specimen. The concept is to scan the beam on the aperture plane, and map
the angular distribution of rays emitted from the source by correlating the ray that

**Fig. 1.9** Quadrupole fields
for astigmatism correction
passes through the aperture with the deflection angle. As a result, the emission pattern of the GFIS can be obtained. The magnification of such recorded emission pattern is determined by the amplitude of the deflection and position of the cross-over point.

**Beam Blanking**

The ion beam can be blanked with a sufficiently large deflection field and a blanking aperture, or “beam stop”. Therefore (1.7) also describes the angular deflection of an electrostatic blanker. The primary consideration for a beam blanker is the speed. The blanking or unblanking time is related to the rate at which a required voltage put on the blanker electrodes, and the required voltage is determined by the blanker sensitivity. Consideration must be given to the geometric design of the blanker in order to provide the speed demanded by the specific applications.

A secondary consideration in blanker design is to manage the blanking tail. The term, blanking tail, refers to the beam movement on the specimen while the blanker changes states (from unblanked to blanked or vice versa) which often results in an unwanted partial beam exposure in an adjacent area. Increasing the blanking speed (e.g. increasing the voltage slew rate) can help to minimize this effect. But even for an infinitely fast blanking, there are also cases in which the transition occurs while the ion is within or adjacent to the blanker during the transition. These effect can be largely mitigated by choosing the location of the blanker so that its center is at or near the cross-over position of the beam. The effects can also be mitigated by increasing the blanking sensitivity so that only a small voltage is required to deflect the beam into the beam stop and prevent it from reaching the sample.
Time Delays in Ion Columns
To form an accurate image or to perform precise patterning, the applied deflection voltage corresponding to a pixel in the image needs to be precisely matched to the physical landing position of the ions. Further, the corresponding landing position needs to be matched to the signal collected in the detector (in imaging applications). Various of dynamic components in the column and signal chain should be precisely orchestrated in time to achieve this spatial–temporal accuracy.

Ions have finite mass and kinetic energy and therefore their time-of-flight (TOF) through the column are non-negligible and often dictate the timing of various dynamic components. In Fig. 1.10a, an ion is shown to start from the end of the blanker, first reaching the center of the scanning deflector after $\Delta t_{TOF1}$, then landing on the specimen after another $\Delta t_{TOF2}$, then the secondary electrons reaching the detector after another $\Delta t_{TOF3}$, and finally a pulse of signal is received at the end of signal chain after a detector delay $\Delta t_{DET}$. In some column designs where the length of the deflector itself is quite long, the TOF through the different deflection stages should also be considered.

A timing diagram of three key dynamic components in the column and signal chain is shown in Fig. 1.10b to show how these TOFs are used in controlling component timing. As an example, suppose the clock starts when the blanker is unblanked. Then after a delay of $\Delta t_{TOF1}$, the deflection voltage is applied. (For

**Table 1.2** Time-of-flight of He, Ne and Ga ions through 100 mm at various beam energies

<table>
<thead>
<tr>
<th>Ion species</th>
<th>30 keV</th>
<th>10 keV</th>
<th>1 keV</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>83 ns</td>
<td>144 ns</td>
<td>455 ns</td>
</tr>
<tr>
<td>Ne</td>
<td>187 ns</td>
<td>323 ns</td>
<td>1020 ns</td>
</tr>
<tr>
<td>Ga</td>
<td>347 ns</td>
<td>600 ns</td>
<td>1900 ns</td>
</tr>
</tbody>
</table>

**Fig. 1.11** HIM image of the surface of a pancreatic cell membrane (courtesy of Prof. Paul Walther of the Univ. of Ulm, Germany)
simplicity of discussion, the rise time of the deflection voltage itself is ignored.) Then after another delay of $\Delta t_{TOF2} + \Delta t_{TOF3} + \Delta t_{DET}$, the signal received at the end of signal chain is counted as the pixel intensity corresponding to the applied deflection signal. In the patterning application, it is important that the particles going through the deflection see a constant deflection voltage, meaning the pixel dwell time cannot be lower than the TOF through the deflector.

