Instrumentation and contrast mechanisms in scanning helium microscopy

Matthew Bergin

University of Cambridge
Fitzwilliam College

December 2018

This dissertation is submitted for
the degree of Doctor of Philosophy

Abstract

Scanning helium microscopy (SHeM) is a novel form of microscopy that uses low energy (5-100 meV) helium atoms to image the surface of a sample. Since helium is inert and neutral, it can be used to study delicate and insulating surfaces, but that also means it can be difficult to manipulate for focussing and detection. The first reflection mode helium images were recently demonstrated using a pinhole to form a narrow beam of helium, but without any focussing of the beam the low signal levels pose a significant challenge. This thesis aims to advance the field by describing how to improve the instrumentation for a helium microscope and how both previously observed and novel contrast is formed in images.

The thesis begins with an introduction in chapter 1 to motivate the development of a helium microscope. Chapter 2 contains a literature review of beam formation methods and successful implementations of helium microscopy to date.

In chapter 3, the design of the scanning helium microscope is explained, with a description of how images of a surface can be formed with an atom beam. The key aspects of the machine are provided, including a new sample stage that uses a magnetically assisted kinematic mount and measurements of the size of the helium source.

When using a pinhole to collimate the beam, the quality and resolution of images are largely determined by the performance of the detector. In chapter 4, the design, implementation and properties of a new helium detector that has both a high efficiency and low background are detailed. The detector is comprised of a solenoidal ioniser, magnetic sector for mass selection and finally a conversion dynode and electron multiplier for measuring the ion current.

In chapter 5, the origins of the observed contrast in images from a scanning helium microscope are discussed. For rough surfaces, the helium atoms are diffusely scattered from the surface and the resulting contrast mechanisms are investigated. The second half of the chapter focusses on new contrast mechanisms. By using a mixed gas beam, it is possible to simultaneously probe a surface with atoms at different energies and wavelengths, and the first such images are presented. It is well known that atom beams will diffract from well-ordered surfaces, and diffraction contrast on LiF is explicitly shown to be possible for the first time with a scanning helium microscope.

Finally, the optimum geometry to maximise the flux using either a pinhole or zone plate to form the beam is investigated by performing a constrained optimisation in chapter 6. The source properties must also be included due to the strong chromatic aberrations present when using a zone plate. It is shown that the pinhole produces the largest flux for large helium spot sizes, but for small spot sizes the zone plate gives the highest flux and therefore the best signal to noise ratio in images, thus indicating the direction for future research.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. It does not exceed the word limit prescribed by the Degree Committee for the Faculty of Physics and Chemistry.

Matthew Bergin
December 2018

Acknowledgments

Firstly, I would like to thank my supervisor, Dr. Andrew Jardine, for his advice and support. I am also indebted to Dr. Bill Allison, who has provided me with invaluable advice. Both Andy and Bill have supervised me since I was an undergraduate, and many years later I still feel that I have much to learn from them both. I would also like to thank Dr. John Ellis, who has provided helpful input on almost all aspects of my thesis.

I must also thank Dr. David Ward for his help and initial introduction to vacuum science, and Sam Lambrick for the many insightful discussions on all things helium microscopy, as well as his assistance with running the machine and experiments.

Much of the work in this thesis relies on precision instrumentation that was built by many technicians in the Cavendish workshop, and I am extremely thankful for all the work they have done. In particular, I would like to thank Nigel Palfrey for his expertise and teaching in the student workshop.

I would like to express my gratitude to Professor Gil Alexandrowicz and his group from the Technion for their contribution to the design and construction of the detector. I am also grateful to Gil for his proposal of and assistance with the mixed beam experiment. I am also thankful to Dr. Donald Maclaren for both his expertise and for producing the pinholes for the microscope, and Peter Atkin for his assistance in using the white light interferometer.

For financial support, I thank the EPSRC, the NEMI project and The Worshipful Company of Leathersellers.

My thanks extends to the rest of the Surface Physics group including, Fulden Eratam, Dr. Peter Townsend, Dr. Jianding Zhu, Dr. Nadav Avidor, Dr. Susanne Schulze, Dr. Toni Tamtögl, Dr. Irene Calvo Almazan and many other students for the many interesting discussions throughout my time in the group.

I would also like to thank all the friends I have made along the way at Fitz; you have made my time in Cambridge very special. Finally, I would like to thank my parents, Bernadette and Paul, along with Joseph and the rest of my family for all of their love and support.

Contents

Chapter 1

Introduction

Microscopy is an essential tool for studying the properties of materials, and it has been used extensively across all of science because of its ability to resolve structure that is not directly observable by the eye. Micrographia by Hooke in 1665 [1] was the first significant publication on microscopy and was hugely influential, coining the term cell and providing detailed illustrations of various biological specimens, such as a flea. Vast improvements in resolution have subsequently allowed investigations of structures much smaller than insects, with recent studies allowing the direct measurement of the nature of bonding between atoms in a molecule [2].

Microscopes can be divided into two general classes: those that illuminate the sample with a beam, and those that use a physical scanning probe. Scanning probe microscopies [3, 4] raster a small probe across the surface of a sample and monitor the interaction of the probe with the sample. The resolution of scanning probe microscopies is set by the size of the interaction region between the probe and the sample, which can be sufficiently small to allow atomic resolution. Using a physical tip does have some disadvantages though including a reasonably small imaging area and difficulty in observing structures with a large height variation over a small lateral distance.

Microscopes that use a beam have a fundamental limit on the minimum resolvable feature size, d_A,which is given by the Abbe diffraction limit,

$$d_A = \frac{\lambda}{2n \sin \theta}, \tag{1.1}$$

where λ is the wavelength of the probing wave, n is the refractive index of the medium the wave is travelling in and the wave is converging to a spot with a half-angle of θ. While it is possible to overcome the diffraction limit on resolution by using a near field technique [5], it is usually quite difficult. Since the denominator of equation 1.1 will be approximately two at best, the wavelength of the probe is the main factor in determining the minimum resolvable feature.

Given the Abbe diffraction limit, the minimum resolvable feature in optical microscopy is about $1\,\mu\mathrm{m}$. X-rays have a much smaller wavelength than visible light, and therefore can achieve superior resolutions when compared to optical techniques. Projection X-ray imaging is conventionally used in medical applications, but high resolution imaging is achieved through scanning X-ray transmission microscopy (STXM). By focussing the X-rays to a small

point and rastering across a sample, a very high resolution can be obtained with light [6]. The small wavelength of X-rays comes at the expense of a high energy, meaning that X-rays can cause radiation damage. Additionally, X-rays are weakly interacting, so they will primarily interact with the bulk of the sample, making investigations of a surface difficult.

Matter waves have a much shorter wavelength than visible light and therefore can be used to study features below about $1\,\mu\mathrm{m}$. The advent of electron microscopy opened up a plethora of new materials that could be studied due to the enormous improvement in the resolution when compared to optical microscopy. Transmission electron microscopy (TEM) was the first form of electron microscopy that was demonstrated by measuring the flux of electrons transmitted through a thin specimen. Images from a bulk material can be obtained with scanning electron microscopy (SEM), where high energy electrons are incident on a sample and energy is lost through the emission of secondary electrons, photons or high energy backscattered electrons. Any of the emitted species can be collected to form an image of the sample, with each emitted species revealing different information about the sample.

The de Broglie wavelength, λ, of a matter wave in terms of its kinetic energy, E, and particle mass, m, is given by

$$\lambda = \frac{h}{\sqrt{2mE}}. \tag{1.2}$$

So for an electron beam to resolve atomic structure, the wavelength of the matter wave would need be at least $\lambda = 1\,\text{Å}$ and therefore the beam energy would need to be given by $E > 100\,\text{eV}$.

Since electrons are charged, they can be manipulated using electromagnetic optics to form extremely small spot sizes on a surface. Due to the wavelength of an electron beam being low, the Abbe diffraction limit can be pushed below atomic resolution and electromagnetic lenses can be used to form very intense spots on the surface of a material. The optics in a scanning electron microscope usually require the beam energy to be in excess of 1 keV to produce the best quality electron beam.

The high energy of the electrons leads to the electron beam interacting not only with the surface of a material, but forming a tear-drop interaction volume underneath the surface. While the spot size on the surface can be extremely low, the electrons will travel into the bulk and continue interacting with the sample. Transmission electron microscopy avoids the bulk interaction by only using thin samples, however usually the electron beam interacting with the bulk will limit the resolution of a scanning electron microscope.

The electron beam can also damage the material, destroying the structure that is to be studied. Electron beam damage can occur through a variety of different processes that are specimen specific [7]. Examples of beam damage include the breaking of bonds in a polymer to change the structure, knock-on damage in metals that creates a vacancy in the lattice and changes to the surface chemistry.

Additionally, the electron beam is charged, which means that the surface of the sample needs to be conducting to avoid charging artefacts in images. One method to mitigate the issue when imaging insulating surfaces is to use an environmental scanning electron microscope (ESEM) [8]. In conventional SEM, the sample is kept under vacuum to prevent both the primary beam and the secondary electrons from scattering from air molecules. An ESEM maintains the sample in a gaseous environment, which prevents the build up of charge on the surface on an insulator. There are disadvantages to using an ESEM though; a

short path length must be maintained to ensure the electron beam reaches the sample, and the presence of positive ions, to maintain surface neutrality of the sample, may influence the imaging.

Therefore, in general, electrons are often not particularly suited to studying delicate surfaces due to the high beam energy and lack of exclusive surface sensitivity. Instead, an alternative beam is required for the study of surfaces at low energy without disturbing the structure.

Given the de Broglie wavelength in equation 1.2, the same wavelength of a beam can be achieved at lower energies by using a higher mass beam. Therefore, if the objective were to study the surface of a material at a resolution below that which can be achieved with optical microscopy, it may be appropriate to use atoms as a probe instead. While neutrons can be used as a probe, they are weakly interacting with matter and can only study the bulk structure of a material [9]. Neutron imaging also requires a high flux of neutrons, which usually can only be provided by a nuclear reactor or particle accelerator.

Microscopy with atoms is not a new idea and has been demonstrated in a different form in 1969 by King and Bigas [10], who showed that an image of a surface can be obtained by heating a sample and monitoring the atoms that evaporate from the surface. King and Bigas suggest it could be possible to form an image by scattering atoms from the surface, though it was some time before it was demonstrated.

Metastable atoms are atoms that have been excited out of the ground state, but remain in an excited state for a significant period of time. Microscopy using metastable atoms has been shown to be possible, where the metastable atoms release stored internal energy into the surface, releasing electrons that provide information on the electronic structure [11, 12]. The kinetic energy of the metastable atoms means that only the surface electronic structure is probed, but the large energy exchange when the metastable atom de-excites will still perturb delicate sample surfaces.

A beam of helium atoms with a wavelength $\lambda = 1$ Å has an energy of $E \approx 20$ meV, which is about the same as the thermal energy. Using a higher mass than electrons means that it is possible to obtain a beam with a wavelength suitable to probe length scales down to the atomic level with a much lower energy.

Thermal energy helium atom beams are exclusively surface sensitive, giving helium scattering an advantage over other techniques such as electron and x-ray scattering for surface studies. Helium atoms directed at an object will interact with the potential of the surface, which is qualitatively similar to the Lenard Jones potential. At long range, the atoms will be attracted to the surface due to a van der Waals force, but at short distances Pauli repulsion will dominate and the atoms will be repelled. Therefore for the beam energies that are used, the helium atoms will have classical turning points 2-3 Å away from the surface atom cores [13]. The turning point is well above the surface atom cores, meaning that the beam will only interact with the outermost electrons.

Helium atoms also have a giant cross section for scattering from isolated adsorbates on a surface, with helium scattering apparatus having a sensitivity down to 0.001 ML [13]. Helium atoms interact with the electronic density of the surface, so a deformation in the potential on an otherwise flat surface will cause some atoms to be reflected off specular. The extreme sensitivity to defects means that a helium microscope has the potential to study very subtle changes to the structure of a surface.

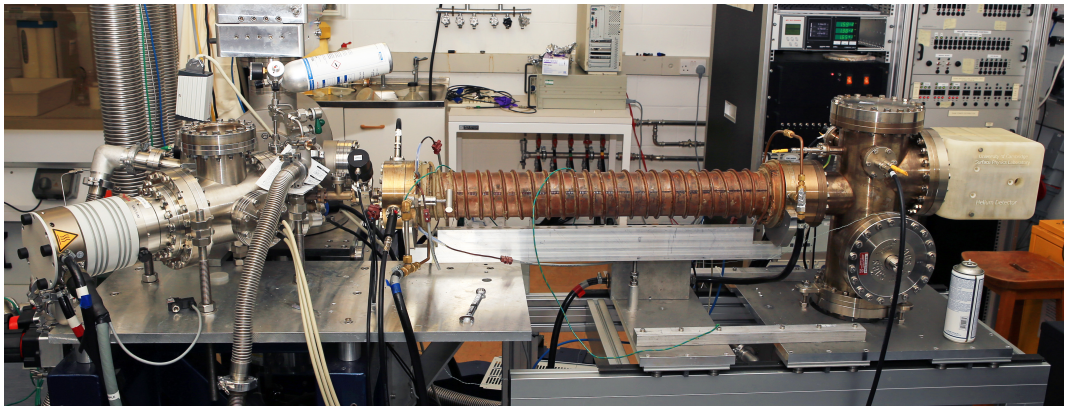


Figure 1.1: Image of the scanning helium microscope (SHeM), showing the source and sample chamber on the left, and the detector on the right. Photograph taken by Paul Bergin.

Helium atom scattering has exploited the properties detailed above to study surfaces in reciprocal space without spatial resolution for some time. The wavelength of a thermal energy helium beam means that diffraction effects are dominant, such that the scattering distribution can be analysed to predict the structure of the unit cell of the top layer of atoms on a surface [14]. Inelastic scattering has been used to study energy exchange processes, such as the creation or annihilation of phonons [15], and provides information about the force constants between atoms. Quasi-elastic scattering is another technique which can be employed to study surfaces using helium scattering. Processes which change the energy of the helium atoms will cause the elastic peak to be broadened. By measuring the specular peak width in the energy domain, information on the diffusion processes taking place on the surface can be extracted [16]. The spin-echo technique [17] utilises an NMR-like spin precession method to measure the quasi-elastic energy broadening and is used to study diffusion phenomena.

However, the majority of the applications of helium atom beams have not been spatially resolved. Usually, a large uniform sample is illuminated with a wide beam of helium to study the surface in reciprocal space. All the properties of helium atom beams that have been useful in reciprocal space studies can be equally applied to a spatially resolved machine.

Thesis overview

Scanning helium microscopy is a novel technique that forms an image of a surface by scattering a narrow beam of helium atoms from a sample surface. The atoms are then detected using a solenoidal ioniser and magnetic mass sector. Figure 1.1 is an image of the machine that has been constructed in Cambridge, with the source and sample chamber on the left of the image and the detector on the right. This thesis addresses the major research issues that are currently outstanding with a microscope that uses atoms.

Chapter 2 is a review of the potential techniques that can be used to build a microscope which uses atoms. In general, neutral atoms can be manipulated through a variety of methods, but it is shown that most are not applicable to ground state helium atoms. Previous efforts to develop scanning helium microscopes are also discussed, with the pinhole method being the most successful to date.

Chapter 3 describes the helium microscope in Cambridge and explains how images are produced from a helium beam formed from a pinhole. The chapter also discusses the sample mount that has been developed for the machine which uses a kinematic mounting principle to provide repeatable repositioning between sample changes. An important parameter that determines the size of the helium beam is the size of the source of helium, and chapter 3 also contains measurements of the virtual helium source.

When using a pinhole to form a helium beam, there is no focussing and therefore there is a low flux of helium in the spot that is incident on the sample. It is therefore necessary to have a high efficiency detector with a low background to measure the small flux of helium that is scattered from the sample. Chapter 4 describes a high efficiency and low background detector for helium atoms based on electron ionisation that was developed for the microscope. The design and implementation of the detector are discussed, alongside the properties of the entire device.

The contrast observed in images obtained with a helium microscope and the mechanisms that generate the contrast are discussed in chapter 5. For rough surfaces, the atoms are diffusely scattered from the surface, leading to contrast from self-shadowing and masking on the surface. Contrast also arises on rough surfaces from diffuse illumination of the surface and variations in the surface normal. The second half of the chapter concentrates on novel forms of contrast that have not previously been observed. A mixed gas beam is used for the first time to simultaneously image a surface with two different atoms with differing wavelengths and energies. The first observation of contrast in a helium image due to diffraction from a surface is observed from a LiF crystal. By varying the distance between the sample and pinhole and comparing the signal obtained from LiF to a normal rough surface, significant structure is observed in the scattering distribution.

Given that the measured signal in a helium microscope is low, optimisation of the geometry of the device is key to maximise the performance of the microscope. Chapter 6 describes a constrained optimisation to maximise the flux produced by either a pinhole or zone plate for a given target resolution. Since the zone plate has chromatic aberrations, and helium sources can either produce bright or monochromatic beams, it is necessary to optimise both the machine geometry and the source properties. A numerical simulation is used to model the properties of a realistic zone plate for a microscope, allowing a model for the beam width and flux obtained from a pinhole and zone plate to be constructed. Lagrange multipliers are then used to perform a constrained optimisation of all the parameters, and it is found that pinholes will produce the highest flux at large beam widths, while the zone plate is the favourable method at low beam widths.

Finally, in chapter 7 the conclusions from this thesis are presented, and an outlook on what future work could be done to continue the development of helium microscopy is also proposed.

Chapter 2

Literature review

A crucial part of producing a scanning helium microscope is forming a narrow beam of atoms to provide spatial resolution of a sample. The existence of lenses in other types of microscopes allows illumination of the entire sample which is then imaged onto a spatially resolving detector. However, as shall be discussed shortly there are no such lenses for ground state neutral helium atoms. So instead, spatial resolution has been obtained by using a narrow beam of helium atoms in a single pixel camera setup. It should be noted though, that there were some initial attempts at developing a spatially resolved detector [18].

Initial work on developing a helium microscope concentrated on developing a method to focus the helium beam to achieve a narrow beam. Due to a lack of sensitive detectors, it was assumed that the best way to achieve a spatially resolved instrument with a high enough flux to be useful, was to focus the beam into a high intensity spot.

An initial review of focussing ground state helium atoms has been performed by MacLaren [19, 20] and there are more in depth reviews for atom optics in general [21, 22]. Since MacLaren performed his review, there have been many significant developments in atom microscopy due to the increased interest in the field.

In order to set the context for this thesis, a review of the existing literature is needed. This chapter begins with a review of the potential methods for focussing ground state helium atoms and discussion of the feasibility of each approach. After, the successful demonstrations of scanning helium microscopy are reviewed, with the relative advantage of each approach considered. Some significant applications of helium microscopy to date are also reviewed.

2.1 Beam formation methods for neutral helium microscopy

2.1.1 Optical manipulation

Optical manipulation of atoms is a fundamental part of cold atom physics and is routinely used to achieve cooling and trapping of atoms [22]. Optical manipulation of atoms relies on photons interacting with an atom with a suitable set of energy levels, such that transitions between the energy levels are possible.

Figure 2.1 is a diagram of the energy levels in helium and illustrates why the electronic structure of helium makes it difficult to manipulate in its ground state using optical tech-

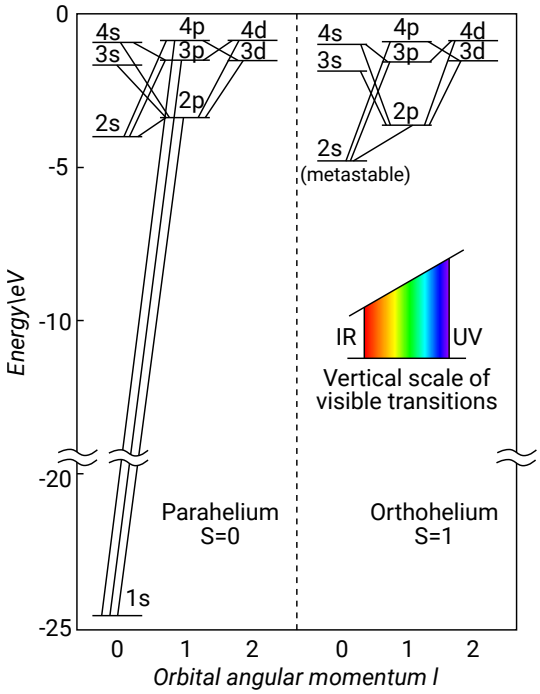


Figure 2.1: A diagram of the energy levels in helium where one electron is in the ground state. The left hand energy levels are for parahelium where the electron spins are antiparallel, while the right is for orthohelium where the spins are parallel. The ground state where both electrons are in the 1s state is much lower than the rest of the states and cannot be accessed by visible photons as illustrated by the added scale. Additionally, the triplet metastable state has also been labelled where decay back to the ground state is doubly forbidden by selection rules. The energy levels are taken from Ahn et al. [24] and the diagram is inspired by a figure on HyperPhysics.

niques. The ground state energy is sufficiently low that a laser would be needed with a wavelength that is deep in the vacuum ultraviolet part of the spectrum ($E > 20\,\text{eV}$, $\lambda < 100\,\text{nm}$) [23]. No appropriate optical source exists that can be used to directly manipulate the helium from its ground state with a high intensity.

Two photon absorption processes can be used to excite molecules into a higher state than it is possible to do so with a single photon, however they are typically orders of magnitude weaker than linear absorption at low intensities [25, 26]. Even with two photons, each would need to have an energy greater than 10 eV and it would still be difficult to find an optical source with a high enough intensity at this energy. It has been proposed that by using a resonant multiphoton process [27], it could be possible to excite ground state helium atoms. Instead of exciting the ground state to a virtual state, a resonant excited state is used, but to date there has been no successful demonstration of the technique.

It is therefore not feasible to use optical methods to manipulate ground state helium atoms. The energy required to excite the helium atom from the ground state to the nearest state is too large for optical techniques. However, if the helium could be prepared out of its ground state through another method it would be possible to optically manipulate the atom.

The metastable triplet state for helium is labelled in fig. 2.1 and can be prepared through electron impact. The metastable state in helium has an extremely long lifetime of about

8000 s due to the direct photon transition to the ground state being doubly forbidden by quantum-mechanical selection rules [28].

Metastable helium can be optically manipulated since there are a set of optically accessible transitions to other energy levels. Noble gases that are exited into metastable triplet states have a much lower energy gap to other energy levels when compared to the ground state. Therefore, it has been shown to be possible to focus metastable helium atoms [29].

However, metastable helium is not suitable for use as a low energy probe since the large amount of internal energy can be released while interacting with a sample. While the incoming particle can have a very low kinetic energy, the metastable atom will typically release its internal energy when it interacts with the surface. Therefore if the objective is to use an extremely low energy probe, metastable atoms are not suitable.

It should also be noted before moving on, that the large energy difference between the ground state and the excited states also precludes focussing with an electrostatic field. Direct focussing of atoms with an electrostatic field is possible if a large enough dipole can be induced in the atoms [30]. However, since the ground state energy is low, the polarisability of ground state atoms is also low, making it difficult to induce an electric dipole in the atom. It was discussed above how the low polarisability prevents the use of the optical dipole force, but it also precludes the use of electrostatic lenses.

Therefore in summary, it is not possible to use optical manipulation of atoms to form a beam for a low energy helium atom beam. The extremely low energy of the ground state makes helium an inert probe, but also leads to it being difficult to manipulate with light because very high energy photons are needed. The large energy difference is connected to a low polarisability of the helium atom which also prevents the use of electrostatic lenses.

2.1.2 Magnetic lenses

Magnetic lenses can be used to focus atoms with a magnetic moment [31, 32]. An inhomogeneous magnetic field will create a force on particles with a magnetic moment which can be exploited to focus an atom beam. However, ground state helium does not possess a magnetic moment and therefore cannot be manipulated using magnetic fields. In the ground state, the electrons have no orbital angular momentum or total spin and the nucleus does not have a nuclear spin.

Excited states of helium can possess a magnetic moment and therefore have been successfully focused with a magnetic field. By using an excited state of helium, the electrons can either have an orbital angular momentum or parallel spins leading to a magnetic moment. A hexapole lens has been used to focus metastable helium to a 3 mm spot [33]. As previously discussed though, metastable helium is not appropriate to be used as a gentle probe on surfaces.

Helium-3 possesses a magnetic moment in its ground state due to the nuclear spin and therefore can be focussed with a magnetic lens. Jardine et al. have used a hexapole lens to produce a 2 mm parallel beam of helium-3 from a diverging supersonic expansion [34].

There are however many issues with magnetic focussing of an atomic beam. The low refractive power of magnetic lenses means that a practical focal length would require a large magnetic field. If the design employed by Jardine et al. [34] were employed, an inconveniently large magnet would then be needed.

Additionally, for a typical hexapole lens, the focal length, f, scales with the atom velocity, v, by $f \propto v^2$ creating a chromatic aberration. Faster moving atoms require a larger impulse to be focussed, but also travel through the lens quicker, causing the lens to have a strong chromatic aberration. While supersonic expansions create a smaller relative velocity spread $\Delta v/v$, the average velocity is much higher so the refractive power is also reduced. Typically, the beam is laser cooled to reduce the chromatic aberrations [32], but as previously discussed this is not possible to achieve with ground state helium.

Another approach that has been proposed to reduce the chromatic aberrations of a hexapole lens is to use a pulsed beam and lens [35]. The pulsed lens ensures that all the time spent in the field is the same for all the atoms, which leads to the focal length linearly depending on the atom velocity ($f \propto v$). By also pulsing the beam and tapering the lens, the field is stronger at the front of the beam where the atoms are moving quickest. The first demonstration of the technique formed a spot size of $(170 \pm 20)\mu\mathrm{m}$ [36], therefore while the spot size is not yet small enough for microscopy, future generations of the lens could become relevant.

However, while it may be possible to focus helium-3 in the future to a suitable spot size, the expense of helium-3 and complexity and size of the lens motivates an alternative method to be developed. Using helium-3 would require the use of a recycling system to be economically viable, which also acts to increase the complexity of the helium source. The progress of hexapole lenses could provide a method for achieving focussing with a low energy atom beam, but using helium-3 will remain expensive and therefore not the ideal option.

2.1.3 Atom mirrors

An atom mirror is another alternative method proposed for forming a focussed helium spot for microscopy. The collection area of the atom mirror can in principle be very large, potentially producing very intense helium spots. In addition, reflection from a surface will not produce any chromatic aberrations, removing the concern regarding about the monochromaticity of the helium beam. There are several potential methods for forming an atomic mirror that shall be discussed in this section.

It is possible to create an atom mirror using either optical [37] or magnetic [38] dipole forces. However, as previously discussed, helium atoms in their ground state cannot be manipulated by lasers and do not possess a magnetic moment. Therefore, these mirrors cannot be used for a helium microscope.

A mirror based on quantum reflections has been demonstrated for helium atoms [39], however there are severe limitations. A quantum reflection occurs when a particle reflects from a potential without reaching a classical turning point. Typically, quantum reflection from a solid surface requires that the wavevector of the matter wave perpendicular to the surface is small which can be achieved either through grazing incidence [39] or at normal incidence with cold atoms [40]. Therefore, there are limits on a mirror that uses quantum reflection and other methods would be more suitable.

The most promising method for produce an atom mirror for helium seems to be from scattering from microscopically smooth surfaces. If a surface is sufficiently smooth, then helium atoms will be specularly reflected. There are strong restrictions on the materials that

can be used as an atom mirror though. On the microscopic level, the surface needs to be flat and ordered enough to produce a strong specular reflection. However, the material also needs to be macroscopically shaped to act as a focussing mirror. Additionally, the surface would ideally not diffract a large flux away from specular so that the efficiency of the mirror is as high as possible.

To demonstrate the difficulty with using a specular reflection from a crystal, the domain size needed for a flat mirror to be useful must be considered. The resolution limit of the mirror will be imposed by the width of the specular peak produced by the atom mirror. Point defects will scatter atoms out of the specular peak causing a reduction in the intensity observed, however such point imperfections will not change the width of the specular peak. The width of the specular peak will be determined by the size of the domain that the atoms coherently scatter from. Following the treatment given by Lapujoulade et al. [41], the scattering domains can be modelled by rectangles of size (l_c, l_c). Coherent scattering from an infinite plane would produce a delta function specular reflection, however the finite domain size will lead to the specular peak being a sinc function. For a given small width of the specular peak, $\Delta\theta$, the domain size, l_c, can be shown to be,

$$l_c = \frac{5.54}{(\Delta\theta)k\cos\theta_f}, \tag{2.1}$$

where k is the wavevector of the incoming beam and θ_f is the outgoing angle from the mirror [41]. The width of the specular peak will need to be given by $\Delta\theta = \sigma/L$, where σ is the desired resolution of the mirror and L is the distance from the mirror to the focal point. Rearranging equation 2.1 then approximately gives,

$$l_c \approx \frac{L\lambda}{\sigma}, \tag{2.2}$$

where λ is the wavelength of the atom beam. So for a room temperature beam with a wavelength of $\lambda = 5.63 \times 10^{-11}\,\mathrm{m}$, a reasonable target resolution of $\sigma = 100\,\mathrm{nm}$ and a distance from the mirror to the focal point of $L = 0.3\,\mathrm{m}$, the domain size is given by $l_c > 170\,\mu\mathrm{m}$. While it may appear easy to relax the condition on the domain size by decreasing the distance from the mirror to the focal point, the reduced distance would require a higher curvature on the mirror making manufacturing of the macroscopic shape of the mirror more difficult.

Therefore to form a mirror that can produce a spot size of less than 100 nm, the mirror must scatter atoms coherently from domains greater than $170\,\mu\mathrm{m}$. Ideally there would be no step edges across the entire domain, however this would require an extremely large, totally perfect crystal structure. Step edges can be present if the path length from the other terrace adds an integer number of wavelengths so that the terraces interfere constructively, which would be a significant experimental challenge and would introduce chromatic aberrations. If the path length is carefully tailored to allow step edges, then the maximum path length difference must not exceed the beam coherence length since then it will not be possible for the beam to interfere. The beam coherence length, D_c, is given by,

$$D_c = \frac{\lambda^2}{\Delta\lambda}, \tag{2.3}$$

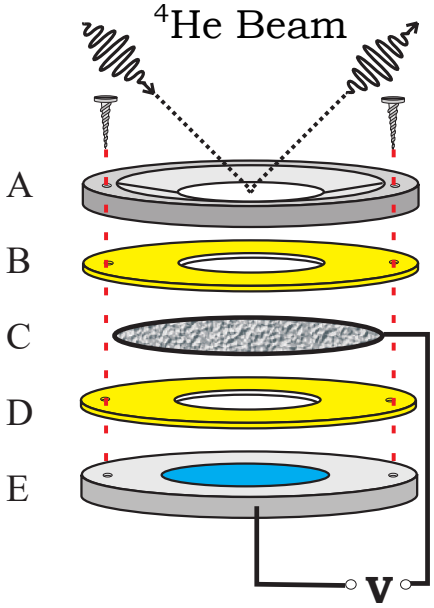


Figure 2.2: Schematic exploded view of an atom-focusing mirror adapted from Fig. 1 by MacLaren et al [20]. Top plate A clamps a thin single crystal (C) between two apertures (B and D). The crystal is grounded and application of a voltage to the lower electrode structure (E) exerts an electrostatic pressure on the free crystal, causing it to deflect through the lower aperture. Boundary conditions of the aperture and electrode dimensions determine the crystal deflection.

where $\Delta\lambda$ is the spectral width of the atom beam [42]. A typical supersonic expansion will give atom beams with $\Delta\lambda/\lambda \approx 100$, therefore the beam coherence length will be approximately 5 nm. So, in order for the domain to produce coherent scattering even if scattering across a step edge is constructive, there must be less than $\mathcal{O}(10)$ step edges across the entire domain. The challenge to create a material with so few steps across a large domain is enormous, given that the mirror also needs to be curved to produce a focus.

Early work by Holst and Allison [43] showed that an atom mirror made from a single crystal of silicon that was $50\,\mu\mathrm{m}$ thick could produce a spot diameter of $(210 \pm 50)\,\mu\mathrm{m}$. Figure 2.2 shows how the surface was deflected into a parabola for focussing by electrostatic deflection thus providing a microscopically flat surface with the macroscopic shape needed. The surface was hydrogen passivated so that it was inert and the reflectivity did not change.

Figure 2.3 shows how the concept was then extended to an ellipsoidal mirror which allowed off-axis focussing [44]. An improved hydrogen passivated silicon surface was used as the mirror surface, which only achieved a specular reflectivity of 0.4%. The focal spot size was significantly smaller than previous attempts, achieving a spot size of $(26.8 \pm 0.5)\,\mu\mathrm{m}$ $\times$ $(31.4 \pm 0.8)\,\mu\mathrm{m}$. Even after extensive development in improving the fabrication of the mirror, a significant number of step edges were still present meaning that it is unlikely the surface is appropriate as an atom mirror [45].

Other candidates for atom mirrors include thin metal crystals [46] and graphene terminated substrates [47]. A method to improve the domain size is to first fabricate a curved mirror, then coat the surface with graphene to smooth over the step edges. As previously seen, it is very difficult to produce a surface with the required roughness but by coating with graphene it may be possible to address the roughness problem. While a higher reflectance

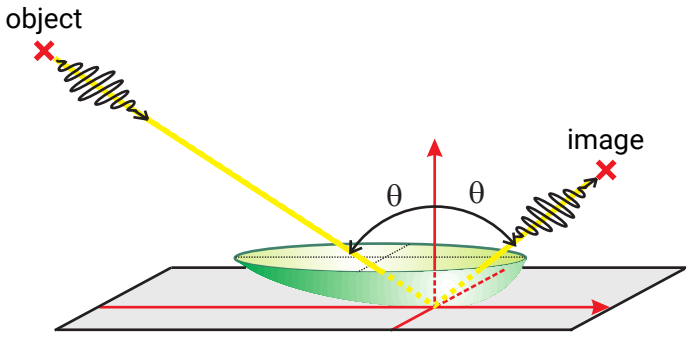


Figure 2.3: A plot illustrating the geometry of the ellipsoidal mirror adapted from Figure 4.3 by MacLaren. [19]. An ellipsoidal Cartesian surface focuses each ray coming from the object point into an image point.

has been measured with metal crystals and graphene termination, there has not yet been a demonstration of an improvement in domain size.

In summary, no viable mirror with the necessary resolution has been produced yet. It has not been possible to produce a surface with a low enough density of defects that can be macroscopically shaped and is inert enough to prevent degradation.

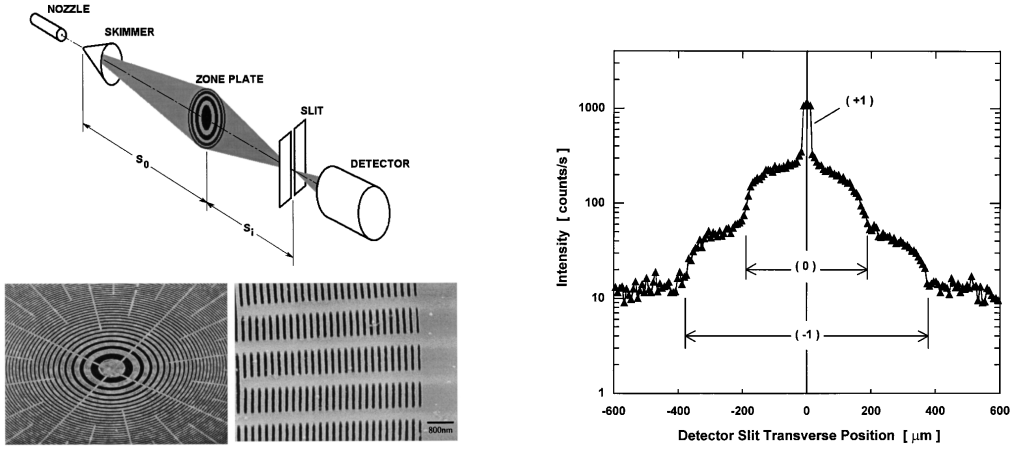
2.1.4 Zone plates

Refraction of a ground state helium beam is not possible and reflection is difficult to achieve due to the extreme requirements of an atom mirror, however diffraction can be exploited to focus helium. A beam of helium behaves as a matter wave and therefore the direction of propagation can be influenced through diffraction. Typically diffractive optic elements will have a lower collection area than a mirror or refractive lens, however they are significantly easier to manufacture and therefore have been applied more successfully.

A Fresnel zone plate is an optics element that uses diffraction to focus waves [48]. More details of how a zone plate works are provided later in chapter 6. The zone plate consists of a series of concentric circular zones that alternate between opaque and transparent. The zones are designed such that the waves that propagate through the transparent zones add constructively at the focal point. Since the transparent zones need to be completely empty, the ideal zone plate cannot be implemented for an atom beam and a support structure is needed as shown at the bottom of fig. 2.4a.

A disadvantage to using a Fresnel zone plate is that only a small part of the incident flux is focussed into a narrow spot. Aside from the fact that only half of the area is open, a significant fraction of the flux will not be focussed into the primary maximum. A Fresnel zone plate with a primary focus at f will possess subsidiary maxima at odd fractions of the primary focus, i.e. at $f/(2m+1)$, where m is an integer. Additionally, a part of the beam will be diffracted out from the focal point leading to a defocused beam and a large portion will be undiffracted and pass straight through. It can be shown that the theoretical maximum flux that is in the primary focus is given by $\eta = 1/\pi^2 \approx 10\%$ [6].

The first demonstration of focussing an atom beam using a zone plate came from Carnal et al. [49], who produced a one dimensional focus of metastable helium atoms. The initial



(a) (b)

Figure 2.4: Figures illustrating the performance of a zone plate from FIG. 1 and 2 by Doak et al. [50] (a) Top: A schematic depiction of apparatus, where the helium is focused by a zone plate then a small part of the beam is picked out by a slit to be detected. Bottom: Scanning electron microscope images of the zone plates that show the circular Fresnel zones and radial support structure. (b) A plot of the scan of the detector slit through the focused He atom beam. The central peak is due to the focused (+1) beam. The other structure seen in the scan is due to the undiffracted (0) beam (also known as the zero order beam) and defocused (-1) beam.

experiment was significant in demonstrating that it was possible to focus neutral atoms with a zone plate. But the focus could only be observed by adding an obstacle to block the undiffracted part of the beam, leading to only obtaining one dimensional focussing (methods for addressing the undiffracted flux are discussed in more detail in a few paragraphs).

Doak et al. were the first to demonstrate focussing of ground state helium in two dimensions using a Fresnel zone plate [50]. Figure 2.4a shows how the diffraction pattern from the zone plate was measured by moving a slit in front of a detector. Figure 2.4b shows the measured diffraction pattern obtained from the zone plate illustrating the focused beam, the undiffracted beam and the defocused beam. While the intensity in the focal spot is very high, the area is small and therefore there is not a large flux of helium in the focus.

Eder et al. showed that sub-micron focusing of a neutral helium beam is possible [51]. The main limiting factor on the spot size in the experiment was the chromatic aberration of the lens. The signal to noise obtained from the zone plate can also be improved by filtering out the undiffracted beam with a zero order stop [52]. By adding an order sorting aperture, the unwanted flux outside of the focal spot can be removed. Methods to mitigate the chromatic aberrations and unwanted flux outside the focal spot are discussed in more detail in chapter 6.

The binary Fresnel zone plate that has been discussed above is the simplest implementation of diffractive focussing. There are many more alternative designs, however most are not appropriate for use with an atom beam. If it were possible to control the phase of the atom beam, then instead of simply blocking the beam in the opaque parts of the zone plate, the phase could be flipped to double the amplitude at the focal point. It is then possible to stack multiple zone plates of different sizes to approach perfect lensing of the wave [53]. However, an atom beam will either pass through open space or collide with a solid surface,

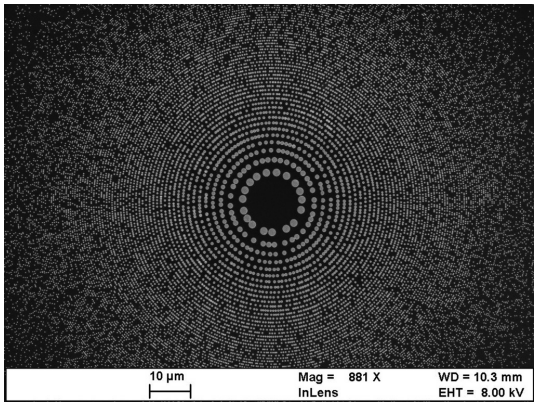


Figure 2.5: Scanning electron microscopy of the atom sieve used to focus helium atoms. Note that the structure resembles that of a Fresnel zone plate, but is now compromised of pinholes. Figure taken from FIG 2 by Eder et al. [55]

so it is not possible to construct these more sophisticated phase modulating zone plates.

A photon sieve was first proposed as a binary structure that can suppress higher order foci and overcome the limit on the resolution set by the outer zone width when compared to a zone plate by Kipp et al. [54]. The same principle was then shown to be possible with atoms by Eder et al. [55]. A photon or atom sieve replaces the concentric circles of a conventional zone plate with a quasi-random distribution of pinholes as illustrated in fig. 2.5. While the outermost pinholes may cover multiple Fresnel zones, they can be designed so that the pinhole will still add constructively at the focus. The maximum size of the sieve structure can then be much larger than a conventional zone plate improving the ultimate resolution limit, however for atom beams the chromatic aberrations and large source size typically dominate the practical resolution limit. By using a random distribution of pinholes within a zone, the symmetry conditions needed for higher order maxima are removed, leading to these additional foci being suppressed. The fabrication of an atom sieve is easier than a conventional zone plate since only pinholes needed to be machined. Optimisation of the design and manufacture of an atom sieve is an ongoing research problem [56].

The Beynon Gabor zone plate is another alternative design that can suppress higher order foci in atom beams [57]. By allowing an azimuthal variation in the binary transmittance function, the average radial transmittance function can be tailored so that it does not abruptly change between the different Fresnel zones leading to a reduction in the higher order foci. While the Beynon Gabor zone plate has been demonstrated with light, it has yet to be tested using an atom beam.

In summary, zone plates have been demonstrated as a viable method for forming a narrow helium beam. The size of a zone plate is limited by the minimum feature size, therefore only a limited area can be collected over. Additionally, only a fraction of the collected flux is focussed into the primary focus of the lens. Therefore, while zone plates can act as lenses for helium, they typically do not produce a high intensity at the focal point.

2.1.5 Pinhole

The simplest method to form a narrow beam is to directly collimate the helium beam to the desired width with an aperture. For a point source, the resolution of a beam formed

by a pinhole is determined by a combination of the geometric shadow of the pinhole and diffraction through it. As the pinhole is made smaller, eventually diffraction will begin to dominate and the beam will be spread out. It will be shown later in chapter 6 that the spreading of the beam due to diffraction determines the ultimate resolution that can reached using a pinhole.

However, it is difficult to realise a practical machine that would achieve the ultimate resolution due to the low flux in a helium beam formed by a pinhole. It will be shown in chapter 6 that for a pinhole, the flux in the spot, F, will scale with the target resolution, σ, as $F \propto \sigma^4$. The severe scaling with resolution means that the pinhole method for forming a beam will not produce a large enough flux to compete with the resolution of an electron microscope, but it is acceptable when the helium beam is greater than about $1\,\mu\mathrm{m}$.

A pinhole is the simplest method to form a beam for a helium microscope since it does not require any complex fabrication. It would be difficult to achieve a very small beam width with a pinhole due to the low flux of helium, but it is simple to implement and therefore has seen the most success in application to a helium microscope. As shall be discussed in section 2.2, the only method that has been successfully used to achieve reflection mode images of a surface with helium is the pinhole.

2.2 Demonstrations of helium microscopy

2.2.1 Zone plate

The first demonstration of helium microscopy was achieved in transmission mode. Helium will not pass through a solid material, therefore a large change in the measured signal is obtained when a sample is placed in-between the source and the detector. By maximising the contrast and using transmission mode, it was much easier to verify the feasibility of the technique.

The first two dimensional images in transmission mode were obtained using a conventional Fresnel zone plate by Koch et al. [58]. Figure 2.6a shows the experimental setup consisting of a zone plate to focus the helium beam down to a spot in the focal plane where a hexagonal copper grating was placed. A detector was then placed behind the copper grating to measure the helium flux that is transmitted through the grating. The sample was mounted on a piezo table, which could be scanned in two dimensions to obtain a two dimensional image of the sample which is shown in fig. 2.6b.

The setup used by Koch et al. [58] with a zone plate did not produce a high enough signal to observe the reflected signal from the surface. Using a zone plate is difficult due to the low focussing efficiency which will be discussed in more detail in chapter 6. Instead, an alternative method for forming a helium beam needs to be used to produce images from the scattered atoms.

2.2.2 Pinhole

The first published demonstration of a two dimensional image formed by helium reflecting from the surface was by Witham and Sánchez who used a pinhole to form the helium beam [59]. Figure 2.7a shows the setup used to collect these first reflection images, where

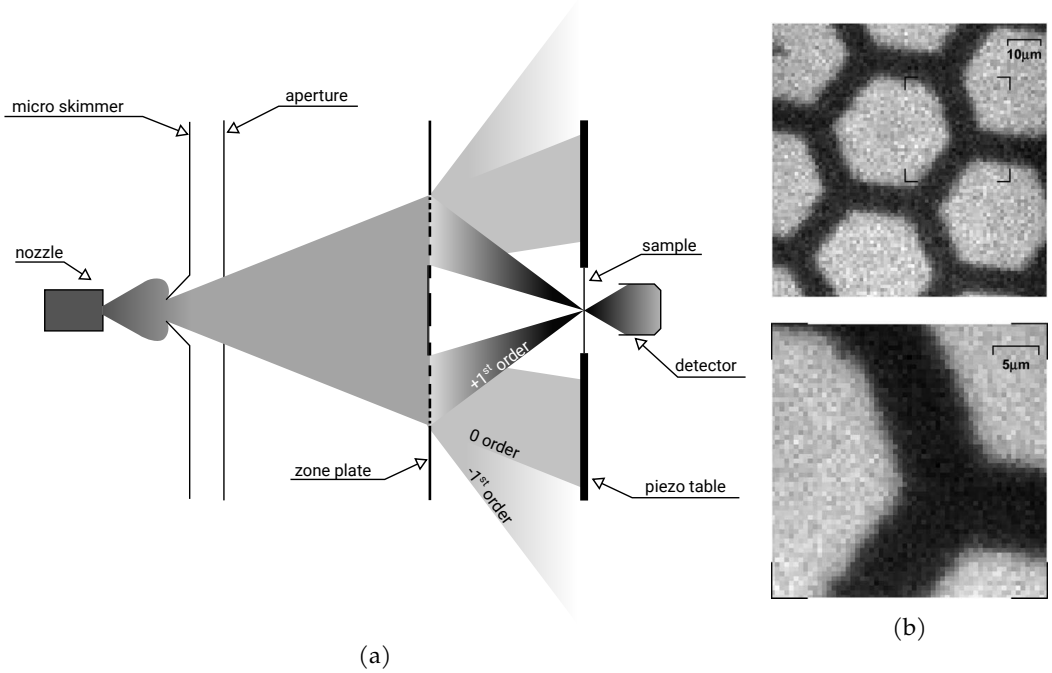


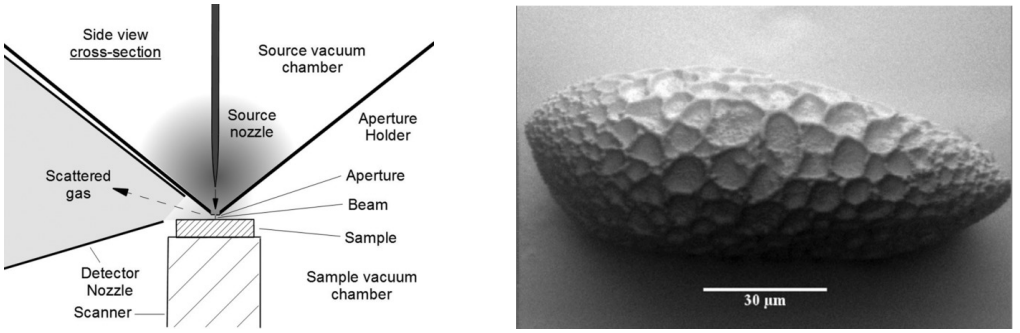
Figure 2.6: Figures illustrating the first transmission atom images from FIG. 2 and 3 by Koch et al. [58] (a) A sketch of the setup used to collect transmission helium microscope images. A zone plate focuses the helium to a spot in the imaging plane where a sample is placed on a piezo table. By measuring the flux that is transmitted through the sample by a detector and rastering the sample, an image can be formed. (b) The first two dimensional images obtained using a helium microscope in transmission mode of a hexagonal copper grating.

a small pinhole is placed very close to a sample and the helium scattered into a large solid angle is fed to a detector. Once again, images are collected by moving the sample around underneath the beam and monitoring how the scattered helium flux changes.

The helium source is not a conventional source due to it not using a normal skimmer and is likely to produce a lower quality beam than what is usually used in atom scattering experiments due to backscattering. Conventional free jet expansions are discussed in chapter 3, but typically the skimmer is designed to collimate an expansion and ensure that the number of atoms that scatter back into the beam path is minimised. Atoms scattering back in the beam will either reduce the flux, or increase the velocity spread of the beam. However, the unconventional skimmer that is used allows the source, pinhole and sample to be close together, maximising the flux in the beam. For a microscope only looking at macroscopic changes in the topography, the beam quality is not important and having a large flux is more desirable.

Figure 2.7b is an example of an image obtained by Witham and Sánchez of a pollen grain, illustrating clear contrast and the good signal to noise that can be obtained. The first images that were obtained demonstrated the feasibility of the technique and provided the first insight into the contrast mechanisms that could be expected. It was seen that there is strong topographic contrast in the images with diffuse illumination in some areas (see chapter 5 for more details).

However, these initial experiments did not reveal any contrast due to diffraction from lithium fluoride (LiF) crystals [60]. Witham and Sánchez used crushed LiF crystals to at-



(a) (b)

Figure 2.7: (a) Schematic of the helium microscope used by Witham and Sánchez, illustrating the arrangement of the source, sample and detector. The source consists of a nozzle placed very close to a pinhole which is unlike the source in a conventional atom scattering experiment. The sample is placed very close to the pinhole and a large solid angle of the scattered gas is collected for the detector. Schematic taken from Fig. 1 by Witham and Sánchez [60]. (b) A helium image of a Crocosmia pollen grain illustrating the topographic contrast that can be observed. Image from Fig. 13 by Witham and Sanchez [60].

tempt to observe diffraction contrast, however typically LiF only forms atomically flat surfaces if the crystal is cleaved. Additionally, the large collection area that maximises the measured signal will also integrate over the diffraction peaks, smearing out any contrast that could be present.

In parallel to the work by Witham and Sánchez, a proof of concept machine named the scanning helium microscope (SHeM) was being developed in Cambridge in collaboration with Dastoor's group from the University of Newcastle [61]. The approach that was adopted was to simplify previous attempts that involved an atom mirror by using a pinhole, but to still use a conventional helium source to produce a high quality beam. Other differences from the Witham and Sánchez design include using a larger sample to pinhole distance, so that a larger variety of samples can be used and to use a smaller collection solid angle, so that it may be possible to observe more subtle contrast. These changes also reduced the total flux in the detector meaning that the collected images were far noisier than the images obtained by Witham and Sánchez. A detailed discussion on the SHeM in Cambridge is deferred to chapter 3.

The microscope was then replicated in the University of Newcastle with improved differential pumping [62]. Figure 2.8 illustrates how the improved differential pumping in the replicated instrument reduced the effusive beam from the pinhole and improved the contrast in images [62]. Panel a) shows an image obtained with the improved differential pumping, while panel b) shows an image obtained after leaking in gas to match the previous differential pumping design. Stagnated gas behind the pinhole produces an effusive beam from the pinhole that smears out contrast from the direct beam. By improving the differential pumping Fahy et al. demonstrated that it is possible to also improve the contrast, which is a concept that will be returned to in chapter 3.

Recently, the images that are produced from a SHeM were critically analysed by Fahy et al. [63]. Since the incoming beam is at 45° to the surface, sample surfaces that are not completely planar with respect to the pinhole plane will become distorted in images. An expression for how the distance between two points in an image corresponds to the actual

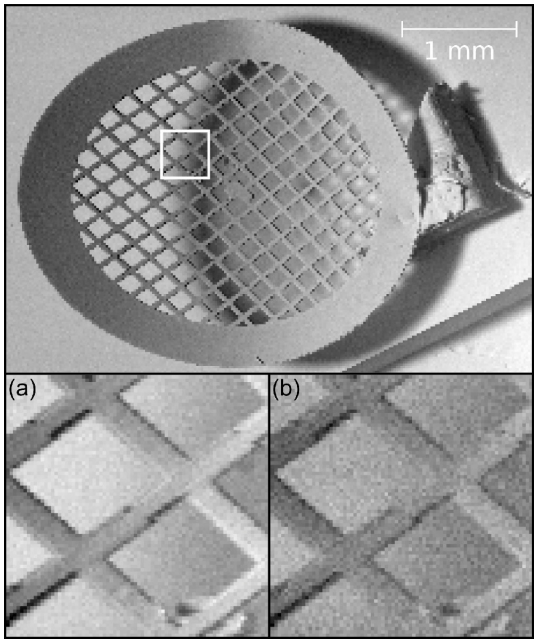


Figure 2.8: SHeM images of a TEM grid illustrating the importance of differential pumping taken from FIG. 5 by Fahy et al. [62]. Top: SHeM image of a TEM grid adhered to a cleaned silicon wafer with a piece of carbon tape. A small section of the grid (as indicated by the white square) was first imaged with a dominant supersonic beam (a), followed by the same supersonic beam along with a significant secondary effusive beam (b). The secondary effusive beam acts to reduce the observed contrast.

displacement of the points in three dimensional space was calculated. The contrast in images was also discussed and the results will be compared to my own work in chapter 5.

The scanning helium microscope has been applied to a variety of different samples including biological samples such as a butterfly and bee wing. Figure 2.9 is an example of an image collected with the machine in Australia of a butterfly's wing demonstrating that it is possible to image biological samples without coating. The microscope has also been used to investigate various other unusual samples such as a sugar cubes and other pollen grains.

Figure 2.10 shows the contrast observed between different metals deposited onto a silicon substrate by Barr et al. [64]. Since the helium atoms should only have enough energy to interact with the top layer of atoms and the samples received no cleaning, the sample should be covered in a layer of adsorbates and no chemical contrast should be observable. The authors suggest that they may be observing subsurface resonances that influence the surface charge density through exotic mechanisms, however a simpler explanation for the observed contrast due to the roughness is provided in chapter 5. Therefore while the authors claim to have observed chemical contrast, it is difficult to truly discern that the observed contrast arises from the different chemistry of the surfaces.

2.3 Summary

In summary, while there are many different methods for forming an atom beam in general, due to the same properties that make helium a useful probe, there are only a few methods for forming a beam of ground state helium atoms. Only two techniques have been demon-

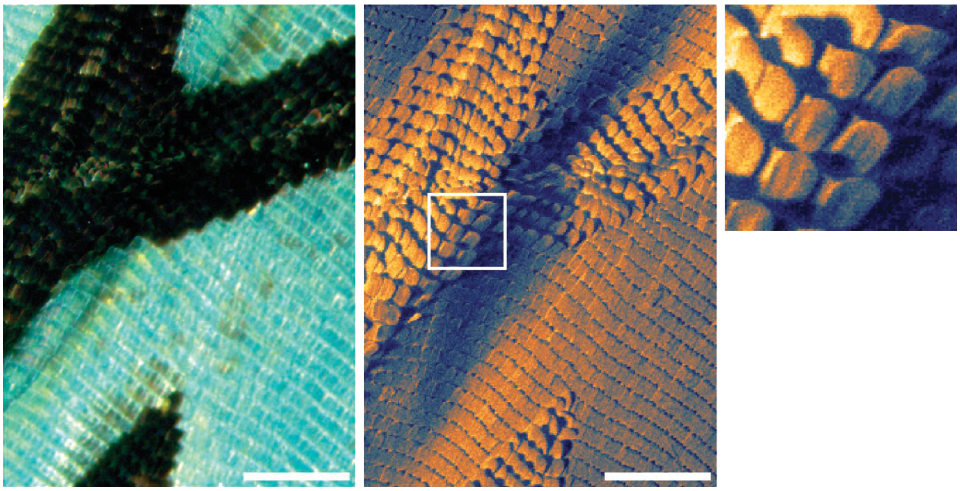
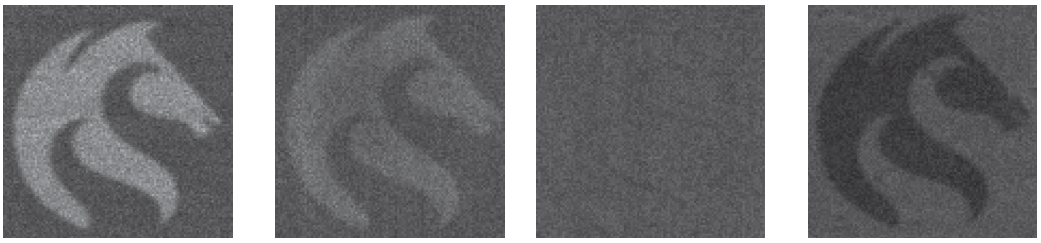


Figure 2.9: Images of a butterfly's wing (Tirumala hamata) taken from FIG. 3 by Fahy et al. [62]. The left image is an optical image, while the other two are helium images of the same area. The centre image was obtained with a $8\,\mu\mathrm{m}$ pixel width, while the right hand image has a $4\,\mu\mathrm{m}$ pixel width. Both scale bars are $600\,\mu\mathrm{m}$ in length.



(a) (b) (c) (d)

Figure 2.10: SHeM images showing the University of Newcastle logo in different metals on a silicon substrate. The metals are: (a) gold (b) nickel (c) platinum and (d) chromium. Each image is about $200\,\mu\mathrm{m}$ wide and have been taken from figure 3 by Barr et al. [64].

strated to be viable, the zone plate and the pinhole. The flux obtained with a microscope that uses a zone plate is low, so it has only been demonstrated in transmission mode to date. Using a pinhole is simpler, but suffers from low fluxes of atoms in the helium beam as the beam width is reduced. The low helium flux poses a challenge for detection to ensure that the signal to noise is high enough to observe contrast.

Helium microscopy that uses a pinhole has been successfully demonstrated with two different designs. Witham and Sánchez have used an unconventional atom source with a small working distance and large collection area to produce images that demonstrate topographical contrast with a good signal to noise. The machine developed in collaboration between the Surface Physics group in the University of Cambridge and Dastoor's group from the University of Newcastle uses a larger working distance and a smaller collection area. Similar topographical contrast has been observed with the alternate design, however there has also been an observation of unusual contrast with metal films on silicon.

Chapter 3

General instrumentation for SHeM

3.1 Description of the SHeM instrument

The scanning helium microscope (SHeM) consists of a source chamber that produces a narrow helium beam, a sample chamber where the helium interacts with a target surface and a detector which measures the scattered flux. A significant proportion of the original construction of the machine was performed by Dastoor's group from the University of Newcastle. The original work was intended as a proof of principle and has been adapted and improved for the work contained in this thesis. Each of the fundamental components of the microscope will now be discussed to explain how images are formed.

The source of the helium beam is a free jet gas expansion and skimmer which forms a beam with a small velocity spread [65] as schematically represented in fig. 3.1. The high pressure gas expands through a small aperture in a nozzle into a vacuum chamber. The pressure difference causes the gas to accelerate in the expansion, reaching sonic speeds at the nozzle exit. The supersonic flow causes the gas to overexpand creating a shock system to recompress the expansion to match the chamber pressure. However, there exists a region known as the zone of silence where the gas is moving supersonically and the shock structure has not perturbed the expansion. Placing a skimmer to extract the gas in the zone of silence produces a high intensity supersonic beam with a narrow velocity distribution.

A method to understand the expansion is to consider the sudden freeze model and the quitting surface illustrated in fig. 3.1. The expansion is initially considered to be a continuum process where, as the gas expands, the pressure drops. Eventually, the pressure is sufficiently low that the gas transitions into a molecular flow regime and the gas particles no longer interact with each other. The sudden freeze model considers a "quitting surface" where the flow abruptly transitions from continuous flow to molecular flow. The quitting surface cannot be experimentally measured, but is a useful construct for modelling.

The virtual source of an expansion is illustrated in fig. 3.1 and corresponds to the area of least confusion obtained when tracing the particle trajectories back to the nozzle plane. At a large distance from the nozzle, the gas particles will be travelling in straight lines since the pressure is too low for interactions. If the trajectories are traced back to the nozzle plane, it

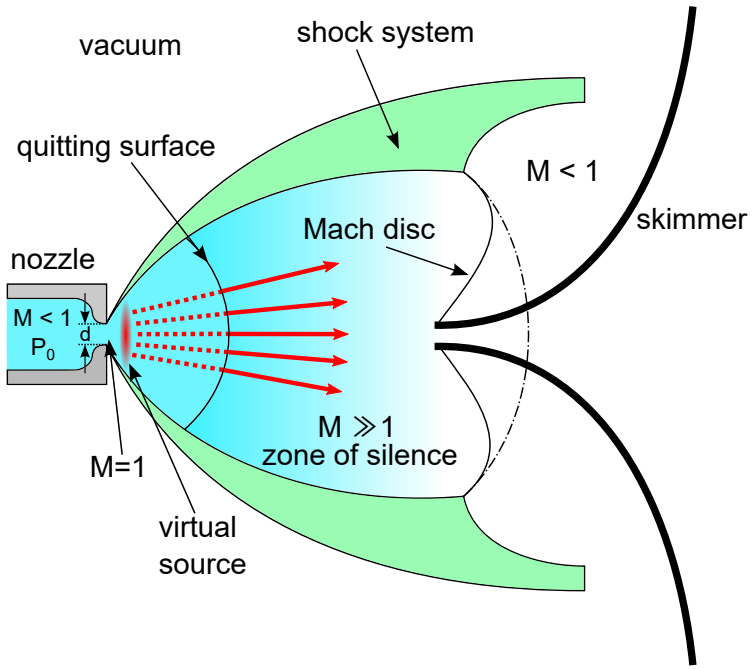


Figure 3.1: Schematic representation of a free jet expansion. Gas exits the nozzle at supersonic speeds and begins to expand adiabatically in the vacuum. Shock structures caused by the over expanding gas meeting the vacuum boundary conditions create the barrel shocks and a Mach disk. A conical skimmer is placed before the Mach disk, to extract a beam of atoms. The imaginary boundary where the flow transitions to free molecular flow is shown as the 'quitting surface'. Trajectories of atoms after the quitting surface are shown in red and can be traced back to a 'virtual source' from where the atoms appear to have originated. Based on fig. 2.1 by Scoles. [65] and fig. 1 by Barr et al. [66]

is possible to define the virtual source for the gas. More details on the virtual source will be provided in section 3.4.

The skimmer illustrated on the right of fig. 3.1 selects the centreline of the expansion forming a beam of helium atoms. When the virtual source of the expansion is larger than the size of the skimmer, the source size will be determined by the skimmer diameter. Hence, smaller skimmers will create narrower helium beams. However, if the skimmer is too small and the flow rate is too high, then backscattering can disrupt the beam [67].

Figure 3.2 illustrates how the helium beam is then used to form an image. The beam is further collimated by a pinhole so that a very narrow beam of helium is incident on the sample. A detector aperture then collects a small solid angle of the scattered helium to be measured.

Once the scattered flux has been measured at one point on a surface, the positioners, illustrated in fig. 3.2, can then move the sample so that another patch is illuminated and the flux is measured there. By measuring a set of points on a surface, the measured flux at each point can be converted to pixels of an image.

Figure 3.3 shows the pinhole plate, which contains both the pinhole and the detector aperture and is mounted to the wall of the sample chamber. The pinhole itself is mounted onto a 3 mm disc that slots into a recess on the pinhole plate. The pinhole is sealed and kept in place using leak sealant to ensure gas can only pass through the pinhole. The pinhole plate determines the scattering geometry and is therefore a key component in determining

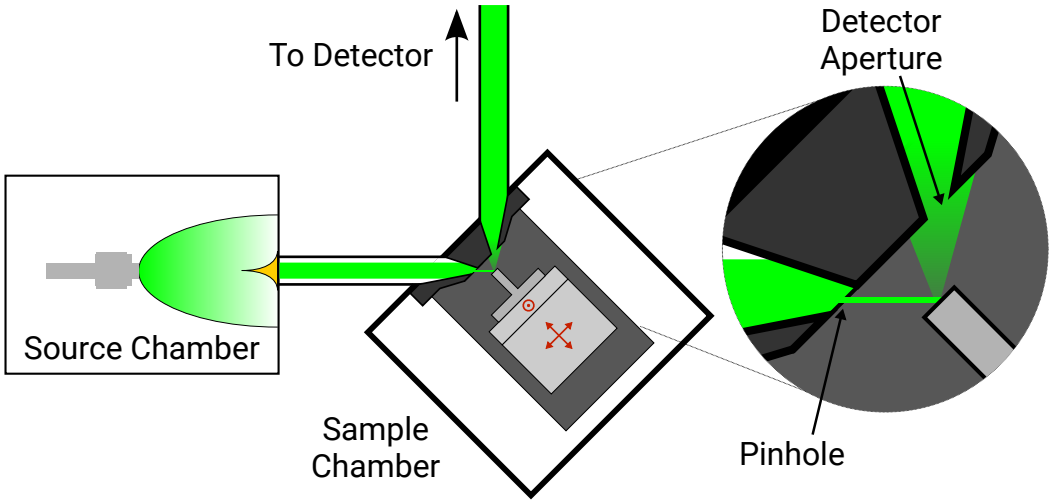


Figure 3.2: Schematic representation of a scanning helium microscope illustrating how images are formed. A helium beam is formed in a free jet expansion on the left and is collimated by a skimmer (shown in gold). The helium beam is differentially pumped and then reaches the sample chamber where it is further collimated down to a narrow beam by a pinhole. The narrow beam is incident on the sample, which scatters the helium into a range of directions. A detector aperture collects a small solid angle of the scattered helium which then leads to a detector. An image can then be formed by moving the sample under the beam using a series of nanopositioners to measure the flux on a grid of different points on the surface.

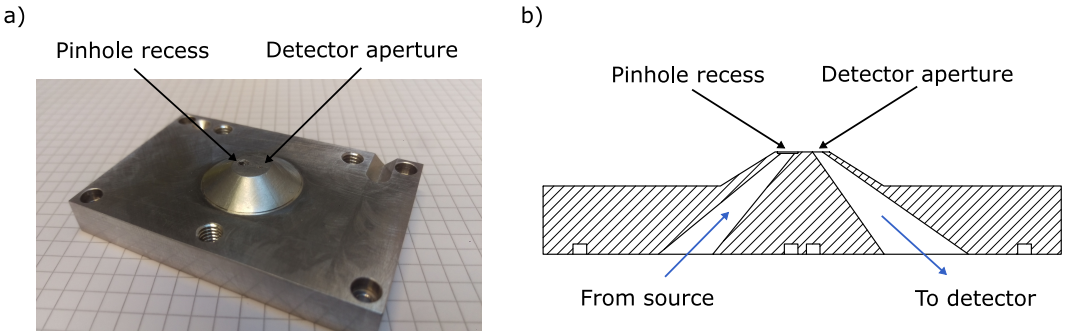


Figure 3.3: (a) Picture of the pinhole plate with the detector aperture and recess for the pinhole labelled. The three taped holes arranged in a triangle are for the sample mount that is illustrated in fig. 3.5. (b) Cross sectional view of the pinhole plate showing how the pinhole plate determines the angles of the incoming and outgoing beams. Gas from the source passes through the pinhole to the sample where it is then scattered into the detector aperture. The squares on the bottom of the diagram are the grooves for O-rings that provide a vacuum seal to the chamber wall. It should be noted that in reality the cones are not smooth and are formed from a series of steps. Based on figures from Lambrick [68]

the contrast that is observed. The implications of the pinhole plate are discussed in more detail in chapter 5.

The helium beam passes through a differential pumping stage before reaching the pinhole to minimise the gas pressure behind the pinhole. Gas that does not pass through the pinhole will scatter from the chamber wall and instead of acting like a beam, will move through Knudsen diffusion by continuing to scatter from the chamber wall. The gas can therefore be described as "stagnating" since it no longer acts like a beam and is instead

behaving like gas normally does in a vacuum chamber.

An effusive beam will be formed by any gas that passes through the pinhole after stagnating in the differential stage which will reduce the observed contrast in images. By differentially pumping the volume, the pressure of stagnated gas can be reduced, which will also reduce the flux in the effusive beam. The implications of the differential stage are discussed further in section 3.3.

3.2 Sample stage

Since the helium beam is neutral and difficult to manipulate, images must be formed by manipulating the sample underneath the stationary beam. It is therefore essential that the sample can be manipulated more accurately than the target resolution of the images, meaning that the sample stage is a key component of the instrument. The design of a sample stage capable of manipulating the sample in three dimensions will now be outlined.

The motion of the sample is achieved using positioners that are piezoelectric motors that use the slip stick phenomenon to achieve macroscopic displacements with nanoscopic accuracy. To achieve full three dimensional motion, three ECS3030/NUM positioners manufactured by attocube systems are used. Each is quoted to have a travel range of 20 mm, repeatability of 50 nm and an absolute accuracy of < 0.01 % of the travel range. The slip stick mechanism utilised by the positioners uses a piezoelectric actuator to control a guiding rod underneath a sliding block. When the guiding rod is slowly moved, the friction between the guiding rod and sliding block cause them to engage and move together. When the maximum displacement of the piezoelectric actuator is reached, it is moved rapidly back to its original position. The inertia of the block causes it to disengage from the rod and remain almost unmoved. By repeating the slow extension and rapid retraction process, the block can be moved large distances, with the potential for precise movements if only the piezoelectric actuator is used over a small range. The positioners also include optical encoders to allow the movement to be controlled using a closed feedback loop.

The base of the positioners must be securely fixed in place to ensure that there is no movement during the rapid acceleration and deceleration of the piezoelectric actuator. Without sufficient clamping, the base of the positioners will drift and the position of the sample will not be set correctly (a process I unknowingly studied in great detail during an undergraduate project in the group).

The necessity for rigidly fixing the base of the stage poses a challenge for also creating a stage that can easily be changed. Unlike with other techniques like SEM where the beam can be electronically rastered, the entire sample needs to be moved and therefore needs to be fixed in place when measuring. Additionally, the stage needs to be repeatably replaced in the same location to ensure that coarse beam alignment with the sample does not need to performed each time that the sample is removed.

A magnetic kinematic mount allows removal and insertion of the sample stage with a high degree of repeatability, while ensuring that during operation it is rigidly held in place. Kinematic couplings use precisely six contact points to constrain all of the degrees of freedom of an object and determine its position and orientation [69]. By defining the contact points carefully, the fixture can be used repeatedly to a high precision. Figure 3.4 illustrates an example of a standard kinematic mount that uses the contact between three vee

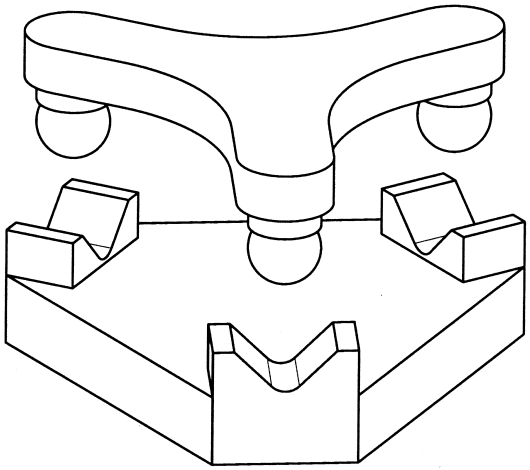


Figure 3.4: A Maxwell kinematic coupling from fig. 1(a) by Hale and Slocum[69] illustrating how three spheres and vee groves can be used to form a precise repeatable fixture. Each sphere and vee groove has two contact points such that all six degrees of freedom are constrained.

grooves (with axes that meet at a single point) and spheres. Each sphere and vee groove will contact in two places, so three sets are needed to totally constrain the fixture. The advantage of using three vee grooves and spheres is that only a limited number of parts need to be manufactured. Additionally, for the application to the sample stage, using a magnet to increase the force between the two pieces in the coupling ensures there is no chance of contact being lost when the positioners are operating.

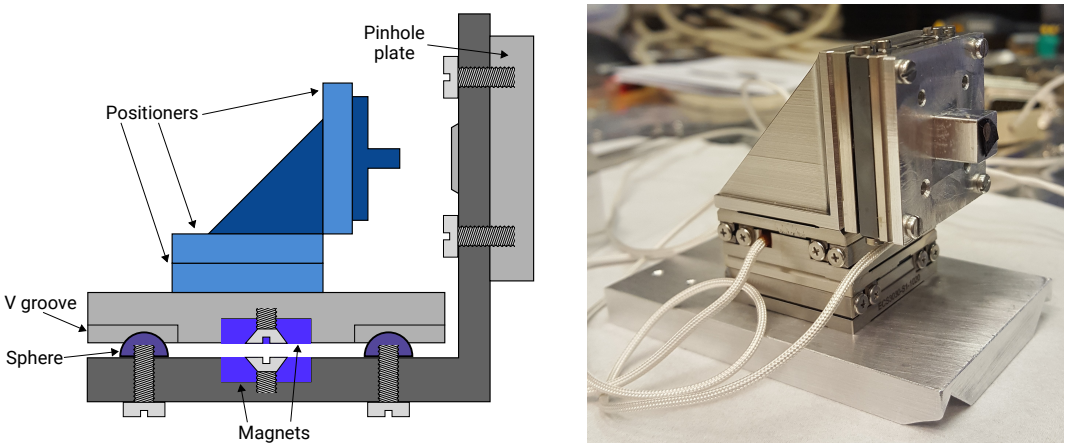
Commercial implementations of magnetically assisted kinematic mounts are not typically vacuum compatible. It was therefore necessary to construct a mount in-house to be used in the microscope, ensuring that all components are vacuum compatible.

Figure 3.5 shows how the kinematic mount was implemented on the sample stage. Figure 3.5a illustrates schematically how the positioners are kinematically coupled to the microscope. An L bracket is fixed to the pinhole plate to provide a horizontal surface with three spheres. The three spheres are 0.25" hardened stainless steel balls with a UNC 2-56 thread (Hitek Hardware KTRB-4). Three 90° vee grooves are machined into the plate that holds the positioners. The magnets are two 15 mm diameter neodymium disks (First4Magnets F338N-2) with opposite polarity, that each exert a 25 N pull force upon a steel surface. The magnets are mounted through a countersunk hole which avoids using glue or any other method that is not vacuum compatible.

The sample stage has been installed onto the microscope and has simultaneously improved the stability of the sample during imaging and made changing samples easier. The installation of the sample stage also allowed the sample to be manipulated in three dimensions, allowing a range of experiments detailed later in the thesis, to be performed.

3.3 Significance of differential pumping

The differential stage between the skimmer and the pinhole is an essential component of the microscope. A substantial flux of gas that is detected arises from stagnated gas behind the pinhole. The effusive beam contributes to a background gas pressure, that leads to a



(a) (b)

Figure 3.5: (a) Schematic of the magnetically assisted kinematic mount for the sample stage. An L bracket that is attached to the pinhole plate provides a horizontal surface for three spheres. The three positioning stages sit on a plate with vee grooves that form a kinematic coupling with the L bracket. Magnets ensure that a large force is present to keep the stages in place during operation. (b) Image of the sample stage plate with a vee groove visible in the bottom corner of the plate.

reduction in the observed contrast in images. In this section, the effect of stagnated gas behind the pinhole is introduced and the differential pumping and its limitations are then discussed.

Figure 3.6 shows how stagnated gas in the differential pumping stage can effusively pass through the pinhole and scatter from the sample over a very large area (represented in blue on the right of the diagram). A stagnated pressure of helium (in blue) can form in the differential stage through either gas from the source chamber diffusing into the differential stage, or by gas in the beam (in green) that does not pass through the pinhole into the sample chamber. A large enough pump directly next to the skimmer, on the differential stage side, will ensure that any gas that diffuses from the source chamber will be removed. Any gas that is not transmitted through the pinhole will scatter next to it, increasing the local pressure. Unfortunately, a narrow pipe connects the differential stage to the sample chamber, causing the pumping speed to be significantly throttled next to the pinhole. The low pumping speed creates a high pressure next to the pinhole and then leads to the formation of a broad effusive beam, in addition to the main direct helium beam.

Stagnated gas that passes through the pinhole will appear as a much wider beam than the direct beam and will blur out any image. Figure 3.7 shows the effect of the effusive beam on images. The wide effusive beam can scatter off the sample even if the direct beam is not incident on the sample surface. Therefore, the easiest place to see the effect of the effusive beam is where the sharp edges of the sample are blurred.

The effusive beam will appear stronger on one side of the sample when compared to the other since, the gas exiting the pinhole will be directed normal to the pinhole, and the main beam does not hit the surface at normal incidence. The right hand side of fig. 3.6 shows how the angle of the pinhole with respect to source leads to the main beam in green not being incident on the surface in the same place as the effusive beam in blue.

Improved pumping next to the pinhole on the differential pump side will reduce the

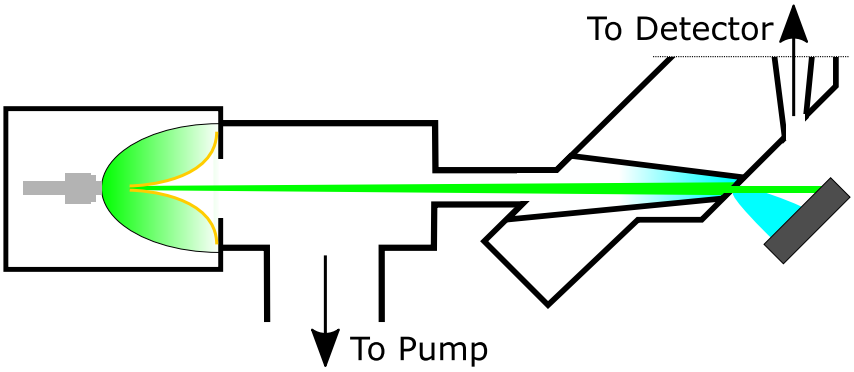


Figure 3.6: Schematic showing the path of the main helium beam through the differential stage. The helium beam is depicted in green and passes from the skimmer to the pinhole on the right. However, gas that does not pass through will stagnate next to the pinhole and increase the local pressure. The stagnated gas is illustrated in blue. The low conductance of the differential stage makes it difficult to pump the gas away. The stagnated gas diffuses into the sample chamber preferentially in the direction given by the normal to the pinhole.

stagnated gas pressure. The conductance of the pinhole will scale as d^2 where d is the diameter of the pinhole, and therefore the ratio between the atoms that pass through the pinhole as a direct beam and the atoms that pass through as an effusive beam is independent of pinhole size. Fahy et al. show that by improving the pumping the effusive component of the beam is reduced and the signal to noise ratio of a large sample is improved [62].

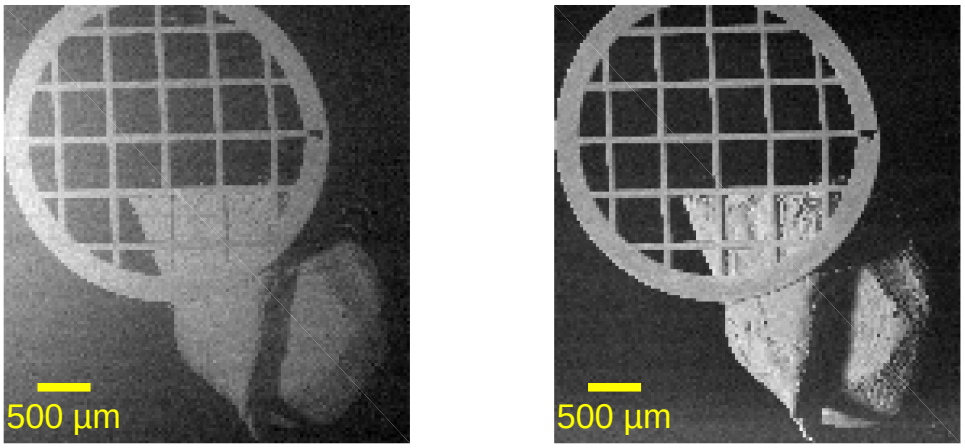
An alternative method to reduce the contribution of the effusive component is to use smaller samples. Reducing the area of the sample will not reduce the total incoming effuse flux but it will reduce the total flux that scatters from the sample. The downside of using a smaller sample, so that a smaller fraction of the effusive beam scatters into the detector, is that it will place an upper limit on the size of the sample that can be used.

Gas can also bounce between the sample and pinhole plate which leads to an increase in the local pressure of helium near the detector aperture. Since the gas will have bounced off multiple surfaces, no useful information will remain and therefore multiple scattering leads to a background signal. Unlike the effuse beam, multiple scattering between the sample and the pinhole plate will only occur if the main direct beam is striking the sample surface.

The background due to multiple scattering can be minimised by reducing the size of either the sample or pinhole plate. By reducing the total area available for the gas to scatter off, the background due to multiple scattering will be reduced. The size of the pinhole plate can be reduced by moving the pinhole and detector apertures closer together, which would make the manufacturing of the pinhole plate more challenging.

The simplest method to minimise both types of background due to the beam is to use the smallest possible sample. Uniform samples that do not need to be totally imaged, can be cut to about 1 mm in size to minimise the area of the sample. In cases where the size of the sample is fixed, the backgrounds cannot be removed without altering the geometry of the machine.

Additionally, to ensure that the positioners and the rest of the sample stage do not cause multiple scattering or present an area for the effusive beam to scatter from, the sample can be mounted onto the end of a narrow stick. By using a stick, the large positioners can be kept far back from where the gas scatters, thus minimising the local pressure around the



(a) (b)

Figure 3.7: SHeM images of a TEM grid held with a PELCO tab on an aluminium sample holder. The images consist of 126 x 136 pixels, each with a size of $30\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 1.125 s at every point. (a) Image acquired with a large $500\,\mu\mathrm{m}$ skimmer, which leads to a high pressure in the differential stage and a strong effusive beam. A significant intensity variation is seen across the image due to an effusive beam. (b) Image acquired with a smaller $100\,\mu\mathrm{m}$ skimmer, the lower pressure in the differential stage leads to more of the flux appearing in the direct helium beam even when the sample is close to the pinhole.

sample.

In summary, a background can appear in images due to either an effusive beam or multiple scattering between the pinhole plate and the sample. Differential pumping can be used to reduce the effect of the effusive beam by pumping away the stagnated gas behind the pinhole. Mounting the sample on the end of a stick and keeping the area of the sample as small as possible will also minimise the background contribution of these effects.

3.4 Determination of the virtual source size

The size of the source of atoms is a crucial parameter to consider when understanding the optics of the system. In order to model the particle trajectories downstream from the source, the size of the source, as illustrated in fig. 3.1, needs to be known. However, its size is notoriously difficult to measure, with measurements often contradicting each other [70].

The difficulty in predicting the source size from a free jet expansion motivates the development of a method to measure the source size during operation of a helium microscope. Direct measurement of the source size removes any uncertainty in the performance of the helium source and informs the design of the optics for the microscope.

Figure 3.8 illustrates how the size of the beam varies with distance from the pinhole due to the source size. The pinhole effectively creates an image of the virtual helium source, with the size of the projection depending on the ratio of the source distance to the imaging distance.

It is therefore possible to estimate the size of the source by measuring how the width of the beam varies with distance from the pinhole. In reality, the beam width will be a convolution of the size of the pinhole and the demagnified size of the source. But provided

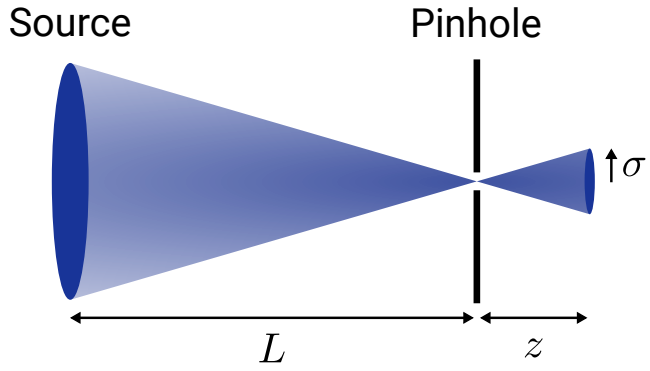


Figure 3.8: A schematic illustrating how the pinhole forms a demagnified image of the source. The demagnification of the pinhole is determined by the ratio of the source-pinhole distance to the pinhole-sample distance. Therefore, the size of the beam, σ, will depend on the distance to the pinhole, z.

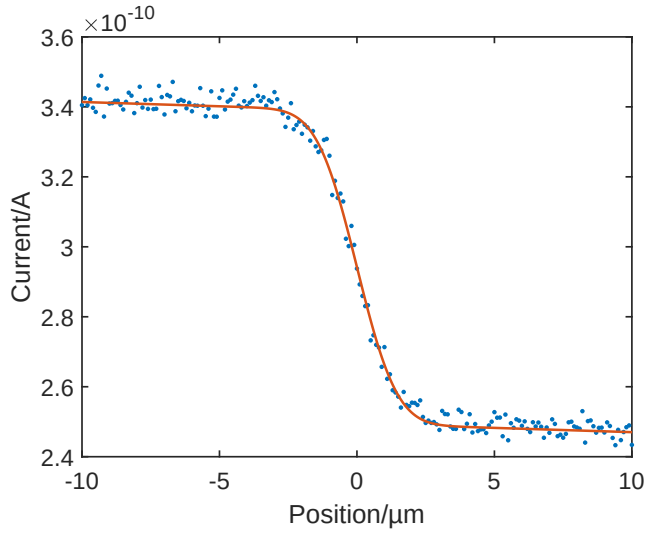


Figure 3.9: A plot with the blue points showing how the current varies in a line scan across a sharp edge. The width of the beam causes the sharp edge to be blurred out. The orange line is an error function fit with a standard deviation of 1.14 ± 0.03 μm.

that the pinhole is sufficiently small, it will be possible to extract the angular size of the source.

The size of the beam can be measured by performing a line scan across a sharp sample edge as illustrated in fig. 3.9. For reasons to be discussed later in chapter 5, the sharp edge must be between a sample and vacuum; any other objects on the beam path can lead to diffuse illumination that complicate the line scans. The edge of a TEM grid was used as a sharp edge and by rastering the sample under the beam, a series of line scans similar to fig. 3.9 were collected under different conditions. It should be noted that the sample also needs to be laterally moved to keep the same area illuminated due to the incoming beam angle being 45°, the procedure needed to keep the same area illuminated is discussed in more detail in section 5.2.2.

The width of the beam can be extracted by assuming the beam shape is a Gaussian function and fitting an error function to the data. The noise on the experimental data meant that direct differentiation of the data to extract the beam shape was not possible. Performing a

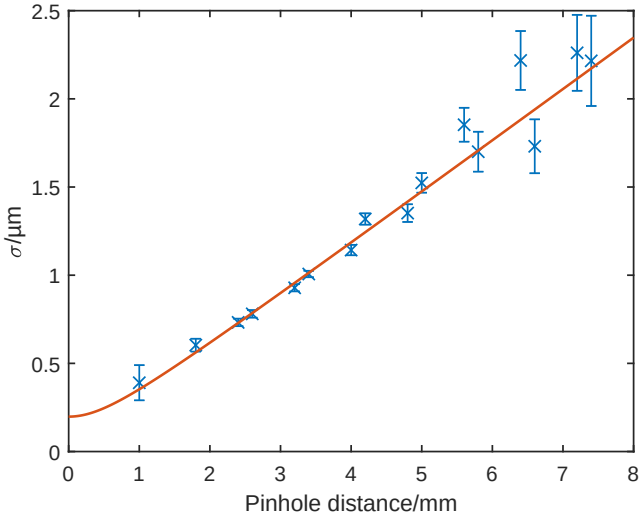


Figure 3.10: A plot showing how the width of beam varies at different distances from the pinhole. Each point is obtained from a line scan across a sharp edge. The orange line is a fit of the form $\sigma(z) = \sqrt{\alpha^2 + (\beta z)^2}$ to estimate the angular source size. In the plotted example, the source is at 100 bar with a $100\,\mu\mathrm{m}$ skimmer and the fitted angular source size is $\beta = (2.92 \pm 0.07) \times 10^{-4}$.

rolling average to the data to smooth out the noise [71] was not possible since the averaging would only be effective if a large number of points were used which would then change the width of the step. Instead, an error function was directly fitted to the data, I, of the form

$$I(x) = A\,\texttt{normcdf}\left(\frac{x-\mu}{\sigma}\right) + Bx + C, \tag{3.1}$$

where x is the stage position and A, μ, σ, B, C are fitted parameters. The function `normcdf` is the normal cumulative distribution function and is closely related to the error function. The parameter μ corresponds to the position of the step and σ corresponds to the standard deviation of the Gaussian beam. The linear term Bx is added to account for any small drifts in the experiment due to a diffuse beam background. An example of the fit is plotted in orange in fig. 3.9. Sources are usually modelled as a double Gaussian [72, 51], but the noise in the data makes it difficult to distinguish the precise shape of the source and it is assumed that the beam can be described by a single Gaussian distribution.

Figure 3.10 illustrates how performing line scans at various distances across a sharp edge allows the angular source size to be estimated from the change in beam width. As demonstrated in fig. 3.8, as the beam is imaged at different distances from the pinhole, the magnification ratio changes and the beam becomes wider. To take account for the convolution of the geometric size of the pinhole and the source size, the beam standard deviation σ from each line scan is fitted with a function of the form,

$$\sigma(z) = \sqrt{\alpha^2 + (\beta z)^2}, \tag{3.2}$$

where z is the distance from the sample to the pinhole and α, β are fitted parameters. The parameter α takes account of the geometric size of the pinhole and β is the angular source size.

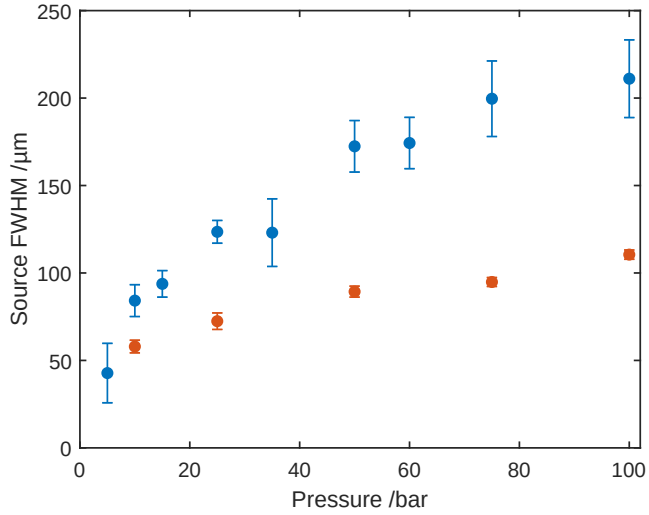


Figure 3.11: A plot showing the FWHM of a helium source with a $500\,\mu\mathrm{m}$ skimmer in blue and a $100\,\mu\mathrm{m}$ skimmer in orange as a function of nozzle pressure. Both sets of data follow the general trend that larger nozzle pressures generate larger virtual source sizes. The $100\,\mu\mathrm{m}$ skimmer collimates the source such that the virtual source size is set by the skimmer diameter.

Once the angular source size has been obtained, it can be used to calculate the full width half maximum (FWHM) of the source. The standard deviation of the source is given by $\beta L/\sqrt{2}$, where L is the distance from the pinhole to the source and the factor of $\sqrt{2}$ is added to account for the $45°$ projection of the source as viewed at the sample. The exact distance to the source is not known due to difficulty in estimating the nozzle to skimmer distance, so the pinhole to skimmer distance is used which gives $L = 0.22 \pm 0.02\,\mathrm{m}$. The error on the pinhole to skimmer distance is estimated due to there being several flanges between the two points. The full width half maximum of a Gaussian distribution with a standard deviation of σ_G is given by $2\sqrt{2\ln 2}\,\sigma_G$, so if the source is assumed to have a Gaussian distribution the FWHM, Δ, is given by,

$$\Delta = 2\beta L\sqrt{\ln 2} \tag{3.3}$$

Figure 3.11 shows the FWHM of the source at different source pressures and with different skimmer sizes. The orange points show the FWHM of the source when using a skimmer that is $100\,\mu\mathrm{m}$ in diameter at different source pressures. Both skimmers show the expected trend that larger nozzle pressures create larger virtual sources. For the smaller $100\,\mu\mathrm{m}$ skimmer, the virtual source size approximately matches the skimmer diameter implying that the skimmer is collimating the source. The blue points are the FWHM of the source as measured using a $500\,\mu\mathrm{m}$ skimmer which does not collimate the source size.

The virtual source size for helium near room temperature has previously been measured by both DePonte et al. [73] and Reisinger et al. [74]. DePonte et al. used a slit as a 1-D pinhole camera, while Reisginer et al used a zone plate to image the expansion. The zone plate had a significant chromatic aberration at low pressures, so comparison at low pressures is not possible. The two sets of data show good agreement, with only small deviations which are not surprising given the differences in the setup, skimmer size and backscattered components.

As previously mentioned, free jet sources are typically modelled using two Gaussian

functions making a comparison to a single Gaussian function difficult. The values presented in fig. 3.11 are between the two widths of each Gaussian distribution given by DePonte et al. [73] for a $9.5\,\mu\mathrm{m}$ skimmer implying that the data here is consistent with the literature.

3.5 Conclusions

In this chapter the method for creating an image with a helium microscope has been introduced. The SHeM uses a free jet source to produce a beam of helium atoms, which is then collimated using a pinhole to a narrow beam. The narrow beam is then scattered from a sample into a detector to be measured. By moving the sample under the beam, it is then possible to generate images of the surface of a sample.

A new sample stage was developed that allows three dimensional movement of the sample using nanopositioners while also providing easy sample changes. The mount uses a magnetically assisted kinematic coupling, so that the stages can be removed and replaced to a high degree of precision while ensuring the sample mount is well constrained during measurements.

It was shown that the differential pumping is very important to generating a narrow beam of helium. Stagnated gas behind the pinhole creates an effusive beam which acts to reduce the observed contrast. The microscope was constructed with a narrow tube leading to the pinhole meaning that the pumping is significantly throttled near the pinhole. Future machines will need to improve the pumping of the area behind the pinhole.

Finally, the virtual source size was measured in the microscope using line scans over sharp edges. By assuming a Gaussian distribution and fitting an error function to the data, it was possible to determine the beam width. The variation of the beam width with distance from the pinhole then allowed the source size to be measured. When a small skimmer is used, the source size was observed to be determined by the diameter of the skimmer. For a larger skimmer, the source size is consistent with other literature.

Chapter 4

Design and implementation of a high efficiency helium detector for SHeM

4.1 Introduction

A high efficiency mass spectrometer is needed to measure the low partial pressures of helium present in the output from a scanning helium microscope (SHeM). Spatially resolved helium scattering requires narrow beams of helium for the best resolution in the images, but narrow beams mean that the flux of helium to be detected is very small. Work on high sensitivity helium detectors is also relevant to many problems beyond scanning helium microscopy and helium scattering in general. Classical helium scattering often involves analysis of diffraction peaks and higher efficiency detectors allows observation of weaker peaks [14]. Quasi-elastic helium scattering can measure surface correlations between adsorbates, and high efficiency detectors allow measurement of weaker correlations [16]. Additionally, helium detectors are important for a variety of other applications such as leak detection and more exotic applications such as neutral gas measurements in space [75] and monitoring fusion reactions [76]. Therefore, there is a clear motivation for producing very high efficiency helium detectors not only for a scanning helium microscope, but also in a variety of other applications.

Unfortunately, many of the properties of helium atoms that make it a useful probe for investigating surfaces also make it difficult to detect. Helium atoms are inert due to the electronic structure shared by the noble gases. The inertness of the helium atoms gives helium scattering the desirable property of strongly scattering with target surfaces at low energies. Direct detection of helium atoms without any excitation is difficult at low fluxes, and the atoms need to be manipulated into a form (e.g. ionised) that is possible to detect. But inert helium atoms are difficult to excite into a state that can be measured due to the ground state of helium having a very low energy. For example, if the helium atoms were to be ionised, it would be difficult since they have the highest first ionisation energy on the periodic table.

The detector that has been developed for use with the helium microscope uses electron

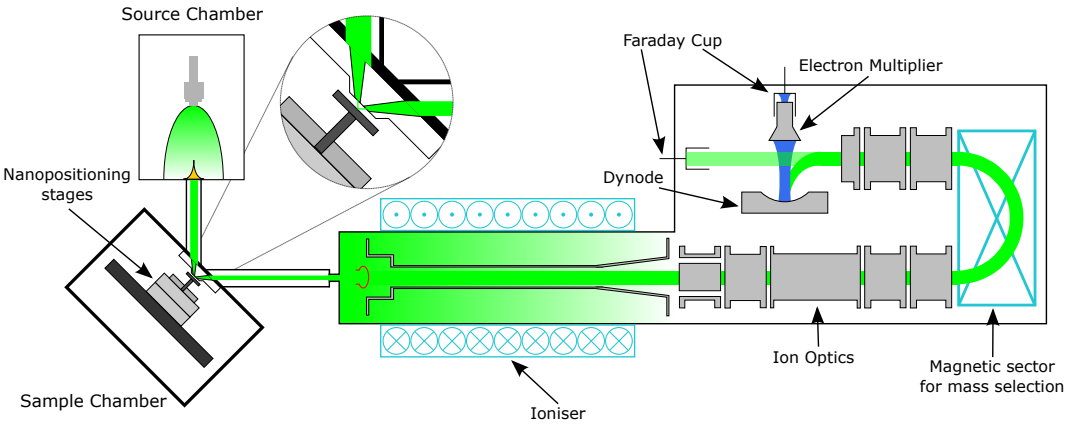


Figure 4.1: Schematic of the scanning helium microscope focussing on the detector design. Helium atoms enter the detector and are ionised by electrons trapped inside the ioniser by electric and magnetic fields. Ion optics are then used to transfer the ions from the ioniser to a magnetic sector where they are mass selected. Ion optics are then used again to transfer the ions to be measured either on axis directly by a Faraday cup or are deflected into a dynode and electron multiplier system.

ionisation and a magnetic sector to measure the helium partial pressure. As illustrated in fig. 4.1, there are three main components to the detector, the ioniser which uses electron ionisation to produce helium ions, the magnetic sector to select the helium ions from the other background species in the chamber and a dynode and electron multiplier to measure the ion current. Ion optics are needed to transfer the ions from each component to the next without significant loss.

The development of an efficient helium detector that uses electron ionisation has been a long running project within the surface physics group in Cambridge. In particular, Dr John Ellis has contributed to almost every aspect of the current detector design. My contributions to the design and implementation of a helium detector are part of a large collaborative effort between many people. In order to understand my contributions, it is necessary to describe the work of others and describe how my research fits into the larger body of work. Additionally, much of the work was undertaken collaboratively with the help of other researchers. Any work described in this thesis that was completed by another researcher will contain an acknowledgement.

In this chapter, I will provide a critical review of the design and implementation of each of the major components of the detector that has been installed onto the helium microscope and the overall performance of the device. First, I will briefly describe the design of the ioniser as implemented by Chisnall [77] for use on the spin-echo instrument and provide a critical review of its operation. I will then go on to explain the assembly and modifications performed by myself and Ward to the Chisnall design that have been made for application of the ioniser to the microscope. Next, I will briefly describe the initial design of the ion optics and mass selection as performed by Ward and Allison and then discuss my implementation in the constructed detector that allows the generation of mass scans for low mass species. I will then go on to describe the design, implementation and results for the dynode and electron multiplier system that I have designed. Finally, I will then show how all these components work together to produce a highly efficient helium detector and provide the key parameters that characterise the detector.

4.2 Ioniser

4.2.1 Review of previous work on helium ionisers

Electron ionisation is the preferred method used to detect helium atoms and is achieved by using high energy electrons to generate an ion current to be measured. Hedgeland [78] provides an overview of the variety of potential methods for detecting helium atoms. Field ionisation [79, 18] involves applying a strong electric field to an atom to ionise it, however usually yields low efficiencies since the volume of the high field region is small. Resonance enhanced multiphoton ionisation [27, 80] has the potential to be an effective ionisation method but has not yet been applied successfully to helium. Exciting the helium atoms into a metastable state for detection is also possible [81], but also suffers from low efficiencies.

Electron ionisation can be achieved through a variety of different ioniser designs. A detailed discussion of the varying designs and the benefits of these ionisers has been performed by Alderwick [82]. A series of helium detectors have been produced in Cambridge using a variety of different ionisation methods. The first was a position sensitive cross beam ioniser implemented by Allison, Bullman [83] and Knowling [84]. Dworski [85] measured the RGA efficiency of the ioniser as $\eta = 6 \times 10^{-6}$. Hedgeland, Jardine and Ellis improved the efficiency by replacing the crossed beam ioniser with a modified Brink ioniser that has an efficiency of $\eta = 6.3 \times 10^{-5}$ [78].

The application of solenoidal ionisers for atom beam experiments was first demonstrated by DeKieviet et al. [86], producing an ioniser with a reported efficiency of $\eta = 7 \times 10^{-3}$. A similar ioniser was constructed by Kalinin et al. [87] with an efficiency of $\eta = 1 \times 10^{-3}$. The solenoidal ioniser creates a scalable ionisation region using a combination of magnetic and electric fields to trap electrons and extract ions. By using a solenoidal ioniser, it is possible to create a large ionisation volume and increase the device efficiency. A more detailed explanation of how this design of ioniser works will be given in section 4.2.2. The first iteration of a solenoidal ioniser was implemented in Cambridge by Maclaren, Jardine, Ellis and Alderwick [82]. A second larger iteration was then designed and implemented by Jardine and Chisnall [77]. However, previous designs for solenoidal ionisers have had issues with erratic and hysteretic behaviour in the electron emission characteristics.

The goal of an ioniser for helium-4 detection is to efficiently produce singly ionised helium with a mass to charge ratio of $m/z = 4$, while avoiding the creation of multiply ionised species with the same mass to charge ratio. Electron ionisation often leads to fragmentation of larger mass atoms and can also create multiply ionised species either through single or multiple ionisation events [88, 89]. Rapid extraction of ions minimises the interaction time between the high energy electrons and ions, which reduces the probability of multiply ionised species being created through multiple interactions.

4.2.2 Review of ideal operation of a solenoidal ioniser

The electrons for ionisation are produced by thermionic emission from a hot filament. The filament is made from iridium wire with a diameter of $200\ \mu\mathrm{m}$ that has been coated in thoria. By passing a large current through the wire, it is heated up to very high temperatures and electrons are released. The thoria coating lowers the work function at the surface, allowing

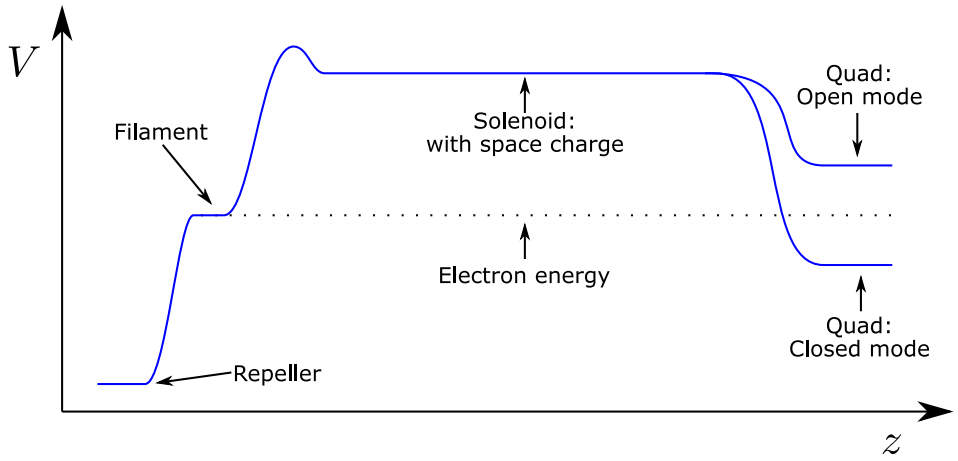


Figure 4.2: Schematic of the centre line potential throughout the ioniser. Electrons created at the filament are magnetically constrained to the axis and then accelerated into the solenoid liner. The electrons can then either escape the solenoid at the end or be reflected and trapped inside the solenoid.

similar electron emission currents to be obtained at lower cathode temperatures. Operating the filament at lower temperatures minimises degassing and keeps the local pressure around the filament low, which will ultimately lead to a lower background.

Electrons are then injected into a solenoid and confined using electric and magnetic fields to create a high electron density. The solenoid creates a strong magnetic field, which confines the electrons to helical trajectories around the magnetic field lines, so confining the electrons along the axis of the solenoid. The filament is held at a lower potential than the liner of the solenoid, so that the electrons are accelerated parallel to the magnetic field lines into the liner. The precise position of the filament will influence the electron trajectories down the solenoid and therefore will change the performance of the ioniser. Moving the filament a large distance will be discussed later in section 4.2.4, but Alderwick [82] showed that even removing the filament and replacing it in approximately the same position will alter the performance.

The two modes of operation of the ioniser are illustrated in fig. 4.2, where either the electrons are trapped inside the solenoid or can pass directly through the ioniser. The first lens element after the ioniser can be used to confine the electrons longitudinally inside the solenoid. If the potential of the first element is above the potential of the filament, the electron beam can exit the ioniser after passing through once which we describe as "open mode". If the potential of the first element is below the filament potential, the electrons will not have enough energy to leave the ioniser and will be reflected and move back down the solenoid, which we describe as "closed mode". In closed mode the electrons will perform reflexive trajectories and in the absence of a loss mechanism will run up and down the ioniser until re-adsorption at the filament.

In a region of high electron density, space charge is formed which acts to reduce the local potential. The electrons passing through the solenoid liner will reduce the potential below the potential of the liner wall. If the electrons are assumed to be confined in a hollow cylinder, due to being released by a circular filament and magnetic confinement, then it is possible to derive analytic expressions for the potential inside the solenoid liner [90].

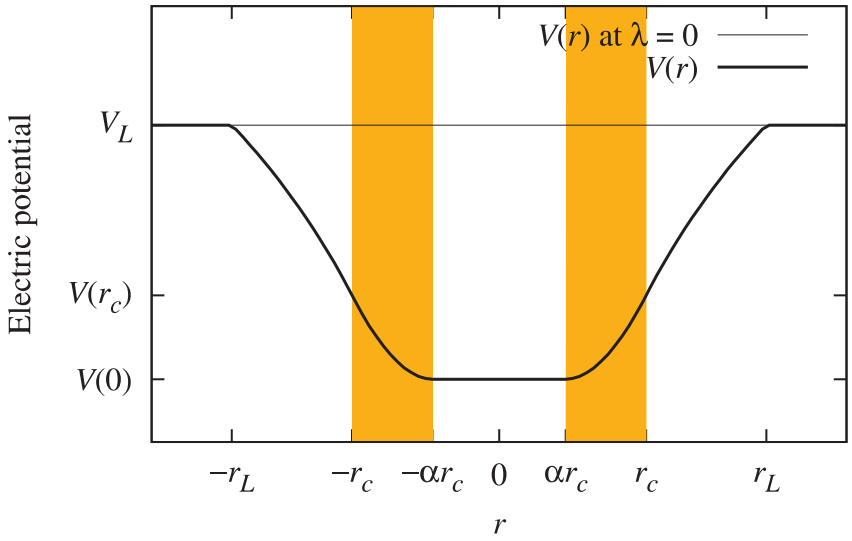


Figure 4.3: Radial variation of the electric potential, $V(r)$, inside a long solenoid ion source from fig. 2 by Alderwick et al. [90]. The thin line is the potential inside the solenoid without space charge, while the thick line is the potential inside the liner with space charge as given by eq. 4.1. In this example the electrons form a cylindrical cloud between the shaded region ($\alpha r_c < r < r_c$).

The space charge well depth inside the solenoid is determined both by the charge density of the beam and the distance from the beam to the potential boundary. Alderwick [90] used Gauss's law to show how the potential varies radially in a solenoid ion source. Electrons emitted from a circular filament will create a hollow tube of electrons, which we describe using the outer radius of the electron cylinder, r_c, and the ratio of inner to outer radius, α, $(0 < \alpha < 1)$. If the hollow cylinder has a uniform volume charge density, $\rho = \lambda / \left[\pi r_c^2 \left(1 - \alpha^2\right)\right]$, where λ is the charge per unit length (negative for a beam of electrons), then the potential $V(r)$ at radius r is

$$V(r) = V_L + \begin{cases} \frac{\lambda}{2\pi\epsilon_0} \ln \frac{r_L}{r}, & r_c \le r \le r_L; \\ \frac{\lambda}{2\pi\epsilon_0} \left(\ln \frac{r_L}{r_c} + \frac{r_c^2 - r^2}{2r_c^2(1-\alpha^2)} + \frac{\alpha^2}{1-\alpha^2} \ln \frac{r}{r_c} \right), & \alpha r_c \le r < r_c; \\ \frac{\lambda}{2\pi\epsilon_0} \left(\ln \frac{r_L}{r_c} + \frac{1}{2} + \frac{\alpha^2}{1-\alpha^2} \ln \alpha \right), & 0 \le r < \alpha r_c. \end{cases} \tag{4.1}$$

This potential is illustrated in fig. 4.3. A key result from this is that the radius of the liner, r_L, enters into the expression as $\ln r_L$. So if the distance to the potential boundary is increased the space charge well depth will also increase. The dependence of the space charge on the distance to the boundary can be understood in less rigorous terms as increasing the distance the electric field exists over and therefore also increasing the potential difference.

Ions that are produced inside the liner are trapped by space charge and cannot hit the solenoid liner. In the absence of subsequent extraction, the ions will be trapped inside the liner for a long period of time without sufficient energy to reach the liner wall and be neutralised.

In the original designs by Alderwick and Chisnall, an ion barrier ring at the entrance prevents ions escaping from the front of the ioniser. By adding a small step in the liner as shown on the left hand side of fig. 4.4, the potential in this region can be raised to block ions.

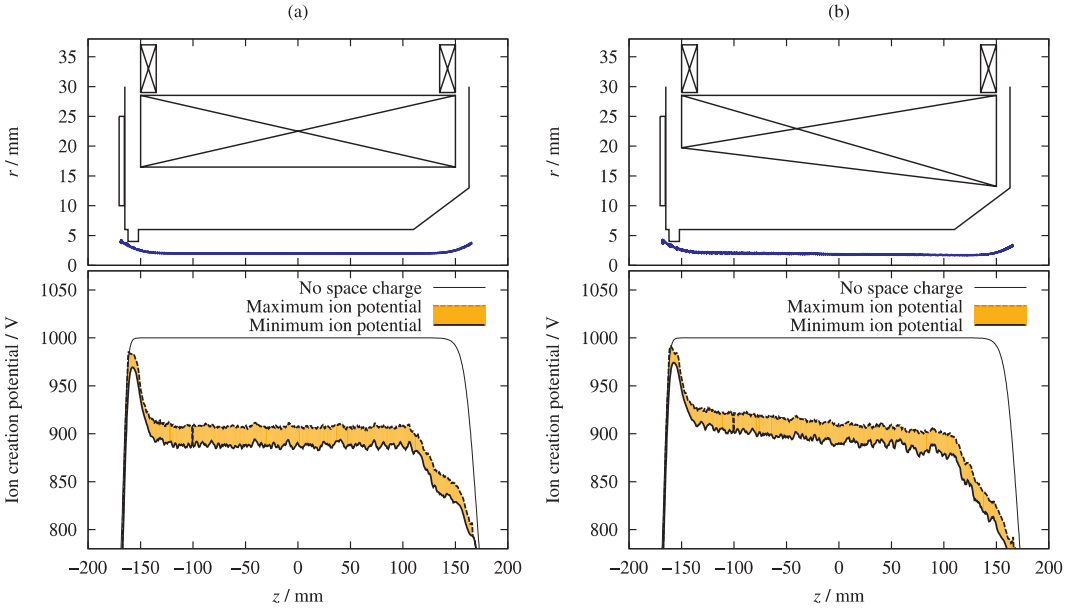


Figure 4.4: Ioniser geometries, electron trajectories and electrostatic potentials from fig. 6 by Alderwick et al. [90]. The upper panels show the solenoid geometry and simulated electron trajectories (blue lines). The lower panels show the corresponding electrostatic potential on the axis, both with and without electron space charge. (a) A uniform solenoid winding pattern with correcting coils at each end and a tapered exit channel in the liner. (b) The same geometry but with a uniform magnetic field gradient.