Table 1.2 shows some example TOFs of He, Ne and Ga ions travelling a distance of 100 mm. It is clear the for heavier species, such as Ne or Ga, the TOFs at medium to low energy is quite signification. Careful timing control is needed to accurately manipulate the landing positions of those beams.

### 1.7 Beam Induced Damage

The HIM is known to provide valuable images because of its small focused probe size, its long depth of focus, and the variety of detectable particles which convey unique contrast mechanisms. Figures 1.11 and 1.12 below show some of the high quality, high magnification images.

On the sample surface and beneath the surface, the helium beam is known to produce damage (surface sputtering, charge transfer, implantation, and lattice damage) [19]. In light of this, it is regarded as very important that the imaging system be optimized to gather the most information for a given exposed dosage of ions. In most circumstances, the adverse effects accumulate with each incident ion and are dependent upon an areal dosage, $\sigma$, (sometimes just called dosage) with units of ions/cm$^2$. Some effects that dissipate with time (e.g. charging or thermal effects) may have dependence upon the areal dose rate which usually expressed in
ions cm\(^{-2}\) s\(^{-1}\), but it is generally understood that most adverse effects are rate independent. Thus, for the GFIS imaging systems, the goal of optimizing the detector’s efficiency is described as maximizing the signal to noise ratio SNR for a given areal dose, or maximizing the “information to damage ratio”. The same goal applies to detectors for SEMs and TEMs, but the damage there is less pervasive due to the lower mass of the incident particle. The same goal applies to gallium FIBs, but is less critical since these instruments are used explicitly for their damage capabilities and are seldom employed for high resolution imaging. So we find that for the HIM, the efficiency of the detector, and the signal chain, are matters of special importance and worthy of optimization.

Regardless of the type of detected particle (e.g. secondary electron, secondary ion, backscattered helium, transmitted helium, etc.), producing a “good” image requires that there be a sufficient number, \(N_d\), of detected particles per pixel in the final image. There are several probabilistic factors that contribute to the detection of the \(N_d\) detected particles: The arrival of ions from the source is believed to follow Poisson statistics. The emission of many secondary particles (such as secondary electrons or secondary ions) is also governed by Poisson processes. Further, the detectability of the particle and the signal generation process all introduce their own probability functions. Appealing to the central limit theorem, we suppose that all of these random factors can be combined to give an overall image signal that is roughly Gaussian with a mean value that is proportional to \(N_d\) and a standard deviation that is proportional to \(\sqrt{N_d}\). Thus, the signal to noise ratio, SNR is presumed to follow the general form:

\[
SNR = \frac{\text{signal}}{\text{noise}} \sim \frac{N_d}{\sqrt{N_d}} = \sqrt{N_d}.
\] (1.9)

If we define \(N_i\) as the average number of incident ions per pixel, and \(Y\) as the effective yield of the detected particle per incident ion, such that \(Y = N_d/N_i\), then

\[
SNR \sim \sqrt{N_i} \sqrt{\frac{Y}{1 + Y}}.
\] (1.10)

The simple interpretation of this result is that the SNR ratio is principally limited by the smallest number in the chain of events that leads to the detected signal. For example, if the yield, \(Y\), is appreciably larger than 1, as for secondary electrons, then the SNR is primarily determined by the Poisson statistics of the ion arrival events, and further increases in the yield do not significantly affect the SNR. Whereas in another situation, if the yield of backscattered ions is much less than one, then the sensitivity to yield is still significant.

As a rule of thumb, the SNR should be about 5 to have an easy to interpret image. In the case of SE imaging, where \(Y\) is typically 3–5, then this requires about \(N_i \sim 25\) incident ions per pixel. Alternately, if the backscatter yield is 0.1, then the SNR rule of thumb would suggest that 250 incident ions are required per pixel. The dependence of areal dosage, \(\sigma\), on the field of view, \(FOV\), is expressed as:
The general scaling of this relation is shown in Fig. 1.13. Eliminating $N_i$ in favor of the SNR with help of (1.10) we can arrive at a form which is arguably more important to the microscopist:

$$\sigma = \frac{N_i}{FOV^2}. \quad (1.11)$$

The microscopist must satisfy the condition, $\sigma < \sigma_{\text{critical}}$, so that the areal dosage is kept below the level that will result in undesired effects. The threshold, $\sigma_{\text{critical}}$, depends very much on the materials in question, but even more so upon the particular application’s tolerance to damage. For example, for materials that are already amorphous, it is probably not important if they are amorphised by the incident ion flux. And when imaging thin sections, sub-surface implantation is not a concern.