Since the distance between the electron beam and the liner wall is slightly smaller in this region, the space charge well depth is shallower and the potential is raised when compared to the rest of the liner. Ions deep in the liner will therefore not have enough energy to exit the ioniser from the front.

A tapered liner exit creates a way for ions to exit out of the trap for detection as illustrated on the right side of the potentials in fig. 4.4. The tapered liner increases the distance between the electron beam and the liner wall, which will increase the depth of the space charge. The potential gradient can then be used to extract ions from the solenoid.

Extra shim coils are added at either end of the solenoid to maintain a uniform field over the ionisation volume. Diverging magnetic field lines will spread the electron beam out and could create a space charge well at either the entrance or exit to the liner. By adding small correcting coils, the magnetic field is kept as uniform as possible at the ends of the solenoid to avoid ion traps from forming.

Ions in the bulk of a uniform solenoid will not be attracted to the exit, but instead ions will be trapped until they reach the field at the exit. When ions are created they are constrained to the axis by the radial space charge well. The ions can then either drift along the solenoid axis towards the tapered liner and be extracted, or drift towards the front where the ions will bounce off the potential wall and then drift towards the tapered liner and be extracted.

An alternative method is to use a small electric field in the solenoid to actively extract ions and ensure that they are detected. This can be achieved either by a liner with a small gradient as used by Kalinin et al.[87] or a varying magnetic field as being constructed by Martens et al. [91]. Both methods cause the distance between the electron beam and the

liner wall to vary along the length of the solenoid. The change in distance between the electron beam and the wall leads to a varying space charge well depth along the axis, providing a small extraction field along the entire length of the solenoid.

While space charge is necessary for the operation of the ioniser, a potential well that is too deep will make mass separation more challenging. The existence of a space charge well ensures that ions that are created do not collide with the liner wall. However, since ions will be generated at different locations in the potential well, the ions will be produced with a spread of ion energies approximately equal to the space charge well depth. As shown later in section 4.3.1, mass separation requires a narrow spread of incoming ion energies, therefore extraction methods that increase the space charge depth will degrade the performance of the mass filter.

Active extraction of ions from inside the solenoid with an electric field will increase the spread of ion energies, therefore making mass filtering more difficult. Manipulating the space charge to create an extraction field inside the ioniser will increase the spread of ion energies which is undesirable. The solenoid that is described in this chapter does not use a field down the solenoid to ensure optimal mass filtering. However, more rapidly extracting ions from the solenoid could reduce the importance of ion space charge and therefore could be used in the future if ion space charge is thought to be significant.

Other communities are also developing powerful ionisers to produce large ion currents. Electron beam ion sources (EBIS) are used to produce bare nuclei of high mass species such as uranium [89]. These ion sources are typically run by trapping ions using both electric and magnetic fields to keep them confined in a volume. The ions are then bombarded with a high electron current to multiply ionise the gas and produce high charge state species. The electron guns that have been developed to inject very high current densities into the solenoid could be used to improve the detectors for helium-4. Producing extremely high electron emission currents will help improve the efficiency of the detector if it is found that the space charge inside the ioniser is not the limiting factor holding back the device.

Maximising the efficiency of a detector is not the only design objective, since in order to obtain an acceptable signal to noise ratio the background must also be minimised. Electron beam ion sources produce many ions that have been multiply ionised by design which would lead to a large signal at the same mass to charge ratio as singly ionised helium ions. Directly using an ioniser for an electron beam ion source is therefore not appropriate for helium-4 detection. Any ioniser that is built for detection of helium-4 must rapidly extract ions from the ionisation region to prevent ions from interacting with the electron beam multiple times to create multiply ionised species. Since the helium flux into the detector will be small, a device is needed with a very low background level.

4.2.3 Ioniser performance with an externally mounted filament

Figure 4.5 shows the design for electron injection as used by Chisnall. A ring filament is positioned at the entrance to the solenoid, where the magnetic field is strongly increasing, causing the electron beam to be magnetically compressed due to the electrons following the B field lines with helical trajectories. By magnetically compressing the beam, much higher current densities can be produced which means more electrons can be injected into the ionisation volume. Converging magnetic fields can act as magnetic mirrors for low energy

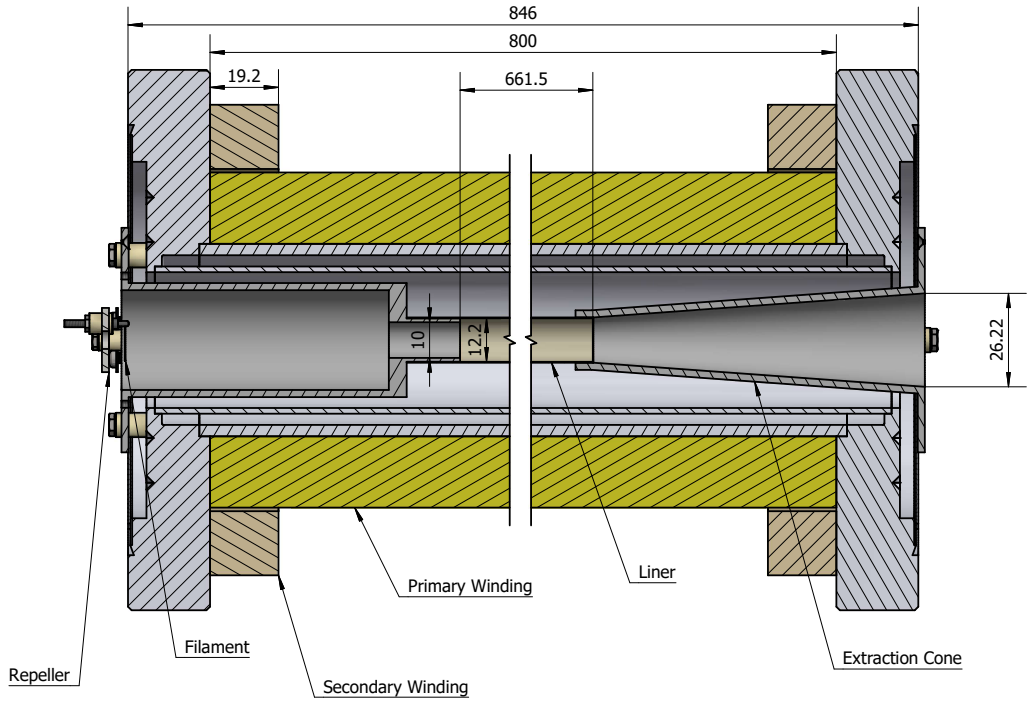


Figure 4.5: Cross sectional view through the solenoid with an externally mounted filament. The solenoid is shown in yellow and the additional secondary windings are illustrated in a darker shade on the ends. The filament sits just outside the solenoid on the left hand side of the diagram. The solenoid liner is 12.2 mm in diameter through most of the solenoid with a step at 10 mm to prevent ions exiting the left of the ioniser. The liner is then tapered on the right hand side to allow ions to be extracted.

charged particles, so care must be taken to avoid reflecting the electrons off the solenoid.

The filament reaches the space charge limit due to the weak field around the filament as shown by the emission properties in fig. 4.6. Placing the filament outside the solenoid leads to magnetic compression of the beam, but also leads to a weaker electric field around the filament. The weak electric field means that the space charge limit is quickly reached and no further emission can be extracted from the filament.

Electron beam ion sources (EBIS) often use magnetic compression to form the electron beam and can produce current densities $j > 1000\ \mathrm{A\,cm^{-2}}$ [89], however this is typically at much higher electron energies and so is not comparable with the present situation. However, recently the EBIS community has started to look at electron sources that have energies as low as 60 eV [92]. Ideally a larger emission current into the liner will lead to higher ionisation efficiencies meaning that in the future these high current electron sources could be extremely relevant.

The electron injection into the solenoid was investigated using a COMSOL simulation. COMSOL uses a finite element method to solve for the electromagnetic fields in a volume. Initially the magnetic field from the solenoid is solved, and the initial electric fields are solved. Using these fields and a suitable algorithm it is possible to trace the trajectory of the electron beam through the solenoid. Once the electron trajectories have been calculated, COMSOL can then calculate the space charge effects of the beam and then recalculate the

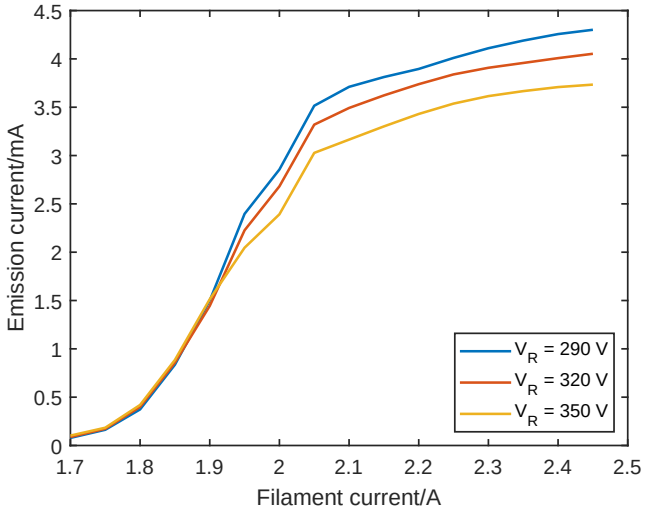


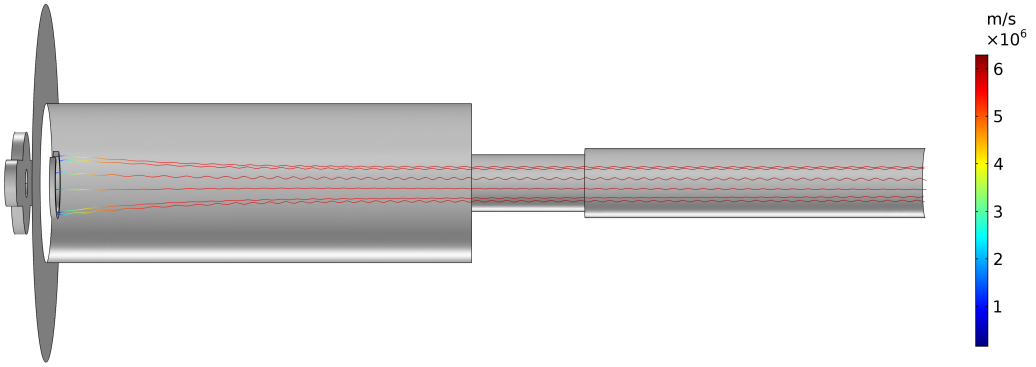
Figure 4.6: Emission characteristics for the externally mounted filament showing that the space charge limit is reached. As the filament current is increased the electron density around the filament increases until further emission is prevented. The repeller has a weak effect on the space charge limit around the filament by slightly changing the local electric field.

new trajectories. The process of recalculating the electric field and trajectories was iterated to ensure that the calculated space charge was correct.

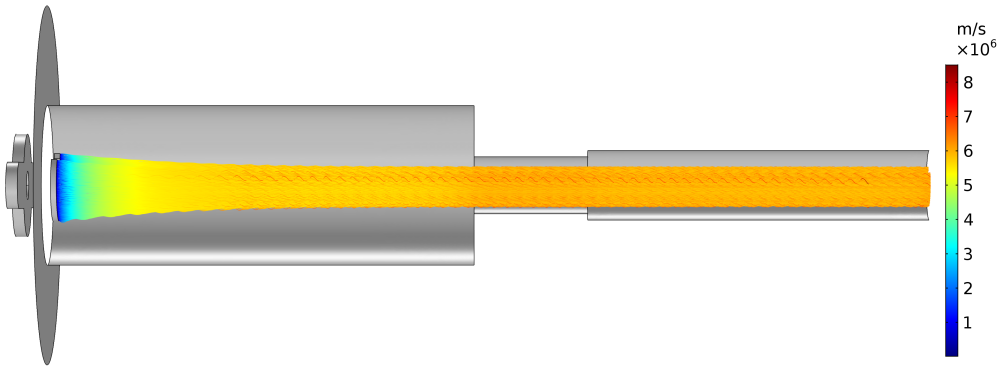
When the volume is approximated with a tetrahedral mesh, it is difficult to model thin objects accurately without a high density mesh. To reduce the complexity of the problem, the $200\,\mu\mathrm{m}$ iridium filament was approximated as a square wire that has a 1 mm width that follows a circular path with a radius of 5.75 mm and a 290° circular arc. If the space charge limited emission were to be calculated for this approximate filament, the current would be very inaccurate due to the much larger size in the simulation. Instead, a 3 mA emission current from the front face of the approximated filament is used, since this is approximately the largest emission current that could be experimentally generated.

The simulated electron trajectories in open mode are shown in fig. 4.7. Figure 4.7a shows a few electron trajectories and illustrates the helical trajectories followed by the electrons as they are pinned to the magnetic field lines. Figure 4.7b shows the full electron beam and clearly demonstrates the magnetic compression of the beam as it enters deep into the solenoid. In order to simulate the performance in closed mode the trajectories would need to be calculated down the entire length of the solenoid and then allowed to reflect multiple times which would be computational expensive. Instead, the performance is assessed in open mode and the electron trajectories are assumed to pass straight through the solenoid. Since only the open mode performance is being investigated the trajectories are ended when the particles enter the narrow part of the liner keeping the computation time down.

The potential on the centreline of the ioniser with and without space charge is shown in fig. 4.8, illustrating that a disadvantage to mounting the filament externally is that it creates a space charge well at the entrance of the liner. Since the liner wall is further from the electron beam at the entrance, the space charge well is deeper at the entrance than in the rest of the ioniser. The ion barrier ring is also shown to be behaving as expected from the step in potential between 55 mm and 65 mm from the start of the solenoid in fig. 4.8 which



(a)



(b)

Figure 4.7: Simulated trajectories of the electrons exiting the filament and entering the solenoid in open mode. The electrons are accelerated from the filament into the solenoid, but are also confined by the strong magnetic field to helical trajectories. The converging magnetic field leads to compression of the electron beam into the narrow liner. The particle speed is represented with the colour of the trajectory, the effect of the space charge well can be seen through the small change in velocity as the electron beam moves into the narrow part of the liner. (a) A plot of a few calculated electron trajectories exiting the filament and entering the solenoid. The electrons can be seen to follow helical trajectories around the magnetic field lines. (b) A plot of the full calculated electron beam as it exits the filament and enters the solenoid. The converging magnetic field can be seen to compress the electron beam it enters the solenoid.

creates the barrier preventing ions created to the right of the barrier from escaping to the left. Figure 4.9 shows the calculated space charge created by the electron beam and illustrates that the well depth is determined by the distance from the electron beam to the potential boundary at the liner wall.

Figure 4.10 illustrates an instability seen under certain conditions in the ion current which is one of the major issues with previous designs of the solenoidal ioniser. Figure 3.6 from Chisnall's thesis [77] shows hysteresis in the measured emission characteristics on the solenoidal ioniser used on the Cambridge spin-echo instrument. While simulations can be used to predict the behaviour of the ioniser, the complex interactions between the electron and ion space charge will be difficult to accurately predict. The deficiencies in simulations will have greater effects when the ioniser is operated in closed mode, as the electrons will have longer residence times inside the ioniser. It is therefore essential to experimentally

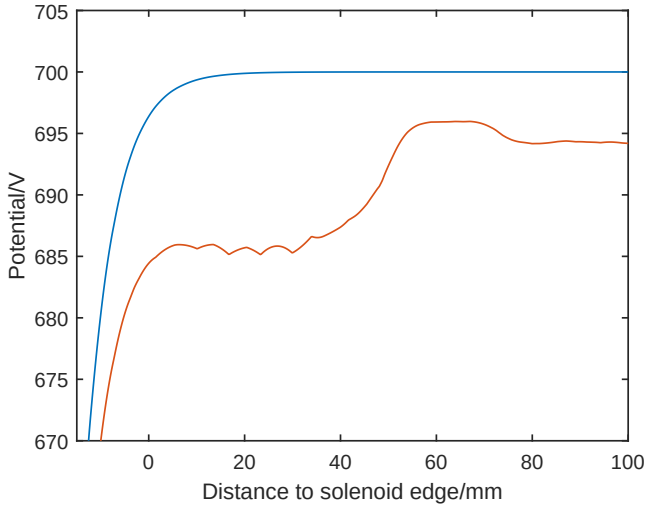


Figure 4.8: Potential on the centreline of the ioniser with and without space charge. The blue line shows the potential through the ioniser in the absence of an electron beam, while the orange line includes the space charge from a 3 mA electron beam. A potential well at the front of the ioniser will trap ions and could perturb the electron beam as it enters the ioniser. The oscillations seen in the orange curve is due to the finite size of the mesh used to calculate the space charge.

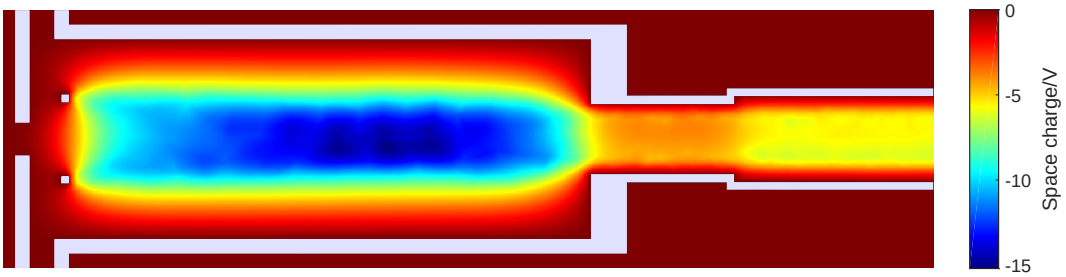


Figure 4.9: Space charge in 2D slice through the ioniser entrance. The space charge well gets deeper the further away the potential boundary is from the electron beam, so a large potential well is created at the entrance to the ioniser.

verify the ioniser characteristics.

It is likely that in the absence of a loss mechanism for the trapped electrons, a population of low energy electrons will be created inside the ioniser. As shown in fig. 3.6 from Chisnall's thesis [77], electron currents onto the repeller have been experimentally measured implying that electrons with an energy larger than their initial energy are being generated. One proposed mechanism for the electrons to gain energy is that the electrons are interacting with each other and exchanging energy. Therefore, when some high energy electrons are created there will also be a population of low energy electrons. The ioniser design has intentionally removed as many loss mechanisms as possible, with only two known loss mechanisms remaining. The first is readsoprtion at the filament, which is only possible if the electron has not lost energy. The second is collision with the liner wall after colliding with other particles and exchanging momentum as illustrated in fig. 4.11. If an electron had its velocity perpendicular to the magnetic field reversed after a collision, it would restart a helical trajectory up to two Larmor radii away from its previous trajectory centre without changing energy. Repeated collisions could allow the electron to reach the wall of the liner.

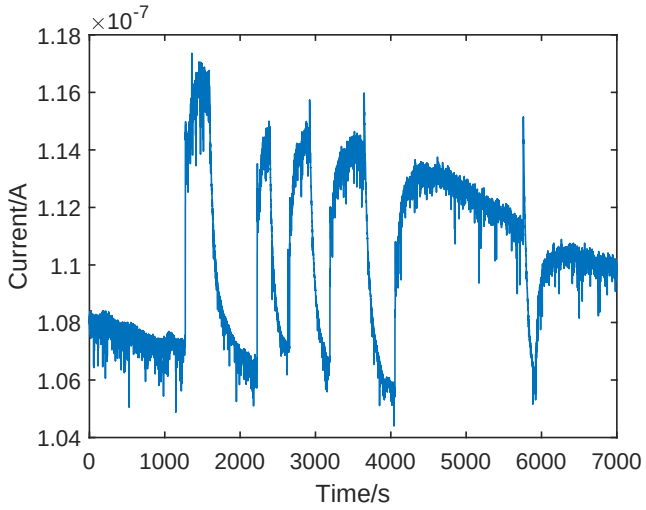


Figure 4.10: A plot of the variation of the mass selected ion current with time without changing any of the detector parameters. The plot illustrates how instabilities in the ioniser can produce strongly fluctuating ion currents that make measurements very difficult.

The diffusion of electrons across a strong magnetic field is an effect known to be present in radiation belts, however the Larmor radius is insignificant on astronomical length scales and therefore is not a dominant effect [93].

There are three main possibilities for the presence of instabilities and hysteresis in this ioniser design, the ion trap next to the filament that was demonstrated in fig. 4.8, the trapping of low energy electrons and fluctuations in a non-symmetric electron beam. The presence of an ion trap where the electrons enter the solenoid initially seems unimportant. Ions that are stuck in the trap will accumulate until the potential is raised back up such that the ions are no longer trapped and an equilibrium is reached. However, an ion trap provides a mechanism for hysteretic behaviour since the potential near the filament will depend on the previous electron and ion densities. Another theory for the presence of instabilities arises from there being a population of low energy electrons trapped in the ioniser which contribute to the space charge in the ioniser but do not contribute to ionisation and can only escape out of the liner by slowly moving towards the liner wall by collisions. Finally if the electron beam is not axially symmetric, the potential will also not be axially symmetric. The electron beam can either be axially asymmetric due to the filament design, or due to cold spots on the filament not emitting electrons. If the electron beam fluctuates, the potential will also fluctuate which could lead to parts of the ioniser not producing ions, thus creating an instability.

Additionally, placing the filament outside the solenoid will create a relatively weak electric field next to the filament preventing large emission currents. Ideally the ioniser will have a larger ionisation efficiency if more electrons are injected into the ioniser. If the local electric field around the filament is weak then the space charge limit will be reached at low emission currents.

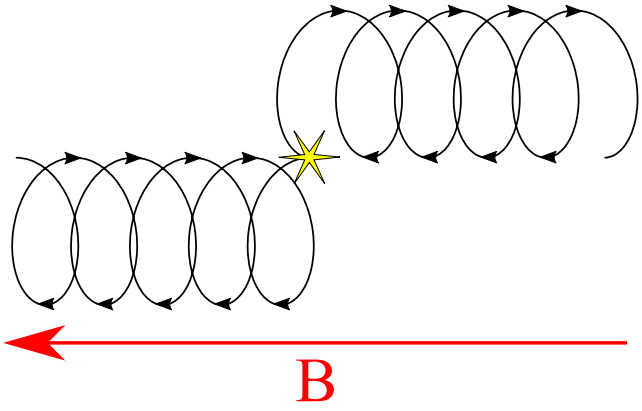


Figure 4.11: Schematic illustrating how an electron can move across a magnetic field without an energy change. An electron from the left of the figure is trapped by the magnetic field to move axially along the field. In the centre of the figure, the electron collides elastically with another particle and the velocity perpendicular to the magnetic field is reversed. The electron now restarts another helical trajectory, with the centre of the orbit shifted by two Larmor radii.

4.2.4 Ioniser performance with an internally mounted filament

The filament was mounted inside the solenoid to remove some of the potential issues in the ioniser. Figure 4.12 shows how the filament was mounted inside the solenoid. The filament was moved much further into the liner to remove the space charge well. The field between the filament and the liner is much stronger because they are closer together, so the well is eliminated. Since the filament is closer to the main tube of the liner, a grid needs to be added to ensure that the filament potential does not penetrate into the liner and pull ions out the wrong end of the ioniser. The addition of a grid also acts as a loss mechanism for the electrons, so that the electrons are removed even if they have lost energy. The close proximity of the filament to the liner increases the local electric field around the filament allowing much higher emission currents to be produced.

The filament geometry was also changed so that it is more compatible with the new position. Since the electrons are emitted deep in the solenoid, there is no magnetic compression of the electron beam and the electrons will travel in approximately straight lines parallel with the magnetic field. Therefore, to ensure that the electrons do enter the liner, the outer diameter of the filament needs to be smaller than the narrowest part of the liner $(10\,\mathrm{mm})$. To counteract the loss of filament area, the previous geometry of a wire loop was replaced with a spiral to maximise the emission area. The spiral filament should also create a more axially symmetric space charge potential inside the ioniser.

Figure 4.13 shows a COMSOL simulation of the new filament geometry approximating the filament and grid as flat plates inside the solenoid and using an emission current of 15 mA. The on axis potential no longer has a potential well due to the strong local electric field near to the filament. Similarly to before, it is difficult to model thin objects and so the grid and filament were approximated with flat plates. The emission current was chosen to approximately match typical emission currents obtained experimentally. The simulation used a flat plate as an approximation for the grid next to the filament, so the potential reaches the liner voltage at the grid on the right of fig. 4.13.

Mounting the filament internally leads to a lower emission current at a given filament

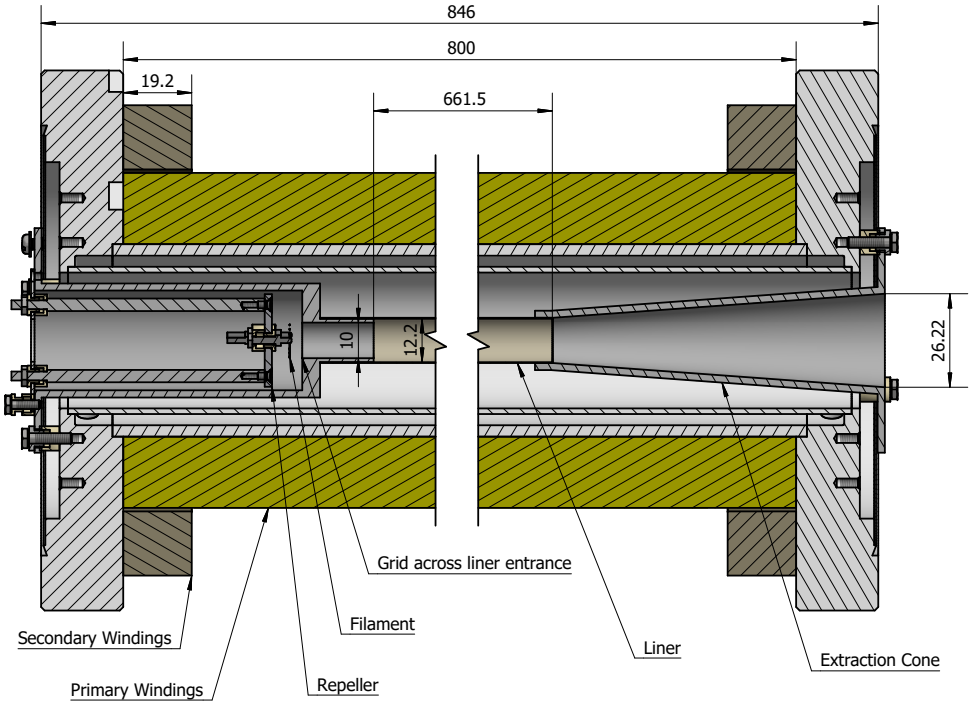


Figure 4.12: Cross sectional view through the solenoid with an internally mounted filament. The filament is now mounted so that it is inside the solenoid, and there is a small gap between the filament and the entrance to the liner to create a strong field and eliminate any potential traps. Long metal rods support the filament from the liner and ceramics are used to then insulate the filament mount from the liner.

current, but the stronger field at the filament means larger emission currents can be obtained before reaching the space charge limit. The emission properties of the internally mounted filament design are plotted in fig. 4.14. The emission characteristics in fig. 4.14a were measured by measuring every other point while decreasing the filament current and then filling in the gaps by increasing the current. It is thus clear that there is very little hysteresis in the data. The power supply used to drive the emission current has a maximum current of 20 mA so the emission current cannot be greater than this value. Figure 4.14b shows that larger electron energies create a stronger electric field to remove electrons from the filament, therefore creating a larger emission current.

4.2.5 Extraction

Ions are extracted from the ionisation volume by having a potential lower than the liner at the exit. Electrons can either be extracted or trapped inside the liner depending on the potential at the exit to the liner. Figure 4.2 illustrates the potential through the centreline of the ioniser for either when the electrons are trapped or allowed to pass straight through and hit the subsequent lens elements.

In closed mode with a filament mounted external to the solenoid, the electrons can only escape by reabsorption onto the filament, or by collision with a component at the liner poten-

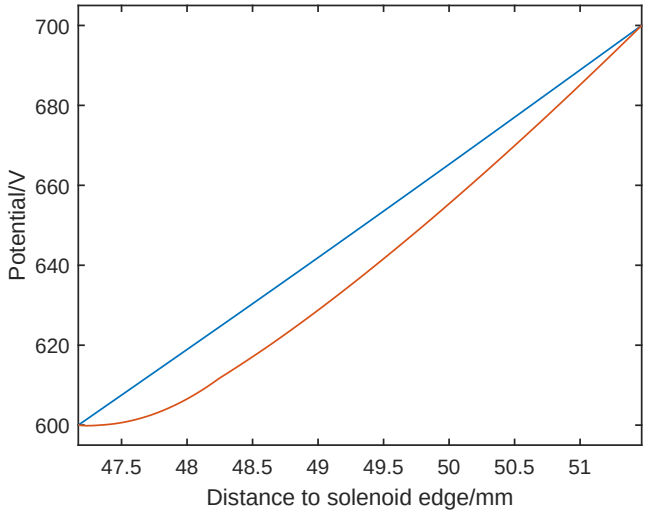


Figure 4.13: Potential on the centreline of the ioniser with and without space charge for the new filament geometry and a 15 mA emission current. Even with a much larger emission current, the potential is only slightly perturbed from the potential with no space charge. The grid was approximated with a flat plate so the potential reaches the liner voltage on the right of the figure.

tial. When operating the ioniser in closed mode, the only components with a high enough potential for the electrons to collide with are the liner or the filament itself. Electrons which are reabsorbed into the filament will not contribute to the emission current. Instead, the electrons that create an emission current eventually collide with the liner wall. As previously discussed, the electrons are pinned by the magnetic field and can only reach the liner by collisions. If filament is mounted internally, a grid is present inside the solenoid liner, at the liner potential, that will remove a fraction of electrons on each pass down the solenoid.

Referring back to fig. 4.2, the potential at the end of the liner is controlled by the first lens element after the solenoid which is a quadrupole lens element. Although, the effect of the quadrupole is discussed later, in section 4.3.4, the mean voltage determines the potential where the ions are extracted from the solenoid, and therefore controls whether the electrons are trapped inside the solenoid. The mean voltage applied to the quadrupole will be referred to as the mean extraction voltage.

Figure 4.15 shows how the emission current and current measured on the liner vary as the mean extraction voltage is changed with the filament mounted externally to the solenoid. Initially the electrons are trapped and all of the current flows onto the liner. After switching to open mode the electrons can escape from the liner, and the vast majority escape. Since there is no grid inside the solenoid when the filament is mounted externally to the solenoid, an insignificant emission current is measured onto the liner in open mode implying the electrons are successfully transferred through the entire solenoid by the magnetic field.

The maximum emission current that can be produced in open mode is 1 mA, it is therefore possible to compare the two operation modes at low emission currents. More detail on what sets the maximum emission current in open mode is provided in section 4.3.4.

With the internally mounted filament design at low filament currents, since the emission is no longer space charge limited, there is not a very large change in emission current

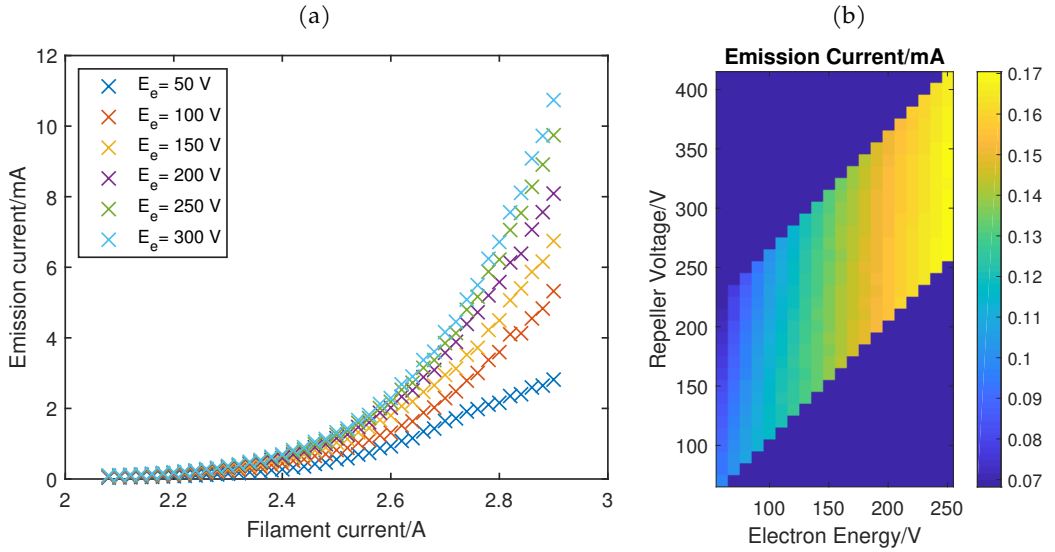


Figure 4.14: Emission characteristics of the internally mounted filament design demonstrating that the emission current is well behaved. (a) A series of plots of how the emission current varies as the filament current is increased for a range of electron energies. The space charge limit is approached when the electron energy is low, but larger electron energies allow much larger emission currents. The error bars are smaller than the plot symbol with the maximum fractional error given by 1%. The emission current is not measured to the maximum limit to avoid power supply damage and the pressure changing too much between the different runs. (b) A plot of how the emission current varies as the electron energy and repeller voltage are varied while the filament current is kept constant. As the field around the filament is made stronger by increasing the electron energy the emission current is increased. The repeller voltage has a weak effect on the emission current.

between open and closed mode. However, a significantly higher ion current is obtained in closed mode since the electrons are trapped inside the ioniser.

It is possible to switch between open and closed mode by either changing the electron energy or by changing the mean extraction voltage. The transition between open and closed mode occurs when the electron energy is equal to the mean extraction voltage. If the liner voltage is given by V_L, the electron energy with respect to the liner is given by V_e and the mean extraction voltage is given by αV_L, then the transition between open and closed mode will occur when

$$V_L = \frac{V_e}{1-\alpha} \tag{4.2}$$

Figure 4.16 shows how the ion current varies as the electron energy and repeller are varied. As the electron energy is increased the detector switches into open mode at the transition given by equation 4.2. A sharp decrease is seen in the efficiency of the detector when the ioniser switches to open mode confirming that there is a distinct difference between the two operating modes of the detector.

4.2.6 Modifications for a stagnated beam

Typically in atom scattering, the scattered helium beam has a well defined velocity and will propagate with minimal divergence. Helium that has diffracted from a surface will be travel with a very small velocity spread and forms a tight atom beam. Since there is little diver-

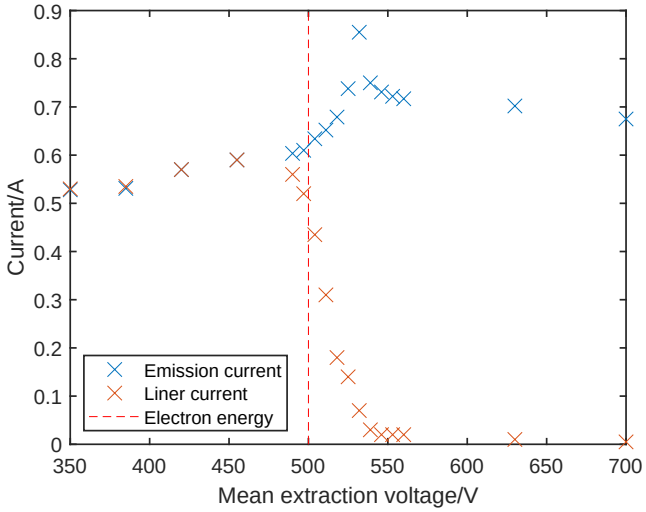


Figure 4.15: A plot of how the electron current varies as the ioniser is switched between open and closed mode. When the mean extraction voltage is lower than the electron energy, the electrons cannot escape from the liner and will eventually collide with the liner wall and the detector is in closed mode. When the mean extraction voltage is above the electron energy, the detector is in open mode and the electrons can escape from the ioniser. When the detector is in open mode the electrons pass straight through the solenoid and the current onto the liner drops to approximately zero. In the data, the electron energy is 500 V above ground due to the liner being 700 V above ground and the filament being held 200 V below the liner potential.

gence in the beam, the detector can be placed relatively far from the sample to achieve high angular resolution. Crossed beam ionisers allow the atom beam to travel straight through the ioniser, and an electron beam is directed into the atom beam to create ions.

The detector installed on the helium spin-echo machine in Cambridge "stagnates" the beam in centre of the ioniser to improve the detector efficiency [77]. If the atom beam is directed straight through the ioniser, the atoms will only have a short residency time inside the ionisation volume. However, if the atom beam collides with the liner wall in the middle of the ioniser, the atoms scatter from the rough surface into a variety of directions and then have to diffuse out of the ionisation volume. The stagnation of the beam increases the residence time of the atom in ionisation volume, since the diffusion of the atoms out of the narrow tube is slow.

Since the helium atoms are in a beam in the detector for the spin-echo instrument, the front of the ioniser can be well pumped to remove the background gas without influencing the helium beam. The beam travels straight through the front of the ioniser, before stagnating in the middle of the ioniser. However, the background gas will be stagnated in the front of the ioniser and by pumping the front of the detector, a significant fraction of the background gas can be removed.

In helium microscopy, the scattered atoms have a wide velocity distribution so the gas stagnates before reaching the detector. It is therefore no longer possible to pump the front to reduce the background gas. Instead care must be taken to ensure that the gas flows through the ionisation volume to ensure that it can be ionised and detected.

To ensure that the maximum amount of gas is ionised, all pumping paths that do not pass through the ionisation volume must be removed including the major pumping path

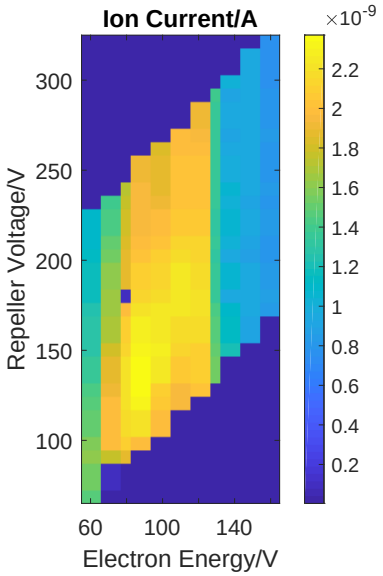


Figure 4.16: A plot of the ion current as the electron energy and repeller voltage are varied showing the transition from open to closed mode as a loss in ionisation efficiency. In this case, the quadrupole mean voltage is set 128 V below the liner, with each element set within a few volts of each other so that the extraction voltage is approximately 128 V. If the electron energy is lower than the extraction potential, then the electrons are trapped inside the solenoid and a large emission current is measured. If the electron energy is above the extraction potential then the electrons will only pass through the ioniser once, leading to a lower ion current from the ioniser.

on the spin-echo instrument. Figure 4.17 shows the potential pumping paths down the detector that have been removed. There is a path down the solenoid between the liner and the solenoid where no ionisation occurs that needs to be blocked. A ring made of insulating material (PEEK), was installed to eliminate the additional pumping path[1].

For a stagnation detector, the pumping speed of the detector is a compromise between the efficiency and time constant. In reality, the pumping speed and therefore pressure varies down the length of the ioniser, but let us consider the front of the ioniser as an illustrative example. A fixed flow rate of helium, Q, enters the ioniser and the pump removes the gas at a flow rate given by $Q_P = Sp$, where S is the effective pumping speed of the front of the ioniser and $p(t)$ is the pressure at the front. The pressure at the front of the ioniser is then given by,

$$V\frac{dp}{dt} = Q - Sp. \tag{4.3}$$

Equation 4.3 is a first order differential equation that can be solved through standard methods. If the pressure at time $t = 0$ is given by $p(0) = p_0$, the pressure as a function of time is given by,

$$p(t) = \frac{Q}{S} + \left(p_0 - \frac{Q}{S}\right) e^{-(S/V)\,t}. \tag{4.4}$$

The pressure will vary exponentially reaching a final value in the long time limit given by,

$$\lim_{t\to\infty} p(t) = \frac{Q}{S}, \tag{4.5}$$

[1]I am grateful to S Lambrick for his contribution to the installation of the block to the pumping path

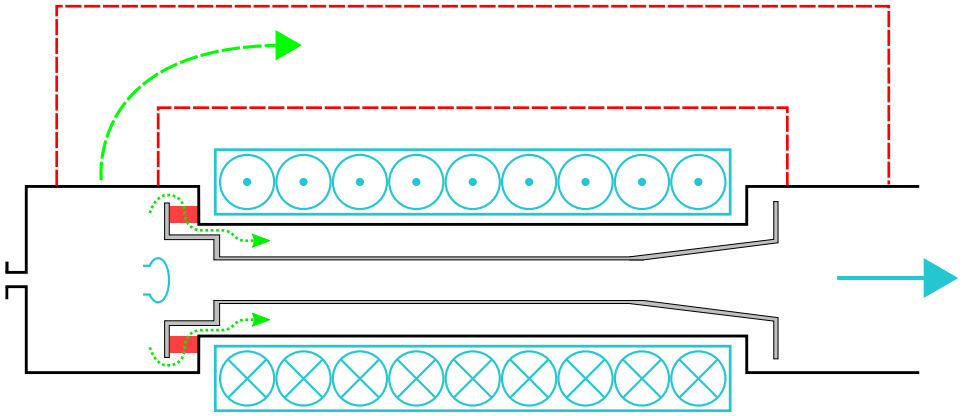


Figure 4.17: Schematic of the pumping paths through the detector that don't lead to ionisation of the helium atoms. In the spin-echo detector a large tube, which is illustrated with a dashed red line, that connects the front and back of the detector needs to be removed. An additional path exists between the liner and the solenoid which can be blocked with a PEEK ring as illustrated in solid red.

showing that lower pumping speeds will result in higher pressures in the detector and therefore a higher ioniser efficiency. However, the pressure will vary exponentially with a time constant given by

$$\tau = \frac{V}{S}. \tag{4.6}$$

So lower pumping speeds will also mean that the response time of the detector will also be lower. It is therefore important to balance the efficiency of the ioniser and the response time to achieve the maximum performance.

Ward [94] designed the front chamber of the ioniser with a small volume to minimise the volume and keep the time constant as low as possible. By filling in the front chamber with material it is possible to reduce the chamber volume further to keep the response time of the instrument as low as possible

Ward also designed a small input tube to the detector to improve performance, however the small input tube also acts to prevent atoms from entering the detector. The reduced pumping speed of the tube connecting the detector to the microscope does ensure that helium that reaches the detector is trapped. However it also acts to prevent helium from entering the detector and increases the response time as discussed in section 4.5.1.

4.3 Mass filtering and ion optics

4.3.1 Theoretical ideal operation

Once the ions have been produced, they need to be separated by mass. The vast majority of the ions produced will be due to other background gas in the vacuum such as hydrogen, so the helium ions need to be separated from the other masses. Magnetic sectors are often used in helium scattering experiments as mass filters since other techniques such as quadrupole mass filters [95] typically have a lower transmission probability. Since the ioniser is operating with a large electron emission current, a space charge well is created that introduces a spread in the ion energies produced. To ensure the ions are kept close enough in energy

to be separable, the ions are accelerated up to a higher energy. Quadrupole mass filters are widely used, but are typically only suitable for analysing ions with a low kinetic energy, but there is work on producing filters for ions with a large kinetic energy [96].

Mass separation is achieved using a permanent magnetic field and varying the ion energy to select different masses to pass through the filter. The force, $\mathbf{F}$, on an ion travelling through an electromagnetic field is given by the Lorentz force,

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{u} \times \mathbf{B}\right). \tag{4.7}$$

$\mathbf{E}$ is the electric field strength, $\mathbf{B}$ is the magnetic flux density and $\mathbf{u}$ is the velocity of the particle. Let us consider an idealised situation in which the ions enter a uniform magnetic field from a field free region. In the absence of an electric field, the force on the ions will always be perpendicular to their velocity, and the ions will undergo circular motion inside the magnetic field region. Equating eq. 4.7 to the centripetal force needed to produce circular motion, $F_{\mathrm{CM}} = mu^2/R$, gives the radius of the ion circular motion as,

$$R = \frac{mu}{qB}. \tag{4.8}$$

$u = |\mathbf{u}|$ is the speed of the ion, m is the mass of the ion, q is the charge of the ion and $B = |\mathbf{B}|$ is the magnitude of the magnetic flux density. The kinetic energy of the ions is supplied from acceleration of the ions out of the liner to the ground potential of the chamber. If the liner is at a potential V, the kinetic energy of the ions will be $T = q\,V$, so the speed of the ions is given by,

$$u = \sqrt{\frac{2qV}{m}}, \tag{4.9}$$

and therefore the radius of curvature of the ions in terms of the liner potential will be,

$$R = \frac{1}{B}\sqrt{\frac{2mV}{q}}. \tag{4.10}$$

Therefore, by only extracting ions with a given trajectory through the magnetic field we are able to distinguish ions with different mass to charge ratios m/q. Varying the liner potential, V, allows control over which mass to charge ratios are transmitted through the magnetic sector, any ion that does not perform a half circle will collide with the chamber wall.

The difficulty in using a magnetic sector to perform the mass filtering is it requires a uniform magnetic field along the ion trajectory. A crucial component in mass selections is constructing a uniform magnetic field to perform the filtering, but also keeping the fringe fields to a minimum so that as close to perfect circular trajectories are performed.

Ions exiting the liner will be accelerated into the chamber and will diverge, so ion optics are required to transfer the ions from the liner to the magnetic sector and minimise losses. Additionally, since there are no electric fields in the magnetic sector, the beam will be diverging and so ion optics are required to manipulate the beam.

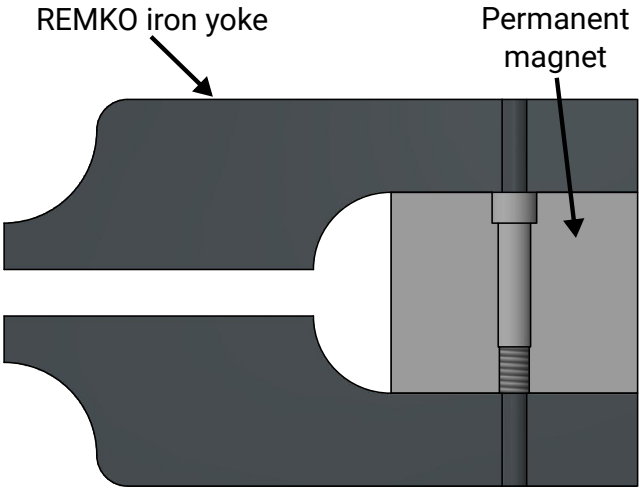


Figure 4.18: Cross sectional view of the magnetic sector taken in a plane parallel to the field. The permanent magnet is lightly shaded on the right of the figure and the soft iron yoke is shown in a darker shading. The yoke is designed to provide a uniform field in the air gap on the left hand side of the figure to provide mass filtering of the ions passing through the magnetic sector. Note that in this rendering the magnet is shown as homogeneous but is not solid neodymium throughout.

4.3.2 Magnet design overview

A uniform field is created by using pole pieces to direct the field from a cylindrical permanent magnet. The pole pieces were designed by Ward and are shown with the cylindrical magnet in fig. 4.18. The magnet has a diameter of 80 mm and is 65 mm long, however the magnet is not homogeneous throughout. The main bulk of the magnet is made from a hollow cylinder of N42 neodymium (NdFeB) that has an outer diameter of 76 mm, an inner diameter of 28 mm and a length of 48 mm, the rest of the cylinder is made from a ferritic stainless steel and is encased in 2 mm of 316 stainless steel. An additional brass insert to adapt the thread size was placed inside the magnet to account for an error in the size of the holes in the pole pieces, it is not included in the diagram. The pole pieces are made from an extremely pure, soft magnetic iron that has the trade name REMKO.

The magnetic sector needs careful assembly due to the strong pull force of the magnet. The manufacturers quote the pull force of the magnet from a steel surface as 155 kg, so the magnet poses a significant risk. To assemble the pole pieces and magnet, the pole pieces were attached to the plastic boxes they are housed in, loosely aligned with an alignment rod to simulate the magnet, and then bolted down hard to an Instron universal testing machine which allowed the components to be brought together in a controlled manner. The pole pieces were then separated and the magnet was slowly brought into contact by moving down a rod of studding. Once the magnet was in contact with the first pole piece, the second pole piece was lowered down back into position using the Instron. By following this careful procedure, the magnet and the pole pieces can be assembled with full control at all stages.

Once the magnetic sector was constructed, the field strength was measured as shown in fig. 4.19. The experimentally measured magnetic field strength is very uniform deep inside the air gap. Figure 4.19 also shows the magnetic field that was simulated by Ward in Lorentz

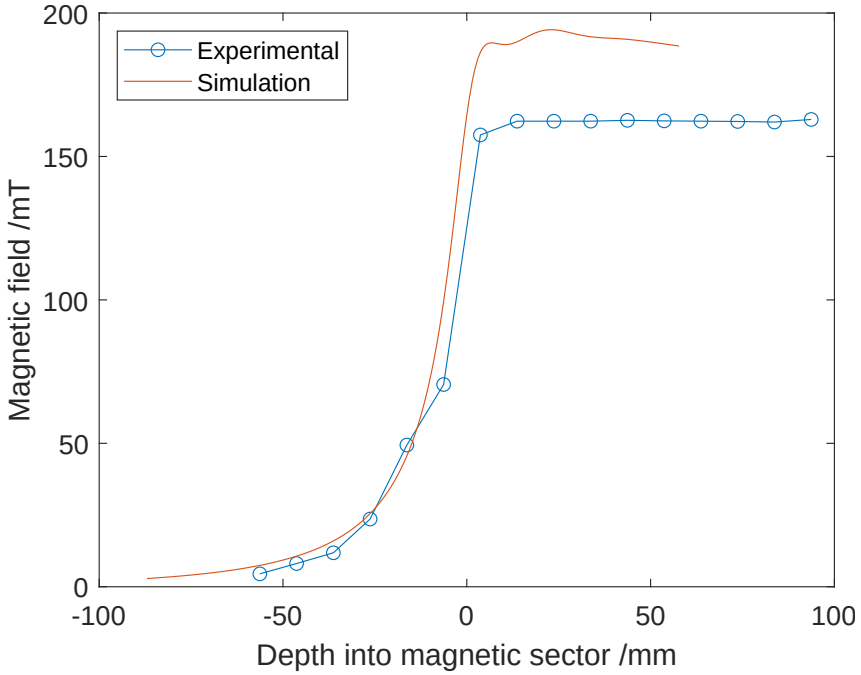


Figure 4.19: A plot of the experimental and simulated values of the magnetic field through the air gap in the magnet. The experimental values are lower than the simulated values due to the magnet being used not being a continuous cylinder as designed. The experimentally measured magnetic field in the air gap is very uniform which is ideal for mass filtering. The experimental magnetic field was measured with a gaussmeter that was rotated to find the maximum value of the magnetic field to avoid misalignment.

for comparison. It is clear that the experimentally measured magnetic field is weaker than predicted from simulations. Ward assumed that the magnet is a solid cylinder, where as previously discussed the magnet is not a homogeneous block of permanent magnet. It is therefore unsurprising that the experimental magnetic field is weaker since there is less magnetic material in the magnet that has actually been used.

Figure 4.20 shows the flight tube that sits in the magnetic sector. A serrated edge was added to reduce the transmission probability of species that hit the chamber wall. Ideally, once the stray parts of the beam have hit the wall of the chamber they are immediately neutralised and then removed from the vacuum chamber. However, there is a small probability that the ions will reflect from the chamber wall. The serrated edges were added to the chamber wall to maximise the number of collisions these stray ions must take.

4.3.3 Ion optics design overview

Ion optics are required to transfer the ions from the ioniser to the magnet and then from the magnet into the dynode assembly. The ion optics are used to extract ions from the liner, and then guide them towards the magnetic sector. The transfer of ions is achieved by using a quadrupole lens to extract and direct the ions, followed by a lens to correct the trajectories to bring the trajectories to being close to parallel. A second lens is then used to focus for transfer around the magnetic sector. The ion optics were designed by Allison and Ward, with the mechanical installation design, assembly and testing of the ion optics performed by myself.

Figure 4.20: Image showing the interior of the flight tube that is placed into the air gap of the magnet during assembly. After the magnet is placed over the chamber, ions that enter the circular arc will be deflected by the magnetic field and ions with a specific m/z ratio will make it around the semicircular path without colliding with the chamber wall. The serrated edges prevent ions that collide with the chamber wall from reaching the exit of the magnetic sector. Photograph taken by workshop technician Kevin Mott

A quadrupole lens is used at the ioniser exit to not only provide the extraction voltage, but also to focus and direct the beam in two dimensions. For a magnetic sector that deflects the ions by $180°$, the ions are focused to first order in the radial direction, but the magnet has no effect on the velocity component parallel to the magnetic field [97]. Therefore the magnet introduces astigmatic focussing of the beam. By using a quadrupole lens it is possible to correct the astigmatism and hence obtain a circular beam profile on exit from the magnet maximising the transmission probability. The quadrupole lens can also be configured to act as pusher-plates, which allows the beam to be shifted in any direction to account for any misalignments in the ion optics stack.

The total ion optic stack consists of a quadrupole lens and three ion optic lenses. After the quadrupole lens extracts the ions, the ions are strongly diverging so the first lens is used to bring the trajectories back to close to parallel. The second lens is used to focus the ions so that they can pass through the magnet gap and be mass selected. After exiting the magnet, the ion trajectories again diverge and need to be focused back to parallel to be transferred to the dynode assembly.

Figure 4.21 shows LORENTZ simulations by Ward for trajectories of the ions up to the exit from the magnet. As noted earlier, the magnetic field strength does not match the actual field strength, however the simulation does illustrate the general behaviour of the ions. It can be seen that the masses are separated in space, however there is limited space between them, so the mass resolution is expected to be relatively poor without adding further apertures. The simulation shown in fig. 4.21 has a transmission probability from the liner to the

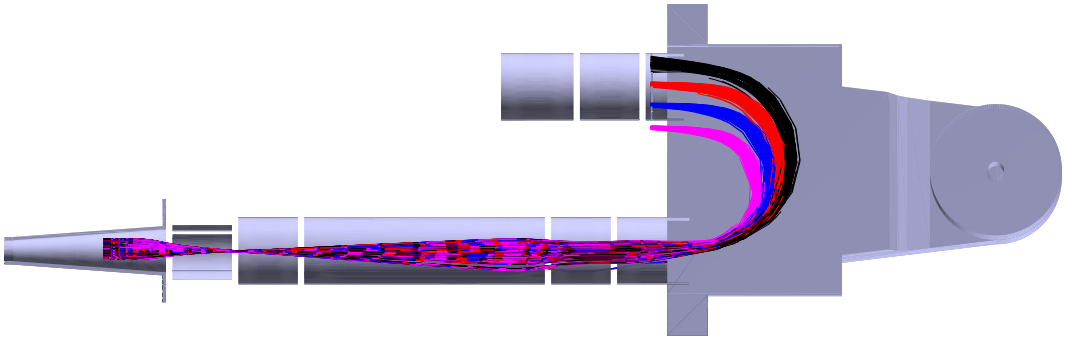


Figure 4.21: Simulation of ion optics to transfer the ions from the ioniser to the magnet and then subsequent mass filtering. Simulation performed by Ward in LORENTZ and replotted for this thesis by myself. Trajectories correspond to singly charged ions of mass 2 amu (purple), 3 amu (blue), 4 amu (red) and mass 5 amu (black). The magnetic field is calculated from a 3D model of the permanent magnet, but as shown in fig. 4.19 this is not an accurate field to use. The lens on the ion optics stack outgoing from the magnet and the dynode assembly are not included in the simulation.

magnet exit of 40% due to the ions not being focussed correctly through the magnet in the direction parallel to the magnetic field. The poor focussing leads to a large proportion of the ions being lost after hitting the chamber wall.

The ion optics stacks were mechanically installed in the detector and electrical connections made ensuring that the lenses are insulated from each other and the chamber wall. The ion optic lenses are made from perforated stainless steel sheets that were rolled into cylinders that are held in stainless steel rings. The lenses are separated from each other by ceramics to ensure they are electrically isolated. Each ion optics stack has four ceramic rods running down their length, two of which are hollow with small studding threaded down the middle. Nuts on the studding are used to compress the lenses together and hold them in place. On the ion optics stack from the liner to the magnet, the two solid ceramic rods slot into holes in the liner and support the weight of the optics. The other end of the ion optics slots into a support structure that rests on the chamber wall as shown at the bottom of fig. 4.22a and installed in the chamber in fig. 4.22b. Two of the quadrupole elements are held in place by the studding that holds the lenses together, the remaining two quadrupole elements are fixed in place by PEEK clamps that clip the quadrupole elements to the first lens.

4.3.4 Voltage scaling

The detector is electronically controlled by a series of Applied Kilovolts modules. The power supply set up is illustrated in fig. 4.23, where all power supplies except those involved in the filament are ground referenced. The filament setup is floated up to the liner energy, with the electron energy set relative to the liner energy so that the emission current can easily be measured from the filament. The filament current and emission are controlled using a Linduino microcontroller. For completeness, the circuit diagram for the printed circuit board used to measure the emission current is included in appendix B.

A preload resistor similar in design to the one installed on the spin-echo detector [77] is installed to prevent damage to the liner power supply since it cannot sink current. Fig-

(a) (b)

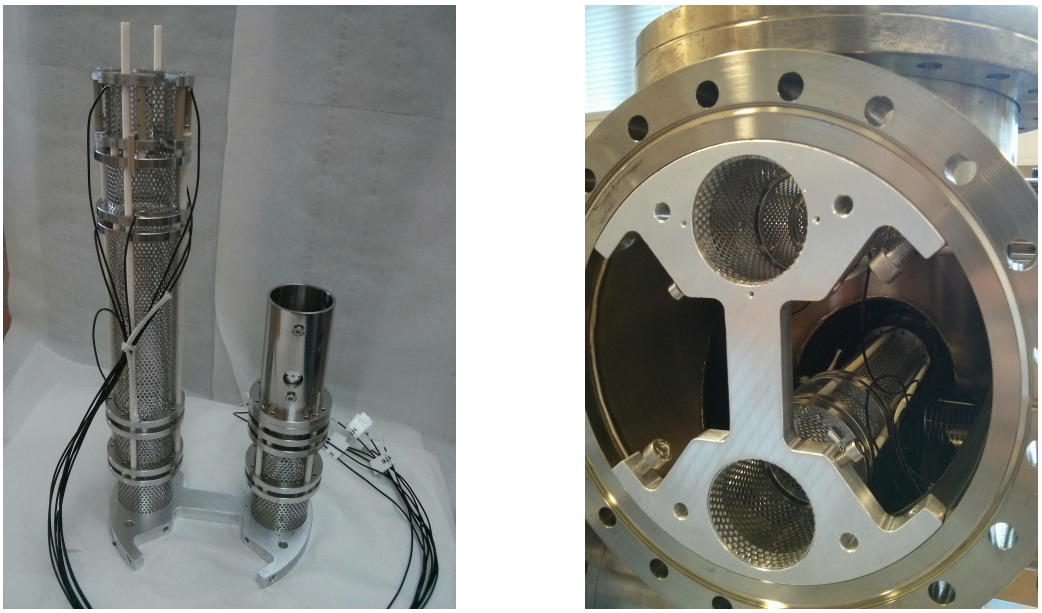


Figure 4.22: Images of the ion optics on the bench and installed inside the machine. (a) Assembled ion optics on the bench with the electrical connections in place. The left hand ion optics stack runs from the ioniser to the magnet and from top to bottom consists of a quadrupole, first lens element, long ground element, second lens and a ground lens. The right hand ion optics stack runs from the exit of the magnet to the dynode and from bottom to top consists of a ground element, the third lens element and the dynode assembly. (b) Ion optics installed inside chamber before adding the chamber splitter as described in section 4.5.2. The support structure screws against the chamber wall to keep the ion optics in place inside the chamber. The $180°$ magnetic sector is then added to move the ions from the bottom ion optics stack up to the top.

ure 4.24 shows how electrons leaving the filament can travel back to ground, requiring a large conventional current to flow back through the liner power supply. By adding in a resistor, the current flows through the resistor instead, preventing any possible damage to the power supply. The rest of the system should only have small ion currents of the order $1\,\mu$A flowing into them, so it is assumed the power supplies can deal with these smaller currents. A series of resistors were used to create a 440 kΩ resistance with a maximum power dissipation of 24 W, which should provide 1.6 mA when the liner is at 700 V. Therefore an emission of greater than 1.6 mA cannot be used in open mode, since that will require the power supply to sink current. The maximum power that the power supply can provide is 10 W, which at 700 V corresponds to 14 mA of current. Decreasing the preload resistance will not significantly extend the emission range, since the power supply would not provide enough current for the detector to operate in open mode.

Since the system is designed to have varying ion energies, the ion optics also need to vary so that the focusing is achieved at any mass. The detector varies the mass being detected by varying the ion energy, but as the ion energy changes the optics also need to be altered for the trajectories to remain unchanged. In addition, the voltages near the filament will also need to be altered with the liner voltage to keep the same field strength near the filament.

The lens voltages need to be varied multiplicatively with respect to the liner voltage to keep the same lens properties [98]. As the ions' speed through the optics is increased with

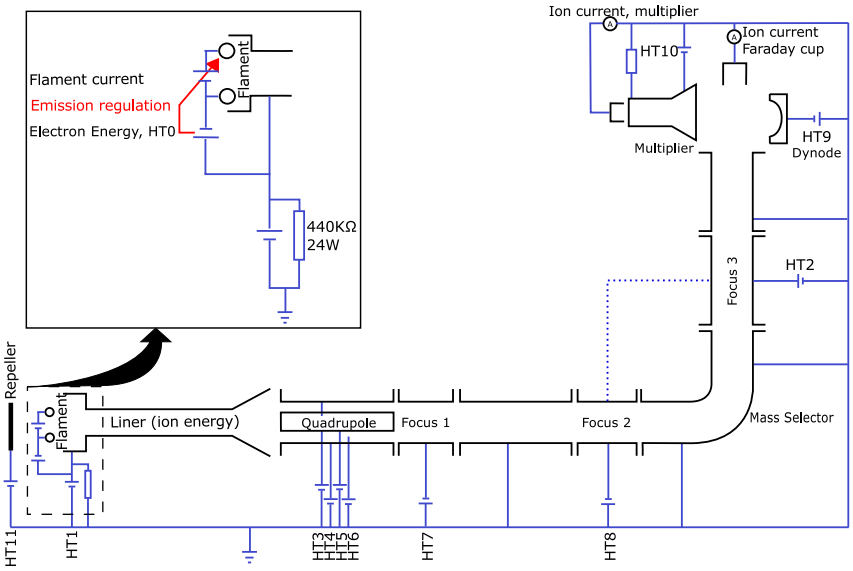


Figure 4.23: Power supply schematic illustrating how the electrical components in the detector are operated. The liner power supply includes a 440 kΩ preload resistor to allow electrons to flow in open mode. The filament is floated by both the liner voltage and the electron energy power supply. The electron multiplier has the voltage set on the back by a voltage divider that uses the internal resistance of the multiplier itself. Adapted from fig. 8 in Cambridge NEMI report [94].

the liner voltage, a stronger field will be needed to deflect the ions by the same amount.

Additionally, the voltages of the filament and repeller will need to be scaled to maintain a constant field and electron energy as the liner voltage is changed. These voltages need to be scaled linearly with respect to the liner voltage to keep the same electric field across the filament.

The voltages are automatically scaled in software by using a set of control variables to define the voltages. Performing a mass scan would be difficult if the user had to directly set each power supply voltage relative to ground. Instead a control system is used where the voltages are all adjusted with respect to the liner voltage automatically according to the lens type. Figure 4.25 illustrates how the different lens types are operated. Offset lenses, such as the repeller, are set by the potential difference between the lens and the liner voltage. As the ion energy is varied, the voltage on an offset lens is varied with a fixed offset. Proportional lenses, such as the ion optics lenses, are set as a proportion of the liner voltage. As the liner voltage is varied, the voltage on a proportional lens is varied multiplicatively with respect to the liner voltage. Fixed lenses are unaffected by changing the liner voltage.

The control variable system can also be used to control the quadrupole lens elements in a more intuitive way. Instead of directly setting the four voltages on each element of the quadrupole, the elements can be set via control variables. Figure 4.26 illustrates how the quadrupole can be set using four control variables, the mean variable, the quadrupole variable, a push x variable and a push y variable. The mean variable sets the mean voltage of the quadrupole elements and is used to control the extraction from the liner. The quadrupole variable sets the strength of the quadrupole, fig. 4.26(a) illustrates how the quadrupole works with the elements above and below the beam being set with an opposite voltage to the elements to the left and right of the beam. As then shown in fig. 4.26(b) the size of this

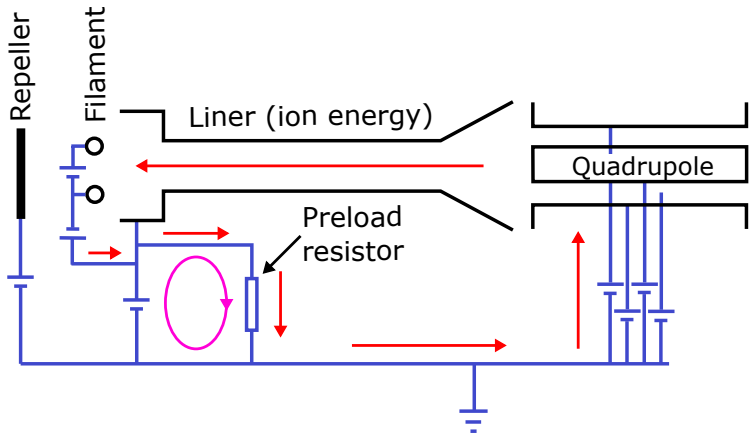


Figure 4.24: A circuit diagram illustrating how a preload resistor is used to prevent the liner power supply from sinking current, diagram inspired by figure 3.2 in Chisnall's PhD thesis [77]. Red and purple arrows show the directions of conventional current which is opposite to the direction the electrons flow in. In open mode electrons will travel from the filament to the elements at the other end of the solenoid. To ensure that the liner power supply does not sink current, a resistor is added to ground so that current instead flows around the power supply. The current in purple is the supply current from the power supply and as the emission current increases in open mode, the supply current shown in purple will decrease.

voltage offset is then varied by the quadrupole variable, which will then set the shape of the beam to counter the astigmatic focusing of the magnet. Finally, the push x and push y variables are illustrated by fig. 4.26(c) and are used to move the beam in any direction to align the beam within the rest of the system.

The voltages for each of the m power supplies can then be generated from the n control variables. Note that, in general $n \neq m$, but for the current implementation there are thirteen supply outputs and control inputs. In order to calculate the output voltages, the i^{th} setup variable (x_i) is converted into a control voltage $\left(V_i^C\right)$ and this conversion depends on the type of variable it is (i.e. proportional, offset or fixed). If x_0 is the variable that sets the ion energy (liner voltage), then table 4.1 details how the control voltages are generated. The scaling parameter defines the type of the variable and is defined in a configuration file.

Variable Type	**Control Voltage**	**"scalings" value**
Proportional	$V_i^C = x_0 \times x_i$	-1
Fixed	$V_i^C = x_0$	0
Offset	$V_i^C = x_0 - x_i$	1

Table 4.1: Table to show how the control voltages are generated from the control variables. The "scalings" value is the code used internally in the program to determine the variable's behaviour.

The control voltages, $\left(V_i^C\right)$, can then be used to generate the output voltages (V_j^S) which are used to set the voltages on the power supplies,

$$V_j^S = \sum_i \alpha_{ji} V_i^C. \tag{4.11}$$

Where α_{ji} is an input matrix determined from configuration files. Combining all of this

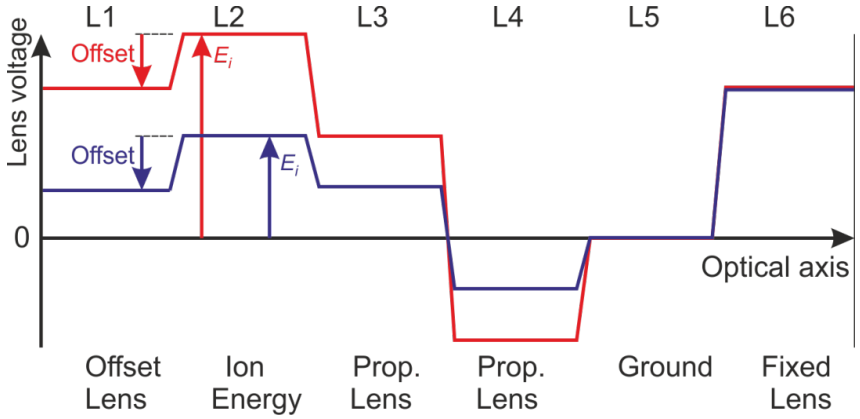


Figure 4.25: Schematic illustrating how the different types of lenses have their voltages scale with the liner voltage. Taken from fig. 9 in Cambridge NEMI report [94]. Offset lenses have a fixed offset from the liner voltage, proportional lenses are set as a control variables multiplied by the liner voltage, while the fixed lenses do not vary with liner voltage. The blue and red lines show examples of how the potential down the optic axis vary for two different liner voltages.

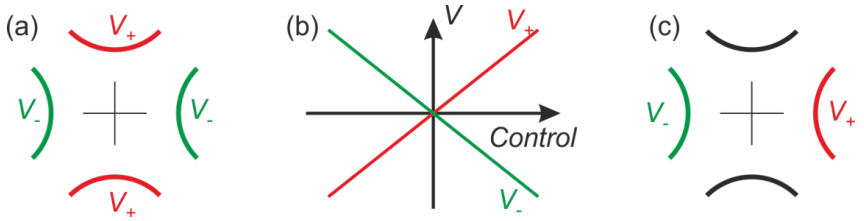


Figure 4.26: Schematic illustrating how the quadrupole voltages are set using control variables instead of directly setting the voltage. Taken from fig. 10 in Cambridge NEMI report [94]. (a) shows a quadrupole element with a quadrupole field generated using the voltages V_+ and V_- applied to opposing electrodes; (c) shows the voltages, V_+ and V_-, to create a field pushing the beam in a horizontal direction; (b) illustrates how the voltages, V_+ and V_-, relate to a single control variable, labelled "Control".

together allows mass scans to be performed.

Using the voltage scaling allows mass scans such as the one shown in fig. 4.27. As the liner voltage is varied different masses are transferred around the magnetic sector. Approximating the magnetic field as a uniform field with a magnetic flux density of 162 mT and the transmission path through the magnet as a semicircle with a radius of 50 mm then we can rearrange equation 4.10 to estimate the liner voltage needed to allow transmission of ions with a $m/z = 4$. Substituting in the above values predicts that $m/z = 4$ will be transmitted at 790 V, which is close to the experimental value of $\approx$ 700 V considering the crudeness of the assumptions. In figure 4.27 we also see the $m/z = 3$ peak at about 900 V, the edge of the $m/z = 2$ peak above 1100 V and the edge of higher mass peak beyond 500 V. While the peaks are not perfectly resolved from each other, it is clear that there will be an insignificant number of ions bleeding into the $m/z = 4$ at the centre of the peak.

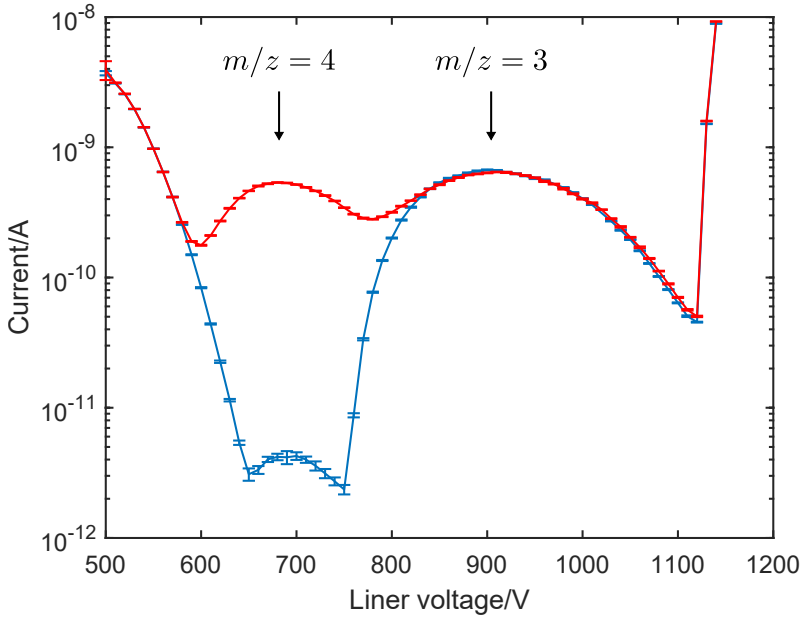


Figure 4.27: A plot showing two mass scans with (red) and without (blue) a helium leak into the detector. The mass 4 peak appears around 700 V and increases in size when the helium gas is bled into the system.

4.4 Conversion dynode

4.4.1 Benefits and design criteria of a conversion dynode

The helium ion current is generally too low to be measured directly, so the signal needs to be amplified so that an ammeter can measure the current. The noise floor of a reasonable good ammeter is typically about 10^{-12} A to 10^{-14} A. Electron multipliers are used to amplify the current such that it is possible to measure the current without the ammeter background being an issue.

Electron multipliers multiply input currents using the process of secondary emission of electrons. Secondary emission occurs when a primary particle with a high energy collides with a surface. Secondary particles can be emitted with a low energy (usually < 10 eV) with directions given by a cosine distribution from the surface normal [99]. Electron multipliers are designed in a way where an incident ion or electron will cause a large number of electrons to be generated upon impact with the surface, these electrons will then be accelerated into another surface so that they will each generate further electrons. The multiplication process is repeated many times inside the multiplier so that the number of electrons exiting will be far higher than the ion current incident on the entrance. Originally the electrons inside the multiplier were accelerated between individual surfaces with specified voltages in a design called a discrete-dynode multiplier. However, more recent electron multipliers are made from glass tubes, with a potential difference applied down the length of the tube to accelerate the electrons. Such electron multipliers are referred to as channel electron multipliers, or more commonly known under the trade name "channeltron" [100].

Electron multipliers typically fail through either of two mechanisms: degradation of the surface at the input or degradation of the surface at the output. The reasonably large

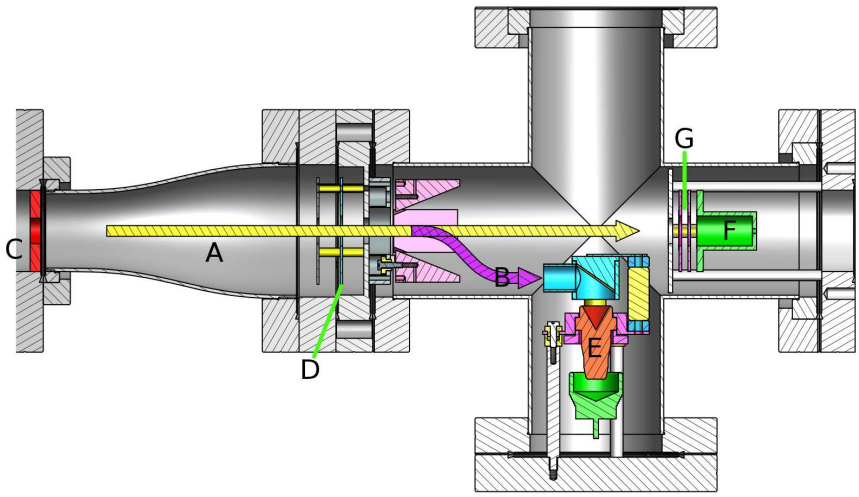
electron current at the multiplier output can damage the surface and reduce the secondary emission yield at the exit. The larger the output current, the faster the surface will degrade and so the lifetime of the multiplier is often measured as the total charge produced from the multiplier. The lifetime output charge produced from a channeltron can be as high as 30 C under ideal conditions [100]. The second failure mechanism is through degradation of the input surface due to collisions with high energy particles that can cause chemical changes in the glass or sputter the thin coating from the surface. Instead, we can preserve the input surface by using a conversion dynode to replace the high energy ions with electrons. The lower mass of the electrons means that collisions will cause less damage to the input surface of the multiplier. Conversion dynodes can also be used to improve the sensitivity to high mass species in mass spectrometers [101].

Moving the electron multiplier off axis prevents high energy neutral species from being measured, which can contribute significantly at low ion currents. Neutral high energy particles can be generated in the ioniser, and it is well known that line of sight needs to be blocked between the ion source and the electron multiplier. The multiplier is also sensitive to light, so light emitted from the hot filament can also cause a background signal from the electron multiplier. Neutral species transmission from the ioniser should be suppressed by both the 180° bend needed for particles to pass through the mass filtering and the serrated edges added to catch chamber wall as described in section 4.3.2. To further eliminate the chances of detecting a neutral species, the electron multiplier is moved off axis so that only charged species will strike its surface.

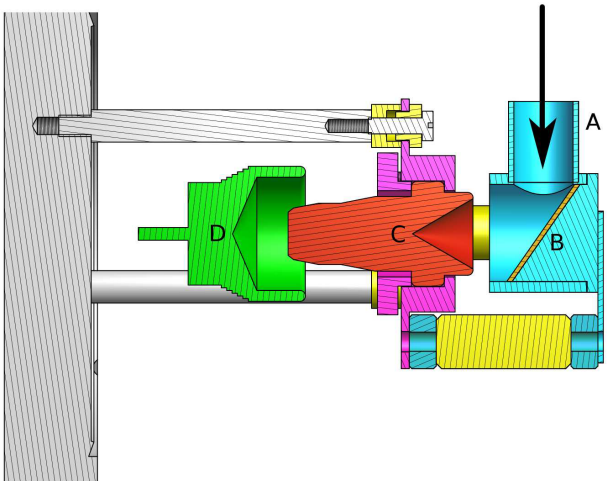
Previous conversion dynode designs in the group move the entire beam off axis before accelerating the beam into a conversion dynode, then under ideal operation all the secondary electrons should be easy to guide into a multiplier. The design used in the Cambridge spin-echo apparatus is shown in fig. 4.28. The difficulty in this design is in guiding the incident ions into the dynode assembly and down the narrow tube, but once the ions strike the conversion dynode it is expected under ideal conditions that an overwhelming majority of the secondary electrons that are created will be detected by the electron multiplier.

The previous work by Chisnall on the dynode produced unusual results, where the optimum configuration was obtained when the conversion dynode and the front of the electron multiplier were at the same potential. As previously mentioned, secondary electrons should be produced with a low energy, so it seems unlikely that electrons produced at the conversion dynode will be efficiently detected by the electron multiplier. It was expected that by lowering the potential of the conversion dynode it would be possible to accelerate the electrons into the electron multiplier so that they were easily detected. The unusual result that optimum ion current was obtained with the conversion dynode and electron multiplier at the same voltage indicates that there is a flaw in the design of the dynode assembly.

In this section, I will present an alternative design for a conversion dynode that was used in the detector for the helium microscope. The alternate design uses the same field to move the ions off axis and to push the secondary electrons into the multiplier. By using a single field, there is a high certainty that most of the ions will strike the dynode, but it is then more difficult to control how the secondary electrons will behave next. In testing of the new dynode design a flaw in the implementation of electron multipliers was found that explains the unusual behaviour observed by Chisnall. The design, simulation, implementation and



(a)



(b)

Figure 4.28: Schematic diagrams of the dynode assembly on the spin-echo detector. (a) Diagram illustrating the ion current measurement section of the spin-echo detector. The ion beam (A) can either be pushed off-axis (B) into the dynode and multiplier assembly (E) or pass straight through to the Faraday cup (F). Figure taken from fig. 5.7 in Chisnall's PhD thesis [77]. (b) Design of dynode and electron multiplier and conversion dynode used in the spin-echo apparatus. Ions enter the dynode assembly (A) and collide with a gold surface (B) creating secondary electrons. The secondary electrons then travel to the electron multiplier (C), where they are multiplied and the output current measured at the faraday cup (D). Figure taken from fig. 4.6 in Chisnall's PhD thesis [77].

testing of the dynode was performed by myself. The mechanical design and manufacture of the dynode was performed in collaboration with Alexandrowicz and his group at the Technion.

4.4.2 Design simulations

The key objective of the dynode is that after impact by an ion, the secondary electrons that are produced are directed to an electron multiplier with a high detection probability. The dynode will therefore need to be at a large enough voltage below ground to ensure that a large number of secondary electrons are produced upon ion impact. Once the electrons are produced, they will need to be directed onto the electron multiplier.

Another important objective is to make the output current measurement possible at ground, thus avoiding the need to float the ammeter and implement more complex communication between the ammeter and the experiment control computer. Requiring the output to be at ground makes the dynode design more difficult, since it requires that the front of the electron multiplier is held a large negative voltage to ensure electrons inside the multiplier cascade and are multiplied correctly. Electrons created at the conversion dynode will be attracted to more positive voltages, such as ground and therefore will be repelled by the negative potential on the front of the multiplier.

The final restriction on the dynode design is that the potential on the dynode is less than 5 keV below ground. By keeping the potential difference between the dynode and the chamber less than 5 keV, we can use SHV connectors for the feedthrough and cables which makes the design of the system simpler and keeps costs down.

There are many potential methods for ensuring electrons are directed into the multiplier, but the design used in this detector uses a curved dynode to direct the beam as illustrated by fig. 4.29. An alternative method is to float the entire case that the dynode and electron multiplier sit in, so that the secondary electrons from the dynode can then be accelerated into the multiplier even if the front of the multiplier has a negative voltage on it. Floating the entire case requires an extra power supply and is difficult to achieve in the allocated space if the dynode cannot have a very large voltage applied to it. A set of additional elements could be installed around the multiplier to repel the electrons and ensure that the field lines point towards the multiplier. Again though, additional elements require more power supplies and added isolation between components. Using a curved dynode can direct the beam to the multiplier and does not require additional parts or power supplies, keeping the design simple. The spherical surface has two effects on the secondary electrons, one is that the secondary electrons that are generated will be directed towards the multiplier. The normal to the surface always points towards the multiplier and since the electrons are emitted with a cosine distribution, they are generally emitted in the direction of the surface normal. The other effect the curved dynode has is that the electric field will also point along the surface normal, so electrons are also accelerated towards the multiplier once they have been emitted. By using a large voltage on the dynode, it is possible to have the initial acceleration from the dynode overwhelm the repulsion at the multiplier.

The geometry shown in fig. 4.29 also allows the use of an axial Faraday cup to measure the ion current directly. By not applying a voltage to the multiplier or electron multiplier, the ions will pass straight through the dynode assembly. A Faraday cup can then be placed

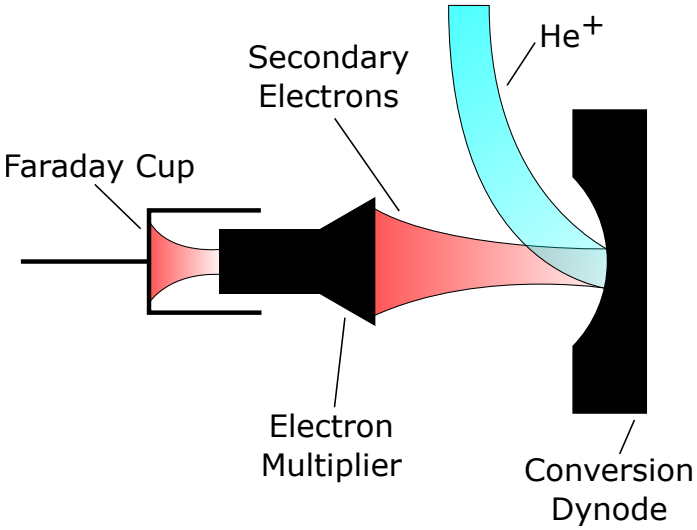


Figure 4.29: Schematic of ideal behaviour of the dynode and electron multiplier. The helium ions are sketched in blue and are accelerated to the dynode. After collision with the conversion dynode, secondary electrons are generated and are accelerated into the electron multiplier as shown in red. The electron multiplier then amplifies the electron current so that it can be measured by a Faraday cup.

on axis to collect these ions and measure the current.

The physical size of the dynode is much larger than the beam to create a uniform field for accelerating the secondary electrons. Since the dynode will be held at a potential much lower than the ground of the case it sits in, large fringing fields will exist at the boundary between the dynode and the case. By using a large dynode, the fringing fields at the dynode edges will have a minimal effect on the secondary electron trajectories and the field in the region between the dynode and the multiplier will be as uniform as possible along the electron trajectories.

An aperture is used to shield the ions from the strong field until they enter the dynode assembly to ensure that the ions strike the centre of the dynode. A large potential on the dynode ensures that as many secondary electrons are produced as possible, and that the secondary electron focusing from the curved dynode is effective. However, the ion beam will be strongly attracted to the dynode if a very large potential difference is applied without any shielding, which limits the size of the dynode that can be used and therefore the uniformity of the field. By using an aperture to shield the ions from the strong field, the ions will travel further into the dynode assembly, allowing a larger dynode to be used and therefore ensuring the field is as uniform as possible.

Particle induced electron emission is a complex topic which is important to a large range of fields including electron and ion microscopy [102]. The processes controlling the emission of electrons determines the interaction volume under the beam and the brightness of each pixel. Simulated results for a 30 keV helium beam incident on a flat aluminium target show a peak in the energy of the emitted electrons at about 3 eV and the angles of the emitted closely following a cosine distribution [102]. The dependence on incident angle is shown to closely follow $\cos^{-1}\theta$ where θ is the incoming angle for the high energy case in simulations by Ishitani et al. [103].

There are two main mechanisms that describe the behaviour of particle induced electron

emission, kinetic emission and potential emission [104]. Kinetic emission is caused by the transfer of kinetic energy from the incoming particle to any of the electrons, atomic cores or adsorbates at the solid surface. Potential emission caused by an ion arises from the potential energy released upon neutralisation of the ion providing enough energy to free electrons from the solid. For singly charged ions the Auger neutralisation process is the dominant factor in potential emission [103] and is independent of the incoming projectile velocity. So at low projectile velocities, the potential emission processes should dominate, while at very high velocities the kinetic emission processes should dominate.

Mechanisms for particle induced electron emission rely on the precise state of the surface of the solid which the particle is colliding with. Modern studies on particle induced electron emission require a clean surface with an overall surface contamination less than a few percent of a monolayer [104]. The surface state of the conversion dynode used in this detector will be unknown and does not undergo any cleaning before use, so it will be difficult to make any accurate predictions about the properties of the electrons emitted from the dynode surface. However, general trends that may be applicable to contaminated surfaces are that ions generally produce more secondary electrons at high impact energies and grazing incidence.

Aluminium was chosen as the material for the dynode due to the favourable secondary electron yield of the clean surface, low cost, low sputtering yield and ease of manufacture. The secondary electron yield for helium on clean aluminium surfaces has been measured various times, by Baragiola et al [105] and the distribution of the emitted electrons has been measured for polycrystalline aluminium by Riccardi et al [106]. In the later paper, we see that even for ions incident at 78° from the normal at very high energies, the maximum energy of the emitted electrons very rapidly falls off after 15 eV with the maximum in energy being below 5eV, supporting the theory that the electrons are produced at the surface with very low energies. Since no cleaning processes are used on the conversion dynode, this literature will only act as a guide and it is not possible to clearly state which surface will be ideal.

4.4.3 Simulation method and results

In order to investigate the performance of the dynode, charged particle trajectories and electric fields need to be simulated. LORENTZ was used to simulate the behaviour of charged particles in the dynode assembly. The electric field was solved by applying a boundary element method to a CAD model of the dynode assembly and the particle trajectories were calculated using an adaptive Runge-Kutta 5 method to solve the ODE for the motion [107].

The helium ion trajectories are first calculated to find where the ions strike the dynode surface. An incoming beam of helium ions travelling parallel to the optic axis, from the magnet, is released with a width of 10 mm and an initial energy of 700 V. The ions are released with 700 V since that was the measured ion energy required for helium ions to pass through the magnetic sector as illustrated in fig. 4.27.

At each position where the ions strike the dynode, secondary electrons are released and the trajectories of these electrons are simulated. A MATLAB script takes the positions where the ions strike the dynode, then generates a set of secondary electrons with randomly generated energies and directions. Using the literature discussed in sec. 4.4.2 as a guide, the

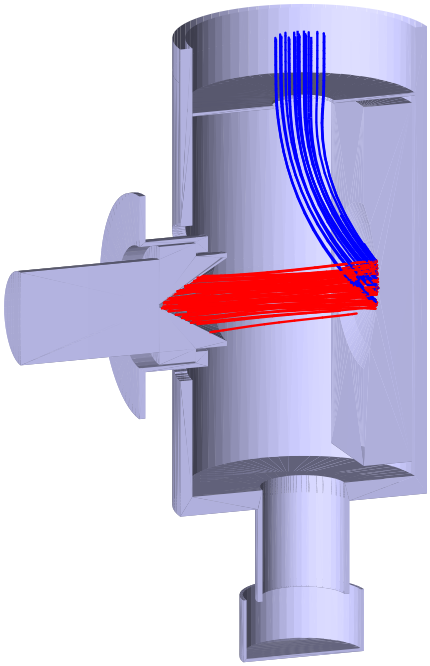


Figure 4.30: Simulated trajectories of helium ions and secondary electrons inside the dynode assembly. The blue lines indicate the trajectories of helium ions, while the red line indicate the trajectories of the secondary electrons. In this simulation, the dynode has a voltage of -2500 V, the electron multiplier has a voltage of -1300 V and the remaining surfaces are at ground.

electron energy is sampled from a Gaussian with a mean of 4 eV and a standard deviation of 4 eV while resampling any energies below zero so that all energies are positive. The direction the electron travels in is sampled from a $\cos^{-1}\theta$ distribution where θ is the angle between the local surface normal on the curved dynode surface and the direction the electron travels in. Following the description by Greenwood [108], the polar angle, θ, is obtained from a random number, x, generated on the interval $[0, 1]$ using

$$\theta = \sin^{-1}\left(\sqrt{x}\right). \tag{4.12}$$

At each position that an ion strikes the dynode, four electrons are generated to ensure that the potential secondary electron trajectories are well sampled.

Figure 4.30 shows examples of helium ion and electron trajectories in the dynode assembly. The helium ions are plotted as blue lines and are bent around by the large negative voltage on the dynode. The red lines are the generated secondary electrons, and these are then pushed onto the electron multiplier by the same field that pulled the helium ions around. In fig. 4.30, 99.3% of the electrons make it inside the multiplier cone before colliding with a surface, indicating that the dynode is functioning correctly in these simulations.

4.4.4 Mechanical design

Once theoretical performance of the dynode is shown to be possible, a method is needed to mechanically implement the design. The dynode is fixed inside a case that can be fitted onto the end of the ion optics to provide the mechanical support needed. The multiplier

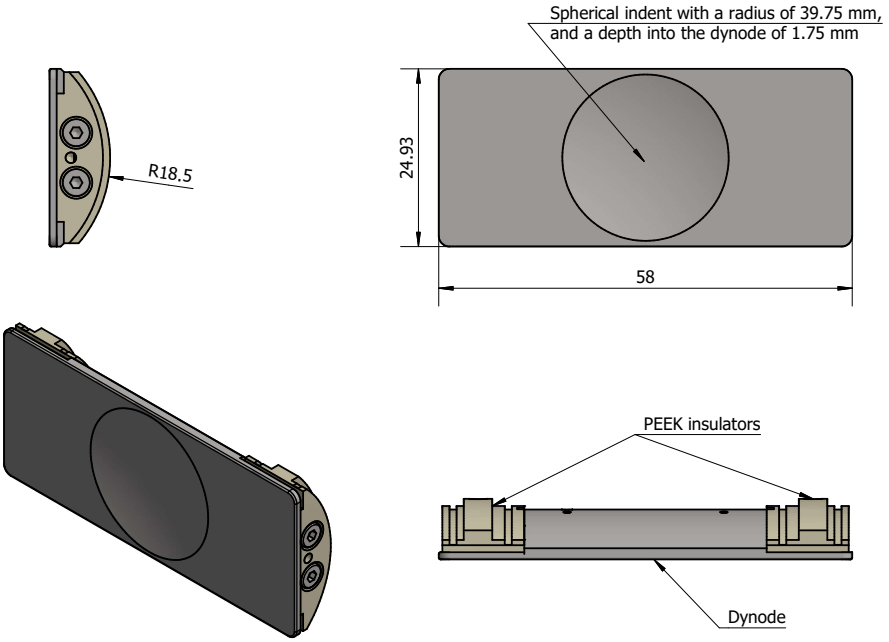


Figure 4.31: Rendered images of the dynode and the insulating PEEK components for mounting. The dynode has a spherical dynode machined into it to improve the dynode performance. The PEEK insulators, as seen in a beige colour, can be seen to have a series of raised section and grooves to increase the linear distance between the dynode surface and the grounded case. By maximising the linear distance between the dynode and the grounded case, breakdown across the surface is prevented. The radius of curvature of the largest part on the PEEK insulators match the curvature of the inside of the case that the dynode mounts to.

is mounted off a separate feedthrough to ensure that it is easy to change the multiplier if needed.

Figure 4.31 shows how the dynode is made as a sector of a cylinder, with PEEK components used to mount and insulate the dynode. In order to achieve the voltages that are needed the components need to be isolated from each other. The most likely mechanism that will allow breakdown to occur is through a surface breakdown, where charge will flow across the surface of an insulator [109]. Figure 4.31 shows a rendered cross sectional view of the dynode and its insulators, where the sides of the aluminium dynode are replaced with PEEK to provide an insulating surface to mount the dynode to the grounded case. To increase the path length over the surface between the dynode and the assembly case, grooves and raised sections were added to the PEEK which ensures that no breakdown can occur. A series of vented screws then allow the PEEK mounts to be attached to the dynode from the side. The final mechanical design was designed in collaboration with Alexandrowicz and his group at the Technion.

The dynode then mounts inside a cylindrical case, which is then mounted to the end of the ion optics on the outgoing end of the magnet. Figure 4.32 shows how the dynode is fixed to the cylindrical case using the PEEK components described earlier. It was found after construction that the cylindrical case fitted into an ion optics ring with a tight fit, so

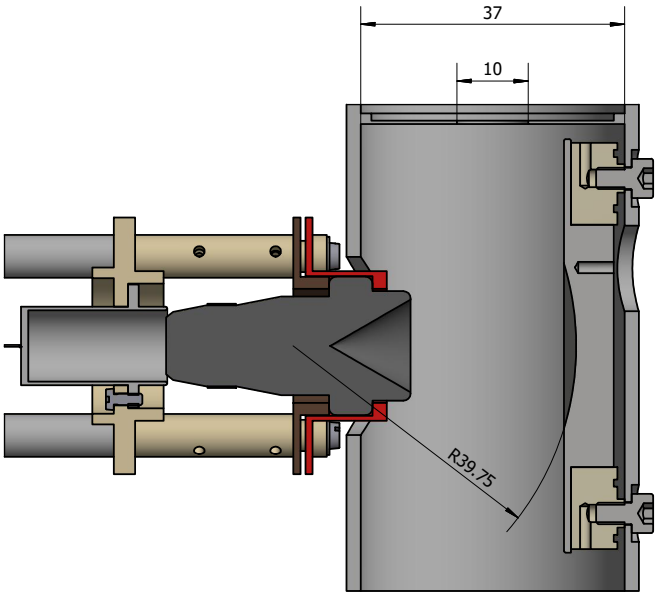


Figure 4.32: Cross-sectional view of the full dynode assembly illustrating how the dynode is attached to the cylindrical case. The electron multiplier is assembled on a separate feedthrough and protrudes into the dynode case as seen in the figure. The red and brown top hat metal pieces on the left of the diagram are pinched together to hold the electron multiplier in place and make an electrical connection. The curvature of the dynode is illustrated to show how the electrons are pushed towards a point inside the multiplier.

did not need spot welding. The ion optics ring on the cylindrical case was attached to the ion optics using the same method that the rest of the ion optics are mounted together. The case also contains the 10 mm aperture to shield the incoming ions from the dynode field.

The electron multiplier is mounted from a separate feedthrough to ensure replacement is easy. The left hand side of fig. 4.32 shows how the electron multiplier fits into the ground case that holds the dynode. The assembled mount for the multiplier is shown in fig. 4.33a, the mount is also used to hold the Faraday cup needed at the exit of the electron multiplier. The multiplier is held in place by two metal pieces shaped as top hats, as illustrated in red and brown in fig. 4.32, that pinch together over the lip at the front of the multiplier. These metal top hats are then suspended from PEEK rods that insulate the electron multiplier from the grounded feedthrough. Figure 4.33b shows the dynode assembly inside the chamber. Care must be taken to ensure that the front of the multiplier does not touch the grounded case of the dynode assembly. To ensure that the ion optics cannot move too much, a mount was added to support the dynode assembly from the Faraday cup port on the detector.

4.4.5 Grid on the front of the multiplier

After installation of the dynode and electron multiplier, fig. 4.34 shows how the voltage applied to the front of the electron multiplier affected the optimum dynode voltage. The optimum dynode voltage should occur when the helium ions strike the dynode in the centre, so changing other voltages near the dynode should have no effect. Figure 4.34 shows the output current from the multiplier as a function of the dynode voltage for a set of different voltages on the front of the multiplier. The potential difference across the multiplier was

(a) (b)

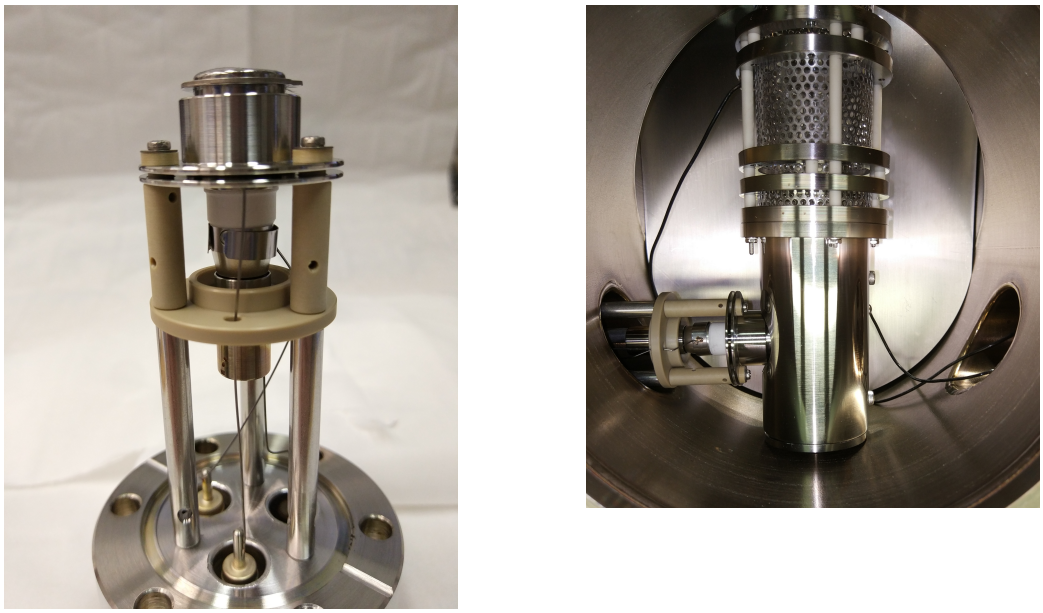


Figure 4.33: Images of the electron multiplier mount and the dynode assembly inside the detector. (a) Image of the electron multiplier assembly. The beige components are PEEK and insulate the front of the multiplier and the Faraday cup from the mounting rods from the feedthrough. Two metal top hats at the top of the assembly pinch the multiplier to keep it in place. (b) Image from the top of the dynode assembly after installation. Note that the chamber splitter is installed below the dynode assembly, which is explained in section 4.5.2.

fixed, so that the gain of the multiplier was also fixed. To achieve a fixed gain, when the front of the multiplier was raised by a voltage V, the back of the multiplier and the Faraday cup were also raised by a voltage V. To raise the potential on the Faraday cup required the ammeter to be floated, which is possible up to 1.5 kV for the picoammeter that was used (9103 USB Picoammeter from RBD instruments[2]). Note that the optimum dynode voltage is heavily affected by the potential on the front of the electron multiplier, which is not the expected behaviour of the system.

The dependence of the optimum dynode voltage on the multiplier voltage is caused by the field from the dynode influencing the performance of the electron multiplier. As previously discussed, electron multipliers operate by triggering a cascade of secondary electrons due to successive collisions of electrons with the multiplier surface generating several electrons at each collision. It is therefore key that when the first electron strikes the surface of the multiplier, several electrons are generated that each go on to generate more secondary electrons. The large electric field from the dynode can interfere with the generation of the first set of secondary electrons at the entrance to the multiplier by either extracting the secondary electron from the multiplier entrance or immediately forcing the electron to collide with the surface such that it does not gain enough energy to produce multiplication.

The multiplier will therefore operate at its highest efficiency when the externally applied electric field is zero. The condition that the field immediately outside the multiplier is zero will depend on both the voltage of the dynode and multiplier. The voltage needed on the

[2]I have uploaded the control script for the ammeter to GitHub which has also been adopted by RBD on their download page

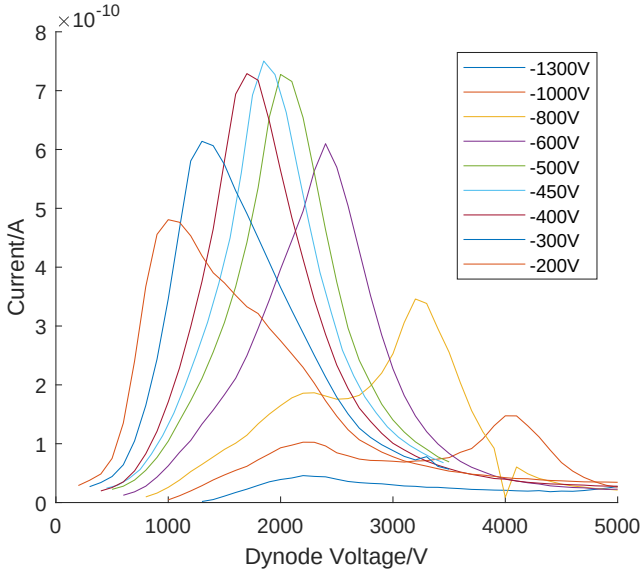


Figure 4.34: Plot showing how the current measured from the electron multiplier changes as the front voltage on the multiplier is changed while holding the gain constant. The optimum dynode voltage is seen to shift as the multiplier voltage is changed due to the condition for no external applied field on the multiplier also varying.

dynode to create no external field outside the multiplier will vary as the electron multiplier voltage varies, which explains the results seen in fig. 4.34.

The potential between the centre of the dynode and the multiplier was simulated using LORENTZ to confirm the voltages needed for there to be no applied electric field outside the multiplier. The blue lines in fig. 4.35 show an example of the simulated potentials between the centre of the dynode and multiplier when the multiplier is at -1000 V and the dynode is at -3000 V, -4000 V and -5000 V. The black line in the figure indicates the start of the electron multiplier cone, so when the potential is not uniform on the left of the black line there will be an applied electric field. Looking at the blue curves in fig. 4.35, we can see that changing the dynode voltage has a strong effect on the field inside the multiplier cone. For a given electron multiplier voltage there will be a specific dynode voltage that gives a uniform potential inside the multiplier cone, for the example in fig. 4.35 when the multiplier is at -1000 V the potential will be uniform when the dynode is at $\sim$-4000 V.

The dynode voltage that experimentally maximises the observed output current in fig. 4.34 is obtained at the same voltage that minimises the electric field inside the multiplier cone as calculated by LORENTZ in fig. 4.35. In fig. 4.34, when the front of the electron multiplier is at -1000 V the maximum current is obtained at $\sim$-4000 V. The calculated potential in fig. 4.35 is uniform inside the multiplier at the same dynode voltage. Therefore, there is strong evidence that the shifting maximum that has been observed in fig. 4.34 is caused by the strong electric field from the dynode interfering with the operation of the electron multiplier.

Adding a grid across the front of the multiplier can shield the strong field being applied by the dynode. It is known that the addition of a grid to the front of the electron multiplier increases the performance [100, 110]. A coarse grid was added to the LORENTZ simulations to test if the field can be successfully shielded, and an example of the results is plotted in red in fig. 4.35. We see that the fields inside the electron multiplier are much weaker, in-

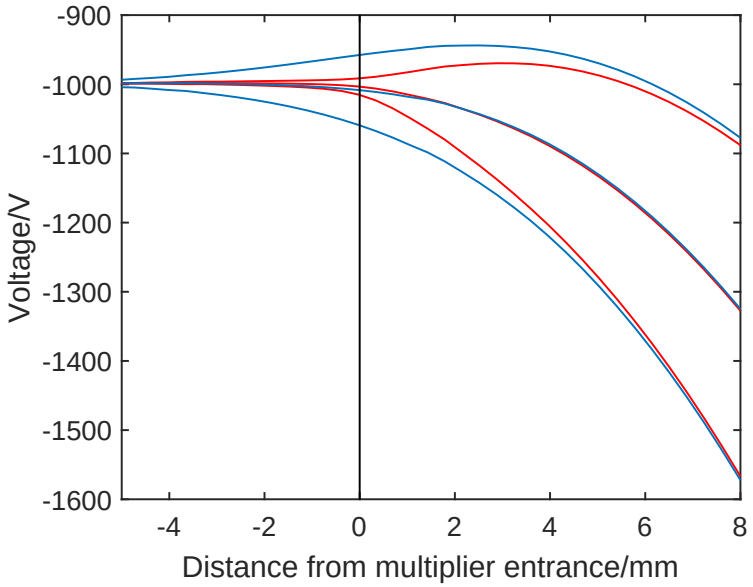


Figure 4.35: A plot showing the voltage along the axis between the centre of the dynode and the electron multiplier for three different dynode voltages. The front of the electron multiplier is at -1000 V and the dynode voltage is either at -3000 V,-4000 V or -5000 V. The solid vertical black line corresponds to the entrance to the electron multiplier cone, with the left side of the figure leading further inside the electron multiplier and the right leading towards the dynode. The blue curves are the voltages without any shielding across the electron multiplier entrance, while the red curves are the voltages with a coarse grid (10 lines/inch and diameter of 0.005") across the multiplier entrance. The electric field inside the electron multiplier will only be negligible when the voltage has a small gradient on the left hand side of the black line. The blue curves illustrate that without a grid, the electric field is only negligible in this example when the dynode voltage is at -4000 V, which matches the experimental result seen in fig. 4.34. The red curves illustrate that with a grid the electric field is negligible at all dynode voltages and the multiplier should correctly irrespective of the dynode voltage.

dicating that the presence of the dynode should no longer interfere with the multiplication of electrons in the multiplier.

4.4.6 Installation of grid

Figure 4.36 shows how a mesh was installed with a wire diameter of 0.001", 40 wires/inch and an open fraction of 92.2% using a circlip with a 16mm external diameter (Accugroup P/N HEC-16-A2) to fit over the electron multiplier. The mesh was cut to the correct shape by placing the mesh over a broken multiplier and cutting the correct shape with a scalpel. Since the mesh has such a large open fraction, it is difficult to see the grid in fig. 4.36 without looking closely at the opening to the multiplier.

Figure 4.37 shows the results of repeating the procedure of varying the dynode voltage at various voltages on the electron multiplier as previously illustrated in fig. 4.34, but with a grid installed and it was confirmed that the dynode voltage that produces the maximum output current is now independent of the voltage at the front of the multiplier. The peak in dynode voltage now remains in approximately the same place, demonstrating that the field from the dynode is being successfully shielded allowing the electron multiplier to function correctly. There is a small shift in the peak positions due to the field from the multiplier af-

Figure 4.36: Image of the top of the electron multiplier with a grid over the entrance to shield the multiplier from external fields. The grid is held in place by a circlip over the entrance of the multiplier.

fecting the bending of the ion beam, however the height of the peak remains approximately unchanged.

Figure 4.38 shows the gain between using the dynode with the electron multiplier and directly measuring the ion current with the on axis Faraday cup. In order to measure the gain between the two modes, we need a signal large enough to be measured directly by the Faraday cup but not so large as to damage the electron multiplier. The gain was estimated by leaking in enough helium to cause the ion current into the Faraday cup to rise by $I_f = 1.73\,\text{pA}$, then the current from the electron multiplier was measured and divided by I_f to obtain the gain. In order to extend the range of the gain measurement, the partial pressure of helium was then reduced leading to the ion current being too small to measure directly with the Faraday cup. The gain with a reduced partial pressure was then obtained by scaling the measured current from the electron multiplier such that the difference in the gain measurements is minimised where the voltages overlap. The gain shows the expected behaviour of strongly increasing as the voltage across the electron multiplier is increased.

Figure 4.39 shows a previous gain measurement performed without the grid that is significantly lower than the gain measured after the grid was added. Figure 4.39 plots two mass scans obtained at $m/z = 3$ using either the dynode or directly with the Faraday cup. When the front of the multiplier is unshielded from the dynode field and is at $V_c = 1300\,\text{V}$, the gain of the dynode and electron multiplier is $G(V_c) = 2490 \pm 180$ which is ~ 15 times smaller than the gain measured with the grid.

The gain measurement plotted in fig. 4.38 is actually a combination of multiple factors, and not directly a measure of the gain of the electron multiplier. In comparing the current measured with the Faraday cup and the electron multiplier, we are measuring the gain, G, which is given by,

$$G = A\gamma\alpha g(V_c). \tag{4.13}$$

Where, A is the proportion of ions that missed the Faraday cup, γ is the secondary electron yield, α is the transmission probability from the dynode surface to the inside of the multiplier and $g(V_c)$ is the gain of the electron multiplier. Since the beam is wider than the

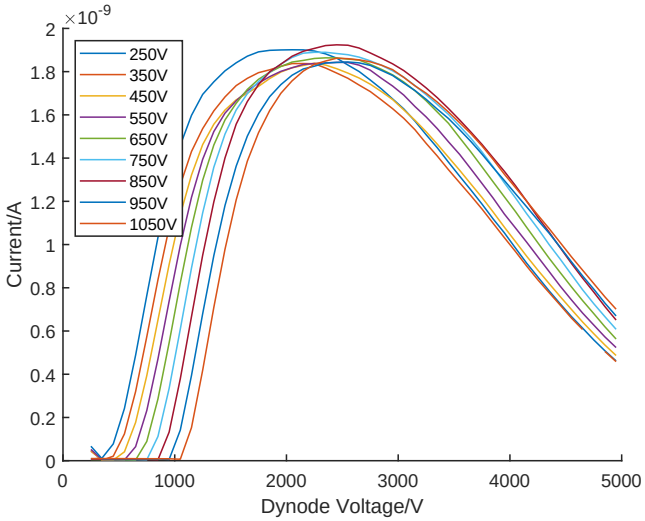


Figure 4.37: Plot showing how the current measured from the electron multiplier varies as the potential on the front of the multiplier is changed while holding the gain of the multiplier constant. Once the grid is installed the front of the multiplier is shielded so the peaks in dynode voltage appear in roughly the same position. The small drift of the peak position is due to the field from the multiplier altering the ion trajectory, so a larger dynode voltage is needed to pull the ions round to the centre of the dynode. One outlier has been removed from the data.