From the previous equation, it is now relatively plain to see that decreasing the FOV by a factor of 3 (say from a 3 to 1 \( \mu \)m) will subject the sample to about $10 \times$ more areal dosage if the SNR is to be maintained. Similarly, it is also evident from the above equation that generating images from low yield particles (secondary ions,
or backscattered helium for example), where $Y \ll 1$ will require either larger fields of view, or lower SNR, or higher areal dosages. Note that averaging techniques, such as line averaging or frame averaging, offer no advantage (with the afore-mentioned provision that dose rate is irrelevant). Note also, that for very small FOVs, the relevant term in the denominator must be increased to reflect the characteristic width of the damage mechanism, $w_{dmg}$:

$$\sigma = \frac{(SNR)^2}{FOV^2 + w_{dmg}^2} \left( 1 + \frac{1}{Y} \right)$$  \hspace{1cm} (1.13)

\section*{1.8 Detectors and Signal Chain}

The above analysis serves to motivate the instrument designer and the operator to ensure that all available information is acquired for a given areal dosage, e.g. maximizing the information to damage ratio. Since secondary electrons (SEs) are known to be produced in abundance from the helium ion beam, and because they can be used to generate high resolution with valuable contrast, the ET detector which collects them should operate with high efficiency. What sets this apart from the same exercise in the SEM is the relatively high helium induced SE yield (typically 3–5), the relatively low probe currents (often 0.5 pA or less), the lower
SE energy distribution (FWHM < 2 eV) [20], and distinctly different nature of the sample charging, and the relatively low contribution of SE2 and SE3 (which rely upon backscattering as an intermediate step).

The in-vacuum components of the HIM’s ET detector are only slightly different from the equivalent detector in an SEM. The distinctions include an ET grid which has a fine and highly transparent mesh, and a scintillator that is more consistent and uniform in terms of its photon yield. The grid is commonly biased to +500 V, and the scintillator is usually set to +10 kV (Fig. 1.14). Modeling and experiments confirm that the collection efficiency can be as high as 90%. The ET grid is also designed to be biased negatively to repel the electron flood beam when it is periodically used to dissipate the accumulated positive surface charging encountered with insulating samples. Generally, the best collection occurs when the electric field from the ET grid is strongest at the beam’s landing position. This occurs for example when the sample is tilted somewhat, and when sample being imaged is at the nearest the edge of the sample holder. Biasing the sample to several volts positive often has an appreciable benefit.

For each electron striking the scintillator it is expected that over 100 photons are sent through the light pipe, and a significant fraction of these are transferred to the opposite end, outside the vacuum chamber (Fig. 1.15). The light is filtered to eliminate any low wavelength components which represent stray light from the flood gun’s filament or the stage encoders. Beyond this, the PMT is selected for its high quantum efficiency photocathode to minimize the chance that any information

Fig. 1.15 The out-of-vacuum components of the ET detector. The PMT provides the necessary gain, and the head amplifier converts the signal to an easily measured voltage pulse.
will be lost at the first strike. The output of the PMT is in the form of a single current pulse (~7 ns width) corresponding to each SE arriving at the scintillator. After passing through a trans-resistance amplifier, the pulse is stretched to about 200 ns FWHM so that it is shaped appropriately for the subsequent sampling at regular time intervals. Since a single incident ion will usually generate a few SEs, and since they arrive at nearly the same time, the resulting pulse height is largely an indicator of the varying SE yield across the sample. Considering that the SE yield is nearly always greater than unity, pulse counting is not of much practical benefit—except to study the emission statistics of the source! So of the two complimentary schemes, pulse frequency provides almost no information about the sample, whereas pulse amplitude provides the richest sample information.

The signal output from the head amplifier is passed to the Scan and Acquisition System (SAS) board where it is first offset with the adjustable brightness (Fig. 1.16). The analog signal is always sampled with a 100 ns sample clock and converted to an 8 bit digital value. This is where it is critical that the pulse shape is well chosen. If the pulse shape is too narrow it could fall between the consecutive samples, and convey no information. If the pulse shape is too wide it will result in too many consecutive samples, possibly causing a smearing in the resulting images. The pixel processor collects all the 8 bit sampled values which fall within the chosen pixel dwell and combine them by either averaging, or integration, or some intermediate blend. For example, if the pixel dwell is chosen to be 20 μs, then 200 samples will be acquired at the 100 ns sample clock and combined to produce a single 8 bit number. The parameter “image intensity” allows the user to determine if the computed result for each pixel is more alike to an average value, or more like an integral. The computed result for the indicated pixel is then stored to memory where it will be transferred to the PC and its displayed image. Optionally, with line averaging or frame averaging enabled, the signal measured for repeated beam visits to the same sample location can be averaged to progressively reduce the overall noise for the corresponding pixel.
in the image. Once the image is transferred to the PC, further digital analysis is available via the histogram, and further manipulations can be conducted via look up tables. When saved, each TIFF image has an integrated metafile that contains all the instrument settings that might be helpful for documenting the imaging conditions.

It is sometimes suggested that the 8 bit analog to digital conversion is not sufficient to represent the dynamic information presented in a given 100 ns interval. But it is important to recognize that each pulse that is converted conveys only one piece of information: the number of secondary electrons arriving at that 100 ns interval. And given the very low probe currents (say 0.5 pA), the vast majority of the 100 ns intervals contain zero signal. When a non-zero signal is eventually measured, the gain would nominally be set so that the largest possible conversion (255) corresponds to an extreme event where say 10 secondary electrons were detected within a single 100 ns interval. By this analysis, the 8 bit A/D conversion is seen to be adequate. The subsequent pixel processing is conducted with 24 bit resolution for each pixel, before yielding the final result that is again reduced to 8 bits. While these final 8 bits are adequate for most static and linear viewing situations, there is a growing desire allow for more dynamic range—beyond what can be conveyed in 8 bits.

Finally, the importance of the “image intensity” can easily be overlooked, but it has considerable impact in the final image quality. And because the typical SEM has no equivalent imaging parameter, it will be described in detail here. The circumstances that are unique to the HIM and require image intensity are (a) the very low probe currents that cause the ions to arrive at an average interval that is much longer than the sample clock, and (b) the secondary electron yield which is normally appreciably larger than unity. Consider the case where the pixel dwell is set to 20 μs, and image intensity is set to 0, so that pixel processor computes a true average of the 200 samples for that pixel. Given the low probe currents, it is likely that perhaps only 40 samples will have measurable pulses. And if the PMT gain is chosen so that the average pulse height is converted to the intermediate digital value of 128, then a true average would give a grey level of just 26 (= 40 × 128/200).

The operator seeing this image, would of course increase the gain by a factor of about 5, causing the signal to rise above the maximum A/D conversion level—effectively saturating the A/D convertor, but providing no evidence in the image to the operator. An astute instrument operator could also interrogate the raw signal by briefly changing the dwell to 100 ns, and the image intensity to 0, so the PMT gain can be properly set to avoid the now visible saturation. The image intensity serves to boost the result of the pixel processor to make the operator more likely to recognize the onset of saturation. Given the low probe currents, it is not uncommon to choose a default image intensity of “23.3%” which corresponds to a signal boost of about 5×. For less frequently detected particles, it is not uncommon to choose higher value of image intensity. Whereas for very high currents (e.g. when the detector pulses are overlapping) it is acceptable to set the image intensity to 0 for a true average.
1.9 Vibrational Considerations

All charged particle microscopes (CPM) that use a scanned, focused probe are susceptible to image vibrations because the image is acquired one pixel at a time. Image vibrations can be caused by all effects that cause the emitted beam to land in the wrong location by some time-varying amount. Given the very small probe size of the HIM (0.35 nm for example), it is important that the system vibrations be small enough to permit the expected high magnification images. For any scanned probe microscope like the HIM, vibrations that are evident in the images are always the result of three necessary factors: There must be (a) a source of energy that feeds the vibrations, (b) a flexing element between source and sample, and (c) a transmission pathway between the energy source and the flexing element. Eliminating any one of these three factors will eliminate that particular image vibration.