Faraday cup, a significant proportion of the ions, $A > 1$, are lost when using the on-axis Faraday cup (the predicted value of $A\alpha > (6.40 \pm 0.09)$ is discussed later in section 4.5.4). Therefore, part of the gain of using the electron multiplier is that the much larger surface area of the dynode will collect all of the beam. As discussed earlier, the secondary electron yield, γ, is difficult to predict but we expect that $\gamma \sim \mathcal{O}(1)$.

The results demonstrated in the current work have a wider significance than just helium microscopy. For example, my observations concerning the value of a grid at the entrance to the multiplier help explain the unusual behaviour of the dynode system on the spin-echo apparatus.

It was noted by Chisnall [77] that the highest output current from the electron multiplier in the Cambridge spin-echo instrument was obtained when the dynode and the electron multiplier are held at the same potential. The design for the dynode used in the Cambridge spin-echo instrument is shown in fig. 4.28b, which includes the dynode in close proximity the electron multiplier. The design was made with the intention to accelerate the secondary electrons produced at the dynode surface into the electron multiplier, however it was experimentally discovered that the optimum current was obtained when there was no potential difference between the multiplier and the dynode.

The optimum current is obtained when there is no field inside the dynode assembly because an external field applied to the electron multiplier will disrupt the performance of the multiplier. As found with the dynode installed on the microscope, if an external field is applied to the multiplier, the multiplier will not function correctly and won't multiply incoming electrons. The problem is exacerbated when the dynode is so close to the multiplier, since the field will be much stronger.

Similarly to the microscope, the addition of a grid over the front of the electron multi-

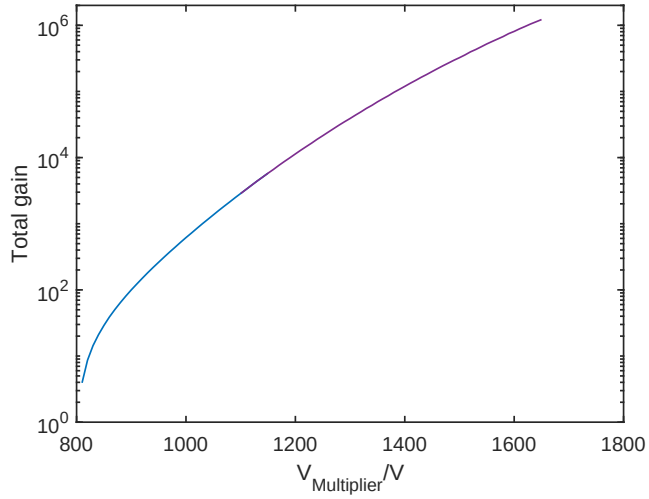


Figure 4.38: A plot of the combined gain of the dynode and multiplier as a function of the voltage applied to the front of the multiplier. The gain was measured in two parts to extend the range of the measurement, with the overlap between the two measurements minimised to calibrate the higher gain data as plotted in purple against the lower gain data plotted in blue. It should be noted that the potential difference between the front of the multiplier and ground is not the potential difference across the multiplier since the back of the multiplier has its voltage set by a potential divider.

plier is likely to significantly increase the gain of the dynode assembly. By adding a grid to the multiplier the field from the dynode can be shielded, and then the electrons can be manipulated as originally designed to be accelerated into the electron multiplier.

4.5 Overall performance

4.5.1 Time constant

A key property of the detector is the response time of the instrument which shall be shown to directly affect the efficiency of the detector. First the experimental measurements of the time constant will be discussed and then a theoretical model for the results will be presented.

The time response of the detector as installed on the microscope was measured by moving the beam over a sharp edge. Since the attocube positioners can move up to 8 mm/s, and the beam width is approximately $\sigma = 2\,\mu\mathrm{m}$, the beam can move over a sharp edge in about $\tau_{\mathrm{edge}} \sim 1$ ms. Therefore if the time response of the detector is much larger than 1 ms, we can assume that the gas flux is instantaneously changed.

Figure 4.40 shows how the time response can be estimated by fitting an exponential decay to the current measured after the input is abruptly changed. A function of the form,

$$I(t) = \begin{cases} A + c & t \leq t_0, \\ Ae^{-(t-t_0)/\tau} + c & t > t_0, \end{cases} \tag{4.14}$$

was used, where A, c are arbitrary constants, t_0 is the time when the stage is moved and τ is the response time of the detector.

To obtain a good estimate of the time response, the measurement was repeated 10 times

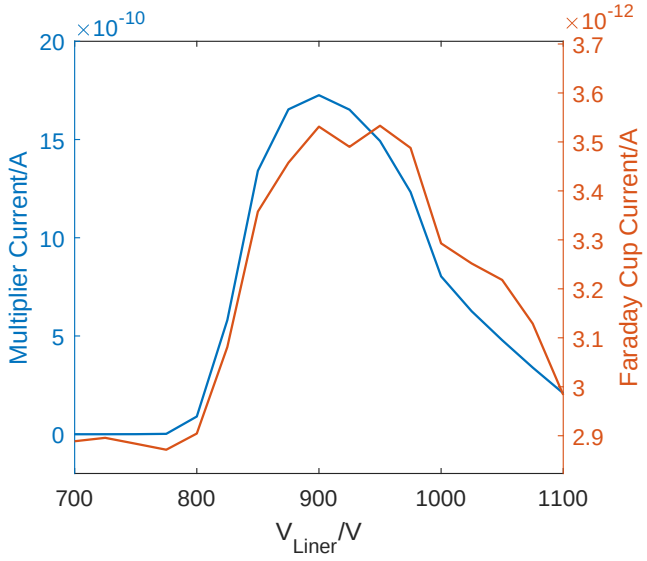


Figure 4.39: Plot showing mass scans across the $m/z = 3$ peak using either the Faraday cup in orange or electron multiplier with V=1300 V in blue. The Faraday cup current was scaled so that the sum of square differences between the two mass scans was minimised, giving an estimate of the gain of the electron multiplier and dynode. When the multiplier is unshielded and has a voltage of $V_c = 1300$ V, the gain of the dynode and electron multiplier is $G(V_c) = 2490 \pm 180$.

and a weighted average was taken. The weighted average and associated error on the mean were calculated using the standard formulae. The result of combining the 10 measurements gives a final estimate of the time constant of the detector as $\bar{\tau} = (278 \pm 13)$ ms.

The gas flow through the ioniser can be modelled to check if the measured time response is reasonable and is arising from the time needed for gas to diffuse through the ioniser. As previously noted in equation 4.6, the time response for a volume being pumped is given by $\tau = V/S$, where V is the volume of the chamber and S is the effective pumping speed of the volume. However, the detector is not a simple volume being pumped, due to the narrow 5 mm entrance to the detector that throttles the pumping on the volume immediately in front of the detector.

In order to calculate the expected time constant of the detector, the pressure distribution inside the solenoid needs to be obtained. In order to calculate the pressure inside the solenoid, the gas flow from the sample to the front of the detector must be modelled first.

Figure 4.41 illustrates the volume at the front of the detector schematically. If the atoms have stagnated, then we can write down the differential equations for the gas flow through to the pump. When the atoms enter the first chamber before the detector, they have just been scattered from the sample and instead of travelling in a random direction the atoms are moving in a preferred direction. However, after scattering from the chamber walls, the atoms behave as an ideal gas and it is possible to use standard formulae (which is discussed in more detail in section 5.2.2 and in a paper by Lambrick, Bergin et al [111]).

For the geometry illustrated in fig. 4.41, conservation of gas flow gives that the pressures p_1, p_2 in the pipe in front of the detector and the front chamber of the detector respectively

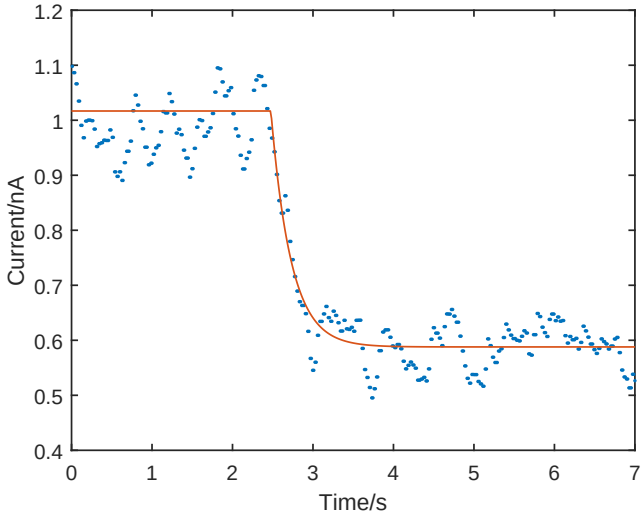


Figure 4.40: Plot of the variation of the detected current with time after the beam moves across a sharp edge. The blue dots are the experimental data points and the orange line is an exponential that is fitted to the jump in current. The small oscillations that are seen are due to noise on the signal. The fit shown here is for a time constant of $\tau = 250 \pm 40$ ms.

evolve in time as determined by

$$V_1 \frac{dp_1}{dt} = Q - C_p\,(p_1 - p_2) - S_s p_1, \tag{4.15}$$

$$V_2 \frac{dp_2}{dt} = C_p\,(p_1 - p_2) - S_i p_2. \tag{4.16}$$

Where V_1, V_2 are the volume of the pipe in front of the detector and the volume of the front of the detector respectively, Q is the gas flow in from the sample, C_p is the conductance of the small pipe into the detector and S_i is the effective pumping speed at the front of the ioniser. Gas can also flow back into the sample chamber with an effective pumping speed of S_s.

The detector parameters were estimated in table 4.2, below, to model how each of the pressures would vary with time. The next few paragraphs justify how the parameters have been estimated so that the theoretical performance matches the experimental data as closely as possible.

The volume of the pipe connecting the source chamber to the detector, V_1, was estimated from it being a CF16 pipe that is approximately 0.235 m long. The volume of the front of the detector, V_2, was approximated by looking at the volume of the front chamber and the liner entrance up to the start of the ionisation region. The effective pumping speed of the detector aperture back into the sample chamber was calculated using that the diameter of the aperture is $d = 1$ mm and the standard formula for the conductance of an aperture $C_{\text{ap}} = A\bar{c}/4$, where A is the area of the aperture and $\bar{c}$ is the average speed of the helium atoms. The conductance of the small pipe leading into the aperture was calculated using $C_p = (\bar{c}\pi d^3)/(12l)$, where $d = 5$ mm is the diameter of the pipe and $l = 38$ mm is the length of the pipe.

The pumping speed down the ioniser, S_i, was evaluated using a Monte-Carlo approach to incorporate the complex geometry of the ioniser. A naive estimate of the ioniser pumping

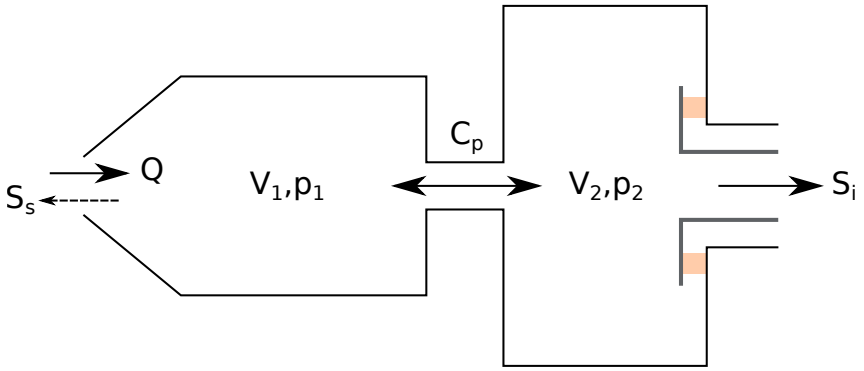


Figure 4.41: Schematic showing gas flow from after scattering from the sample into the detector. A gas flux Q exits the sample chamber and enters a pipe to the detector with a pressure p_1 and volume V_1. A small pipe of conductance C_p connects the pipe to the front of the detector that has a volume of V_2 and a pressure of p_2. The front of the detector has an effective pumping speed of S_i. The gas can also go back into the sample chamber due to an effective pumping speed S_s.

Variable	**Estimated value**
V_1	0.047 L
V_2	0.18 L
S_s	0.24 Ls^{-1}
C_p	1.05 Ls^{-1}
S_i	0.67 Ls^{-1}

Table 4.2: Table to show the estimated variables used to find the detector time constant.

speed using just the narrow part of the liner gives a pumping speed of $\sim 0.9\,\mathrm{Ls}^{-1}$. However, since the filament is very close to the liner entrance, gas needs to move past the filament mount before it can pass down the ioniser. To incorporate the effect of the complex geometry, gas flow was simulated using a Monte-Carlo package called Molflow+ [112]. The full geometry was incorporated by importing the CAD model into the simulation package. The entrance and exit to the ioniser are given sticking probabilities of 1, and then the transmission probability α is calculated. The conductance is then given by $C_m = C_a\alpha$, where C_a is the entrance aperture conductance [113]. The result was that the conductance of the ioniser is S_i =0.67 Ls^{-1}.

The coupled differential equations 4.15 and 4.16 can be solved analytically giving the steady state pressures, but the time evolution does not give a particularly neat solution. The steady state quantities have neat analytic forms and are given in appendix C. The pressure variation over time for the front of the ioniser is plotted in fig. C.1, the time evolution closely follows what is expected from simple gas flow calculations as detailed in appendix C.

Now that an expression for the pressure at the front of the ioniser has been obtained, it is possible to calculate the total pressure and time response of the entire detector. The response time of the output current from the ioniser can then be modelled by solving a 1D diffusion equation down the length of the ioniser. The gas needs to diffuse down the ioniser to reach the pump, with the time variation described by,

$$\frac{\partial p(t,x)}{\partial t} = \frac{S_i L}{A}\frac{\partial^2 p(t,x)}{\partial x^2}, \tag{4.17}$$

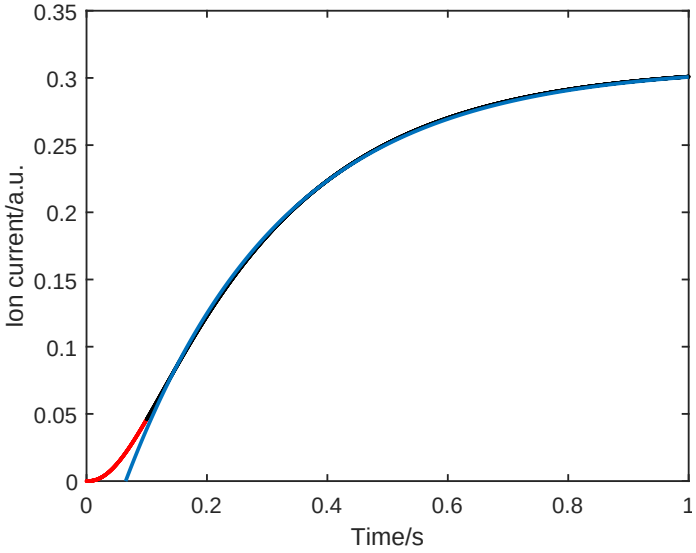


Figure 4.42: A plot showing the simulated time variation of the ion current from the ioniser after the flux is increased. The ion current initially has a slow increase, but then later closely follows an exponential. The blue curve is a single exponential that has been fitted to the data plotted in black with a time constant of $\tau_s = (261.4 \pm 0.2)$ ms. The first 100 ms of simulated data is plotted in red and is not included in the fit to the exponential curve.

where $p(t, x)$ is the pressure at position x down the ioniser at time t, $S_i = 0.67\,\mathrm{Ls}^{-1}$ is the pumping speed at the front of the ioniser, $L = 0.661$ m is the length of the narrow part of the liner and $A = 1.17 \times 10^{-4}\,\mathrm{m}^2$ is the cross sectional area of the liner. The boundary conditions are set so that the pressure at the front of the ioniser is given by $p(t, 0) = p_2(t)$, the pressure at the back is $p(t, L) = 0$ and initially $p(0, x) = 0$. The ion current from the ioniser is then given by,

$$I(t) = a \int_0^L p(x, t) dx, \tag{4.18}$$

where a is the sensitivity of the ioniser.

Figure 4.42 shows how the simulated ion current varies with time after the flux is changed. There is an initial slow rise in the current, but then the current closely follows an exponential. A single exponential is fitted to the data after 100 ms and is plotted in blue in fig. 4.42. The time constant of the exponential is given by $\tau_s = (261.4 \pm 0.2)$ ms, which is slightly less than the experimental value of $\bar{\tau} = (278 \pm 13)$ ms. It is not surprising that the value is less given that the first 100 ms of the simulated performance has been discarded, therefore there is an extremely good agreement between the experiment and simulated response times. The close agreement of the experimental and theoretical performance indicates that there is a good understanding of the origins of the time response of the detector.

The addition of a small tube into the detector, that adds a conductance C_p, acts to decrease the flux of gas passing into the detector and will also slow down the response time of the detector. However, provided the conductance is not too small there will not be a large effect since the effective pumping speed of the front of the ioniser, S_i, will always be poor due to the conductance of the solenoid being low.

4.5.2 Sources of background

There are multiple potential sources of background in a helium detector that add together to provide the total background signal that is measured. The background is defined as an ion current that is measured which is not due to helium atoms scattering from the sample and entering the front of the ioniser. Since the microscope will typically be measuring extremely small partial pressures of helium, the sources of background are an extremely important consideration which will limit the performance.

Poor compression ratio

One potential source of background signal is from helium gas that has leaked in from the backing lines. Turbomolecular pumps can only sustain a certain compression ratio between the pressures at the inlet and outlet of the pump. If there is a large partial pressure of helium in the backing lines behind the pump, gas could leak into the detector and be measured. The Pfeiffer HiPace 300 M pump used on the optics chamber has a compression ratio for helium of $> 10^8$, which may not be adequate if there were a large partial pressure of helium in the backing line. To prevent helium leaking in through the pump, two additional pumps were added (Pfieffer HiPace 80) each with a compression ratio of 1.3×10^7. The background in the detector is reduced when one of the small turbos are turned on, but no change is seen after the second turbo is activated which is consistent with a compression ratio of $\sim 10^8$ not being sufficient. However, adding even one of the compression turbos substantially increases the compression ratio so an insignificant flux of gas leaks in through the backing line.

Stray ions avoiding mass filtering

Ions that escape from the optics can pass straight from the ioniser to the dynode assembly without being mass filtered in the magnetic sector. The incoming and outgoing optics for the magnet are in the same chamber due to the 180° geometry, allowing ions to pass between the two ion optics stacks. Since the ion optics are made from perforated sheets, a fraction of ions that are not correctly controlled by the optics will escape and can move freely inside the optics chamber. If an ion can reach the entrance to the dynode assembly, it will be accelerated inside the dynode assembly and be detected without being mass filtered. Since the partial pressure of helium is normally much lower than the total pressure, if even a small fraction of the ions that are created manage to enter the dynode assembly, the desired helium signal will be swamped by a mass independent background.

Stray particles moving between the optics stacks were confirmed to be a source of background by removing the magnet and still measuring an ion current. Under ideal conditions, if the magnet is removed no particles from the ioniser should be able to reach the dynode assembly without colliding with the chamber wall multiple times. However, after removing the magnet a significant ion current was measured indicating that there are stray ions moving between the ion optics stacks.

A plate was designed and added to block particles from travelling through the optics chamber such that any measured particles must have travelled through the magnetic sector. The plate is supported from the same structure that the ion optics are mounted from as shown in fig. 4.43. Particles are now blocked from travelling through the chamber, and no current is seen on the dynode if the magnet is removed. Additional evidence is provided by

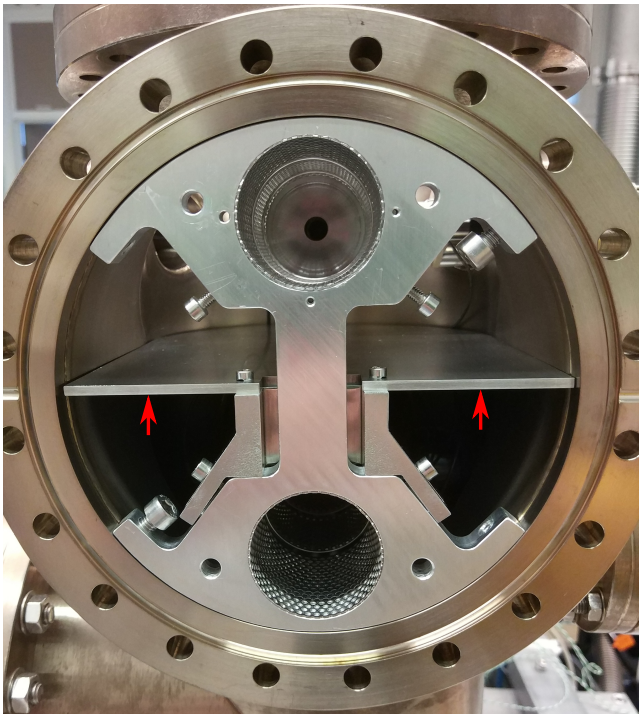


Figure 4.43: Image of the inside of the optics chamber with red arrows pointing to the plate added to block particles travelling from the bottom optics stack to the top. An image of the inside of the optics chamber without the plate is shown in fig. 4.22b. The plate mounts off the same mount used to support the ion optics.

directing the ion beam off axis using the quadrupole so that ions do not pass enter the magnetic sector. When the ions are moved off-axis, no current is measured after the installation of the plate.

Secondary electrons from ion optics elements

An additional source of background is due to ions striking the lens element after the magnetic sector. Ions that are not travelling on axis due to either not having the correct m/z value or not being focussed correctly will strike the ion optics lenses and create secondary electrons. If the ions strike an element that is at ground, the secondary electrons that are generated will not have sufficient energy to reach either the dynode or the electron multiplier. However, if the ions strike an ion optics element that is at a large negative voltage, the electrons that are generated will be accelerated away from the ion optics element and then have sufficient energy to reach electron multiplier.

An unfortunate side effect of the dynode design is that secondary electrons entering the dynode assembly will be attracted to the electron multiplier and may be detected. The dynode voltage that causes electrons to strike the multiplier will depend on the incoming energy of the electrons.

Figure 4.44a is a plot of the output current from the multiplier against dynode voltage illustrating the background signal arising from secondary electrons. The figure shows how the output current depends on the dynode voltage when a negative voltage is applied to the lens element next to the dynode assembly. Figure 4.44a contains a sharp peak around 1000 V

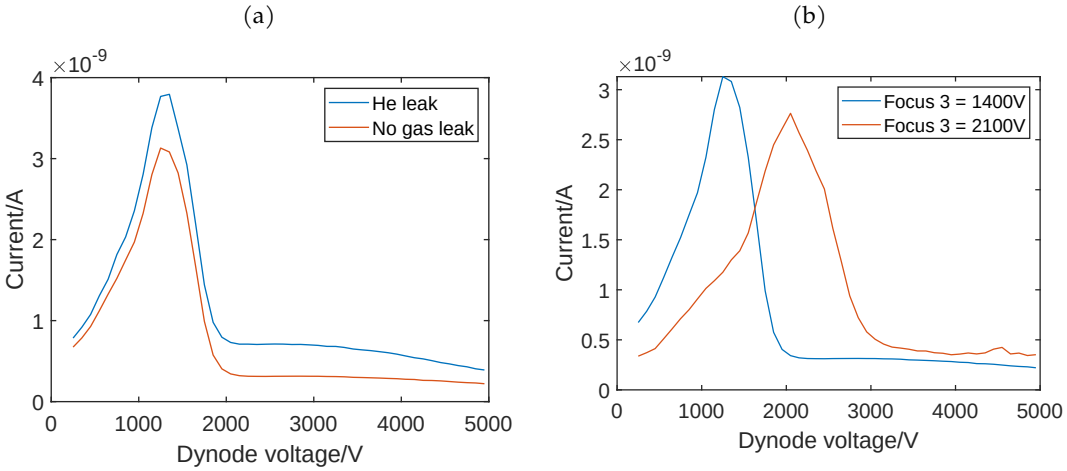


Figure 4.44: Dynode scans illustrating the background arising from secondary electrons entering the dynode assembly.(a) Plot of how the current from the multiplier varies with dynode voltage with and without a small helium leak. The large peak appearing around 1000 V can be seen to not scale as expected when additional helium is added, so is due to a background signal. (b) Plot of how the background current varies with dynode voltage for two different voltages on the third focus lens next to the dynode. The background peak can be seen to shift as the lens voltage is changed due to the energy of the secondary electrons that generate the background peak also shifting.

which is not expected from the curve shape seen earlier in fig. 4.37 where the output current slowly varies and peaks around 3000 V. By leaking in a small amount of helium, the output current is increased slightly across the entire range of dynode voltage values as expected from fig. 4.37. Therefore the sharp peak seen at 1000 V must arise from a background source and not from $m/z = 4$ ions.

Figure 4.44b shows how the background peak due to secondary electrons shifts as the voltage on the lens element next to the dynode assembly is changed. As the voltage on the lens element is changed, the energy of the secondary electrons that enter the dynode assembly is also changed. Therefore it was possible to confirm that the sharp peak seen in fig. 4.44 comes from particles colliding with the lens element and generating secondary electrons that are being counted by the multiplier.

Fortunately, the background from secondary electrons can be eliminating by using a lens element with a positive voltage instead. If ions strike a lens element that has a voltage that is positive, the electrons that are generated will not have enough energy to interact with the dynode or electron multiplier. Three element electrostatic lenses can be used with either a positive or negative central element voltage to manipulate the ions. Using a positive voltage will slow down the ions in the lens, which means that the fractional spread in energy will increase and chromatic aberrations of the lens will become more significant. The positive lens also acts as a high pass filter for any species travelling down the optic axis, which could help eliminate any unusually low energy ions.

Multiple ionisation

Now that the stray ions and secondary effects have been eliminated, any remaining background signal must be from a species that has $m/z = 4$, which is confirmed by the obser-

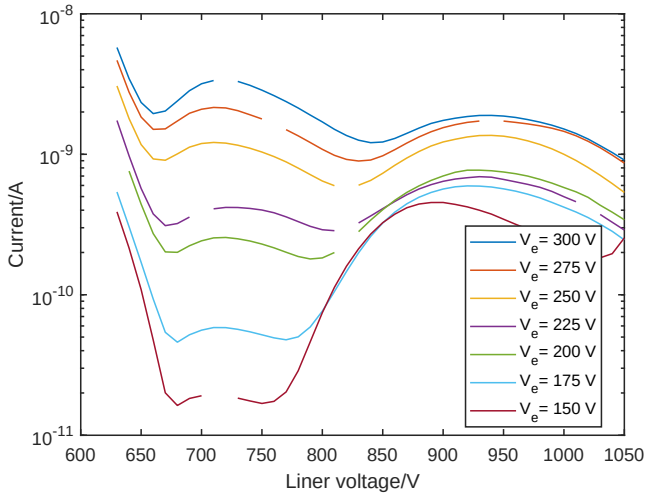


Figure 4.45: A series of mass scans of the background while changing the electron energy. As the electron energy is reduced the relative proportion of $m/z = 4$ ions are reduced. Outliers have been removed to improve the readability of the graph.

vation that the background signal now varies with dynode voltage and quadrupole push variables in the same way as the helium signal does. Therefore, these ions must be produced in the ioniser with the same mass to charge ratio as helium.

A background from multiply ionised species is present when a large electron energy is used; the mass scans of the background in fig. 4.45 illustrate how the mass scan changes with electron energy. Once ions are produced inside the ioniser, they should be rapidly extracted such that the probability of ions further interacting with electrons is low. However, it is still possible to generate multiply ionised species in a single interaction with a high energy electron.

The difficulty in investigating how the background varies with electron energy is that the total ionisation efficiency also varies when the electron energy is changed. The ratio of the height of the $m/z = 3$ peak around 900 V and the $m/z = 4$ peak around 700 V is plotted in fig. 4.46. By looking at the ratio of the peaks it can be seen that the relative proportion of species with $m/z = 4$ is reduced at lower energies with little difference seen below about 100 V.

Two potential ions that have a mass to charge ratio of $m/z = 4$ are O^{4+} and C^{3+}. The cross sections of the ions have not been experimentally measured [114], but have been theoretically calculated [88]. The threshold energy for the production of C^{3+} which is given by the sum of the first three ionisation energies is $E_C = 83.5312 \pm 0.0003$ eV, while the threshold energy for O^{4+} is given by $E_O = 181.0882 \pm 0.0003$ eV [115]. It therefore seems reasonable that at least one of these multiply ionised species are contributing to the background at higher electron energies.

Deuterium

An unavoidable source of background arises from the natural abundance of deuterium in the vacuum chamber. D_2 molecules also have a very similar charge to mass ratio to helium of $m/z \approx 4$. Some mass spectrometers can separate molecules of deuterium from helium by their mass difference [116, 117], but for any high efficiency ioniser the space charge will

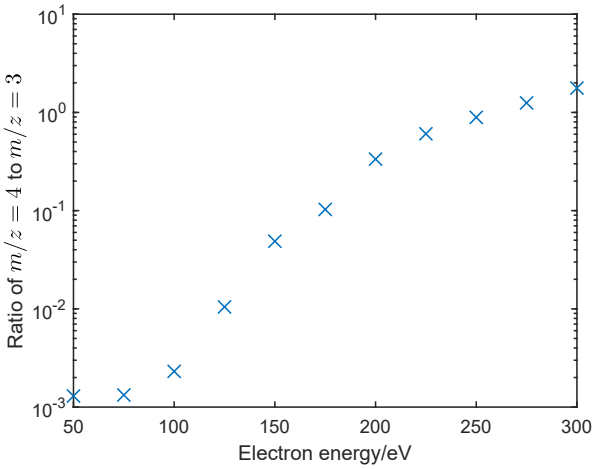


Figure 4.46: Plot showing the peak ratio of mass 4 to mass 3 while changing the electron energy. Below electron energies of 150 V the channeltron voltage had to be increased since the total efficiency was dropping. The peak ratio dramatically falls below about 200 eV and continues rolling off till about 100 eV.

spread the ion energies, out making it extremely difficult to separate the two species. It is also possible to separate the deuterium from helium using the difference in ionisation energy [118], but using low energy electrons would also result in a large loss in ionisation efficiency.

The abundance of deuterium is difficult to determine inside the vacuum chamber, but the accepted "absolute" mole fraction of deuterium to hydrogen atoms from a single natural source is $\alpha_D = 0.00015574(5)$ [119]. The actual abundance inside the vacuum chamber is difficult to estimate, so the best estimate will be to use the abundance found in nature. Additionally, the ion optics and the secondary electron yield are very likely to be different for the isotopes of hydrogen, but these effects will be ignored. If the natural abundance of D is $\alpha_D = 0.00015574(5)$, then the natural abundance of HD is $2\alpha_D$ and the natural abundance of D_2 is α_D^2. Therefore, the ratio of the $m/z = 3$ to $m/z = 4$ peaks is expected to be $2/\alpha_D = 12842 \pm 4$ if the above assumptions for the abundance hold true.

The background was measured to give a representative example, after allowing the whole system to relax for 24 hours. First the efficiency using the Faraday cup was measured as $\eta = (1.75 \pm 0.03) \times 10^{-4}$. The helium leak was then removed and the background was measured using the electron multiplier. Using fig. 4.38, the current from the multiplier can be used to determine what the current incident on the Faraday cup would be without multiplication. It was found that for an ionisation efficiency quoted above, the weighted mean estimate of the background ion current onto the Faraday cup is $I_b = (2.43 \pm 0.02) \times 10^{-16}$ A. The flux of helium into the front of the detector needed to produce an ion current equal to the background current is given by $F_b = I_b/(e\eta) = (8.71 \pm 0.15) \times 10^6\,\mathrm{s}^{-1}$. Alternatively, if we assume that the average pumping speed of the ioniser is given by $S_a = 2S_i = 1.34\,\mathrm{Ls}^{-1}$ from table 4.2, then the average helium pressure needed in the ioniser to match the background ion current is given by $p_b = F_b k_B T/S_a = (2.67 \pm 0.05) \times 10^{-13}$ mbar. The error bar on the effective background pressure does not incorporate the very high error on the estimate for the average pumping speed of the ioniser. The flux is the preferred metric for

estimating the background since it does not rely on estimating the pumping speed.

The seemingly high value for the effective background pressure is due to the poor pumping in the middle of the ioniser and highlights that the sensitivity is a poor metric for measuring performance (which is covered in more detail in section 4.5.3). If the source of gas is at the front of the detector and the average pressure of the ioniser were given by $p_b = 2.67 \times 10^{-13}$ mbar, then the pressure in the optics chamber next to the pump would only be $p_o = 1.6 \times 10^{-15}$ mbar.

The source of the background is difficult to determine, since the peak ratio does not seem to match the expected abundance of deuterium. As can be seen in fig. 4.27, the background ratios of the peaks at $m/z = 3$ and $m/z = 4$ is $\sim \mathcal{O}(100)$ which is about two orders of magnitude from the expected value for deuterium. When the emission properties are being varied, the mass peak ratio does get to within about an order of magnitude of the expected ratio as seen in fig. 4.46. However, after letting the detector pressure settle the ratio does not seem to be consistent with a background solely from D_2.

To summarise this section, many of the contributions to the background were due to the instrument design and have been eliminated. The remaining background that is measured is due to ions being created in the ioniser with a mass to charge ratio of $m/z = 4$. At high electron energies there is a significant background due to multiple ionisation of a higher mass species. At low electron energies, the ratio of the peaks at $m/z = 3$ and $m/z = 4$ and are between one and two orders of magnitude from the ratio expected if the peaks are due to HD and D_2. The origin of the background is therefore either due to the ratio of HD and D_2 in a UHV vacuum chamber being pumped with getter pumps diverging from the natural abundance, or there is another $m/z = 4$ due to multiple ionisation that still exists at low electron energies.

4.5.3 Efficiency

In typical detectors the gas pressure will be constant throughout an ioniser, so the best metric for the performance of the ioniser is the sensitivity often measured as the current per unit mbar of gas in the ioniser. However, for the solenoidal detector the pressure varies across the ionisation region and there is no way to directly measure the pressure inside the ioniser. Instead, the total current can be measured and compared to the flux of gas passing through, to give an ionisation efficiency.

The efficiency is obtained by leaking gas into the front of the detector, measuring the gas flux at the back and comparing the flux to the ion current obtained on the Faraday cup. Assuming that all the gas that enters the detector reaches the back of the ioniser, the gas flux through the ioniser can be measured from the pressure and pumping speed next to the pump on the detector. The flux of helium down the ioniser Q is given by,

$$Q = p_b S, \tag{4.19}$$

where p_b is the partial pressure of helium at the back of the ioniser and S is the pumping speed at the back of the ioniser. The pumping speed of the Pfeiffer HiPace 300 M pump for helium is $S = 215\,\mathrm{L\,s^{-1}}$. We can also express the gas flux Q as the number of atoms passing

through the ioniser using,

$$Q = \frac{N}{\Delta t} k_B T, \tag{4.20}$$

where N is the number of atoms passing through the ioniser in time Δt, k_B is the Boltzmann constant and T is the temperature of the ioniser. By rearranging equations 4.19 and 4.20, we find the number of atoms N passing through the ioniser in time Δt is given by,

$$N = \frac{p_b S \Delta t}{k_B T}. \tag{4.21}$$

The total number of ions that are detected is given by,

$$N^+ = \frac{I \Delta t}{e}, \tag{4.22}$$

where I is the ion current and e is the electron charge. The efficiency of the ioniser is then defined as N^+/N, which is the fraction of ions produced from the incoming atoms. The efficiency is then given by,

$$\frac{N^+}{N} = \frac{k_B T}{eS} \frac{I}{p_b}. \tag{4.23}$$

Therefore, if the detector is at $T = 298\,\mathrm{K}$, then the efficiency will be given by,

$$\frac{N^+}{N} = 1.194 \times 10^{-3} \frac{I}{p_b}, \tag{4.24}$$

when the current I is measured in amps and the pressure at the back of the ioniser p_b is measured in mbar.

It should be noted that measuring the efficiency using the Faraday cup will only provide a lower bound on the total efficiency when using the electron multiplier. Since the on axis Faraday cup is smaller than the beam, some of the ions are lost. However, the dynode is much larger and will collect a larger fraction of the ions. An estimate of the difference in the collected ion fraction is obtained by investigating the signal noise in section 4.5.4.

It is often assumed that the efficiency of an ioniser scales linearly with emission current, with detector performance often quoted as output current per unit emission current [78]. Although, space charge effects usually mean there is some degree of non-linearity between the efficiency of an ioniser and the emission current [120].

Figure 4.47 shows how the ioniser behaves as the solenoid current is varied. At low emission currents where the efficiency behaves relatively simply, higher solenoid currents give larger efficiencies. However, above an emission current of about 0.2 mA, the efficiency collapses and behaves in a complex non-linear way, with very narrow ranges of emission currents providing much higher efficiencies.

Varying the electron energy also causes the total efficiency to vary in a complex and non-linear manner as seen in fig. 4.48. For low emission currents, there is very little difference between using different electron energies. However, at large emission currents, larger electron energies generally give higher efficiencies.

The observed behaviour is likely to originate from the ion current that is escaping the ioniser. The measurements presented above are of the total efficiency and include the transmission probability from the ioniser, through the mass filter to the electron multiplier. It

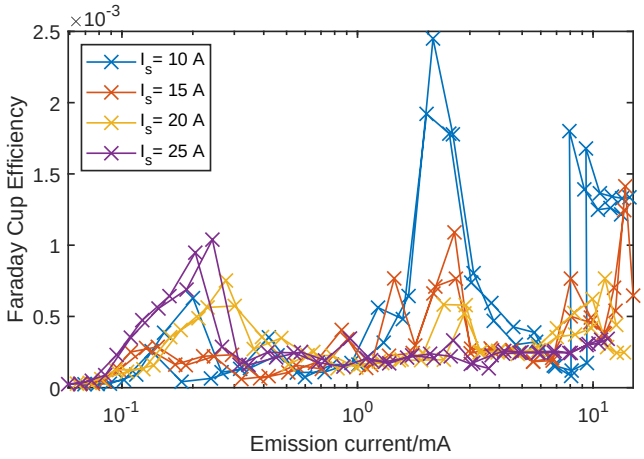


Figure 4.47: Ioniser efficiency as measured by the Faraday cup for different filament and solenoid currents. As the emission current is increased by increasing the filament current, the efficiency does not follow a simple linear relationship and behaves highly non-linearly above about 0.2 mA. The data was collected by decreasing and then increasing the filament current, the structure is repeatable as can be seen from the peaks being present when either increasing or decreasing the emission current.

seems unlikely that the behaviour is generated in the ion optics, therefore a simple explanation is that the complex behaviour is due to the ions not escaping from the ioniser correctly.

One potential explanation is that the space charge inside the ioniser is not behaving as expected and is non-uniform. If the space charge is not uniform, the potential will form wells and peaks that could affect the performance of the ioniser and prevent ions from escaping. The filament that produces the electrons may also have hot and cold spots that will create a non-uniform space charge preventing ideal ioniser operation.

Before an emission current of about 0.2 mA, the space charge is sufficiently small that the potential inside the ioniser has no irregularities. Figure 4.47 shows that increasing the solenoid current increases the efficiency in the low emission current regime. As the solenoid current is increased, the helical trajectories of the electrons have a smaller radius and the path length through the ioniser increases. Since the path length increases, the electrons interact with the gas for longer and the efficiency is increased.

In the highly non-linear regime the behaviour of the efficiency is consistent with potential wells forming inside the ioniser. The helical trajectories have an increased radius when either the electron energy is increased, or the solenoid current is reduced, which will lead to the space charge from the electron beam being smaller since the beam is more dispersed and closer to the liner wall. Figure 4.47 shows that in the non-linear regime, the efficiency is increased by decreasing the solenoid current and fig. 4.48 shows that in the non-linear regime, the efficiency is increased by using a higher electron energy. Therefore, it seems plausible that the complex behaviour is due to potential wells forming in the ioniser that block ions from exiting.

The potential in the region where the ions are extracted is very difficult to accurately model for the geometry in the detector. When the electrons are trapped in the ioniser, they will reflect off the end of the solenoid and will form a space charge well at the turning point. As the electron energy is varied, the electron turning point will shift and the extraction potential will be altered.

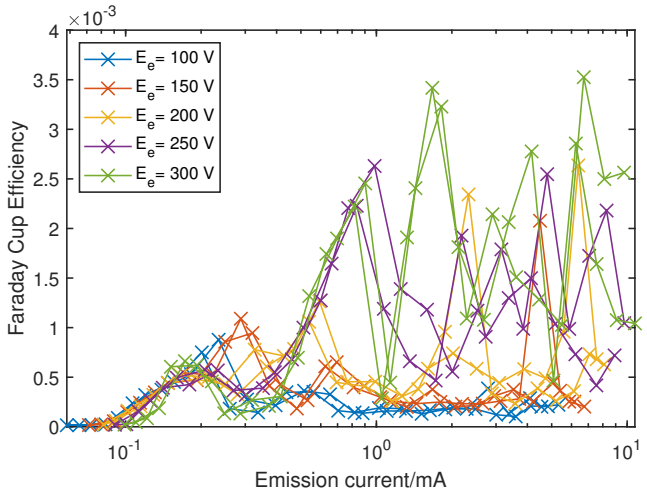


Figure 4.48: A plot of how the efficiency varies with filament current and electron energy. As the electron energy is increase larger efficiencies are possible but still the efficiency varies chaotically with filament current. Similarly to earlier, the peaks that appear are repeatable.

It therefore seems it could be possible to substantially improve the detector performance if the space charge is controlled more precisely. Additional measures to ensure that the potential inside the ioniser is more uniform could lead to much simpler behaviour and a larger efficiency of the ioniser.

4.5.4 Noise analysis

Another critical property of a detector is the noise on the output current from the instrument. The noise on the detector determines the time needed to collect data for at each pixel to achieve the desired signal to noise in the final image. Additionally, it is also possible to estimate the gain of the multiplier using the noise on the signal. Therefore, it is crucial that the noise on the detector output is properly investigated.

Early investigations into the noise on the detector output were performed for the NEMI collaboration and concluded that while closed mode produced a higher efficiency ioniser, the output current was also noisier [94]. It was shown that closed mode produced more noise than operating the ioniser in open mode. However, these measurements were performed using high energy electrons and large fluxes of helium, which are not the conditions used for microscopy. Instead the remainder of this section will focus on the noise obtained when using the detector under realistic conditions for helium microscopy.

The current being measured in the detector arises from the arrival of discrete ions from the ioniser, therefore in the limit of low signal the noise on the measured current should be from shot noise. The number of ions being detected in a given time interval can be modelled as a Poisson process if we assume that the ions are produced at a constant mean rate and ions are produced independently of each other. The statistical fluctuation on the number of ions being detected in a given time interval becomes significant if the ion current is low. If N ions are detected in a time interval Δt, then the shot noise on the signal due to the statistical fluctuations is given by $\sigma = \sqrt{N}$. Therefore, the signal to noise ratio $\beta = I/\sigma$ on a mean current of I is given by,

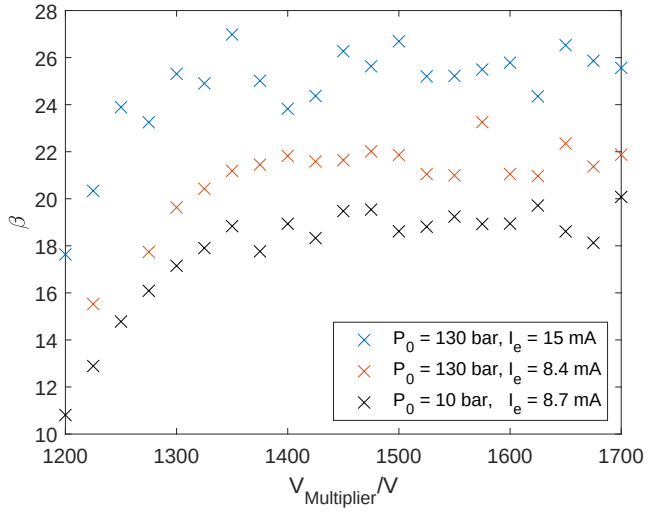


Figure 4.49: A plot of the signal to noise ratio, β, as the potential of the front of the multiplier, $V_{\text{Multiplier}}$, is varied for a series of different input helium fluxes. When the gain of the multiplier is low, the noise floor of the ammeter dominates the signal to noise ratio. As the gain is increased the signal to noise ratio reaches a maximum value given by the shot noise of the helium ion current. Increasing the helium ion current into the dynode assembly improves the signal to noise ratio. It should be noted that despite the large changes in source pressure and emission current, the measured flux does not vary much. Therefore the maximum measured signal to noise ratio is similar across the different curves.

$$\begin{aligned} \beta &= \frac{I\Delta t}{\sqrt{I\Delta t}} \\ &= \sqrt{I\Delta t} \end{aligned} \tag{4.25}$$

Figure 4.49 shows how the signal to noise varies as the gain of the multiplier is increased; as the gain increases the signal to noise ratio reaches a plateau. When the total current from the multiplier is low, the noise is dominated by the noise floor of the ammeter. However, as the gain of the multiplier is increased the signal to noise plateaus due to the shot noise of the small ion current that is being multiplied. The ion current was then varied by either changing the helium source pressure or by altering the emission current. When the ion current being measured is altered, it is seen that the signal to noise does depend on the total number of ions being detected as expected from equation 4.25. The transition between the noise being dominated by the ammeter noise floor and the inherent shot noise of the signal has also been observed on the spin-echo detector by Townsend [121].

The shot noise on a measurement can be used as an independent estimate of the gain of the dynode assembly. If both the number of events that are being detected and the output current from the multiplier are known, then it is possible to calculate an estimate for the gain of the dynode assembly. Since the shot noise depends on the number of events that are being measured, it is possible to predict the number of events from the shot noise. Let the number of electrons exiting the multiplier N_o in a time interval Δt be given by,

$$N_o = \frac{I_c \Delta t}{e}, \tag{4.26}$$

where I_c is the mean output current from the multiplier, and e is the charge of an electron. If the secondary yield coefficient γ is greater than one then it will contribute to the gain of the multiplier, however if $\gamma < 1$ then there will be no contribution to the gain of the multiplier. The number of ions that created the output current N_e is then given by,

$$N_e = \begin{cases} \dfrac{N_o}{g(V_c)}, & \text{if } \gamma \leq 1 \\ \dfrac{N_o}{\gamma g(V_c)}, & \text{if } \gamma > 1. \end{cases} \tag{4.27}$$

If equation 4.25 is assumed to be true then $N_e = \beta^2$, and an estimate for the gain of the multiplier is given by,

$$\begin{aligned} \gamma g(V_c) &< \frac{N_o}{N_e} \\ &< \frac{I_c \Delta t}{\beta^2 e} \end{aligned} \tag{4.28}$$

An estimate of the difference in the number of ions that are detected using either the Faraday cup or the electron multiplier can be obtained by comparing the results obtained by analysing the noise above to the gain measurement through equation 4.13. Figure 4.38 shows the total gain of the dynode assembly $G = A\gamma\alpha g(V_c)$, whereas the gain measurement for the electron multiplier using the signal to noise ratio only estimates $\gamma g(V_c)$. So by combining equation 4.28 and 4.13, it is possible to obtain a bound on the factor of ions that are not collected by the Faraday cup,

$$A\alpha = \frac{G}{\gamma g(V_c)}, \tag{4.29}$$

$$A\alpha > \frac{\beta^2 e G}{I \Delta t}. \tag{4.30}$$

Calculating $A\alpha$ for each point in fig. 4.49 where $V_c > 1400\,\text{V}$ and the noise is given by the shot noise, gives a mean value of the bound as $A\alpha = (6.40 \pm 0.09)$. Therefore, the efficiency obtained by using the electron multiplier is at least 6 times higher than when using the Faraday cup.

In conclusion, while the general noise properties of the detector may be complicated, under the conditions to be used on the microscope the noise on the output current is shot noise. Given that shot noise is the dominant noise source, the noise is improved by increasing the measured helium flux either by maximising the helium flux arriving in the detector or using the highest efficiency possible.

By measuring the shot noise of the detector, it is possible to obtain an estimate of the number of ions that are being detected. Therefore, given the number of ions being measured and the total output current, an estimate of the gain of the electron multiplier and dynode assembly is possible. Comparison of this estimate to the previous gain measurements shows that the dynode collects a helium flux at least (6.40 ± 0.09) greater than the Faraday cup due to its larger size. It therefore follows that any efficiency measured with the Faraday cup needs to be multiplied by this factor to give the actual efficiency that would be measured using the dynode.

4.6 Conclusions

In this chapter, the design of a high efficiency detector for helium microscopy using a solenoidal ioniser has been described. The electron injection into the solenoid has been simplified to increase the maximum emission current and provide a stable loss mechanism for electrons in the ioniser. The performance of the ioniser was found to be more stable and repeatable than the previously implemented ioniser on the spin-echo apparatus.

Mass filtering was achieved using a 180° permanent magnetic sector and varying ion energies. Ion optics with voltages that scale with ion energy have also been developed to allow mass scans that show that helium is successfully filtered.

A conversion dynode was designed and implemented to convert helium ions into electrons and then multiply the signal using an electron multiplier. The shielding of the entrance of the electron multiplier to external fields is essential to its performance, and addition of a mesh to the front of the multiplier is necessary. The importance of shielding the electron multiplier has implications beyond just helium microscopy. The unusual behaviour of the dynode system installed in the Cambridge spin-echo apparatus by Chisnall [77] can be explained by considering the effect of an external field on the electron multiplier.

The noise on the output of the detector was found to be dominated by shot noise. Analysis of the shot noise showed that there is a significant difference between using the small on-axis Faraday cup and large dynode, with the dynode assembly collecting a helium flux at least (6.40 ± 0.09) greater than the Faraday cup. The large increase in collected flux acts to further confirm that the conversion dynode is performing as expected.

The efficiency of the ioniser was measured by an on-axis Faraday cup and shown to not scale as expected with emission current or electron energy indicating significant issues with either ion generation or extraction. A potential explanation for the unusual behaviour is that potential wells are being formed inside the ioniser due to space charge. By using a larger ionisation volume than a commercial instrument, the detection efficiency can be as high as 2% for large electron energies (under normal operation a small electron energy is used) which is several orders of magnitude higher than a commercial instrument. As the signal to noise ratio is dominated by shot noise, images obtained with the new detector have a substantially improved signal to noise ratio when compared to the original images collected with the commercial mass spectrometer.

When attempting to measure very low partial pressures of helium, the background of the instrument is a key parameter. Large backgrounds can occur from insufficient control of the ions in the instrument, leading either to direct measurement of stray ions or indirect measurement through secondary electron generation. Methods have been described that mitigate the background so that the only signal that is measured is from a species with a mass to charge ratio of $m/z = 4$. The remaining background signal is then dominated by multiple ionisation, which can be reduced by using a lower electron energy of about 100 eV. After tackling all the sources of the background, the remaining background is equivalent to a helium flux of only $F_b = (8.71 \pm 0.15) \times 10^6\ \mathrm{s}^{-1}$ entering the detector, which is likely to correspond to either multiple ionisation at an electron energy of 100 eV or from D_2.

Chapter 5

Contrast mechanisms in scanning helium microscopy

5.1 Introduction

One of the major benefits of the scanning helium microscope is that it can be used on any vacuum compatible material. Since the beam is formed of neutral atoms with a low energy, surfaces will not suffer from beam damage and do not need to be conducting, making the technique applicable to a large range of materials.

Historically, the small de Broglie wavelength of a helium beam has allowed diffraction experiments to be performed from the atomic lattices of crystals. The diffraction pattern from a helium scattering experiment can then be used to determine atomic crystal structures. Careful preparation of atomically ordered surfaces provides a periodic potential for the helium to scatter from and form diffraction peaks.

In general, most materials do not form well defined crystal structures on the top layer of atoms of a surface due to either the presence of a contamination layer or the material not forming a signal crystal lattice at the surface. Only some materials will have crystal structures, of which many are polycrystalline and other surfaces will reconstruct and may not possess the crystal structure of the bulk. Most materials that are studied with classic helium scattering require extensive cleaning through sputtering and annealing to remove surface contaminants and form a clean surface.

In this chapter I will discuss the contrast that has been observed in scanning helium microscope images of a variety of different materials. To begin with, I will focus on the contrast that is observed on general disordered surfaces, showing how it arises from the scattering distribution, topography of the surface and the geometry of the instrument. After establishing how contrast is formed on the vast majority of surfaces, I will then move on to looking at some more exotic forms of contrast. Simultaneous imaging with two gases presents an opportunity to image with two atomic beams with different wavelengths, which has a direct analogy to colour in optical images. I will present the first images obtained with a mixed gas beam and highlight the difficulties. Finally, I will present the first observation of contrast in a scanning helium microscope due to diffraction from an ordered surface.

5.2 Contrast from disordered rough surfaces

In order to understand the origin of contrast in images of rough surfaces, first the scattering distribution from a rough surface needs to be understood. The signal measured at each pixel is given by the total flux scattered into the solid angle subtended by the detector aperture. Therefore, the simplest contrast that can be observed from a single scattering event will be caused by a change in the scattering distribution from the surface. Contrast can also occur from the helium atoms having multiple interactions with the surface, but again the underlying scattering distribution will be key to understanding these contrast mechanisms.

Once the scattering distribution has been established in the next section, the contrast mechanisms that exist are then discussed in decreasing order of strength. The strongest contrast that is seen when imaging a sample with three dimensional structure, arises due to masking of the beam and diffuse illumination of shadowed regions. The scattering distribution has a weak preferred direction which produces angular contrast from a sample. A weak height dependent contrast is also present, but is only seen when the sample has macroscopic changes in height and is therefore usually negligible. The final mechanism that is discussed is due to topographic changes of the surface at a length scale below the resolution of the instrument.

5.2.1 Review of scattering distribution from rough surfaces

Modern helium atom scattering has typically been used to study scattering from well ordered crystal structures which leads to a large variety of different scattering distributions that shall be discussed in more detail later in section 5.4. However, the majority of surfaces do not form well ordered flat crystal structures and instead given that the wavelength of the helium beam is ~ 0.5 Å, most surfaces will appear as extremely rough when compared to the wavelength of helium. Instead, earlier work on atom scattering that doesn't include diffraction must be considered.

Early work by Knudsen [122, 123] proposed a cosine law for the angular distribution of atoms diffusely scattering from a surface. Knudsen stated that the probability, ds, of a molecule scattering from a surface into an element of solid angle $d\Omega$, making an angle θ with the normal to the surface is,

$$ds = \cos(\theta)\, d\Omega/\pi. \tag{5.1}$$

Knudsen measured the scattering distribution of mercury from a glass surface, showing that the angular distribution followed a cosine distribution. Knudsen then used the cosine law to calculate the diffusion of gas down cylinders, which determines the behaviour of gas in the molecular flow regime. The success of vacuum conductance calculations provides the major experimental evidence for the cosine law.

Justification of the cosine law originally relied on the system being in equilibrium and using the second law of thermodynamics to show that the average distribution is given by a cosine function [124]. However, Comsa [125] pointed out that there is a fundamental difference between Knudsen's cosine law and the equilibrium cosine law. It is important to note that Knudsen's law is empirical and describes the angular distribution of each scattered molecule and therefore is applicable to an atomic beam experiment where the gas is not in

equilibrium.

Research into diffuse atom scattering was mainly focused on the calculation of molecular flow of gases through vacuum systems [126], since understanding how atoms are reflected from the walls of a vacuum chamber is essential to calculating the gas flow. While it is known that there can be deviations from the cosine distribution [127, 128], these typically only produces small deviations from using the cosine distribution. In fact, more recent work highlights that the precise nature of the scattering distribution of each collision may not be necessary to obtain an accurate prediction of the total gas flow [129, 130]. Therefore, more explicit determination of the scattering distribution is needed to determine the characteristics of the scattered helium.

O'Keefe and Palmer [131] explicitly show that helium atom scattering from rough surfaces produces distributions that are close to cosine distributions rather than through inferring from diffusion calculations. The authors scatter helium from a series of glass samples that were treated to vary the roughness and show that all the distributions are all very diffuse in nature. However, some of the rougher samples show enhanced backscattering indicating a deviation from the cosine distribution.

Knudsen's cosine law has a direct equivalent in diffuse optical scattering known as Lambert's cosine law [132]. Lambert's law is also an empirical law and provides some context to understand the origin of Knudsen's law. A Lambertian surface is one which produces perfectly diffuse scattering that follows a cosine law. A Lambertian surface has a constant radiance, which gives a cosine distribution when the detector angle is varied due to the change in the projected area of the source. The cosine variation that is observed is not due to the helium having a preferred scattering direction, but instead arises due to changes in the projected area of the illuminated surface. The connection with Lambert's law is also useful since it allows optical methods to be applied to helium scattering.

5.2.2 Observed scattering distribution from rough surfaces

As previously discussed, the scattering distribution is the key to understanding the contrast produced in images, but it is difficult to obtain directly with the microscope. Determination of the scattering distribution requires the flux at a range of different scattering angles to be measured. Ideally the sample would be fixed and a detector would be moved through a range of angles to measure the full scattering distribution. A movable detector is not possible in the microscope and instead an alternative method to vary the scattering angle is needed.

Figure 5.1 illustrates how the scattering angle can be varied in the microscope through changing the distance from the pinhole to the sample. Since the detector aperture is fixed at a distance of $h = 4.24\,\text{mm}$ from the pinhole, the angle needed for helium to scatter from the sample into the detector aperture depends on the distance between the sample and the pinhole. The outgoing angle from the sample to the centre of the detector aperture, θ, is given by,

$$\tan\theta = \frac{h - z}{z}. \tag{5.2}$$

Figure 5.1 also illustrates how the sample needs to be repositioned after changing the z distance to ensure the same area of the sample remains illuminated by the helium. Given that the incoming angle is not normal to the sample, the sample needs to be moved as the

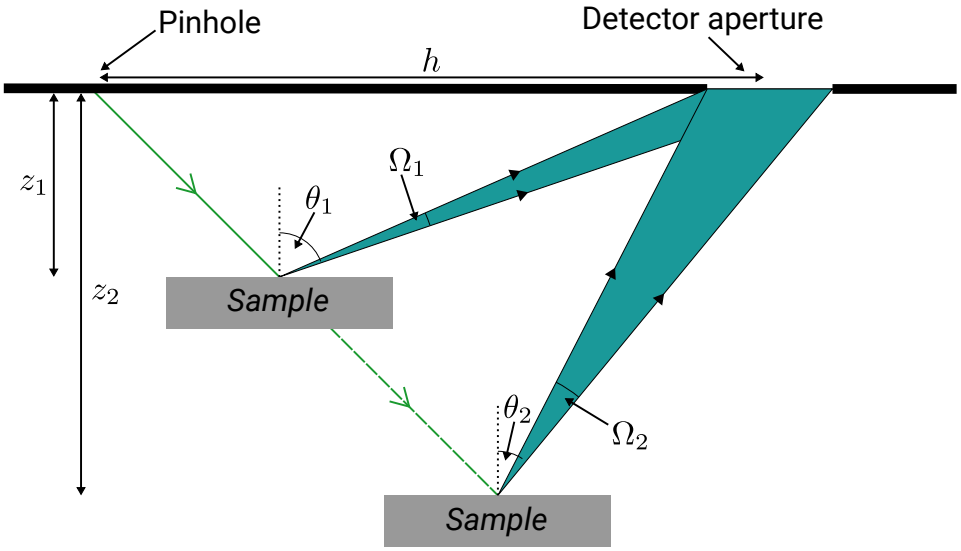


Figure 5.1: A schematic showing how the sample is moved during a z scan and which variables are changed as the pinhole to sample distance is increased. Since the incoming beam is at 45°, the sample needs to be shifted horizontally to ensure that the same part of the surface is illuminated. When the sample to pinhole distance is changed, the outgoing angle of gas that is detected and the solid angle of the detector aperture changes.

distance between the pinhole and the sample is changed. Given that the incoming angle is 45°, if the sample is moved Δz back from the pinhole then the sample must be moved horizontally by $\Delta x \approx \Delta z$. Due to small changes in the angle of the sample mounting, the exact distance that the sample must be moved laterally is not exactly equal to the shift in z position. Similarly, the vertical position of the sample will slightly vary with z position due to the sample mounting imperfectly and backlash on the stages.

However, as the sample distance is varied the outgoing angle is not the only quantity that varies. There are two other important factors that are varying, the solid angle of the detector aperture and transmission probability through the cone to the detector.

Figure 5.2 shows the predicted variation of the measured flux and solid angle when a z scan is performed. In order to understand how the geometry of the pinhole plate influences a z scan, a mote-carlo ray tracing simulation was performed. A uniform emitter was chosen so that the effects of the scattering distribution are removed and only changes due to the geometry are important. The calculation was performed by importing the geometry of the pinhole plate and tube to the detector into Molflow+ [112] and measuring the flux that arrives at the detector for a uniform point source at the sample. The blue line in fig. 5.2 gives the fraction of the total emitted flux that was detected for a uniform emitter at the sample. To support the Molflow data, the variation of the solid angle was independently calculated. The orange curve in fig. 5.2 shows how the solid angle of the detector aperture changes as the sample is moved further from the pinhole and was calculated using equation 52 in the paper by Conway [133]. Comparing the two curves, it becomes apparent that the change in solid angle is the dominant factor determining the change in flux measured by the detector when the sample distance is changed.

Figure 5.3a shows the change in the detected flux not due to the variation in the solid angle of the detector aperture and was obtained by dividing the simulated flux from Molflow

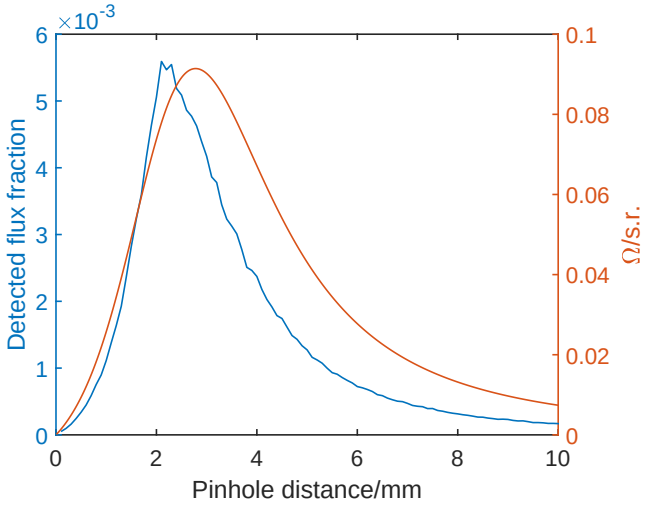


Figure 5.2: A plot where the blue curve shows the expected variation of the fractional detected flux from a uniform point source calculated using Molflow and the orange curve shows the solid angle change during a z scan calculated using equation 52 in the paper by Conway [133]. As the distance between the sample and pinhole is varied, the majority of the change in the detected flux is given by the change in the solid angle of the detector aperture. The remaining variation in the detected flux is given by the change in the transmission probability through the detector cone.

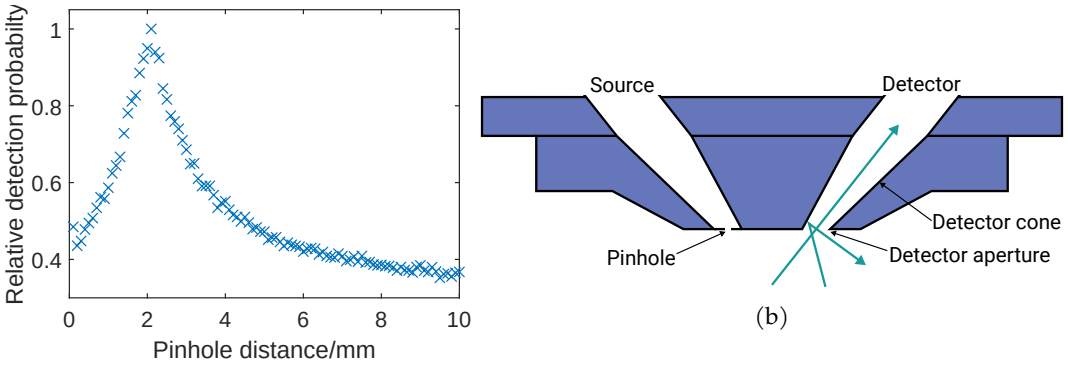


Figure 5.3: (a) A plot showing the variation of the transmission probability through the detector cone during a z scan. The curve is obtained by taking the curves in fig. 5.2 and dividing the fractional detected flux by the solid angle and normalising. The peak in the detection probability occurs when the gas atoms travel straight down the axis of the detector cone. If the gas enters the detector cone at a large angle, the gas is more likely to diffuse back into the scattering chamber and not be detected.

by the solid angle as calculated in fig. 5.2. Any variation seen that isn't due to the solid angle must be due to the transmission to the detector depending on the incoming angle of the gas entering the detector aperture. Figure 5.3b illustrates how the detector aperture connects to the detector through a cone in the pinhole plate. Gas travelling down the cone axis will travel much further before scattering than if the gas enters at a large angle. The maximum in the detection probability is observed at $z = 2.12\,\text{mm}$, which is the distance at which the gas would travel directly down the axis of the cone. The detection probability was also independently calculated by Lambrick and the results agree with the calculation presented here [111]. In reality, the detector cone is actually made from a series of discrete steps and therefore the function calculated here is actually just an approximation.

The direction dependent transmission probability illustrated in fig. 5.3a can be tuned to either improve the depth of field or background rejection rate. By changing the shape of the cone, it would be possible to change the direction correction factor. A larger opening angle will produce a higher transmission probability for angles far from specular, whereas as the cone approaches a narrow tube, the transmission will peak around the specular scattering condition. Accepting a wider range of angles allows images to be collected with the outgoing angle being further from specular, thus the instrument depth of field and signal level is increased. However, the increase in transmission probability for large angles will also mean that any background helium that is entering the detector aperture far from the specular will also have a larger transmission probability. Future instruments will need to tailor the precise shape of the cone to ensure that the maximum depth of field is obtained while maintaining an acceptable rejection rate for stray background atoms.

Figure 5.4a shows an experimentally collected z scan from a roughened area of a LiF surface. The full signal obtained during a z scan, $I(\bar{\theta})$, is given by,

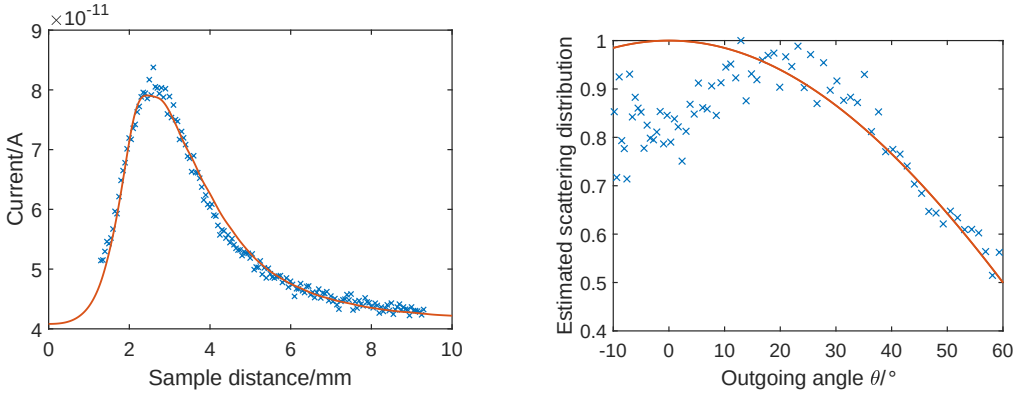
$$I(\theta) = \int_A f(\theta')T(\theta')d\Omega, \tag{5.3}$$

where $f(\theta)$ is the scattering distribution of the sample, $T(\theta)$ is the transmission probability to the detector, $d\Omega$ is a solid angle element and the integral is performed across the whole area of the detector aperture. An extremely crude first order approximation is to assume that neither the scattering distribution or detection probability vary across the detector aperture and take an average value giving,

$$I(\theta) = f(\theta)T(\theta)\Omega(\theta). \tag{5.4}$$

The orange curve in fig. 5.4a is the result of smoothing the data obtained from the Molflow simulation which is approximately equal to $T(\theta)\Omega(\theta)$ and multiplying it by a cosine distribution to approximate equation 5.4. The result is very close to the experimental data despite the gross approximations involved by not performing the integral or correctly modelling the detector cone as a series of steps.

It is then possible to obtain a crude estimate for the scattering distribution by taking the experimental data and dividing it by the simulated Molflow data which is equal to $T(\theta)\Omega(\theta)$. Figure 5.4b shows an example of the scattering distribution that was obtained, a cosine distribution is shown in orange for reference. The distribution that is obtained is very wide and therefore supports the theory that the helium is scattered diffusely. The peak position is



(a) (b)

Figure 5.4: (a) A plot showing a z-scan collected on a heavily distorted LiF surface in blue points. The orange curve is a rough estimate of the expected behaviour if the scattering were given by a cosine distribution as calculated using a smoothed version of the Molflow data in fig. 5.2 and equation 5.4. There is a good agreement between the data since the variation is dominated by the change in solid angle and transmission probability. (b) The approximated scattering distribution obtained from the z-scan data in blue and a cosine distribution for reference in orange. The statistical error bars are smaller than the markers, however the real uncertainty is much larger particularly at negative outgoing angles where the detector noise is dominating. The distribution is wide indicating that diffuse scattering is occurring. A peak in the distribution exists, but the location is difficult to interpret due to the heavy approximations involved.

not at $0°$ as expected from a cosine distribution, however the significant approximations in obtaining the estimated scattering distribution are very likely to influence the peak position.

The difficulty with obtaining a scattering distribution through analysing data from a z scan is that the variation of the solid angle and direction dependent transmission probability is so strong that the structure of a broad distribution will be smeared out. Any error in the calculation of either the solid angle or transmission probability will heavily affect the estimated scattering distribution. Therefore, the best that can be stated with the z scan data is that it is consistent with a broad cosine distribution, but it does not definitely prove it.

In summary, the literature suggests that the angular scattering distribution of atoms from a general surface should give a cosine-like distribution. In the microscope, the outgoing angle of the atoms that are collected can be varied by changing the distance between the pinhole and the sample. However, the major contributor to the change in signal is due to the variation in the solid angle and detection probability, not the change in the scattering distribution. In section 5.2.5, the angular contrast that is observed provides another piece of evidence that is consistent with a cosine-like distribution. However, before continuing to discuss the scattering distribution, some stronger effects due to the local geometry must be introduced.

Now that the scattering distribution from a rough surface has been established, the remainder of this section of the chapter is devoted to discussing the contrast mechanisms that have been observed. First, in section 5.2.3, the contrast that arises due to the geometry of the machine is detailed and the concepts of shadowing and masking are explained which can produce very strong contrast in images. Next, in section 5.2.4, since the scattering distribution is very diffuse, it is possible to diffusely illuminate other areas of the sample

after scattering leading to a weaker contrast mechanism. The scattering distribution has a preferred direction, meaning that the angle the surface makes to the detector will generate contrast in images which is discussed in section 5.2.5. In section 5.2.6, an additional very weak contrast mechanism is detailed from the distance between the sample and the pinhole, however this is only relevant in samples with macroscopic changes in height. To end this part of the chapter, in section 5.2.7 the contribution of structure smaller than the beam width is discussed.

5.2.3 Scattering geometry contrast

The strongest contrast mechanism that is observed from a diffuse scattering distribution is caused by the geometry of the incoming beam and outgoing detector aperture. Figure 5.5 illustrates the two principal effects caused by the machine geometry, shadowing and masking.

Shadowing of a three dimensional structure, as shown in fig. 5.5, leads to parts of an image that are not directly illuminated by the helium beam and therefore only have a weak influence on the helium images that are formed. Since the incoming beam is at 45° from the surface normal, some areas of the surface will be hidden from the helium beam. The only way these areas can then can then contribute to image formation is if through diffuse illumination which is discussed in section 5.2.4.

It should be noted that shadows do not appear in scanning helium microscopy images in the same manner as they manifest in more conventional images. Optical cameras and our eyes form images by illuminating a target with a wide beam and then use a position sensitive detector, areas of the sample that are not illuminated and are in shadow will therefore appear as dark patches on the image. However, the scanning helium microscope is a single pixel camera that moves the sample around under a narrow beam of helium, therefore areas that are not illuminated do not contribute to the scattered signal and shadowed regions simply do not appear in the image.

Masking occurs when the detector aperture is not directly visible from the scattering point and produces very strong contrast in images. An example of masking in SHeM images is seen in fig. 5.6. Again looking at fig. 5.5 it is possible to see how masking occurs on a three dimensional surface. Since the incoming angle of the helium beam and outgoing angle to the detector are not the same, it is possible to illuminate parts of the surface with helium that do not have line of sight to the detector aperture. Helium scattering from a masked region can only reach the detector through multiple scattering and therefore there is a large reduction in the intensity seen in these regions.

The size of the masked region will vary as the angle from the sample to the detector aperture is varied. A machine using normal incidence for both the incoming and outgoing beams would have no masked regions and contrast from masking would be eliminated.

One way to interpret the images formed in the helium microscope is to consider the position of the detector aperture to be the effective source of the helium illumination, and the position of the pinhole to be the effective detector for the image [134].

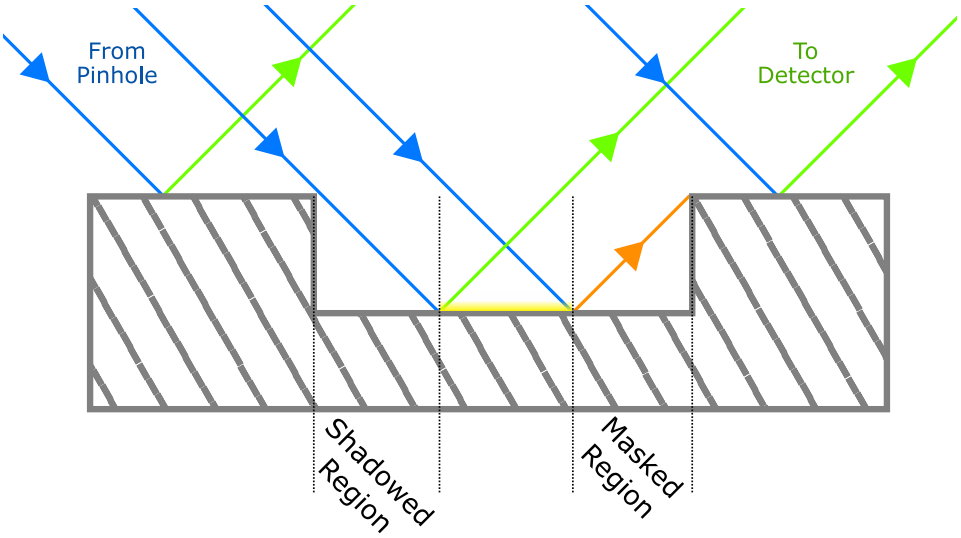


Figure 5.5: Schematic showing how the scattering geometry creates shadowed and masked regions in the images. Since the beam is incident at 45° to the normal, regions of the sample are not directly illuminated with helium leading to them not appearing in images. These parts of the sample that are not directly illuminated are referred to as shadowed. Some parts of the sample will not have direct line of sight to the detector aperture leading to the beam being masked. The masked regions will appear darker due to the helium being blocked from entering the detector.

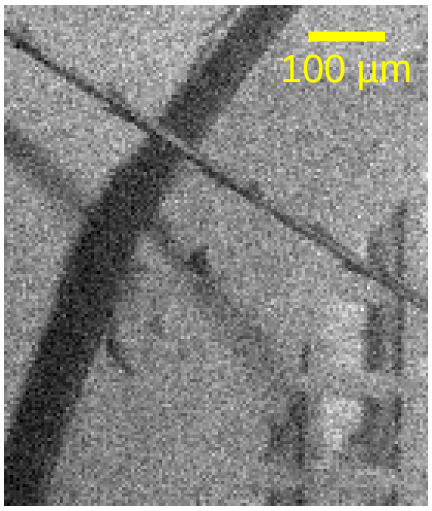


Figure 5.6: A SHeM image of a dandelion pappus on top of a TEM grid illustrating masked regions. The dandelion pappus blocks line of sight from parts of the surface below it, leading to a masked region across the image. Additionally, the edge of the TEM grid creates a dark masked region across the image. The image was collected using the original Hiden HAL/3F PIC detector, which has a low sensitivity leading to the shot noise in the image. The image consists of 137 x 162 pixels, each with a size of $4\,\mu\mathrm{m}$, that were measured with a dwell time of 4 s at every point.

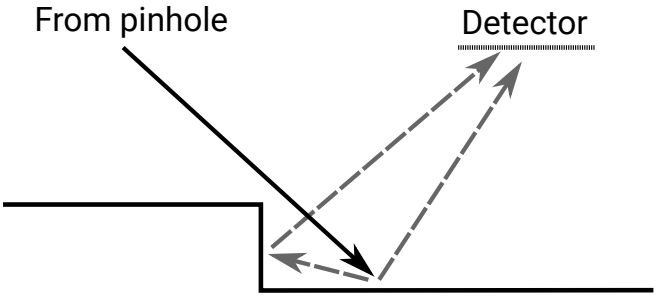


Figure 5.7: Schematic illustrating how diffuse illumination arises in an image. When helium scatters from a rough surface it will have a wide scattering distribution and a significant flux will not be directly detected. It is possible to gas to scatter off another surface nearby and then enter the detector. The primary incoming helium beam is therefore diffusely illuminating the other structure which was not directly visible with the primary beam. Figure taken from fig. 6(b) by Lambrick, Bergin et al. [111]

5.2.4 Diffuse illumination

Figure 5.7 illustrates a secondary effect of having such a wide scattering distribution is that helium that has been scattered from a three dimensional sample can diffusely illuminate other parts of the sample producing another contrast mechanism. As discussed in section 5.2.3, shadowed regions of samples are not directly illuminated and therefore don't directly appear in the images formed by the microscope. However, after the gas has been initially scattered from the sample some of the scattered gas could interact with structure in the shadowed regions producing signal.

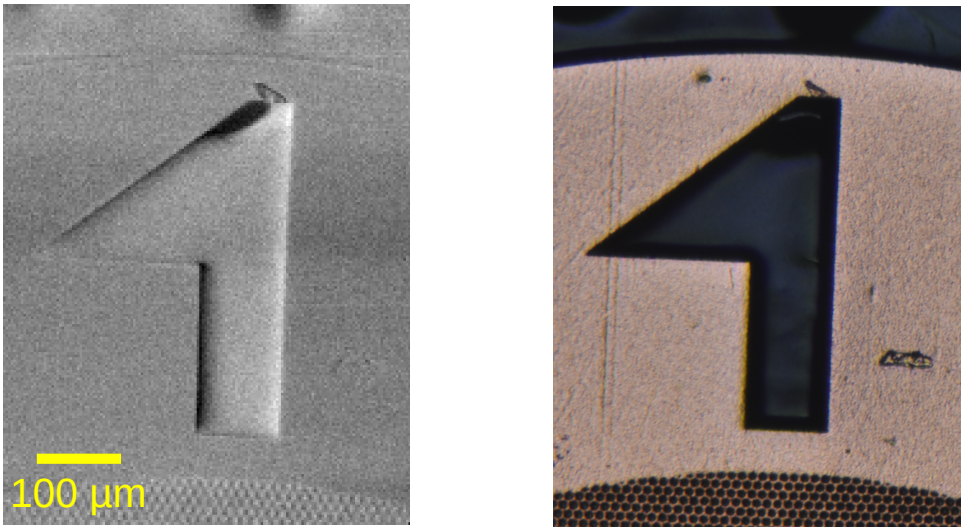
The effects of diffuse illumination are easiest to observe when looking at a sharp step down in height of the surface. Figure 5.8a shows a SHeM image of the calibration mark for TEM grid (G2000HA Fine Hexagonal Mesh from Gilder Grids Ltd.). Inside the calibration mark, the material that is visible is the tape beneath the TEM grid that is holding the TEM grid to the sample mount. In the image there are examples of both a masked region and diffuse illumination. On the left hand side of the calibration mark there is a masked region so the signal drops. On the right hand side of the calibration mark the helium that is scattered from the tape diffusely illuminating the side of the TEM grid and increasing the signal measured here.

A more detailed discussion on the contrast observed in fig. 5.8a can be found in a recent paper by Lambrick, Bergin et al. [111]. By using a ray tracing simulation, it is possible to produce a simulated image of a sample that reproduces the contrast seen in fig. 5.8a.

5.2.5 Angular contrast from rough surfaces

The scattering distribution from rough surfaces leads to a contrast mechanism that is observed is due to the angle the normal of a surface makes to the detector aperture. Since the scattering distribution has been shown to be a cosine-like distribution, the helium that is scattered from the surface has a preferred direction and contrast can be seen when the surface normal changes.

Figure 5.9 is an example of an image of the compound eye of a fly, which is made up of many smaller spherical lenses without any coating, demonstrating the contrast arising from



(a) (b)

Figure 5.8: Images of the calibration mark on a TEM grid (G2000HA Fine Hexagonal Mesh from Gilder Grids Ltd.). (a) A SHeM image of the calibration mark from a TEM grid mounted onto a PELCO tab. The left hand side of the arrow mark illustrates the loss of signal in a masked region, while the right hand side illustrates an example of diffuse illumination increasing the signal. Note that in the image the pinhole is illuminating the sample from the right. The image consists of 160 x 205 pixels, each with a size of $3\,\mu\mathrm{m}$, that were measured with an initial pause of 0.75 s and a dwell time of 1.2 s at every point. (b) An optical image of the same TEM grid that has been rotated and cropped to match the SHeM image.

the change in the surface normal. When the normal to the sphere points towards the detector aperture the signal is at a maximum, with the signal falling across the spherical lens. Figure 5.9(c) also contains a good example of a masked region underneath the contaminant.

The observed angular dependence to the contrast is consistent with a cosine-like scattering distribution. When the total scattering angle is 90°, the outgoing normal from the surface of a sphere will point towards the detector when the incoming beam is at a tangent to the sphere. Therefore, since there is no visible peak on the sphere surfaces, the peak in the scattering distribution must occur either at or beyond the surface normal which is consistent with a cosine scattering distribution. While the angular contrast is not direct evidence that the scattering distribution is a cosine distribution, it does add to the evidence that the scattering distribution is given by Knudsen's Law.

5.2.6 Height dependent contrast

An additional very weak contrast mechanism that occurs is due to changes in the distance between the sample and the pinhole. Looking at fig. 5.4a, we see that the signal from a surface varies as the distance between the pinhole and the sample is changed. As previously discussed, the variation with distance is a combination of the change of the outgoing angle, the solid angle of the detector aperture and the detector response function.

The variation with working distance is a smooth slowly varying function, so is only relevant when there are large macroscopic changes in height on the order of $100\,\mu\mathrm{m}$. The height change must be sufficient to substantially alter the scattering angle and solid angle of the

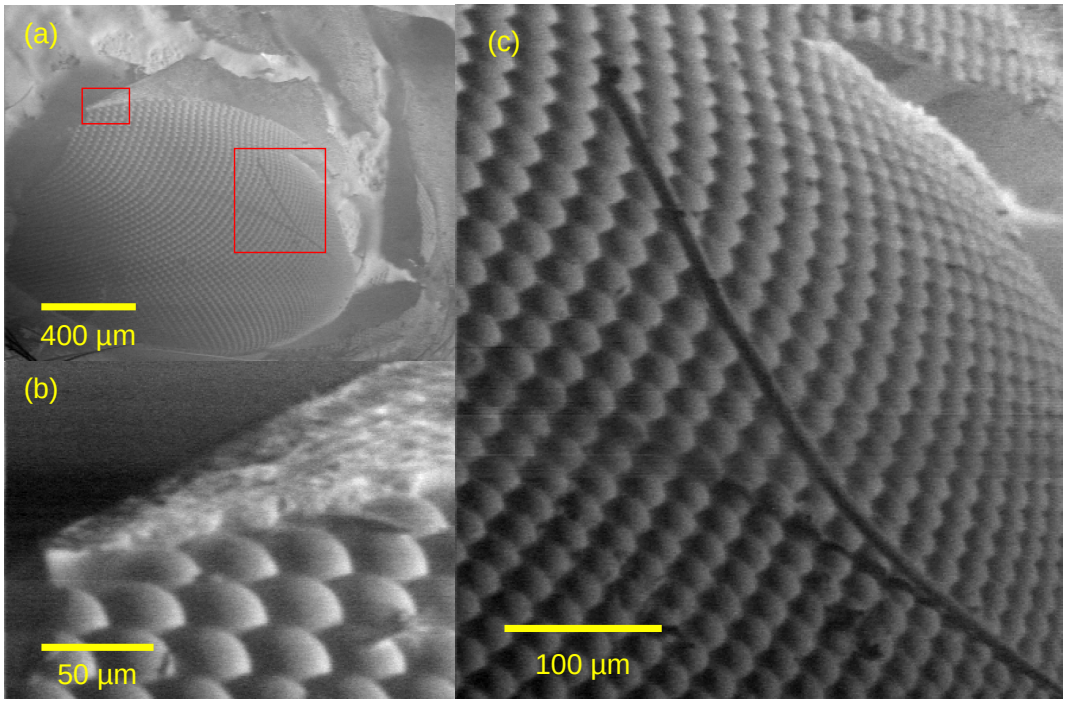
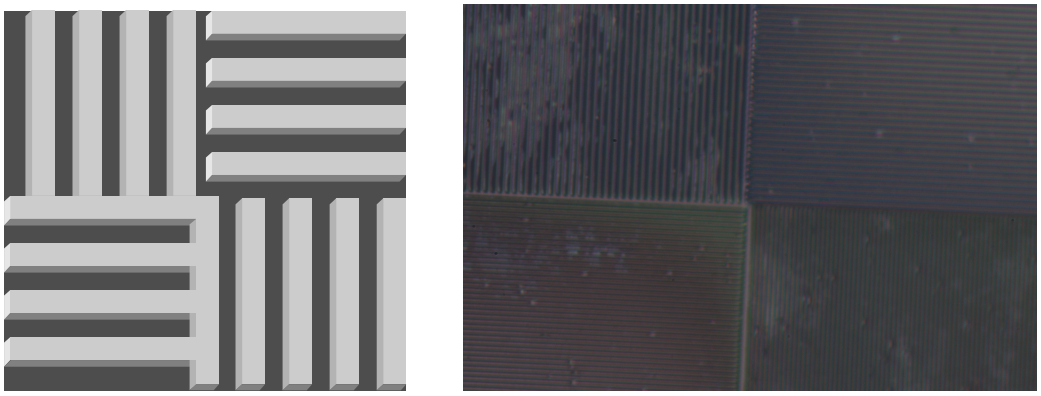


Figure 5.9: SHeM images of a fly compound eye. All images were post processed by subtracting the median height difference between the rows so that the detector instabilities are not visible. (a) Large scale SHeM image of the entire eye with the red boxes indicating where the other two images are obtained. The image consists of 380 x 300 pixels, each with a size of $5\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 0.5 s at every point. (b) A high resolution image of the edge of the eye. The helium signal is highest when the surface normal points towards the detector. The image consists of 333 x 250 pixels, each with a size of 600 nm, that were measured with an initial pause of 0.5 s and a dwell time of 0.5 s at every point. (c) A composite image of the eye with a contaminant that creates a mask on the surface of the eye. Due to issues with the detector, three images have been combined together to produce a complete image. The image consists of 385 x 440 pixels, each with a size of $1\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 0.5 s at every point.

detector aperture and therefore must be macroscopic in size. For samples with such an egregious height variation, other techniques will be more suitable for measuring the difference in height.

Fahy et al. recently published work where a 3D printed sample with height variation was used to explicitly show the height dependent contrast [63]. The authors also outline a Monte Carlo ray tracing simulation that was used to support the experimentally measured Michelson contrast values. Fahy et al. use the model to predict that the minimum resolvable step height is $(67 \pm 5)\,\mu\mathrm{m}$, confirming the notion that height dependent contrast will only be significant for macroscopic samples.

Fahy et al. predict that the contrast will invert at the specular position, however this is in contradiction to the z scan presented earlier in fig. 5.4a since the signal does not peak at the specular condition. The Michelson contrast is proportional to the differential of the z scans presented earlier. As previously discussed, the obtained signal is a combination of the solid angle of the detector aperture, the transmission probability and the scattering distribution. Therefore, there is no reason for the signal to peak on the specular condition and it would



(a) (b)

Figure 5.10: (a) Schematic of the plastic samples with stamped trenches. The surface consists of a series of trenches running in either a horizontal or vertical direction. (b) Optical image of the plastic trenches illustrating the different directions of the trenches.

purely be coincidence if it did. It is not clear whether the discrepancy is because the authors have not carefully defined the distance to the pinhole or there is a difference in our methods.

5.2.7 Sub-beam contrast

In reality, the observed scattering distribution is due to an average over the area of the surface that is illuminated by the beam. The mechanism that creates the scattering distribution that is measured may be due to structure that cannot be directly spatially resolved with the microscope. Knudsen's cosine law arises from the surface being considered rough when compared to the helium wavelength which is typically $\lambda \approx 0.05\,\mathrm{nm}$. Therefore, while it may be possible for a surface to locally have a cosine distribution, when averaged over a large area the effective scattering distribution could diverge from Knudsen's law.

Topographic contrast is not limited to changes on length scales larger than the beam width, changes in the topography can be seen even if they are smaller than the beam width. Provided that the basic scattering distribution remains a cosine distribution, all the contrast mechanisms that have been described so far are independent of the size of the beam. Therefore any structure that is significantly larger than the atoms, such that the local distribution is given by Knudsen's cosine law, but smaller than the beam width, will generate contrast through the previously described mechanisms, but the structure will not be spatially resolved.

Figure 5.10a illustrates a series of trenches that can be used to demonstrate how structure that cannot be spatially resolved can still generate contrast. A plastic sample with a series of trenches stamped into it were provided NIL Technologies through the NEMI consortium. Figure 5.10a is a schematic showing how the directions of the trenches are in different directions at different points on the surface. The trenches are $\sim 1\,\mu\mathrm{m}$ apart, and can be seen to just be resolvable with an optical microscope in fig. 5.10b.

Figure 5.11 is a SHeM image of the plastic trenches illustrating contrast between the regions with alternating trench directions. The individual trenches cannot be resolved, however there is a clear contrast between the different trench directions. The Michelson contrast for the image in fig. 5.11 is 0.05, although this does include a significant background.

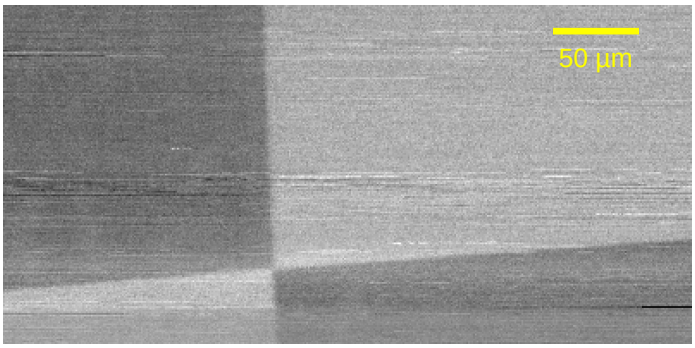


Figure 5.11: SHeM image of a plastic sample with a series of trenches stamped into the surface. When the trench direction is horizontal in the image the beam is not masked and the pixel is brighter. However when the trenches are vertical, the beam is partially masked and the measured signal is lower. The image has had the median difference between the rows removed to remove detector instabilities from the image. The resulting image has a Michelson contrast between the regions of 0.05. The image consists of 400 x 197 pixels, each with a size of $1\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 0.8 s at every point.

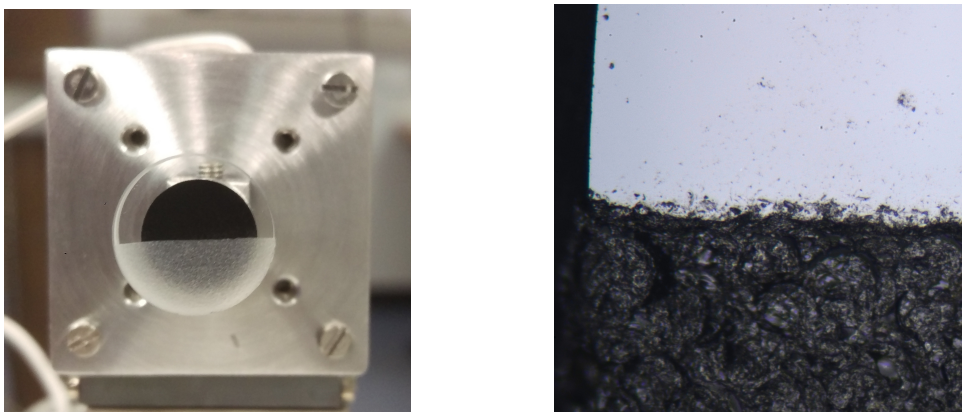
The observed contrast in fig. 5.11 arises from the difference in beam masking as the trench direction switches from vertical to horizontal. As previously discussed in section 5.2.3, when the detector aperture is not visible from the scattering point on the surface the beam is masked and a reduction in signal is observed. When the trenches are horizontal, the detector aperture is visible from the scattering point. But for the vertical trenches, the detector aperture will not be visible and the beam will be masked. Therefore, even though the width of the beam that is larger than trenches, parts of the beam will be blocked from entering the detector and a reduction in signal is observed when the trenches are vertical.

Figure 5.11 provides clear evidence that even if topography cannot be spatially resolved, it can still significantly contribute to the contrast that is observed. By using a series of well controlled trenches, it was possible to experimentally prove that structure below the width of the helium beam can significantly alter the observed contrast.

The roughness of the surface can also influence the observed contrast, since it will affect the topography of the surface below the beam width. If the structure of the surface is flat to an atomic level, then diffraction will occur and a separate contrast mechanism will be seen. However, any variation in roughness between the atomic scale and the beam width will also lead to contrast and will be very difficult to experimentally verify.

Single crystals that have been cleaned produce a strong specular peak, indicating that a very flat surface can scatter atoms specularly. For clean surfaces with some roughness, the specular peak will become broadened. As the surface roughness is continued to be increased from perfectly flat to rough, there is a transition from specular reflectance to diffuse scattering. Therefore, the general trend as roughness is increased is that the reflected atoms shift from the specular condition towards the normal to the surface.

Both optical and atomic scattering show that surfaces with a very large roughness produce an enhancement in the flux that is backscattered. Extremely rough surfaces will have significant self shadowing and masking, which will reduce the flux scattered into large angles away from the surface normal. The enhancement of the backscattered flux has been observed for both optical scattering [135] and as previously mentioned in atom scatter-



(a) (b)

Figure 5.12: Images of the optical flat after bead blasting. (a) Image of the optical flat after the bottom half of the surface has been bead blasted and attached to the sample mount. The bead blasting roughens the surface enough that the flat is no longer transparent. (b) An optical image of the edge of the optical flat.

ing [131]. The backscattered component has also been seen in recent theoretical treatments of scalar waves scattering from rough surfaces [136].

The importance of roughness on contrast was tested by taking an optical flat and bead blasting one half to increase the roughness. The optical flat is a very well polished glass surface that can be used as near perfect single surface optical mirrors. The flat that has been used is a $\lambda/10$ fused silica optical flat (#43-319-000 from Edmund Optics), with the $\lambda/10$ grade implying that the maximum peak to valley deviation of the surface is $\approx 63\,\text{nm}$. Figure 5.12 shows two optical images of the sample showing that one half is unprocessed and the other half has been roughened[1]. Bead blasting creates an extremely rough surface such that it is no longer transparent.

Bead blasting is a very destructive tool that introduces roughness to the sample on all length-scales. After bead blasting, the surface will become macroscopically rough, which will be spatially resolvable with the helium microscope. However, the roughening will also take place on length scales smaller than the beam width and therefore will not be spatially resolved. Figure 5.13 shows an example of a helium image where the surface is macroscopically varying in topography.

Since the roughness of the sample can be spatially resolved, a large area of the sample needs to be averaged over to simulate the effect of sub-beam roughness. Averaging over a large area produces a virtual beam that is much wider than the actual beam size. The virtual wide beam cannot spatially resolve the large scale changes in topography, therefore providing a way to test how roughness will influence contrast if it cannot be spatially resolved.

The histogram in fig. 5.13 shows that contrast is seen between the rough and flat parts of the sample. By looking at the distribution of values over a large area in the flat and smooth areas we can see that there is a difference in the mean scattered helium intensity. Since the roughened area also has macroscopic variation in the topography, the distribution of values for the rough area is wider than the flat area.

A large proportion of the measured signal is a background signal due to the diffuse he-

[1]I am grateful to S Lambrick for his contribution to the preparation of the sample

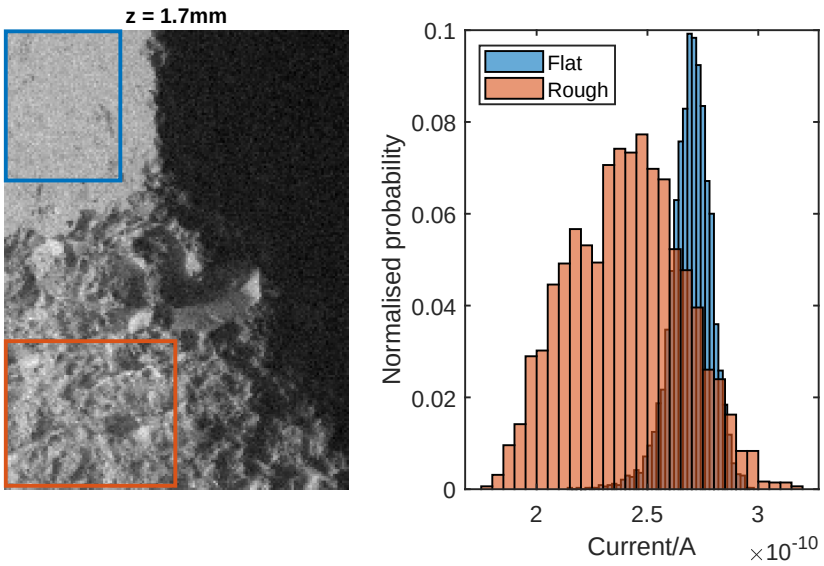


Figure 5.13: SHeM image of optical flat with a histogram for the flat and rough regions. The blue box is over the undistorted part of the optical flat, while the orange box is across the heavily distorted part of the surface. The bead blasting has introduced a large amount of roughness to the surface that can be spatially resolved by the helium beam. The histograms on the right show the distribution of the measured currents on each pixel. The distortion of the surface leads to a wider distribution of pixel values over a large area. However, in this example the flat region is clearly brighter on average than the rough region indicating contrast between the two areas of different roughness. The image consists of 150 x 200 pixels, each with a size of $2\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 1.125 s at every point.

lium beam coming through the pinhole. In order to investigate how the helium signal varies between the flat and rough areas of the sample, it is necessary to eliminate the background sources that will swamp the observed contrast. The optical flat sample has a diameter of 12.7 mm and therefore there is a significant contribution to the background from the diffuse beam exiting the pinhole. By taking the ratio of the background to the signal on the flat region, it is possible to predict what the background due to the diffuse beam is in any image.

Figure 5.14 shows the background corrected Michelson contrast between the rough and smooth surfaces as the scattering angle is altered. The average measured current was measured for both the rough and smooth parts of the surface, as previously shown in fig. 5.13, for a range of different sample to pinhole distances. The predicted background was subtracted from each of these values to produce the corrected current at each of these values. The Michelson contrast, M, is then calculated at each sample distance using,

$$M = \frac{I_f - I_r}{I_f + I_r}, \tag{5.5}$$

where I_f and I_r are the average currents on the flat and rough parts of the surface respectively. The fractional error on each contrast value was calculated by adding the fractional error of the sum and difference of the average currents together in quadrature. At small sample distances where the outgoing scattering angle for the helium is large, the flat region

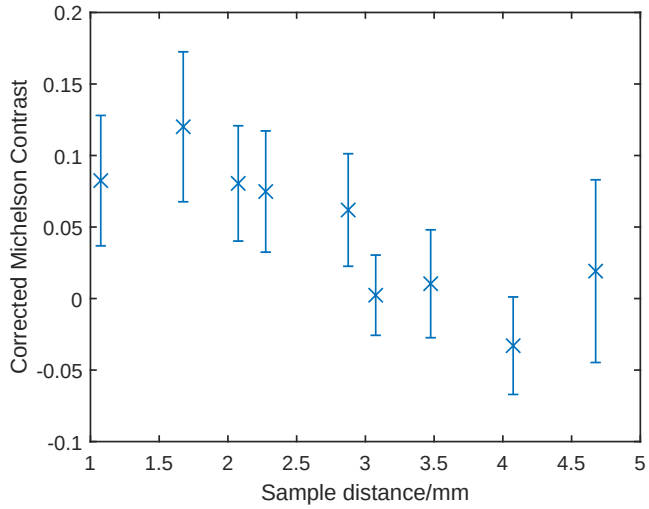


Figure 5.14: A plot of the background Michelson contrast between the flat and rough regions on the optical flat sample for a range of different sample distances. When the sample is close to the pinhole, the scattering angle for the atoms is large and there is a large effective contrast difference between the two regions. As the scattering angle is decreased (larger sample distances), the observed contrast also decreases. These observations are consistent with the idea that the rough region self shadows and masks at large scattering angle reducing the flux that can enter the detector.

of the sample scatters significantly more helium than the rough part of the sample. As the scattering angle moves towards the normal of the surface, the noise on the measurements becomes too large to make any significant statement about the contrast between the flat and rough areas.

The observation that there is a reduction in the helium that is scattered into large angles from an extremely rough surface is consistent with the literature. For very rough surfaces the surface will be significantly self shadowing and masking when the scattering angle is large, leading to a reduction in the measured intensity [137]. The self shadowing and masking are linked to the previously discussed enhancement in backscattered flux when the sample is very rough.

In conclusion, there is strong evidence that the topography of a surface below the beam width can significantly alter the effective scattering distribution. A set of trenches in a plastic sample is used as an initial demonstration of the effect of structure below the size of the beam. The roughness of a sample is then shown to change the flux that is scattered from a sample, indicating that the roughness can create a contrast mechanism.

5.2.8 Implications for previously observed chemical contrast

Barr et al. observed contrast between a silicon substrate and several metals that were deposited onto the surface [64]. The authors created a set of samples of thin layers of metal on a silicon (100) substrate with a native oxide layer. The metal layers were applied using electron beam lithography and a 3 nm wetting layer of titanium was also used. The surfaces have not been cleaned in vacuum and therefore it is likely there is a layer of adsorbates on top of the surface. The images obtained from the experiment were shown back in fig. 2.10.

The explicit observation of sub beam topography influencing the contrast has large im-

plications for the "chemical contrast" that has been previously reported. In section 5.2.7, it was established that contrast can be influenced by non-spatially resolved structure. Therefore, verification of contrast arising from a mechanism that is not topographic would require that the surfaces have precisely the same topographic properties down to the atomic level. Experimentally generating such a set of surfaces would be extremely challenging and is likely to only be achieved by using well defined single crystal surfaces.

Barr et al. tried to eliminate the possibility of the contrast arising from changes in the topography of the surface using AFM. The measured rms roughness values of the surfaces are within 2 nm of each other and do not follow the trend in helium contrast that was measured. However, without also measuring the coherence length of the roughness it is difficult to make any strong statements on whether the roughness of the surfaces is a significant contributor to the observed contrast.

It is therefore possible that the contrast observed by Barr et al. [64] between the different metal surfaces arises from the change in the topography of the surface below the beam resolution. The adsorbate layer on top of the surface means it is unlikely that the changes in the subsurface surface chemistry are creating significant contrast variations. Instead a much simpler explanation is that the topography is varying between the different samples and creates a contrast variation.

5.3 Microscopy with a mixed gas beam

Helium is often used in atom scattering due to a variety of different reasons. Helium beams formed in a supersonic expansion have an unusually high speed ratio and therefore are almost monochromatic [138]. Helium atoms have a large cross section for diffuse scattering from adatoms and defects [13] and therefore are useful in studying very low coverages of atoms. Finally, helium is usually detected through ionisation and the composition of the background gas in a vacuum leads to a low background signal at $m/z = 4$ making it easier to measure very low signals of helium. However, atom scattering is possible with other gas species including Ne, Ar, H_2 and D_2 [13]. Different species will scatter from the surface differently due to the changes in mass, de Broglie wavelength and available internal degrees of freedom in the case of molecules.

Witham and Sánchez have performed microscopy with a krypton beam [60], but did not observe any contrast that was significantly different from using a helium beam. The krypton beam is also formed in a supersonic expansion, meaning that the beam energy will be the same and the wavelength will differ by the mass ratio of krypton and helium. Given the two beams have very similar properties and the scattering from the surfaces is mostly diffuse, it is not surprising little contrast difference was observed between using krypton and helium.

However, it is possible to mix gas species in a supersonic expansion by seeding one expansion with another species [65]. If one species is only a small fraction of the total beam, then the velocity of the two atoms will be the same, giving the heavier atom an energy greater than its thermal energy. Using seeded beams can produce high energy atom beams without the need for changing the source temperature or other more complex methods.

By using a gas mixture and measuring the scattered flux of each gas from the sample, it is possible to acquire two images of a surface simultaneously. After the gas mixture is

scattered from the sample, either one detector can be used to sequentially mass filter each gas and measure the flux, or the gas can be split and two detectors can simultaneously measure the gas flux of each species. Normally the exit from the detector cone leads directly to the detector, however by installing a tee it is possible to split the gas from the detector cone into two separate detectors.

By having two species with significantly different wavelengths in the beam, it is theoretically possible to observe different scattering from surfaces. As discussed earlier in the chapter, the wavelength of the atom beam will determine the scattering distribution. If the surface is sufficiently well ordered for diffraction, a shift in the wavelength will cause the diffraction pattern to change. Alternatively, if the wavelength of one gas species were smaller than the surface roughness of a sample then diffuse scattering would be observed, but then the other gas species could have a wavelength larger than the roughness and specular scattering could occur.

Argon was chosen as the second gas to be used with helium due to its previous use in seeded beams and inertness. The second gas in the mixture needs to be inert to ensure that no damage occurs to the surface, which limited the selection of gases to the noble gases. Previously Witham and Sánchez [60] used a krypton beam and did not see any significant contrast when the gas was collected over a large solid angle. The previous lack of contrast observed with krypton motivated using a different noble gas. Argon has previously been used in seeded beams and it is easy and cheap to acquire, therefore is a prime candidate for use.

To acquire the images simultaneously, the original Hiden HAL/3F PIC mass spectrometer was used simultaneously with the new helium detector that was discussed in chapter 4. By using the two detectors simultaneously, there is no need to rapidly switch between two different masses on a single spectrometer and waste time waiting for each signal to measure. The downside to using two detectors is that the gas is split such that each detector measures a lower flux of each gas when compared to the total incoming flux.

Vacuum systems will contain a large proportion of the regular contaminants such as H_2O, CO, CO_2 and H_2 and these do appear in a large abundance when using the Hiden mass spectrometer. However, there is a family of peaks near $m/z = 40$ which has also been previously observed by Davies et al. [118] that is not simply due to argon leaking in from the atmosphere. The manufacturer of the mass spectrometer has designated the peak at $m/z = 40$ to be due to some organic compound that is in the vacuum chamber.

Figure 5.15 shows that there is a significant background at $m/z = 40$ when using the commercial mass spectrometer. While the peak at $m/z = 40$ may seem insignificant in the scale used in fig. 5.15, when the detector sensitivity is increased to measure the extremely small atom fluxes used when performing microscopy, the background will be extremely significant.

The significant background at $m/z = 40$ makes the experiment more challenging since spatially resolved atom scattering involves measuring extremely low fluxes of gas. Using a pure helium beam, the standard pinhole setup given earlier in section 3.1 and optimised settings on the Hiden mass spectrometer gives a helium signal of 1000 cps. Given that the mixed gas beam will be composed of 10% argon and 90% helium and the difference in gauge correction factors between argon and helium is ≈ 7, the expected argon signal will be 700 cps. Using the same settings on the Hiden mass spectrometer gives a background

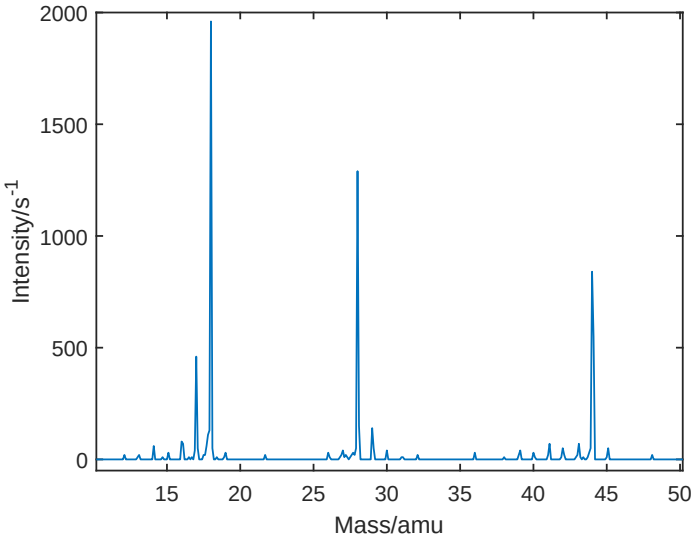


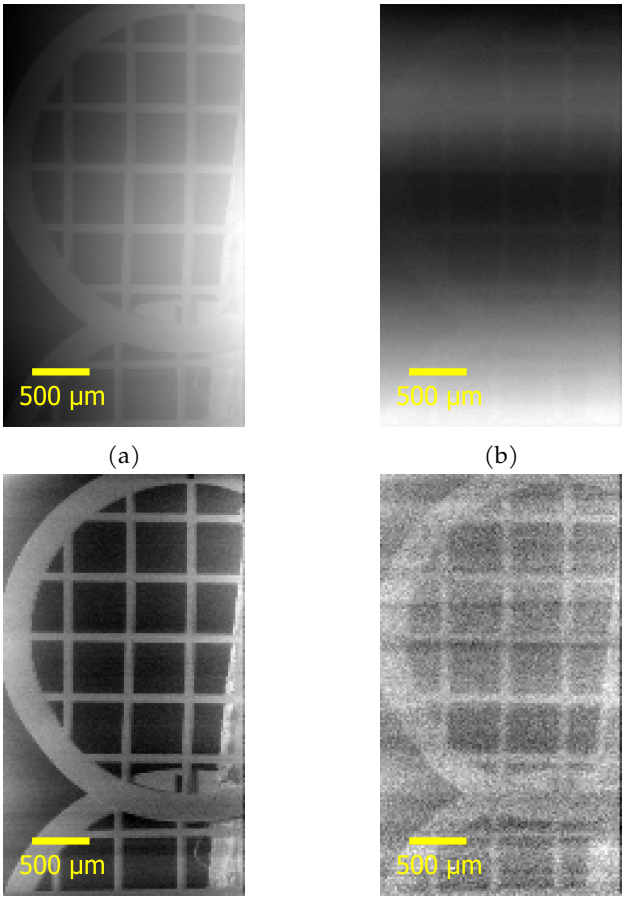
Figure 5.15: A plot showing a mass scan from a Hiden HAL/3F PIC mass spectrometer. The large peaks at 18 amu, 28 amu and 44 amu correspond to H_2O, CO and CO_2. The small cluster of peaks near 40 amu are thought to be caused by an organic compound in the vacuum.

signal at $m/z = 40$ of $\sim 10^6$. The background signal is far too large to measure the argon signal using the same setup that has been used when collecting helium images. However, the size of the atom beam incident on the sample can be increased since the resolution is not particularly important when testing the viability of using a mixed gas beam.