For the GFIS gun in particular, vibrations at the source are a significant issue since the total emitted current and the angular current density are relatively low, and hence the option of discarding current to achieve source demagnification is generally not viable. Under typical conditions, the overall column magnification from source to sample is in the range of 0.3–1.5 meaning that vibrations at the emitter are directly visible in the images, just as sample vibrations would be. (In contrast, this situation is quite different for a Schottky thermal field emitter where the source is typically demagnified by ×20 or more.)

The first approach for vibration immunity is to design the GFIS gun, column, and sample stage to be as rigid as possible. Unfortunately, the GFIS gun cannot be rigidly mounted to the underlying vacuum chamber, but rather is attached with the flexible vacuum seals to allow it to be shifted and tilted in order to aim the emitted beam along the axis of the column. In effect, the GFIS emitter is mounted on a translation and tilt stage with four degrees of freedom at the top of the microscope. At the opposite end of the microscope, the sample is usually mounted on its own 5 axis stage making a total of 9 degrees of freedom from source to sample, and leaving the system vulnerable to flexing components. As a further engineering challenge, the GFIS gun must be supported by relatively long, thin, thermally insulating supports to reduce the heat transport from the outside world (at room temperature) to the GFIS emitter (at about 70 K).

Addressing the energy sources is another approach. A good microscope will avoid having sources of vibrational energy too close to the critical components of the microscope. Cooling fans, circulated cooling water, 50 or 60 Hz power, and boiling cryogens can be avoided. However the demanding requirements for ultra-high vacuum despite a high flux of noble gases requires the use of turbo pumps. These necessarily introduce vibrations into the system at either the rotor spin frequency (typically 800 Hz to 1 kHz), or its harmonics. The newer models of magnetically levitated turbo pumps are generally preferred, but tend to introduce additional vibrations across a broader frequency range. Energy sources that arise from the environment in the form of floor vibrations, AC fields, and acoustic energy are additional risks. The instrument has a limited tolerance for these, since they can
excite image vibrations, so the installation requirements strictly prescribe a site with low energy levels to achieve the highest image quality.

As a last resort, there is the general approach of interrupting the energy source before it can excite the flexible element between the source and the sample. The commercial HIM instruments have all been built upon a massive granite slabs that are floated on an air isolation system. This helps to intercept and attenuate any floor vibrations, and helps to add inertia to the microscope making it less susceptible to acoustic energy sources. In turn, the granite supports the chamber/column/gun stack by way of another air isolation system. If the site survey measurements indicate that the desired installation site is unsuitable, the customer has other options to help mitigate the facility’s problems. If 50 or 60 Hz magnetic fields are an issue, there are commercial packages that can apply opposing magnetic fields to create a field-free region in the vicinity of the column. If floor vibrations are a problem, they can generally be attenuated by introducing an intervening passive or active support structure that is commercially available through third parties. And if the acoustics are deemed to be unacceptable at a specific site, there are acoustic enclosures already designed for the HIM that can greatly diminish to the sound pressure waves and their impact on the image.

1.10 Conclusion

Like the proverbial reference to “shoulders of giants” the HIM relies upon a foundation of several key technologies and sub-systems. The HIM concept was certainly recognized by Muller shortly after 1955 when he recognized that the spots of light on his phosphor corresponded to the individual atoms on the apex of his pointy needle. For the next half century, the critical sub-systems and enabling technologies matured to they could sufficiently support the instrument we’ve come to call the HIM. These critical technologies outlines in this chapter include cryogenics, high vacuum, gas purity, vibration control, ion optics, and detectors.

References

1. The HIM is presently commercialized by Carl Zeiss Microscopy LLC, as the ORION, ORION Plus, and ORION NanoFab


16. In the world of GFIS, the imaging species are regarded as noble and good, while the residual vacuum constituents which corrupt the emission process are jokingly regarded as “peasant atoms”. Hence the idiom, “adatoms are bad atoms”.


Helium Ion Microscopy
Hlawacek, G.; Gölzhäuser, A. (Eds.)
2016, XXIII, 526 p. 320 illus., 204 illus. in color., Hardcover
ISBN: 978-3-319-41988-6