Given several improvements to the signal were required, the pinhole, skimmer and detector aperture were all increased in size to increase the flux of argon so that it can be measured above the background level. The pinhole diameter was increased in size from $1.2\,\mu\mathrm{m}$ to $20\,\mu\mathrm{m}$,such that the expected total gas flux would increase by a factor of ~ 270. The skimmer diameter was increased from $100\,\mu\mathrm{m}$ to $500\,\mu\mathrm{m}$ to ensure that the full expansion is being imaged by the pinhole. The detector aperture was also increased from a diameter of 1 mm to 2 mm to maximise the collected flux of argon, which would naively give a factor of four increase in the measured flux. However, the increase in size of the detector aperture also increases the pumping speed back into the scattering chamber from $S_s = 0.24\,\mathrm{Ls}^{-1}$ to $S_s' = 0.97\,\mathrm{Ls}^{-1}$. Since the pumping speed down the main detector is expected to be $S_D \sim 0.41\,\mathrm{Ls}^{-1}$, the actual increase in the predicted collected gas flux in the main detector is calculated using equation C.2 to only be 1.9 times higher than when using the smaller detector aperture. The gas ratio was also changed to 20% Ar to 80% He to further increase the argon flux.

A major disadvantage to increasing the size of the skimmer is that the diffuse background gas that exits the pinhole is increased. Since the skimmer is larger, the flux of gas stagnating just behind the pinhole is also increased leading to a larger flux of gas that will diffuse the pinhole. As discussed in section 5.2.4, the gas will form a very wide beam that will reduce the observed contrast in the image from sharp features.

The final beam composition did not match the expected ratio of helium to argon, which could have been caused by the gas ratio being altered during the expansion. After making the changes to the instrument and detuning the detector to prevent damage to the electron multiplier, the helium signal was $\approx 5000\mathrm{s}^{-1}$ on a negligible background. The argon signal



(a) (b) (c) (d)

Figure 5.16: Raw and processed images that were acquired using a mixed gas beam of helium and argon. The top row of images (a,b) are the raw images, and the bottom row (c,d) are the processed images. The left hand images (a,c) were obtained using helium, the right hand images (b,d) were obtained using argon. The significant background present in the helium image (a) is from the large diffuse beam due to using a larger skimmer than usual, and the background in the argon image (c) is from both the large diffuse beam and the large background present at $m/z = 40$. The background was removed by fitting a two dimensional polynomial to the images which is first order in the horizontal direction and eleventh order in the vertical direction. The images consist of 100 x 175 pixels, each with a size of $20\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 4.5 s at every point.

was $\approx 15500\mathrm{s}^{-1}$ on a background of $\approx 27500\mathrm{s}^{-1}$. The argon signal is much higher than expected and conversely the helium signal is much lower than the expected change after altering the pinhole plate. One explanation for the unusual gas ratio is that the heavier argon atoms pushed the helium to the outside of the expansion, leading to a higher flux of argon through the pinhole [65, 139].

Figure 5.16 shows raw and processed mixed gas images collected of two TEM grids. Panels (a) and (c) show helium images, while panels (b) and (c) show argon images. The raw images on the top row are difficult to interpret directly due to the large backgrounds present. Specifically, the helium image contains a large diffuse component making it difficult to see sharp features and the argon image has a large temperature fluctuation in the background that is present across the whole image. The images on the bottom row were

processed by removing a polynomial background that is first order in the horizontal direction and eleventh order in the vertical direction.

Figure 5.16 shows the first dual gas images to be acquired and illustrate that is possible to use a mixed gas beam. The largest challenge facing the acquisition of images with multiple gases at once is finding another species that is not only inert, but also does not share a mass with any vacuum contaminants and can be used in a supersonic expansion. The background at the mass of the probing species needs to be extremely low to be used in microscopy, so while argon initially seems like a good candidate it is unlikely to be easy to use.

Looking forwards, using dual gas beams to observe small features on a surface would be difficult. However, it is entirely feasible to use a dual gas beam if there is no need to investigate small features and there would be interest in using atoms to investigate the surface. For example, in the next section diffraction is discussed and having two beams with differing wavelengths and energies would be useful.

5.4 Contrast from ordered surfaces

It is well known that helium beams can diffract from well ordered surfaces due to the wave nature of atoms. Estermann and Stern [140] originally used the diffraction of helium from the surface of lithium fluoride as evidence for the wave nature of matter. The use of helium scattering for structure determination has significantly progressed since 1930 and can now be used to investigate the complex structure of adsorbates on surfaces [14].

Diffraction from a surface would significantly change the scattering distribution and therefore create strong contrast in images. Being able to use diffraction as a contrast mechanism would open up a large range of possibilities, allowing the possibility to be sensitive to changes to a surface on the atomic scale. Additionally, diffraction peaks are typically very sharp, potentially leading to strong contrast in images. In this section, the aim is to demonstrate that it is possible to observe diffraction contrast using a helium microscope.

The difficulty in atom scattering experiments is forming clean well-ordered surfaces so that diffraction occurs. The helium beam is usually only sensitive to the outer electrons of a surface since the helium beam has a very low energy and the classical turning point of the helium atoms is ~ 2 Å above the surface atom cores [13]. Without careful preparation, most materials will possess a layer of contamination above the surface that will hide the periodic structure of the substrate. Typically in atom scattering experiments, large amounts of effort is put into cleaning the sample to achieve order over the full beam size and obtain diffraction patterns.

Currently the microscope does not have a facility to perform in-situ cleaning of the sample for atom scattering. It is therefore difficult to study diffraction of atoms from the surface without making upgrades to the hardware to allow some surface preparation. However, there are a few materials that can be studied without cleaning to allow initial investigations into how spatially resolved atom scattering could be used to observe surface diffraction.

Lithium fluoride that has been cleaved in air has been known to form atomically clean surfaces without any surface preparation in vacuum. LiF therefore provides an ideal sample to test how scattering from an ordered surface will appear without making any costly upgrades to the machine. After the crystal was cleaved, the sample was rapidly placed into the vacuum chamber and pumped down to minimise contamination.

Following the procedure outlined earlier in the chapter, the scattering distribution can be measured directly in the microscope by performing a z scan. However, the relatively large solid angle of scattered helium is collected for detection varies throughout a scan, meaning that it will be difficult to directly interpret the results of a z scan if it showed signs of diffraction.

It is possible to perform a simple calculation to predict the characteristics of a z scan from a surface that is diffracting. There have been many experiments looking at the diffraction pattern obtained from scattering helium from LiF [141, 142], but there do not exist any experimentally measured two dimensional diffraction patterns with the exact conditions needed. While it is possible to predict the full diffraction pattern using computational methods, the peak locations can be identified using less complex methods.

The aim of the simple calculation performed here is to simply help identify how diffraction contrast would appear in a z scan, not to accurately calculate the signal that would be expected. The difficulty in not fully simulating the system is that the peak heights will not be correct, but the locations of the peaks will be correct which will help identify features in the data.

Firstly, the locations of the diffraction peaks need to be determined for helium scattering from LiF with a 45° incoming angle. Lithium fluoride has an fcc lattice, and the fluoride ions in the (100) plane have a simple cubic structure with the spacing given by $a/\sqrt{2} = 2.84\,\text{Å}$ [142]. Assuming that the corrugation in the electron density is dominated by the positions of the fluoride ions, the diffraction peaks are then found at,

$$\mathbf{K_f} = \mathbf{K_i} + \mathbf{G}, \tag{5.6}$$

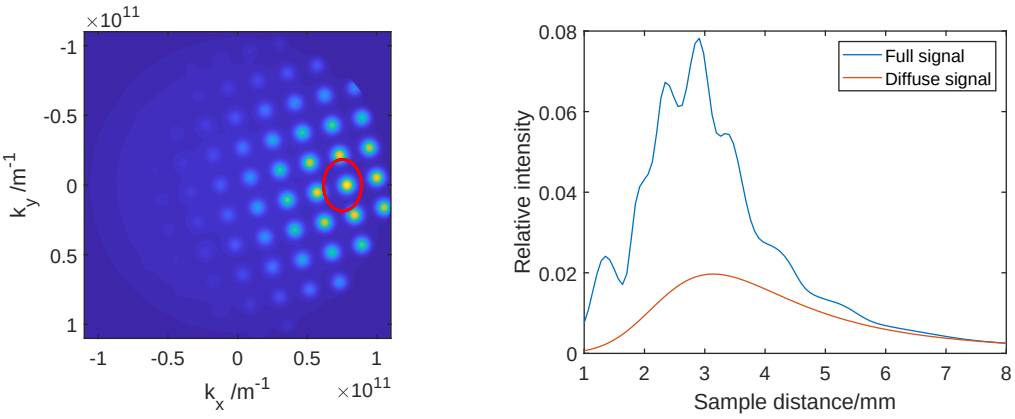
where $\mathbf{K_f}, \mathbf{K_i}$ are the two dimensional final and initial wavevectors parallel to the surface and $\mathbf{G}$ is a two dimensional reciprocal lattice vector for the surface lattice.

An approximation to the diffraction pattern can be produced from the peak locations by making a series of assumptions about the heights and widths of the peaks to match experimental data. Figure 5.17a shows a scattering distribution that is generated by placing Gaussian peaks at each of the diffraction peak locations with a standard deviation of $3.5 \times 10^9\,\text{m}^{-1}$. The standard deviation was chosen to approximately match the widths of the peaks as measured by Boato et al. [141]. The maximum height of each peak also follows a Gaussian distribution from the specular peak so that the peaks are less intense further from specular, with the peaks decaying such that the second order peaks are $\exp(1/\sqrt{2})$ lower than the specular. Finally, a diffuse cosine distribution is added to include the effects scattering that is not due to the diffraction. The scattered helium has half of the flux in the diffuse component and half in the diffracted intensity.

The diffraction pattern can then be integrated over the directions that would enter the detection aperture to compute the intensity of helium that would be measured in the detector. The scattering distribution is projected onto the pinhole plate using,

$$x = \left(\frac{k_x}{k_z} + 1\right) z_p, \tag{5.7}$$

$$y = \left(\frac{k_y}{k_z}\right) z_p. \tag{5.8}$$



(a) (b)

Figure 5.17: (a) Simulated diffraction pattern of helium beam at 298 K with an incoming angle of 45° and the cubic lattice rotated anticlockwise by 14° to match the experimental setup. The peaks are reduced in height from the specular peak to simulate how the heights in a real diffraction pattern would act. The red circle indicates the outgoing wavevectors needed to enter the detector aperture under the specular condition. The reasonably large detector aperture leads to the possibility of multiple peaks entering the detector at once. (b) A simulated z scan on LiF using the diffraction pattern given in fig. 5.17a. The peaks in the z scan correspond to different diffraction orders entering the detector aperture. Without a rigorous diffraction calculation, the peak heights are not correct and therefore the shape of the curve is the most important point. The qualitative behaviour matches the experimental results, indicating that the likely origin of peaks in the z scan arise due to diffraction from the sample.

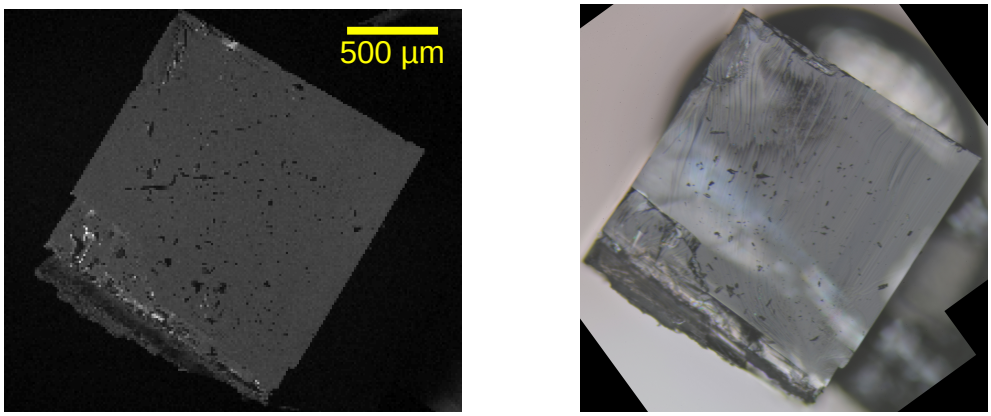
The additional term in the x position arises due to the horizontal shift that is needed in a z scan. Once the distribution has been projected onto the pinhole plate, the intensity distribution is integrated over the values of k that would enter the detection aperture to estimate the signal that is measured. The code to implement the calculation is provided on GitHub and Zenodo[2]

Figure 5.17b is the result of the simulated z scan, showing peaks that appear when each diffraction peak moves through the detector aperture. The diffraction pattern that was generated is symmetric and it is only the variation of the solid angle of the detector aperture that breaks the symmetry of the signal around specular. Experimental results usually show a large asymmetry with the first order peak even being larger than the specular in some cases [141]. Therefore the heights of the peaks are not accurate, but it is only the qualitative shape of the curve that is important.

The simple calculation detailed above has shown that diffraction will appear as a series of peaks in a z scan. Previous discussions on z scans for diffuse scatters showed that they do not have any sharp features in them. Therefore, the appearance of any sharp features in a z scan must be due to the scattering distribution having sharp features which is likely to be due to diffraction.

Now that it has been established how diffraction would appear, the experimental data can be examined. Figure 5.18a is a helium image of the entire LiF crystal showing that the surface is relatively uniform with only a few defects. The cleave was not perfect and the bottom of the crystal did not form a flat surface. Figure 5.18b is an optical image taken after

[2]M. Bergin, Zenodo, `https://doi.org/10.5281/zenodo.2528969`



(a) (b)

Figure 5.18: Images of the entire LiF crystal. (a) SHeM image of the LiF crystal showing that it is mostly homogenous with only a few defects as expected. The bottom of the crystal contains many defects due to the cleaving not being perfect. The image consists of 208 x 183 pixels, each with a size of 12 µm, that were measured with an initial pause of 0.5 s and a dwell time of 1.125 s at every point. (b) Optical image of the crystal further illustrating that the bottom of the crystal is not flat.

the crystal was removed from the vacuum chamber, it is clear to see that the bottom of the crystal has not perfectly cleaved and will not be atomically flat. The orientation of the crystal with respect to the helium beam was determined using fig. 5.18a.

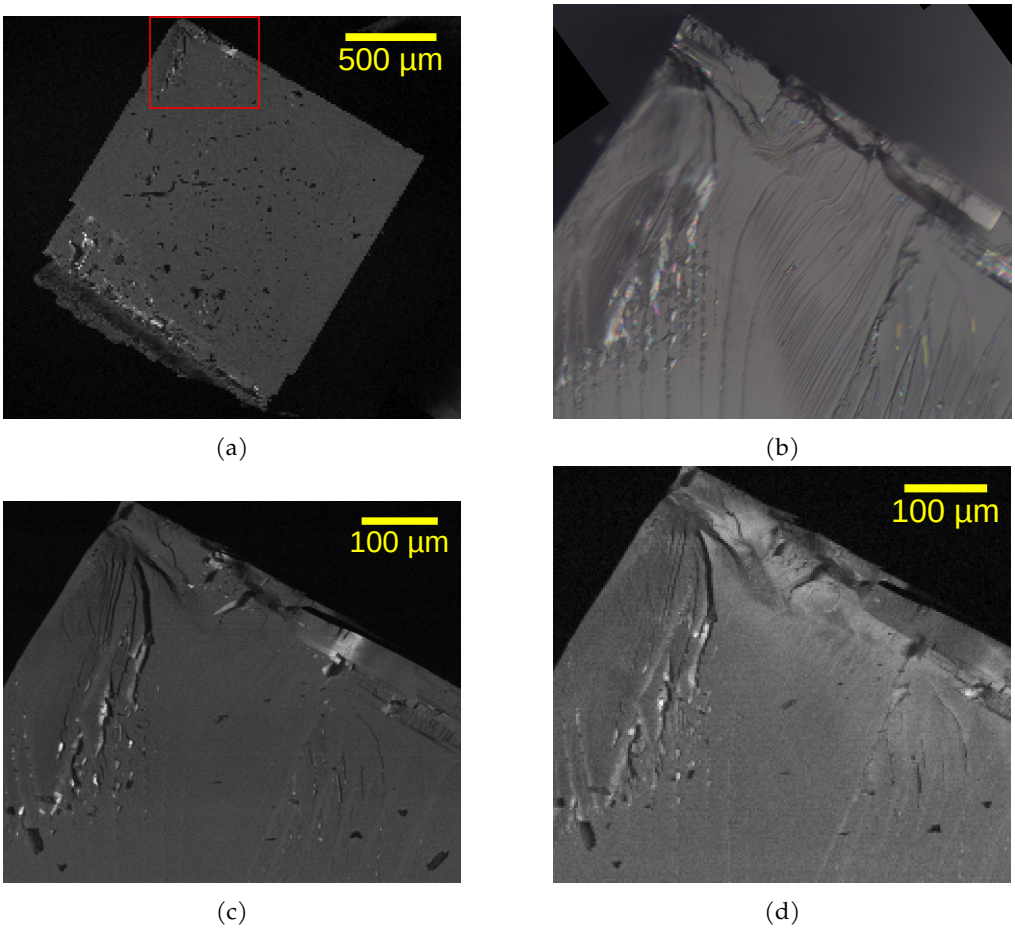
Figure 5.19 shows a series of images taken at the top of the crystal. Figure 5.19a is the same as fig. 5.18a with a box added to indicate which part of the crystal we are looking at in the other figures. Figure 5.19b is an optical image of the area to show what the area looks like using light, it can be seen with light that the surface is not entirely flat.

Figures 5.19c and 5.19d are SHeM images of the top of the crystal taken at 2.5 mm and 3.3 mm from the pinhole. Parts of the surface that are not flat create contrast due to masking in the same way as it appears from diffuse scattering. However, there are very bright features on the crystal that change depending on the observation angle. For example, in panel (c) there is a bright band on the top right of the image that is no longer present in panel (d). In fact, acquiring many images and rendering a video indicates that the bright feature slowly changes as the scattering angle is varied, panels (c) and (d) represent just two frames from the video.

The bright features seen in the images only appear in narrow bands of the outgoing scattering angle, implying that the scattering distribution has sharp features in it. Therefore, the observed bright features are consistent with helium diffraction from the surface. The images can be used to determine changes in the slope of the surface, due to the surface normal change leading to diffraction peaks not being collected by the detector.

To confirm that the dark areas, particularly on the left of the images, are due to masking, a white light interferometer was used to measure the height variation of the sample and results are plotted in fig. 5.20. White light interferometers bounce white light off the sample and a mirror to measure the interference pattern between the two paths, by then varying the height of the sample, it is then possible to measure the topography of the surface from how the interference pattern changes [3]. The surface on the left can be seen to be higher than

[3] I am very grateful to Peter Atkin for performing the measurements on the white light interferometer.



(a) (b) (c) (d)

Figure 5.19: Images of the top of the LiF crystal. (a) SHeM image of the entire crystal with a red box illustrating where the other SHeM images in the figure were collected. (b) Optical image of the top of the crystal to compare with the SHeM images. (c) SHeM image of the bottom of the crystal obtained with a distance between the pinhole and sample of 2.5 mm. The image consists of 400 x 333 pixels, each with a size of $1.5\ \mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 1.125 s at every point. (d) SHeM image of the bottom of the crystal obtained with a distance between the pinhole and sample of 3.3 mm. The bright features seen in the images shift as the outgoing scattering angle into the detector is varied. The image consists of 220 x 200 pixels, each with a size of $2.5\ \mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 1.125 s at every point.

the rest of the surface confirming that masking is occurring.

Figure 5.21 is an array of images collected from the bottom of the crystal where there are a large number of defects. Figure 5.19a is the same as fig. 5.18a with a box added to indicate where the other images have been obtained. Figures 5.19c and 5.19d are helium images of the bottom of the crystal taken at different distances between the sample and the pinhole. Again, very bright features in the images can be removed by varying the scattering angle, indicating that there is a large amount of structure in the scattering distribution. For example, in the top left of panel (c) a small flake appears extremely brightly at one working distance, but then disappears in the panel (d). The large change in signal must arise from the scattering distribution having a sharp feature in it, which is likely to arise due to diffraction from the surface.

Figure 5.21b is an image from the same area obtained using a while light interferome-

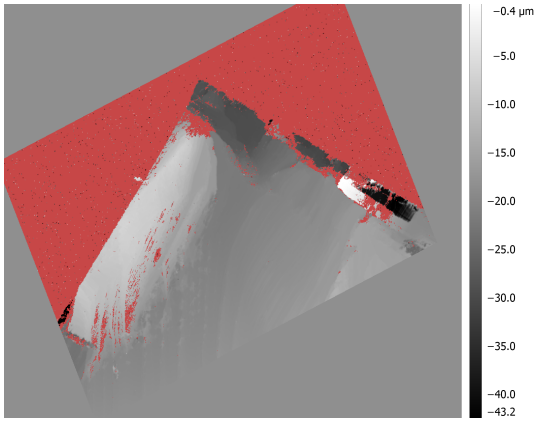


Figure 5.20: White light interferometry image of the top left corner of the crystal showing that the left hand side of the image is higher than the rest of the crystal. The noisy part of the image on the right is likely due to part of the crystal flaking away from the bulk and forming two surfaces.

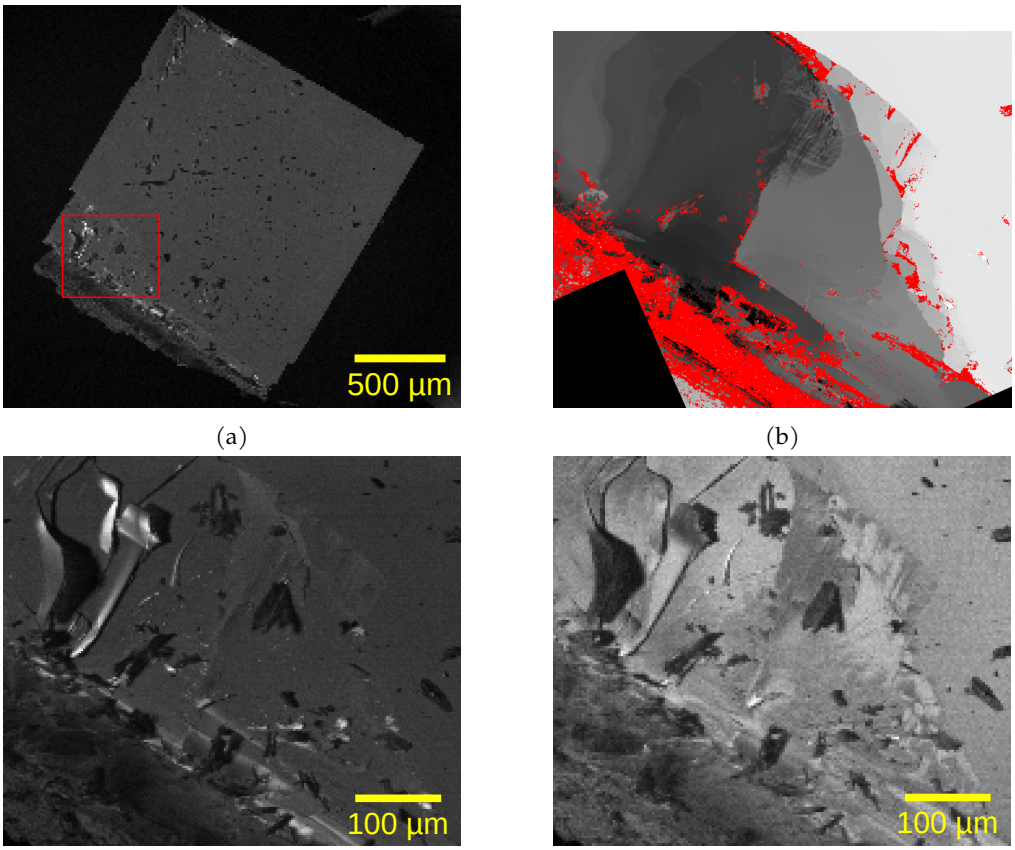
ter to measure the surface topography confirming that the top of the imaging area is flat but the bottom is heavily distorted. The red areas are sufficiently distorted that a measurement of the height could not be made. The top of the image is seen to be extremely flat, corresponding the area which is most likely to produce diffraction.

Figure 5.22 shows how z scans over different areas of the surface produce different shapes depending on the nature of the surface. Figure 5.22a illustrates three areas of the surface that were selected for z scans, the blue circle is over a heavily distorted region of the surface, the purple circle is over a lightly distorted region and the red circle is over a flat region. The resulting z scans are then plotted in fig. 5.22b, with the heavily distorted region in blue smoothly varying and following the expected lineshape for a rough surface, but the other two curves contain significant structure.

The z scan for the flat region where diffraction is most likely (red circle) is replotted separately in fig. 5.23 to illustrate the peaks that are observed. As discussed previously in section 5.2.2 a typical z scan from a rough surface arises due to a cosine scattering distribution and the variation of the solid angle of the detector aperture. The line shape of the z scan qualitatively has the same behaviour as the z scan obtained from the simple calculation performed earlier in fig. 5.17b. Both figures show peaks that appear due to the diffraction peaks moving across the detector aperture.

Comparison with the simulations therefore provides evidence for diffraction contrast in the microscope and a step towards a machine capable of spatially resolved diffractometry measurements. The detector aperture is reasonably large to ensure that the maximum signal is obtained for normal diffuse samples. However for ordered surfaces that diffract, a smaller detector aperture would allow the resolution of individual diffraction peaks making it easier to investigate the surface properties. Additionally, once the peaks are individually resolvable a rotation stage is likely to be needed to ensure that the scans are performed directly down the principal directions of the surface lattice.

Another system that is known to form atomically clean surfaces with minimal surface preparation is highly ordered pyrolytic graphite (HOPG). The same procedure was repeated with HOPG, however no structure was seen in the pattern. The surface may not



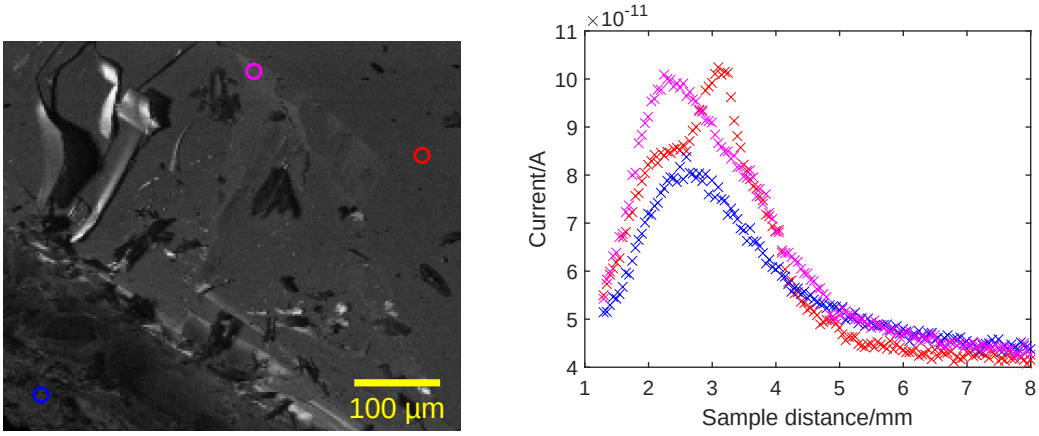
(a) (b) (c) (d)

Figure 5.21: Images of the LiF crystal. (a) SHeM image of the entire crystal with a red box illustrating where the other SHeM images in the figure were collected. (b) Digitally stitched image of the bottom of the crystal obtained using a white light interferometer. The colour bar ranges from $-60\,\mu\mathrm{m}$ to $0\,\mu\mathrm{m}$ with the red areas being places were no height could be determined which is likely due to the point not being sufficiently flat. The top of the image is extremely flat as expected and the bottom seems to be too rough to be measured by the interferometer. (c) SHeM image of the bottom of the crystal obtained with a distance between the pinhole and sample of 2.3 mm. The image consists of 212 x 180 pixels, each with a size of $2.5\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 1.125 s at every point. (d) SHeM image of the bottom of the crystal obtained with a distance between the pinhole and sample of 3.1 mm. The image consists of 176 x 150 pixels, each with a size of $3\,\mu\mathrm{m}$, that were measured with an initial pause of 0.5 s and a dwell time of 0.75 s at every point.

have shown evidence of diffraction either because it required some crude level of cleaning, or the grain size was too small and several domains were being averaged over.

The initial work on LiF opens the questions of whether heating other samples will have a large effect, since many of the adsorbed contaminants could be removed. There are a great deal more substrates that can diffract helium and the addition of some sample preparation will allow these other samples to be studied using a spatially resolved diffractometer that is sensitive to atomic structure.

It could be possible to investigate polycrystalline materials and determine the different orientations of the domains using diffraction. If the diffraction pattern could be determined inside the helium microscope, then the process could be repeated at different positions to



(a) (b)

Figure 5.22: A series of z scans collected from the LiF surface. a) The SHeM image of the LiF surface from fig. 5.21c indicating the positions z scans were collected. The blue circle is around the extremely rough area of the surface, and the red circle is on a flat part of the surface. The pink circle is around a slightly distorted part of the surface. b) A series of z scans from the LiF surface. The red scan from the flat part of the surface shows there is a large degree of structure in the scattering distribution. The other scattering distribution are much closer to the z scans seen from general rough surfaces, with little structure in the scans.

map out the domain structure of a material. Obtaining a diffraction pattern would then definitely require the upgrades mentioned above including, sample preparation to remove contamination, a rotation stage and a smaller detector aperture.

More generally, flat surfaces create scattering distributions that peak around the specular, however heavily distorted surfaces will create distributions that follow Knudsen's law and peak around the normal to the surface. There should be some transition between these two regimes when the surface roughness is at some intermediate value. Explicitly showing the transition would demonstrate high sensitivity to the surface structure on lengthscales much smaller than the beam. However to achieve the clear transition in roughness, a surface would need to be created that is flat on macroscopic and mesoscopic lengthscales, but then controllably rough on nanoscopic lengthscales.

5.5 Conclusions

Scanning helium microscopy has been shown to be a viable technique to measure the topography of a surface, irrespective of its conductivity. It has been shown that when scattering thermal energy helium from a surface, most materials exhibit a wide cosine-like scattering distribution due to the extreme roughness of most surfaces when compared to the wavelength of helium. Since the helium has a preferred scattering direction, contrast was observed in helium images of rough samples when the surface normal changes.

It was also shown that a significant part of the observed contrast on a general surface comes from the scattering geometry and local topography. Masking of the beam can occur when the detector aperture is not directly visible from the scattering point, leading to a reduction in the observed signal. Due to the wide scattering distribution, the helium scattered

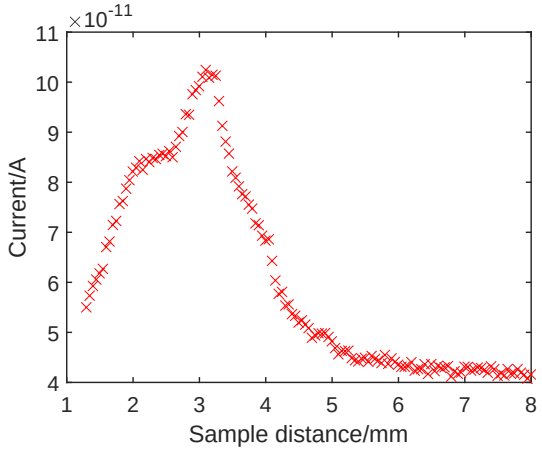


Figure 5.23: The experimental z scan of a flat region of the crystal from fig. 5.22b. There are many qualitative similarities between the simulated and experimental z scans indicating that diffraction is likely the source of the structure in the z scan.

from the surface can diffusely illuminate other parts of the surface leading to an increase in the observed signal.

Topography of the surface that is not directly spatially resolvable was shown to still lead to contrast in images. Arrays of trenches in plastic were used to directly prove that structures smaller than the beam width can lead to strong contrast in images. The surface roughness was also shown to be an important factor in the contrast that is observed in images.

It has been shown for the first time that it is possible to image a surface with two gases simultaneously using a mixed gas beam. The experiment was performed using helium and argon, however the background at $m/z = 40$ made it difficult to obtain high quality images using argon.

Diffraction contrast in a helium microscope was explicitly shown to be possible for the first time in a helium microscope using a LiF crystal. Data from LiF shows significant structure in the scattering distribution which is highly likely to be due to diffraction. Further development of the microscope could lead to a machine capable of measuring diffraction patterns at different points of a surface, giving spatially resolved structure determination.

Chapter 6

Constrained optimisation of the design of a scanning helium microscope

6.1 Introduction

In all current implementations of helium microscopy, the available signal level is a significant limitation due to the nature of the source and difficulty in ionisation and detection of helium atoms. Limited signal affects the signal to noise in images, and since the signal varies strongly with microprobe size, the ultimate usable resolution. Determining an optimised microscope geometry, which maximises signal levels while keeping the beam size to a minimum, is required to produce a design for a realistic helium microscope with an improved resolution.

Figure 6.1 shows a schematic view of the main atom-optical elements in a normal incidence scanning helium microscope. The beam of atoms is formed through the supersonic expansion of helium into vacuum. The beam is collimated initially by a skimmer (typically 0.1-0.5 mm), then collimated or focussed by the 'main' optical element, which for this chapter shall be considered to be either a pinhole or zone-plate. The sample is then rastered in front of the helium microprobe, and atoms scattered in a particular direction are transferred to a detector, which is usually a high-sensitivity mass-spectrometer.

Previous attempts to optimise the design of a SHeM typically treat the source and atom optics as separate problems that need solving. Palau et al. [143, 144] use Sikora's model [145, 65] for the centerline intensity downstream from the source. Optimal parameters for a zone plate and pinhole microscope are found by computing the intensity for a set of atom optics variables and finding the maximum. Alternatively, Kaltenbacher has used a multi-objective optimisation approach to design a complex atom optics setup involving two zone plates and a pinhole [146]. Kaltenbacher has used large approximations for the intensity and size of the beam, and using a multi-objective optimisation leads to Pareto-optimal solutions rather than direct solutions for a specific resolution.

However, a combined optimisation of both the source and atom optics is essential when considering a zone plate due to the speed ratio of the beam affecting both the source bright-

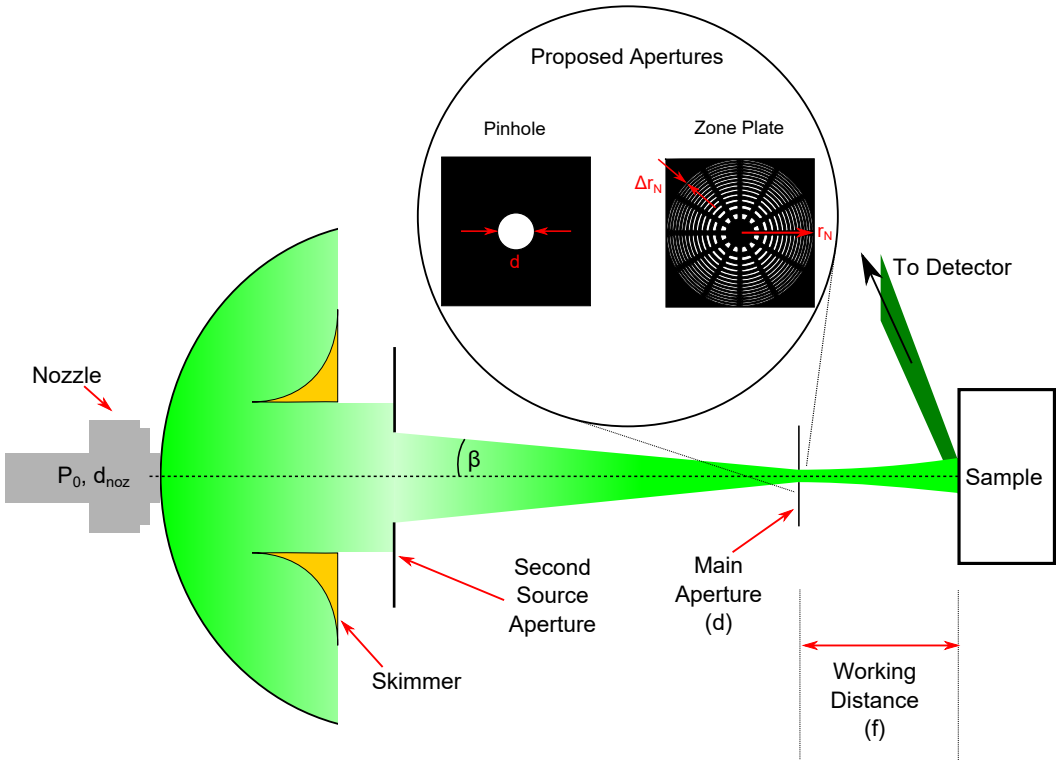


Figure 6.1: Schematic of a normal incidence scanning helium microscope. The beam is formed in a free jet nozzle expansion (left) and is initially collimated by a skimmer (centre-left). The beam is further collimated by an (optional) second aperture, in conjunction with suitable differential pumping. The beam passes into the sample region through either a pinhole or a zone plate (centre-right) and scatters from the sample (right). Atoms scattered into a particular solid-angle are collected and counted.

ness and the chromatic aberrations. Later in the chapter, I will show that increasing the brightness of the source and decreasing the helium beam size are directly opposing objectives. Therefore, the only way to find an optimal design for a machine is to use a model that includes both the source and atom optics.

In this chapter, I will outline a method to perform a simultaneous constrained optimisation of both the source and the atom optics, for a particular target resolution. After assuming a model for the intensity obtained from the source, the method produces a set of design criteria for future instruments aiming to achieve a particular resolution. Section 6.2 describes a diffraction calculation that is used to model the performance of a zone plate. Section 6.3 introduces the key theory needed to develop a model for the beam size and flux, in a machine with a normal incidence beam formed using either a pinhole or zone plate. Section 6.4 explains the analytic and numerical constrained optimisation procedure that is used. Section 6.5 presents key results from applying the constrained optimisation to the models for the beam. Section 6.6 provides a summary and conclusions that can be used as design criteria for future instruments.

6.2 Diffraction model

In order to model the performance of a zone plate, a diffraction calculation is needed to calculate the intensity obtained at the focal point. To begin with the integral that calculates the diffraction pattern from an arbitrary aperture is derived. The integral cannot be solved analytically, so it is then converted for use in a discrete numerical calculation. The calculation is then used to model the efficiency and chromatic aberration of a realistic zone plate that could be used in the microscope.

6.2.1 Defining the Integral

The derivation of the diffraction integral that is presented closely follows the treatment by Goodman [147]. It is assumed that the helium beam can be described with a scalar potential, ψ, that obeys the wave equation,

$$\nabla^2\psi - \frac{1}{\nu^2}\frac{\partial^2\psi}{\partial t^2} = 0. \tag{6.1}$$

If it also assumed that the problem is time independent, then the solution of the wave equation can be written as

$$\psi(\vec{r},t) = U(\vec{r})e^{-i\omega t}, \tag{6.2}$$

meaning that if the wavenumber is given by $k = \omega/\nu$, the wave equation can be recast into the Helmholtz equation,

$$(\nabla^2 + k^2)U = 0. \tag{6.3}$$

In order to proceed further, Green's Theorem needs to be used in the form,

$$\int\int\int_V (U\nabla^2 G - G\nabla^2 U)dv = \int\int_S (U\frac{\partial G}{\partial n} - G\frac{\partial U}{\partial n})ds. \tag{6.4}$$

The function G is an arbitrary auxiliary function, however only a good choice of G will lead to solving the diffraction problem. The function G is set as the Green's function for the problem so that it satisfies,

$$(\nabla^2 + k^2)G = \delta(\vec{r} - \vec{r_p}) \tag{6.5}$$

However, since G is discontinuous at $\vec{r_p}$ some care must be taken. Let the total volume to be integrated over be broken into two parts, V_ϵ and V' where V_ϵ is a sphere of radius ϵ that surrounds the discontinuity and V' is the rest of the volume to integrate over.

Substitution of equations 6.3 and 6.5 into Green's Theorem and performing the integral over V' gives us

$$\int\int\int_V{}' (U(-k^2G) - G(-k^2U))dv = \int\int_{S'} (U\frac{\partial G}{\partial n} - G\frac{\partial U}{\partial n})ds. \tag{6.6}$$

$$0 = \int\int_{S'} (U\frac{\partial G}{\partial n} - G\frac{\partial U}{\partial n})ds. \tag{6.7}$$

Where S' is the surface of the volume V' and is made from two parts, the outside surface around the volume enclosing the discontinuity S_ϵ and the outside surface of the whole volume S so that $S' = S + S_\epsilon$. Note that the outward normal from S_ϵ is actual towards $\vec{r_p}$.

Splitting S' into the two surfaces it is made from gives,

$$-\int\int_{S_\epsilon}(U\frac{\partial G}{\partial n}-G\frac{\partial U}{\partial n})ds.=\int\int_{S}(U\frac{\partial G}{\partial n}-G\frac{\partial U}{\partial n})ds. \tag{6.8}$$

The left hand side can be solved by substituting $G=\exp(ik\epsilon)/4\pi\epsilon$ and taking the limit as ϵ goes to zero. Note that the green's function will always take this form when very close to the discontinuity, whatever the boundary conditions. Performing the integral gives the key equation,

$$U(\vec{r_p})=\int\int_{S_1}(U\frac{\partial G}{\partial n}-G\frac{\partial U}{\partial n})ds \tag{6.9}$$

Where S_1 is the surface just behind the diffracting screen, the other surface S_2 has been ignored since the wave must obey the Sommerfeld radiation condition [148]. To continue two assumptions must be made, the first is that at the aperture, U and $\frac{\partial U}{\partial n}$ are exactly as they would be without the aperture screen being present. The second assumption is that in the plane immediately behind the screen, the integrand is identically zero in the geometric shadow of the aperture. With these assumptions equation 6.9 can then be written as,

$$U(\vec{r_p})=\int\int_{\Sigma}(U\frac{\partial G}{\partial n}-G\frac{\partial U}{\partial n})ds. \tag{6.10}$$

Where the integral is now over the plane in the aperture which is open Σ, and U is the amplitude of the scalar field due to the source with no disturbance caused by the aperture.

Kirchoff imposed the second condition described above by setting both U and $\partial U/\partial n$ to zero. However, setting both the value of the field and its gradient to zero will lead to the field being zero at all positions [148]. Kirchoff's boundary conditions therefore lead to mathematical inconsistencies, so instead a Green's function which has either G or $\partial G/\partial n$ equal to zero shall be used. For reasons to be discussed later, the Green's function that is used is chosen to be identically zero in the aperture plane.

Figure 6.2 illustrates the geometry of the diffraction problem and the key parameters to be used. Using the method of images gives the function,

$$G_-=\frac{\exp(ikr)}{4\pi r}-\frac{\exp(ik\tilde{r})}{4\pi\tilde{r}}, \tag{6.11}$$

where $r^2=(x-x_p)^2+(y-y_p)^2+(z-L)^2$, $\tilde{r}^2=(x-\tilde{x}_p)^2+(y-\tilde{y}_p)^2+(z+L)^2$, (x_p,y_p) is the co-ordinate of the point the amplitude is to be evaluated at and $(\tilde{x}_p,\tilde{y}_p)$ is the point the image source is placed. By setting $\partial U/\partial n$ to be zero directly behind the aperture, the integral now becomes,

$$U_I(\vec{r_p})=-\int\int_{\Sigma}U\frac{\partial G_-}{\partial n}ds. \tag{6.12}$$

The final remaining step is to then evaluate $\partial G_-/\partial n$ at the aperture plane. Taking G_- from earlier gives,

$$\frac{\partial G_-}{\partial n}=\frac{\partial G_-}{\partial r}\frac{\partial r}{\partial n}+\frac{\partial G_-}{\partial\tilde{r}}\frac{\partial\tilde{r}}{\partial n}, \tag{6.13}$$

$$\left.\frac{\partial r}{\partial n}\right|_{z=0}=\frac{1}{2}\frac{2L}{r}=\cos(\theta_p), \tag{6.14}$$

$$\frac{\partial G_-}{\partial r}=(ik-\frac{1}{r})\frac{\exp(ikr)}{4\pi r}, \tag{6.15}$$

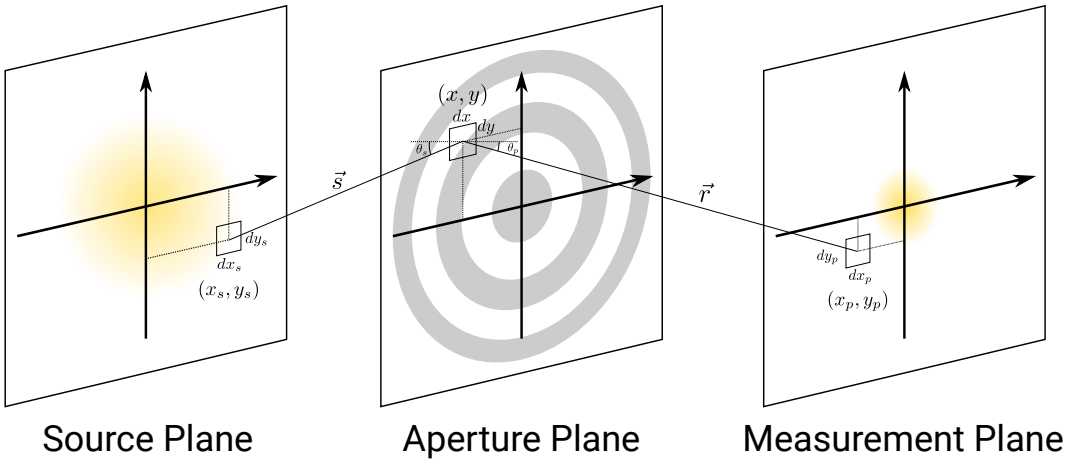


Figure 6.2: A schematic of the parameters needed to calculate the diffraction integral. The source plane is where the source of helium is located, the wave then travels to the zone plate where it is diffracted in the aperture plane. The diffracted intensity is then calculated in the measurement plane.

$$\left.\frac{\partial G_-}{\partial n}\right|_{z=0} = \cos(\theta_p)(ik - \frac{1}{r})\frac{\exp(ikr)}{2\pi r}. \tag{6.16}$$

Since the helium is arriving from a point source, the incident field is taken to be a spherical wave so that $U(\vec{r}_{S_1}) = a_p \exp(iks)/s$ where $s^2 = D^2 + (x - x_0)^2 + (y - y_0)^2$. The amplitude at a point (x_p, y_p, L) behind an aperture with a transmission function $h(x, y)$, is obtained by evaluating

$$U_I(\vec{r}_p) = \frac{a_p}{2\pi}\int\int h(x,y)\frac{\exp(ik(s+r))}{sr}\left(\frac{1}{r} - ik\right)\cos(\theta_p)dxdy \tag{6.17}$$

By making the further small assumption that $r \gg \lambda$, the amplitude at a point (x_p, y_p, L) is given by evaluating,

$$U_I'(\vec{r}_p) = \frac{a_p}{i\lambda}\int\int h(x,y)\frac{\exp(ik(s+r))}{sr}\cos(\theta_p)dxdy. \tag{6.18}$$

It should be noted that the result stated above can be rapidly achieved by considering Huygens' principle and using an obliquity factor of $K = -\cos\theta_P = L/r$. It is not possible to guess the obliquity factor before beginning the derivation and in fact it is altered by choosing different boundary conditions for the problem. However, using Huygens' principle is a useful tool to qualitatively understand the origin of the terms in equation 6.18.

Since $r(x - x_p, y - y_p)$ is a function of $(x - x_p)$ and $(y - y_p)$ only and $s(x, y)$ is a function of x and y only, the integral in equation 6.17 can be rewritten as,

$$U_I(\vec{r}_p) = \int\int a(x,y)b(x - x_p, y - y_p)dxdy \tag{6.19}$$

Where the functions a and b have been defined as,

$$a(x,y) = h(x,y)\frac{a_p\exp(iks(x,y))}{s(x,y)}, \tag{6.20}$$

$$b(x-x_p, y-y_p) = \frac{L\exp(ikr(x-x_p, y-y_p))}{2\pi r(x-x_p, y-y_p)^2}\left(ik - \frac{1}{r(x-x_p, y-y_p)}\right). \tag{6.21}$$

Expressing the diffraction integral using a and b shows how the integral is in fact a convolution. The earlier choice to use a Green's function that is zero in the aperture plane has led to the final integral being a convolution, a different choice of boundary condition would not have produced a convolution. Using the convolution theorem, a convolution can be calculated using fast Fourier transforms making less computational expensive than through direct evaluation.

6.2.2 Extended Sources and Non-Monochromatic Sources

A real helium source is not a point source and therefore finite sized sources need to be accounted for. An extended incoherent source can be modelled as a sum of point sources. Since the source is incoherent, the total intensity is calculated by summing the intensities produced by each point source giving,

$$U_S^{tot}(\vec{r_p}) = \int_{Source} f(\vec{r_S})\left(\int\int a(x,y)b(x-x_p, y-y_p)dxdy\right)d\vec{r_S}, \tag{6.22}$$

where the extra integral added is over the entire source, $\vec{r_S}$ is a point on the source and $f(\vec{r_S})$ is the intensity distribution of the source. For simplicity, the intensity distribution is taken as uniform over a circular area so that the source size is the only important parameter.

Supersonic helium sources produce helium beams with a very narrow velocity distribution, but the non-monochromaticity of the source still needs to be accounted for. The effect of a non-monochromatic beam can be evaluated using a similar procedure to above and integrating over the diffraction pattern obtained for a series of monochromatic beams with different wavelengths. The total intensity is then given by,

$$U_\lambda^{tot}(\vec{r_p}) = \int g(\lambda)\left(\int\int a(x,y)b(x-x_p, y-y_p)dxdy\right)d\lambda. \tag{6.23}$$

The distribution of wavelengths $g(\lambda)$ is assumed to be Gaussian and is given by,

$$g(\lambda) = A\exp\left(-S^2\left(\frac{\lambda-\overline{\lambda}}{\overline{\lambda}}\right)^2\right), \tag{6.24}$$

where $\overline{\lambda}$ is the average wavelength of the helium beam, S is the speed ratio of the helium beam and A is an arbitrary constant for normalisation. When performing the integration, any value of $g(\lambda)$ that is less than 5% of the maximum value is ignored to ensure that the computation time is kept to a minimum. It should be noted that the derivation of the above expression used the fact that the speed ratio is related to the standard deviation of the wavelength, $\Delta\lambda$, through,

$$\frac{\Delta\lambda}{\overline{\lambda}} = \frac{1}{\sqrt{2}S}. \tag{6.25}$$

Some care must be taken when using speed ratios to differentiate between the standard deviation and the full width half maximum of the wavelength distribution.

Performing the full calculation for a finite source size and non-monochromatic beam would include a large number of integrals and therefore only one case is considered at a

time. While the core of the computation to compute the diffraction integral is a reasonably efficient convolution, integrating over all the source sizes and wavelengths would take a long time. The full calculation could be performed, but would likely require a high performance computer to obtain an answer in a reasonable time.

6.2.3 Moving to a Discrete System

Since the diffraction integral cannot be analytically evaluated, the problem needs to be numerically evaluated using a discrete set of points. The convolution theorem applies in the discrete case, but it assumes that the convolution is in fact a circular convolution. Therefore, the matrices involved will need to have cyclic boundary conditions and some care needs to be taken to ensure that the answer is accurate.

To begin, the matrix $\mathbf{a}$ is computed over a N_a x N_a grid using a pixel width of Δ, and then the matrix $\mathbf{b}$ is computed over a N_b x N_b grid with the same pixel width of Δ. N_a and N_b are arbitrary sizes and are not fixed to particular values. Since the matrix a contains the aperture function which will always go to zero at large values, $N_b > N_a$ to avoid clipping. These matrices are then zero-padded to be N_p x N_p grids, called $\mathbf{A}$ and $\mathbf{B}$, where $N_p > N_b$ and $N_p = 2^n$ where n is an integer so that the FFT routines later are quick.

The Fast Fourier Transform (FFT) can then be used on $\mathbf{A}$ and $\mathbf{B}$ to obtain $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ respectively. As the convolution theorem dictates, these matrices are then multiplied elementwise to obtain,

$$\tilde{\mathbf{K}} = \tilde{\mathbf{A}} \odot \tilde{\mathbf{B}} \tag{6.26}$$

The final step is to then Inverse Fast Fourier Transform (IFFT) back to obtain $\mathbf{K}$. The earlier zero padding introduces boundary effects with the convolution, so the edges of $\mathbf{K}$ are not correct. The central $|N_b - 2N_a - 1|$ x $|N_b - 2N_a - 1|$ grid[1] from $\mathbf{K}$ is the best estimate for $\mathbf{U}$, the actual diffraction pattern on the grid that has been defined.

So by implementing the routine described above, it is possible to obtain the diffraction pattern from any arbitrary aperture function. The next section will describe how the method has been applied to study a particular diffractive element, the Fresnel zone plate.

6.2.4 Fresnel zone plate

Using the framework set out by Huygens' principle, an ideal lens is an element that will alter a wave so that secondary waves produced in an aperture plane will add constructively at one point in the measurement plane. If all the secondary wavelets that are produced in the aperture plane add constructively at one point, then the total flux that is incident on the aperture plane will be focussed to a single point in the measurement plane.

By making the Fresnel approximation, it is possible to obtain a simple expression for how the phase of a wave changes when passing through different points in the aperture plane. Looking at fig. 6.2, the optical path length between a source and an on axis point in the measurement plane that passes through a position (x, y) in the aperture plane is given

[1] For the actual implementation the full expression that deals with odd grid sizes is actually $-(\text{floor}(Na - Nb/2)) + 1)$ where the function floor() rounds the value down the nearest integer.

by,

$$\begin{aligned} s + r &= \sqrt{D^2 + x^2 + y^2} + \sqrt{L^2 + x^2 + y^2} \\ &= D + L + \frac{x^2 + y^2}{2D} + \frac{x^2 + y^2}{2L} + O(x^4, y^4). \end{aligned} \tag{6.27}$$

If the term $\frac{1}{R} = \frac{1}{L} + \frac{1}{D}$ is introduced and the Fresnel approximation is made to ignore the higher order terms then the optical path length is given by

$$s + r = \text{const.} + \frac{x^2 + y^2}{2R}. \tag{6.28}$$

Therefore, in the near field the phase of a wave passing through the aperture plane and reaching the optic axis varies quadratically with the distance of the intersection point in the aperture plane with the centre axis.

An ideal lens would therefore vary the phase quadratically across the element, which would normally be achieved with light by using a spherical lens. In conventional optics, a variation in the refractive index between two media can be used to introduce the quadratically varying phase difference that is needed. The existence of transparent materials with refractive indices that are different from air permits the construction of lenses for light.

When considering helium waves, there are no solid materials that are transparent to helium and therefore conventional lenses cannot be used to focus helium. Any lens element will have to be binary with respect to helium transmission, either allowing helium to pass through or to be blocked.

A Fresnel zone plate is an optical element made only from opaque and transparent regions that can achieve focussing of a wave to a point. Given the optical path length in equation 6.28, the aperture screen can be divided up into a series of concentric circles where the phase of waves passing through each circle increase by π. The phase will change by π when the optical path changes by $\lambda/2$, so the radii of the boundary between the n^{th} and $(n+1)^{\text{th}}$ Fresnel zones is approximately given by,

$$r_n^2 = n\lambda R. \tag{6.29}$$

R is given by $\frac{1}{R} = \frac{1}{L} + \frac{1}{D}$, which when compared to the thin lens equation gives $R = f$, where f is the effective focal length of the lens.

Normally, the contributions from each Fresnel zone add together destructively to give a small amplitude at the focal point, since the phases are opposite between each alternate zone and there is destructive interference between each zone. However, by blocking out alternate Fresnel zones, each open zone would have the same phase, so the total amplitude at the observation point would increase due to constructive interference.

Figure 6.3 shows a realistic design for a Fresnel zone plate for helium where the zones have been placed using equation 6.29 and a support structure has been added. Since helium will not pass through any solid surfaces, the zone plate must be free standing and the open areas of it must have no material in them. Therefore, the zone plate design requires some scaffold pieces to hold the concentric circles together. Blocking parts of the zones that should be transparent leads to a degradation in the performance of the zone plate when compared to the theory. The reduction in performance will be discussed in more detail later.

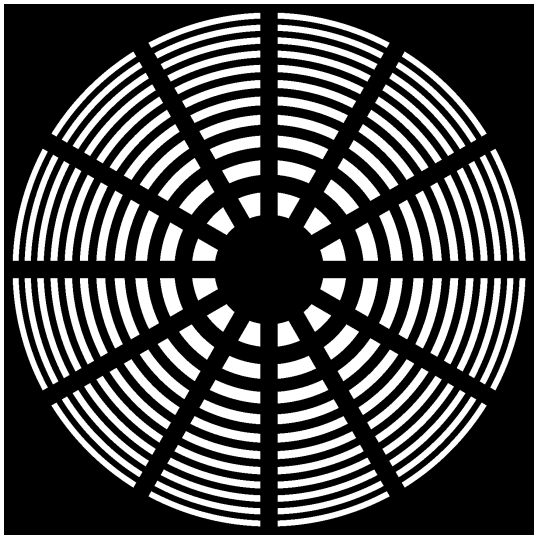


Figure 6.3: Image of the Fresnel zone plate designed to focus helium with a wavelength of $\lambda = 5.63 \times 10^{-11}\,\mathrm{m}$ at a focal length of 1 mm. The minimum feature size of the zone plate is set to be greater than $\Delta r_N = 25\,\mathrm{nm}$, and a set of bars are added with a width of $75\,\mu\mathrm{m}$ to ensure that the zone plate could be manufactured. The radius of the N^{th} zone is proportional to $\sqrt{N}$ such that the area of each zone is a constant.

Like in classical optics with a lens, the diffraction pattern from a point source has a finite width due to the finite angular size of the zone plate. Similarly to standard lenses, the numerical aperture (NA) of the Fresnel zone plate is a crucial parameter and determines what the size of the point spread function is. Larger zone plates will produce smaller spot sizes and vice versa.

The size of the zone plate is controlled by the minimum feature size that can be manufactured, which therefore means that the spot size produced by a point source is determined by the minimum feature size. Smaller minimum feature sizes allow larger zone plates to be made with larger numerical apertures. The point spread function is found to be an Airy function for a zone plate with its first minimum at $\Delta r_{\mathrm{Rayleigh}}$. Following the treatment given by Attwood [6], it can be shown that the size of the point spread function ($\Delta r_{\mathrm{Rayleigh}}$) is related to the minimum feature size of the zone plate (Δr) by ,

$$\Delta r_{\mathrm{Rayleigh}} = 1.22 \Delta r. \tag{6.30}$$

A realistic value of the minimum feature size was chosen to be 25 nm, since free standing zone plates have previously been demonstrated with a 50 nm outer zone width [51] but nanofabrication has progressed since then and zone widths of 30 nm have been achieved in X-ray microscopes [6].

Figure 6.4 shows an example of a calculated diffraction pattern obtained in the focal plane from the zone plate in fig. 6.3. The diffraction pattern was calculated using the procedure described above in section 6.2.3. A sharp spot is formed where the helium is successfully focussed and the width of the spot is close to the value given by equation 6.30. There is a significant flux that also passes directly through the zone plate without being focussed which is seen by the large circular ring around the focal point that has the same size as the zone plate. The effect of the support structure is largely seen through small changes in the

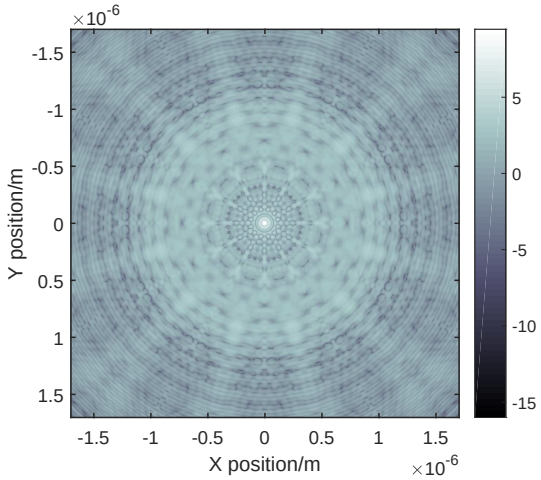


Figure 6.4: Image showing the logarithm of the calculated diffracted intensity pattern from helium with a wavelength of $\lambda = 5.63 \times 10^{-11}\,\mathrm{m}$ (which corresponds to a nozzle temperature of $T = 300\,\mathrm{K}$) at a distance of 1 mm from the Fresnel zone plate shown in fig. 6.3. The sharp spot in the centre indicates the helium is being focussed by the zone plate at the focal distance. There is also a large flux pass directly through the zone plate and not being focussed.

diffraction pattern outside of the focal spot.

Figure 6.5 is a plot of the calculated diffraction pattern at different distances from the aperture plane and illustrates how Fresnel zone plates do not focus all of the flux that is incident into a single point, but instead diffract into a series of different orders. The condition that the wave from each open zone of the zone plate adds together constructively does not exclusively occur at the principal focal point. If the primary focal length of the zone plate is given by f, then there are also a series of subsidiary focal points at odd fractions of the primary focal length given by $f/(2m+1)$, where m is an integer. Figure 6.6 is a one dimensional slice through the centre of fig. 6.5 illustrating the subsidiary focal points.

Figure 6.5 shows another example of how a significant portion of the flux is transmitted straight through the zone plate in a zero order beam. It can be seen in the figure that the undiffracted beam creates a large circular disc of flux that has the same radius as the zone plate that passes straight through. Despite the central zone of the Fresnel zone plate being opaque, a significant flux is observed in the middle of the diffraction pattern.

Often an aperture, known either as an order sorting aperture or zero order stop, will be placed into the diffracted beam to block the zero order beam from reaching the focal plane [52]. Placing an additional aperture in the beam path will lead to further diffraction through the aperture, which will lead to an increase in the spot size for a point source. However, all experimental zone plate setups that use atoms have been limited by either the chromatic aberrations or source size, therefore the zero order stop has not altered the size of the focal spot.

A significant portion of the flux can also be seen in fig. 6.5 to be diverging from the zone plate. The presence of the diverging part of the beam highlights the importance in correct normalisation of the calculated intensity. Summing the total intensity in a given diffraction pattern will only correctly normalise the result if a very large area of the diffraction pattern is used that includes the diverging parts of the diffraction pattern.

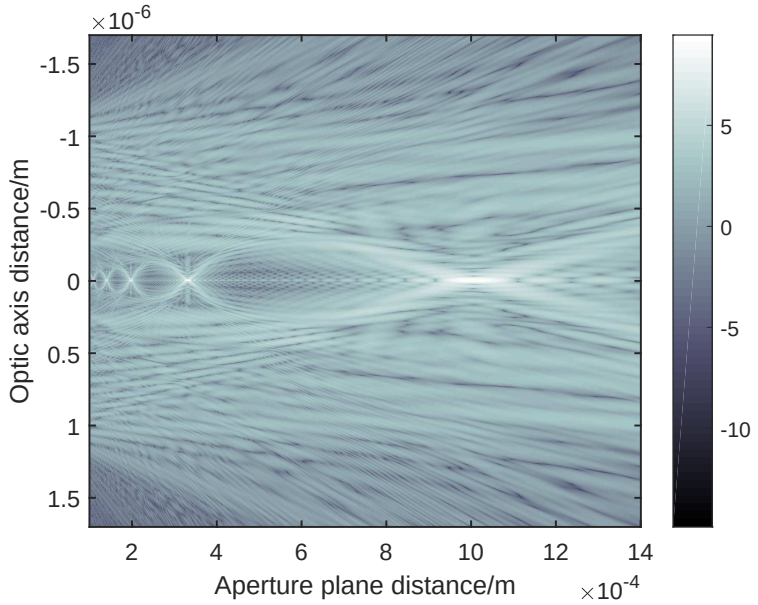


Figure 6.5: A plot showing the logarithm of the intensity of helium after being diffracted through a Fresnel zone plate at different distances from the zone plate. On axis the subsidiary foci at $f/3, f/5, f/7, f/9$ are visible. While the intensity of the helium at the focal point is high, it can be seen that there is a significant flux of helium passing straight through in the zero order beam or diverging out from the zone plate in the "-1^{th}" order of the beam.

The theoretical efficiency of the first order focus of a zone plate is $\eta = 1/\pi^2 \approx 10\%$ [6]. The efficiency of a real zone plate has been experimentally measured by Reisinger et al. to be about 7% of the incident flux [149]. The remaining transmitted flux appears as a wide beam which contains a much higher flux than the first order focus.

The efficiency of the zone plate is influenced by the design of the support structure, therefore in order to evaluate the efficiency of a realistic zone plate, a representative example needs to be calculated. The zone plate that is used is shown in fig. 6.3 and has a focal length of $f = 1\,\text{mm}$, a minimum feature size of $\Delta r_N = 25\,\text{nm}$ and is designed for a wavelength of $\lambda = 5.63 \times 10^{-11}\,\text{m}$ which corresponds to a nozzle temperature of $T = 300\,\text{K}$. The support structure has been chosen to ensure that it would be feasible to manufacture the zone plate.

Figure 6.7 shows the integrated flux in the central region, including and around the central peak, for two different speed ratios. The results are normalised to the incident flux and plotted cumulatively as a function of radius. The simulation shows that in both cases the central focus from the zone plate contains about 5% of the incident flux. The difference from the theoretical efficiency arises due to the conservative support structure design that deviates from the ideal zone plate design. In subsequent calculations $\eta = 0.05$ shall be used.

Figure 6.8 illustrates the effect of a finite source size on the diffraction pattern. Figure 6.8a is a plot of the full width half maximum (FWHM) of the diffracted spot size as the source size is increased. The plot very closely follows a straight line indicating that the zone plate is acting like a perfect lens for reasonably large source sizes of $100\,\mu\text{m}$ that are 15 cm away. Figure 6.8b is a plot of how the integrated flux varies for a small and large source size. For the larger source size in orange, the spot size is larger but the same flux is contained inside the spot.

Helium at different wavelengths will be focused to different points due to the strong

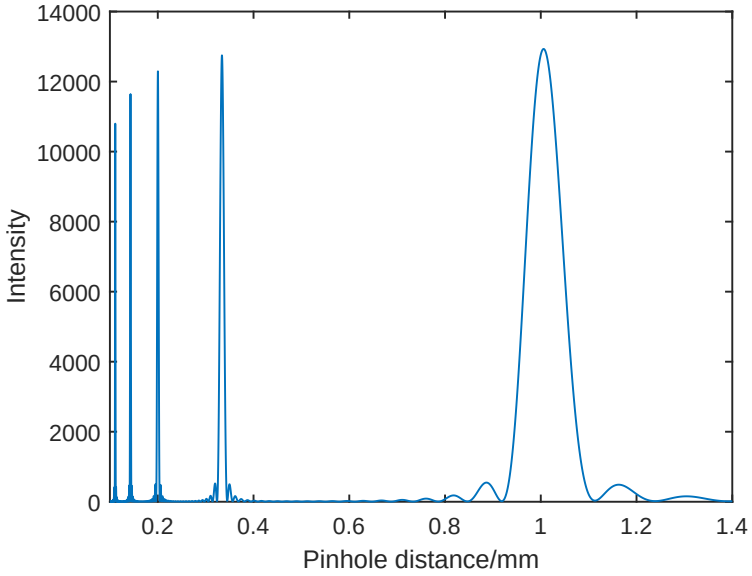


Figure 6.6: A plot showing on axis intensity of helium diffracted from a Fresnel zone plate. Subsidiary maxima are present at odd fractions of the focal point, i.e. at $f/3, f/5, f/7$ etc. The reduction in the peak height for the high order peaks close to the zone plate is due to the peak not being sufficiently sampled to obtain the correct maximum value.

dependence of the focal point on wavelength as seen in equation 6.29. It can be shown that chromatic aberrations are significant when,

$$\frac{\Delta\lambda}{\lambda} > \frac{1}{2N\sqrt{2\ln 2}}, \tag{6.31}$$

where $\Delta\lambda$ is the standard deviation of the wavelength distribution, λ is the average wavelength and N is the total number of zones [6].

Equation 6.31 can be rearranged to show that chromatic aberrations are typically significant when $S < N$, where S is the speed ratio. Therefore, larger zone plates will typically suffer from stronger chromatic aberrations. Equation 6.31 provides a useful guideline for the magnitude of the chromatic aberration, but to a more accurate model is needed to proceed.

The chromatic aberrations from a zone plate are difficult to predict accurately, and previous attempts have used geometric optics to obtain a crude estimate. By considering the focal points of two extreme wavelengths and assuming the beam travels as simple rays, it is possible to predict the chromatic aberrations, σ_{chrom}, through geometric arguments [150, 51] to be,

$$\sigma'_{\text{chrom}} = \frac{r_N}{\lambda/\Delta\lambda}, \tag{6.32}$$

where r_N is the radius of the zone plate. In order to obtain a better estimate for the chromatic aberrations of a zone plate, the diffraction pattern from a representative example of a zone plate is needed. By taking the zone plate design that has been discussed already in this chapter and investigating the specific aberrations that are simulated, the modelled chromatic aberrations can then be extended to a general zone plate.

Figure 6.9 is a plot of the simulated intensity profile across the spot for two different speed ratios. The apparent full width half maximum of the central spot does not change

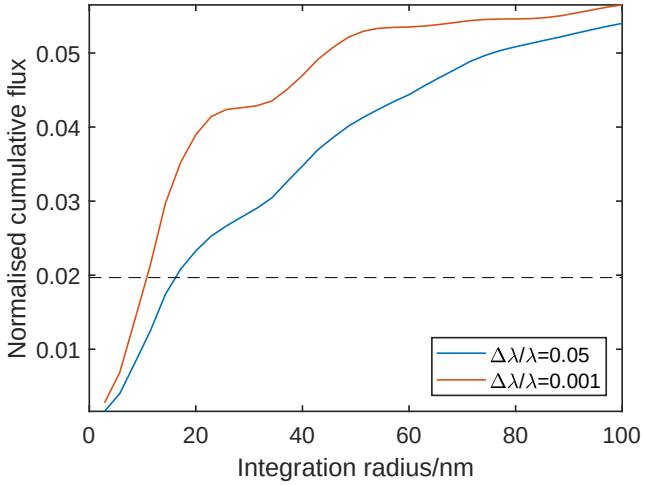


Figure 6.7: Simulation of the integrated flux from the centre of the focal spot produced by a zone plate with $f = 1\,\text{mm}$, $\lambda = 5.63 \times 10^{-11}\,\text{m}$, $\Delta r_N = 25\,\text{nm}$ and 12 support bars that are 75 nm wide. For both speed ratios the flux levels off, showing that there is about 5% of the total flux in the central spot. The lower speed ratio produces a beam that has a larger distribution of flux due to the chromatic aberrations of the zone plate. The dashed line is the flux contained within one standard deviation of a Gaussian beam normalised to have a total flux of 5%, which is used to estimate the beam standard deviation.

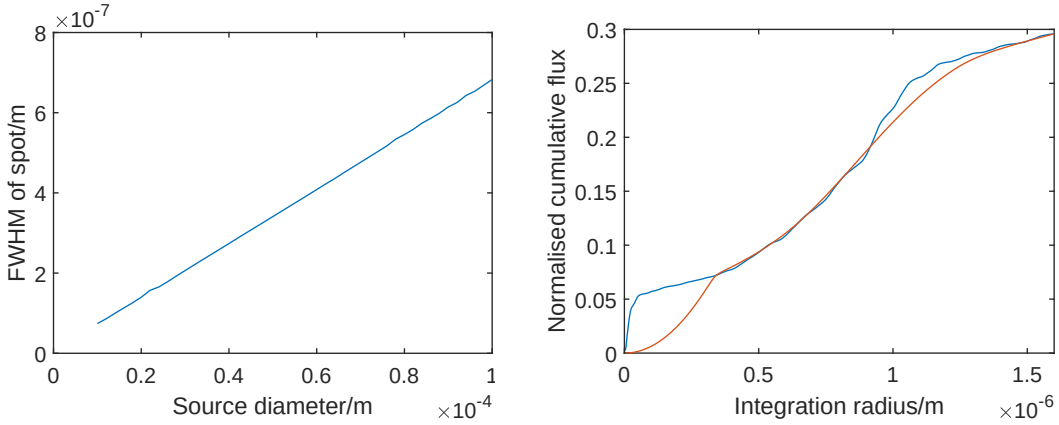
significantly when the speed ratio is increased due to the very sharp peak that is formed when the beam is correctly focussed. However, there is a significant redistribution of the intensity, with more flux appearing in the central region and a reduction in the surrounding areas.

To progress with any optimisation of the machine design, the intensity distribution needs to be represented with a single representative standard deviation. One method is to use the integrated flux from the centre of the spot, as shown in Fig. 6.7, and estimate the standard deviation of the spot by comparison with an integrated 2D Gaussian function. Even though it is a substantial simplification, it should still be representative of the overall imaging resolution in a microscope. The integral of a normalised 2D Gaussian from the centre is $X(R) = 1 - \exp(-R^2/2\sigma^2)$, so when the radius of the integrated area is equal to the standard deviation, $1 - e^{-\frac{1}{2}}$ of the flux is enclosed. The standard deviation of the beam size can therefore be estimated by finding when the integrated normalised flux is equal to $\left(1 - e^{-\frac{1}{2}}\right) \times 0.05 \approx 0.020$.

The variation of the estimated standard deviation of the beam with speed ratio is plotted in Fig. 6.10. The beam standard deviation is a convolution of the perfect diffraction pattern from a point source and the chromatic aberration to be estimated. A function of the form,

$$\sigma_{\text{chrom}} = \sqrt{a^2 + \left(b\frac{\Delta\lambda}{\lambda}\right)^2}, \tag{6.33}$$

is fitted to the data to estimate the beam size. The result is that $b = (1.985 \pm 0.007) \times 10^{-7}$. Previously the formula obtained with geometric optics predicts a chromatic aberration for a zone plate with a working distance of $f = 1\,\text{mm}$ and a minimum feature size of $\Delta r_N =$



(a) (b)

Figure 6.8: (a) A plot showing how the full width half maximum of the difraction pattern of helium varies as the diameter of the source is changed. Fitting a straight line to the data gives a gradient of 0.01353 ± 0.00002, which agrees with the expected gradient of $2f/D = 1/75$. (b) A plot showing how the integrated flux changes for two different source sizes. The blue curve is for a source size with a diameter of $2\,\mu\mathrm{m}$, while the orange curve is for a helium source with a diameter of $100\,\mu\mathrm{m}$. While the size of the spot significantly changes as the source size is changed, the total flux inside the spot remains approximately the same.

25 nm of

$$\sigma'_{\text{chrom}} = 1.14 \times 10^{-6} \frac{\Delta\lambda}{\lambda}, \tag{6.34}$$

since the zone plate has a radius of $r_N = 1.14 \times 10^{-6}\,\mathrm{m}$. It should be noted that the size of the zone plate deviates from the continuum result ($r_N = (\lambda f)/(2\Delta r_N)$) significantly since there are so few zones. The result from the diffraction simulation gives a value that is (5.066 ± 0.011) times less than this expression, so a modified form of equation 6.34 for the chromatic aberration that can be used is given by,

$$\sigma_{\text{chrom}} = \frac{r_N}{5.066\sqrt{2}S}. \tag{6.35}$$

6.2.5 Elliptical zone plates

Zone plates can also be used when the wave is not normally incident by using elliptical zones instead of circular zones. For a zone plate which is not directly facing the wave source, the Fresnel zones form ellipses over the aperture plane [151]. The current geometry of the helium microscope uses an incoming angle of 45° and therefore would either require that the mounting for the aperture be altered so that a regular Fresnel zone plate can be used or that an elliptical zone plate is used.

The focal length of the elliptical zone plate is given by the distance the helium would travel from the zone plate to the focal point, which is different to the perpendicular distance between the aperture plane and the measurement plane. The focal length of the zone plate will be larger due to the wave no longer being normally incident, which leads to the outer radius of the zone plate being increased.

Figure 6.11 shows the design of an elliptical zone plate for use with an incident beam at 45° and the same properties as the Fresnel zone plate that has previously been discussed.

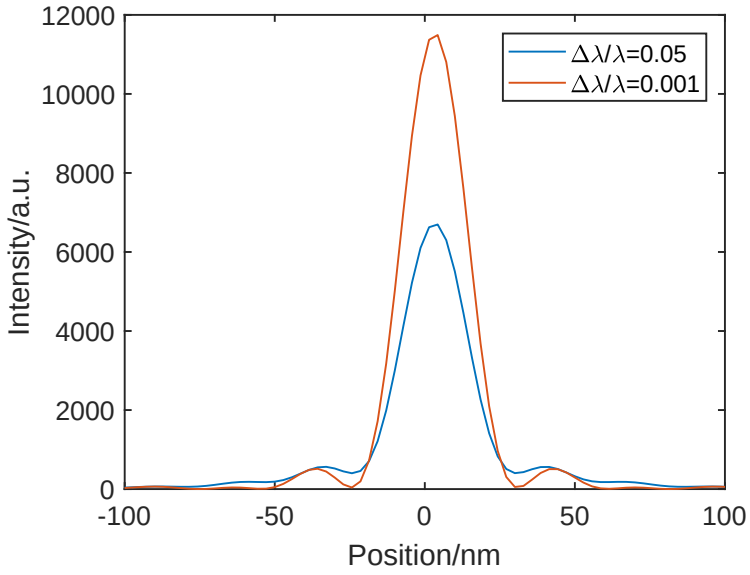


Figure 6.9: Section of the intensity in the spot produced at different speed ratios. The different speed ratios produce central spots with approximately the same full width half maximum. However, there the total flux in the central peak is lower for the lower speed ratio and more of the intensity appears outside where the first zero occurs for the monochromatic beam.

The elliptical zone plate has very similar properties to the standard Fresnel zone plate, only the point spread function produces an airy function is spread into an ellipse. The vertical direction still has the first minimum in approximately the same position of $1.22\Delta r$, but the horizontal minimum is stretched to occur $\sqrt{2}$ further from the centre.

For source sizes up to $100\,\mu\mathrm{m}$, the zone plate still acts as a lens as previously seen with the Fresnel zone plate. The circular source size will produce a demagnified circular spot in the focal plane. The calculated size of the demagnified source formed by the zone plate also still scales with the expected gradient of $2f/D = 1/75$.

The chromatic aberrations of an elliptical zone plate are similar to the regular Fresnel zone plate. By performing the same procedure as before, except now integrating the flux inside ellipses centred at the focal point it is possible to estimate the chromatic aberrations. The correction factor on the standard chromatic aberration is found to be 3.015 ± 0.008, where the radius of the zone plate was taken to be the semi-minor axis (vertical radius).

Elliptical zone plates would be more difficult to implement than regular zone plates due to the asymmetry of the design. The axes of the elliptical zone plate need to be carefully aligned with the horizontal and vertical axes of the machine to ensure that the zone plate works correctly. The maximum dimension of the zone plate is likely to only be about $1\,\mu\mathrm{m}$ and therefore alignment would require careful manipulation with a microscope.

While the elliptical zone plate is a feasible method for focussing helium waves, the added complexity does not help in understanding the limits of using a zone plate for microscopy and therefore a Fresnel zone plate for use in a normal incidence machine shall be considered.

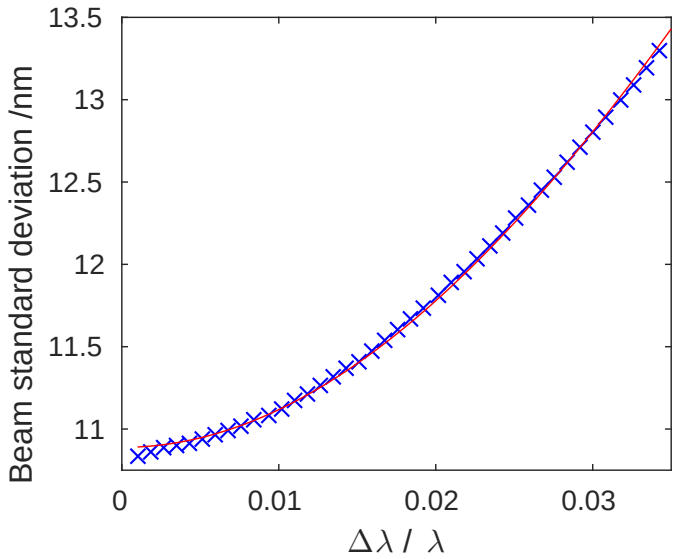


Figure 6.10: Beam size at different speed ratios for a zone plate with a focal length of 1 mm. The blue crosses are where the normalised integrated flux from the centre is equal to 0.020, which is used as an estimate for the beam standard deviation as discussed previously. The red line shows the model $\sigma_{\text{chrom}} = \sqrt{a^2 + (b(\Delta\lambda/\lambda))^2}$, with the parameters optimised using a least squares algorithm to be $a = (10.889 \pm 0.005)$nm and $b = (2.246 \pm 0.005) \times 10^{-7}$. The zone plate radius here is 1.14×10^{-6}, so that gives the correction factor as (5.066 ± 0.011).

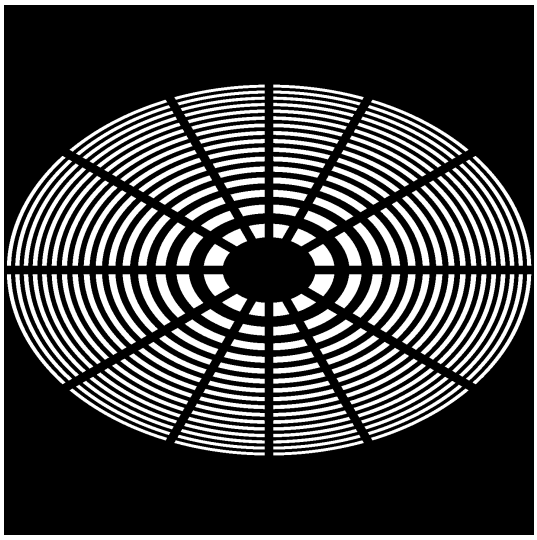


Figure 6.11: An image showing an example of an elliptical zone plate for a helium wave that is incident at 45° with a wavelength of $\lambda = 5.63 \times 10^{-11}$ m at a focal length of $\sqrt{2}$ mm. Since the incoming wavevector is at 45°, the perpendicular distance from the zone plate to the focal point is 1 mm. The minimum feature size of the zone plate is set to be greater than $\Delta r_N = 25$ nm, and a set of bars are added with a width of $75\,\mu\text{m}$ to ensure that the zone plate could be manufactured.

6.3 Principal Components of model

Now that an accurate model for the zone plate has been constructed, it is possible to proceed and model the performance of a microscope using either a zone plate or a pinhole. To construct a full model of the performance of a helium microscope, a model for the performance of a supersonic source is needed.

6.3.1 Free jet expansion

The supersonic helium beam is formed in a free-jet gas expansion where high pressure gas (typically around 100 bar) expands through a fine nozzle (typically $10\,\mu\mathrm{m}$) into a low pressure vacuum chamber. The expansion is schematically represented in Fig. 3.1 and is covered in more detail in section 3.1. The gas is accelerated by the pressure difference, reaching sonic speeds at the nozzle exit. The gas expands as it enters the chamber, but since the flow is supersonic the gas overexpands and a shock system is formed to recompress the system to match the chamber pressure. The resulting shock structure leaves a zone of silence where the expansion can be sampled, using the skimmer, to produce a high intensity beam with a narrow velocity distribution.

As evident from Fig. 6.1, a narrow helium microprobe generally requires a small angular source size, β, and future high resolution instruments will need to achieve a narrow beam without disrupting the expansion. The angular source size can be reduced either by moving the source further away from the main aperture, or by increasing the collimation of the helium expansion. Moving the source is the simplest method, but increases the overall size of the instrument to an impractical degree. Very small skimmers (often known as 'microskimmers' [139]) can be used to reduce the source size, but their narrow geometry means atoms backscattered into the beam from either the leading edge, or the internal skimmer walls, risks affecting the beam intensity significantly. An alternative is to introduce a second aperture behind the skimmer that, provided the volume is adequately pumped, will set the source size of the instrument whilst minimising additional beam interference. It has recently been shown that the second aperture can also be used to rapidly change the resolution of the instrument [152].

The parallel velocity distribution of the atom beam is often defined as a speed ratio, $S = v/\Delta v$, where v is the average speed of the atoms and Δv is the standard deviation of the distribution in the parallel direction. Beams of helium-4 atoms formed in a free jet expansion have an unusually high speed ratio due to quantum effects on the cross-section, which is exploited in many surface scattering experiments. Toennies and Winkelmann measured and simulated the speed ratio for helium-4 beams as a function of nozzle pressure [138]. The later optimisation requires an analytic form for the speed ratio, so an empirical model needs to be created. Figure 6.12 shows an empirical fit to the model that was proposed by Toennies and Winkelmann using the function,

$$\log(S) = a\log(P_0 d_{noz}) + b + \frac{c}{1 + e^{-\delta(\log(P_0 d_{noz}) - \mu)}}, \tag{6.36}$$

where $a = 0.43 \pm 0.03$, $b = 0.76 \pm 0.02$, $c = 0.84 \pm 0.10$, $\delta = 5.2 \pm 0.9$ and $\mu = 1.97 \pm 0.03$ are fitted parameters to the model when $P_0 d_{noz}$ is measured in torr cm.

The apparent spatial distribution of a free jet source can be described by introducing

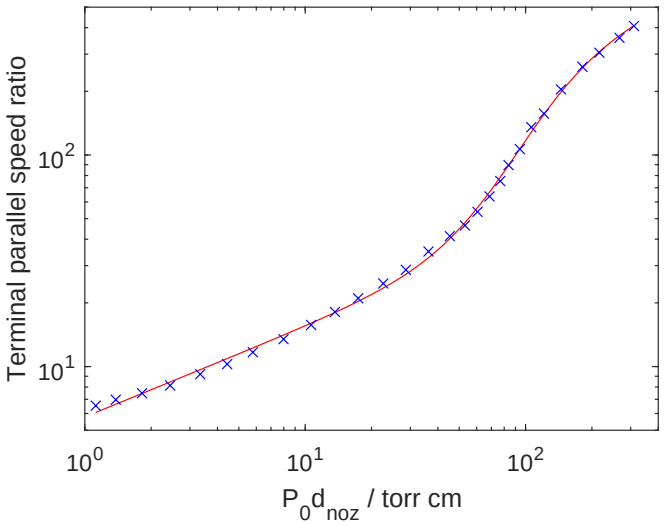


Figure 6.12: Empirical fit (red line) to the speed ratio model proposed by Toennies and Winkelmann [138] (blue crosses). The fit combines a straight line and logistic function to represent the increase in speed ratio due to the onset of quantum cross-section effects.

the concept of a virtual source [72]. At large enough distances from the source, collisions between helium atoms cease and the atoms follow straight line paths in a molecular flow regime. These trajectories can be traced back to an area of least confusion near the nozzle, with radius r_v, which is defined as the virtual source. Collisions of the atoms near the nozzle usually lead to the virtual source being much larger than the nozzle aperture itself, and often larger than the skimmer. Virtual sources of free jet expansions have been shown to have a double Gaussian shape [74] which results in the virtual source having large tails on the end of the distribution. However, the wide component of the source can be neglected if the source is suitably collimated, i.e. only a small area from the centre of the expansion is used to form the final beam.

6.3.2 Virtual Source Brightness

Most conventional atom-surface scattering experiments use wide beams to study a spatially homogeneous sample in reciprocal space. Since the beam is not spatially resolved, the size of the source is unimportant and only the intensity needs to be maximised. However for microscopy, both a narrow and high intensity beam is needed, so the brightness of the source needs to be maximised.

Let the brightness, B, of an atom beam source be defined as

$$B = \frac{I}{\pi r_v^2}, \tag{6.37}$$

where I is the centerline intensity per steradian of the beam and r_v is the virtual source radius. To obtain a complete expression for the brightness expressions for the centerline intensity and virtual source size are required.

DePonte et al. [73] show that the centerline beam intensity is proportional to the total flux through the nozzle. They obtain the constant of proportionality by fitting to experimental

data. Their resulting expression is

$$I = 0.36 P_0 d_{noz}^2 \left(m k_B T_0\right)^{-\frac{1}{2}}, \tag{6.38}$$

where P_0 is the pressure in Torr in the nozzle, d is the nozzle diameter in cm and T_0 is the stagnation temperature of the gas in the nozzle in Kelvin. There is deviation from the simple expression when the flux through the nozzle is high, which is attributed to skimmer interference effects [67]. These interference effects arise at high fluxes, when significant numbers of atoms are backscattered from the outer surface of the skimmer into the expansion, disturbing the trajectories of the expanding atoms.

The virtual source size is a more complex parameter to estimate, since quantum effects can influence the virtual source size and these are difficult to accurately model. However DePonte et al. [73] note that it is possible to obtain an approximate expression for the virtual source size in terms of the speed ratio. By using scaling arguments and fitting to experimental data, the authors obtain the approximate relationship[2]

$$r_v = 0.80 d_{noz} S^{0.52}. \tag{6.39}$$

By combining equations 6.37, 6.38 and 6.39 it is possible to obtain an approximate expression for the brightness (count rate per steradian per unit area of the source) of a helium source,

$$B = 0.18 \frac{P_o}{S\sqrt{m k_B T}}. \tag{6.40}$$

Figure 6.13 shows the variation of brightness with nozzle pressure for $5\,\mu\mathrm{m}$, $10\,\mu\mathrm{m}$ and $20\,\mu\mathrm{m}$ diameter nozzles, obtained by combining the brightness model in equation 6.40 with the empirical fit for the speed ratio plotted in Fig. 6.12. Quantum effects in the helium expansion cause the virtual source size to increase suddenly between 30 and 100 bar (depending on nozzle size), so the brightness decreases at higher nozzle pressures. The brightness is therefore expected be maximised by either using a pressure just before the onset of quantum effects, or by using a much higher nozzle pressure as shown in Fig. 6.13. Despite apparently being able to obtain the highest brightness at very large nozzle pressures, it would be difficult to practically realise a source that runs in that regime; to use gas fixtures at greater than 1000 bar would pose significant technical challenges. An instrument operating near the lower pressure peak, which for a $10\,\mu\mathrm{m}$ nozzle is at about 60 bar, is considered for the remainder of the work.

Experimental results from DePonte et al. [73] show the peak in brightness appearing at $P_0 d \approx 10$ torr cm, which is lower than the expected value of $P_0 d \approx 44$ torr cm from equation 6.40. The authors suggest that quantum effects on the perpendicular speed ratio become significant at lower values of $P_0 d$ than in the parallel speed ratio, which leads to larger than expected virtual source sizes. Data on source parameters is notoriously difficult to measure and often different measurements contradict each other [70]. It also is possible that the difference in measured values arises due to interference effects from backscattered gas. In the absence of further data, the brightness as defined in equation 6.40 will be used for the subsequent calculations.

[2]This expression differs from that given by DePonte et al. [73] due to a potential error in using the source diameter instead of the source radius.

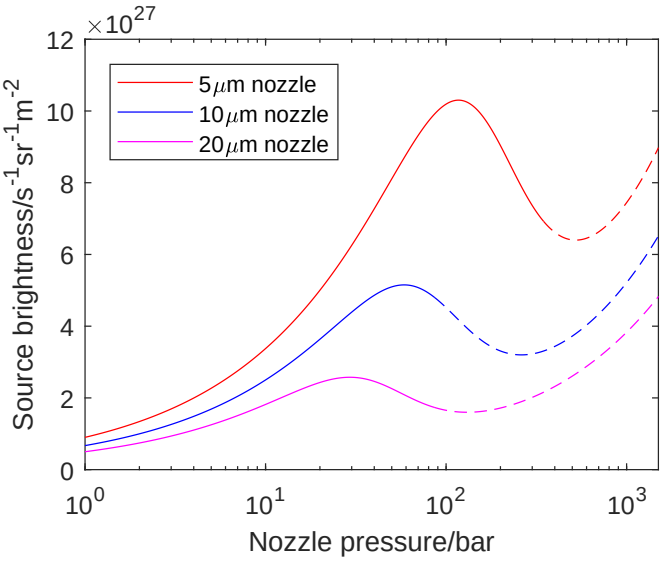


Figure 6.13: Brightness of a helium source modelled as a function of pressure for $5\,\mu\mathrm{m}$, $10\,\mu\mathrm{m}$ and $20\,\mu\mathrm{m}$ diameter nozzles. The sudden increase in the virtual source size causes the brightness to drop at about 50 bar for the $10\,\mu\mathrm{m}$ nozzle. The model used by Toennies and Winkelmann is only supported by experimental data up to $Pd \approx 100$ torr cm, so the extrapolation to higher pressures is indicated by a dashed line.

6.3.3 Intensity and beam size in pinhole and zone plate geometries

There are two established methods for creating a narrow beam of helium atoms: collimation with a small pinhole and focusing with a Fresnel zone plate. For the pinhole case, in addition to the geometric broadening shown in Fig. 6.1, the beam of atoms diffracts after the aperture, leading to additional broadening of the beam. The effect is a weak function of the speed ratio of the beam, so the pinhole can be treated as achromatic, meaning the helium beam production and collimation processes can be treated separately. In contrast, Fresnel zone plates [49] are strongly chromatic lenses [48] and focus with significant chromatic aberrations [51]. Equation 6.40 shows that the speed ratio is related to the brightness of the source, and therefore also the total flux in the beam. Since both the flux and size of the beam produced by zone plate focusing depend on the speed ratio they must be optimised together.

In order to compare these two focussing methods, a model for the intensity and size of the helium microprobe formed is needed for both cases. The Gaussian shape of the source is assumed to not be important, since it was established earlier it is advantageous to limit the source extent using a source aperture as illustrated in Fig. 6.1. Since the aperture that is being considered is small, only the centre of the helium expansion is sampled, where the brightness can be taken to have a constant value.

When the source size is similar in size to the skimmer being used, then the Sikora model is needed to obtain an accurate value for the centerline intensity [145]. The Sikora model can be rederived in the context of a virtual source, the reader can find a derivation of this in appendix D. It also shown in appendix D that when a small point like skimmer is used, the Sikora model simplifies to the product of the brightness and the area of the skimmer.

If the source is assumed to have a virtual source size of $200\,\mu\mathrm{m}$, the fractional differ-

ence between the Sikora model in the form presented in appendix D and the small skimmer limit that is to be derived in equation 6.41 is about 0.3%. Additionally the use of the mathematically awkward Sikora model would make it difficult to obtain the approximate analytic results that are presented later in the chapter. For the remainder of the chapter it is therefore assumed that the virtual source is much larger than the aperture since there is currently a lack of robust experimental data on virtual source sizes available to use the Sikora model and it makes little difference.

First, an expression for the beam intensity at the optical element is obtained in the limit of skimmers that are smaller than the virtual source. Since brightness is the number of atoms per steradian per unit area of the source, the flux of the helium beam per steradian can be obtained by multiplying by the area of the chosen source aperture. For a source aperture of radius x_s, that quantity is simply given by $\pi x_s^2 B$. The intensity (flux per unit area) at the optical element can then be obtained by converting from solid angle to area. Since the angular size of the source, as seen from the optical element, is given by $\beta = x_s/r_s$, where r_s is the distance from the source to the aperture plane, the overall expression for intensity, I, is given by

$$I = \pi \beta^2 B. \tag{6.41}$$

For a pinhole, the total transmitted flux is given by $F_{\mathrm{p}} = I\pi d^2/4$, where d is the diameter of the pinhole. After substituting with the expression for I in equation 6.41, the transmitted flux is given by,

$$F_{\mathrm{p}} = \frac{0.18\pi^2 P_o}{4\sqrt{m k_B \overline{T}}} \frac{d^2 \beta^2}{S}. \tag{6.42}$$

To avoid any confusion over the definition of the term resolution, the size of the helium microprobe will be described with the term beam standard deviation. The beam standard deviation, ϕ, at some distance after a pinhole is given by a convolution of three competing broadening effects. These originate from the pinhole geometry, diffraction at the pinhole, and the finite source size. The total beam standard deviation can be estimated by convolving the beam standard deviation obtained using geometric optics, the extent of the Airy disc diffraction pattern from the pinhole, and the (demagnified) size of the source.

The geometric pattern formed by the pinhole is a circle of diameter d, which can be approximated with a Gaussian distribution with a standard deviation given by $d/\left(2\sqrt{3}\right)$. The choice of standard deviation is difficult since approximating a non-Gaussian function with a Gaussian function will inevitably lead to a loss of detail of the beam, but the approach used by Palau et al. [144] is being used to maintain consistency. The distribution across the source aperture is uniform, which can similarly be approximated by a Gaussian with a standard deviation given by $\beta f/\sqrt{3}$, where f is the pinhole to sample working distance. The Airy function has its first zero at $1.22\lambda f/d$, but using this expression directly would overestimate the broadening. Instead, the central Airy disk can be approximated with a Gaussian distribution [153] with a standard deviation of $0.42\lambda f/d$. Combining these terms in quadrature gives the total beam standard deviation as

$$\phi_{\mathrm{p}} = \sqrt{\left(\frac{d}{2\sqrt{3}}\right)^2 + \left(\frac{\beta f}{\sqrt{3}}\right)^2 + \left(\frac{0.42\lambda f}{d}\right)^2}. \tag{6.43}$$

For a zone plate, the total flux in the beam is given by $F_{\mathrm{zp}} = \eta I \pi (r_N)^2$, where η is the

efficiency of the zone plate and r_N is its radius. By substituting and using the continuum limit approximation $r_N = \lambda f/2\Delta r_N$, where Δr_N is the size of the outermost zone, the total flux in a zone plate focused beam is given by,

$$F_{\text{zp}} = \frac{0.18\pi^2 P_o}{4\sqrt{mk_BT}} \frac{\eta\beta^2\lambda^2 f^2}{S(\Delta r_N)^2}. \tag{6.44}$$

As before, the total beam standard deviation at the sample is found by combining the contributing broadening effects. These are the combination of the point spread function of the zone plate, the demagnified source size and the chromatic aberration. The properties of a zone plate that were calculated in section 6.2 can now be used to predict the performance of a zone plate.

The diffraction pattern obtained from a point source illuminating a zone plate was shown in section 6.2 to approximately be an Airy pattern. The first zero from the diffraction pattern through the zone plate is at $1.22\Delta r_N$ [6], so using the same argument for Airy disks as the pinhole case, the diffraction pattern can be approximated with a Gaussian that has a standard deviation given by $0.42\Delta r_N$.

The extended source is modelled in the same way as the pinhole case, as a uniform disk with an angular size of β. A demagnified image of the source is formed, which is approximated with a Gaussian that has a standard deviation of $\beta f/\sqrt{3}$.

The chromatic aberrations from a zone plate can be modelled using equation 6.35 that was derived earlier using diffraction calculations. Using equation 6.35 instead of the standard equation that has been used that relies on geometric optics should give a more accurate result for the size of the beam.

Together, these give that the zone plate produces a beam that is approximated to be a Gaussian with a standard deviation given by,

$$\phi_{\text{zp}} = \sqrt{(0.42\Delta r_N)^2 + \left(\frac{\beta f}{\sqrt{3}}\right)^2 + \left(\frac{\lambda f}{10.13\sqrt{2}\Delta r_N S}\right)^2}. \tag{6.45}$$

6.4 Geometry Optimisation Method

In a helium microscope, the general aim is to maximise the flux in the beam reaching the sample while also minimising the beam size. However, these are incompatible quantities and there is no simultaneous optimum. Instead, given the helium flux at the sample is a function of many variables, an optimisation process can be developed to select a microscope geometry that maximises flux for a particular (constrained) resolution.

As described above, given a general set of parameters $\underline{x}$, the flux $F(\underline{x})$ and the beam size $\phi(\underline{x})$ produced by a focusing element can be calculated. To constrain the problem, consider a new function $g(\underline{x}) = \phi(\underline{x}) - \sigma = 0$, such that the beam standard deviation is given by the target standard deviation, σ. The best method for solving the constrained optimisation problem is to use Lagrange multipliers, a comparison to the alternative substitution method is discussed in appendix E. Consider the functional

$$\mathcal{L} = F(\underline{x}) - \Lambda g(\underline{x}), \tag{6.46}$$

which incorporates both the flux and beam size of a given geometry. The constrained optimisation problem is then solved by obtaining the maximal values of the Lagrange function, $\mathcal{L}$. To avoid confusion with the wavelength, the standard notation for the Lagrange multiplier is avoided and instead is denoted by Λ. The maximum values are found by solving the system of equations defined by

$$\frac{\partial \mathcal{L}}{\partial \underline{x}} = 0, \tag{6.47}$$

where $\underline{x}$ represents any free variable in the system. Some of these variables do not have maxima and instead practical limitations provide the values of these variables. The set of equations produced may not be analytically solvable, in which case a numerical method will need to be used.

6.4.1 Analytic Solution for Pinhole Geometry

An analytic solution for the optimal flux in the beam from a pinhole can be obtained, since the source and atom optics are separable problems. Let the flux through the pinhole be given by

$$F_{\mathrm{p}} = \gamma d^2 \beta^2, \tag{6.48}$$

where γ is defined by $\gamma = \pi^2 B_{\mathrm{peak}}/4$ and B_{peak} is the optimised peak brightness from the source. Given the Lagrange function for a general geometry,

$$\mathcal{L} = F - \Lambda(\phi - \sigma), \tag{6.49}$$

the Lagrange function for a pinhole geometry is then given by,

$$\mathcal{L} = \gamma d^2 \beta^2 - \Lambda \left(\sqrt{\left(\frac{d}{2\sqrt{3}}\right)^2 + \left(\frac{\beta f}{\sqrt{3}}\right)^2 + \left(\frac{a}{d}\right)^2} - \sigma \right), \tag{6.50}$$

where $a = 0.42\lambda f$ for simplicity.

For the pinhole geometry, the working distance, f, does not have an optimum value and should be minimised, although in reality it is limited by fabrication capabilities. The remaining variables to be optimised are the Lagrange multiplier, Λ, the pinhole diameter, d, and the angular source size, β.

Differentiating equation 6.50 with respect to the Lagrange multiplier and setting it to zero gives

$$\frac{\partial \mathcal{L}}{\partial \Lambda} = (\phi - \sigma) = 0 \implies \phi = \sigma \tag{6.51}$$

Which specifies that the beam standard deviation must be given by the target resolution σ. Differentiating with respect to the angular source size gives

$$\frac{\partial \mathcal{L}}{\partial \beta} = 2\gamma d^2 \beta - \frac{\Lambda}{\sigma}\frac{\beta f^2}{3} = 0, \tag{6.52}$$

and so,

$$\frac{\Lambda}{\sigma} = \frac{6\gamma d_o^2}{f^2}. \tag{6.53}$$

Similarly, differentiating with respect to the pinhole diameter gives

$$\frac{\partial \mathcal{L}}{\partial d} = 2\gamma\beta^2 d - \frac{\Lambda}{\sigma}\left(\frac{d}{12} - \frac{a^2}{d^3}\right) = 0, \tag{6.54}$$

so

$$\frac{\Lambda}{\sigma} = \frac{2\gamma\beta_o^2 d_o^4}{\left(\frac{d_o^4}{12} - a^2\right)}. \tag{6.55}$$

Here, d_o and β_o are the optimum pinhole diameter and angular source size for the given target resolution σ. Equations 6.53 and 6.55 can be combined to eliminate Λ and find an expression that relates the optimum angular source size to the optimum pinhole size

$$\frac{f^2\beta_o^2}{3} = \left(\frac{d_o^2}{12} - \frac{a^2}{d_o^2}\right) \tag{6.56}$$

Substituting back into the definition for σ (equation 6.51) gives the optimum pinhole diameter as

$$d_o = \sqrt{6}\sigma. \tag{6.57}$$

Similarly, the optimum angular source size can be shown to be given by

$$\beta_o = \frac{\sqrt{3}}{f}\left(\frac{\sigma^2}{2} - \frac{a^2}{6\sigma^2}\right)^{\frac{1}{2}}. \tag{6.58}$$

Using equations 6.53 and 6.55 the optimum flux, $F_o = \gamma\beta_o^2 d_o^2$, can be shown to be given by

$$F_o = 3\gamma\left(\frac{3\sigma^4}{f^2} - (0.42\lambda)^2\right). \tag{6.59}$$

6.4.2 Zone Plate Geometry

A similar process can now be applied to analyse the zone plate geometry. In general, the source properties affect the speed ratio, and hence the ultimate resolution through chromatic aberrations. However, these effects are only significant at very small beam standard deviations. At larger beam standard deviations, it is possible to make an approximate separation of the source and zone-plate focussing to obtain approximate analytic expressions for the optimum conditions.

Let the flux through the zone plate be given by

$$F = \gamma\beta^2\eta(2r_N)^2, \tag{6.60}$$

where γ is defined as before. In terms of the minimum feature size, Δr_N, the flux is given by

$$F = \frac{\gamma\eta\lambda^2 f^2}{{\Delta r_N}^2}\beta^2. \tag{6.61}$$

Since chromatic aberrations are negligible at large beam standard deviations, the previous

expression for the total beam standard deviation can be simplified to

$$\phi_{zp} = \sqrt{(0.42\Delta r_N)^2 + \left(\frac{\beta f}{\sqrt{3}}\right)^2}. \tag{6.62}$$

The minimum feature size of the zone plate, Δr_N, is set by the practical limits of fabrication rather than any theoretical limit. Previous zone plates fabricated for microscopy had much larger radii and minimum feature sizes of about 50 nm [58]. Given the zone plates being considered here are much smaller, and the steady progress in nanofabrication technology, the minimum feature size is taken to be 25 nm in the examples below.

Since the speed ratio has been eliminated, the expression for the beam size can be directly substituted into the expression for the flux. Similarly to before, the beam standard deviation, ϕ_{zp}, must be equal to a target resolution, σ, such that $\phi_{zp} = \sigma$. The only remaining free variables are β and f which appear as a pair and can be substituted for. Hence, the expression for the flux in the zone-plate focused beam, in the absence of chromatic aberrations, is

$$F_o = 3\gamma\eta\lambda^2 \left(\frac{\sigma^2}{{\Delta r_N}^2} - (0.42)^2\right). \tag{6.63}$$

However, for the narrow beams required for high resolution helium microscopy, chromatic aberrations are significant and equation 6.63 is inadequate. A full expression for the flux can be obtained by combining equation 6.44 for the flux in the beam, with the empirically fitted function for the speed ratio, given in equation 6.36. It is not possible to analytically solve the resulting set of equations, so a numerical approach must be used. The function `fmincon` in MATLAB was used to perform the numerical constrained optimisation. The function seeks to maximise the flux in the beam, subject to the nonlinear constraint $g = \phi - \sigma = 0$. Instead of directly using the flux and beam standard deviation, the functions are scaled appropriately such that the quantities are $\mathcal{O}(1)$, so the numerical routine can converge more easily. The code to implement this optimisation is provided on GitHub and Zenodo[3].

The numerical approach was validated by comparing with the analytic solution determined for the pinhole geometry; the optimum brightness was substituted into equation 6.59 and used to obtain the optimum flux as a function of the target resolution. The fractional difference between the analytic expression and the fully numerical optimisation is always less than 2×10^{-9}.

6.5 Results and Discussion

6.5.1 Optimised configuration

The results obtained using the optimisation methods above are now presented for a realistic configuration of a scanning helium microscope. To obtain results from the numerical optimisation, values for a specific geometry have been used as justified previously in the work including a working distance of f =1 mm, a zone-plate minimum feature size of

[3] M. Bergin, Zenodo, `https://doi.org/10.5281/zenodo.1227273`

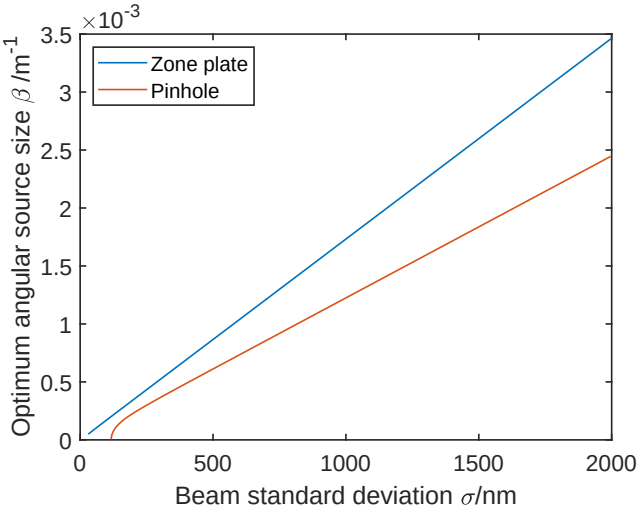


Figure 6.14: Plot of the optimum angular source as a function of beam standard deviation for both the pinhole and zone plate geometries. The pinhole optimum is approximately a straight line at larger beam standard deviations but then drops rapidly as the diffraction limit is reached. The zone plate optimum angular size also follows approximately follows a straight line with a small deviation at small beam standard deviations.

$\Delta r_N = 25\,\mathrm{nm}$ and a zone-plate efficiency of $\eta = 0.05$. Unless otherwise specified, a helium nozzle size of $10\,\mu\mathrm{m}$ is used.

Figure 6.14 shows the optimum angular source size for a zone plate and pinhole, as calculated by the numerical optimisation. The pinhole optimum angular size is the same as the one calculated analytically using equation 6.58. The optimum follows a straight line until diffraction effects become important and the optimum angular size drops more quickly. In Fig. 6.14, the optimum angular source size for a zone plate has a very small deviation from a straight line which occurs at small beam standard deviations due to the chromatic aberrations.

Figure 6.15 shows the optimum pinhole diameter as a function of the target beam standard deviation. The optimum pinhole diameter as calculated from the numerical optimisation fits our analytic expression of $d_o = \sqrt{6}\sigma$ to within a negligible error.

Figure 6.16 shows the variation of the optimum nozzle pressure with target resolution for the zone plate geometry. The horizontal dashed-line is the speed-ratio independent optimum, which applies to both the pinhole and zone plate cases producing a wide beam with a flux given by equation 6.63. The speed ratio independent optimum pressure corresponds to the peak in brightness on the $10\,\mu\mathrm{m}$ line shown in fig. 6.13. The solid curves represent the full numerical solution for the zone plate geometry, which deviate from the approximate expression at small target beam standard deviations. Here, chromatic aberrations become increasingly important in the zone plate focussing, broadening the focussed beam, and requiring a higher speed ratio beam. Achieving that speed ratio requires a higher nozzle pressure, even though the optimum pressure is no longer at the peak brightness.

The predicted optimum beam fluxes are plotted versus target beam standard deviation in Fig. 6.17, for both pinhole and zone plate geometries. The optimum flux for the pinhole setup approximately varies as $F_\mathrm{p} \propto \sigma^4$ (equation 6.59), while for the zone-plate setup approximately scales as $F_\mathrm{zp} \propto \sigma^2$ (equation 6.63). Hence, the pinhole geometry is preferable

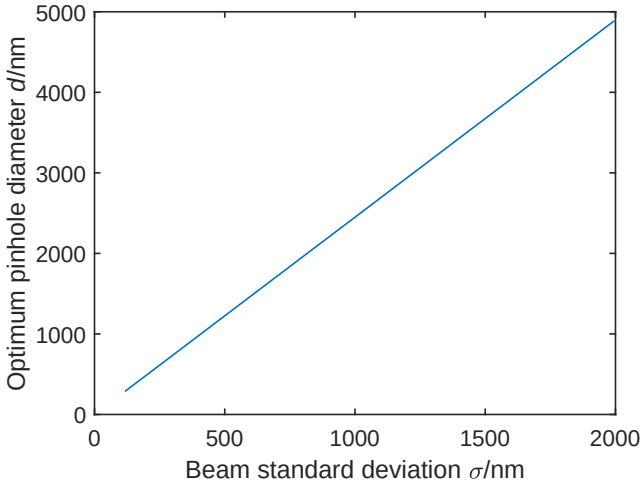


Figure 6.15: Plot of the numerically calculated optimum pinhole diameter as a function of the beam standard deviation for a pinhole geometry. The calculated optimum pinhole diameter follows the analytically expected function $d_o = \sqrt{6}\sigma$ to within 10^{-6} nm.

at larger beam standard deviations, but as the target resolution is reduced the zone plate geometry becomes favourable. The cross over point of the flux can be estimated by equating equation 6.59 and equation 6.63 and solving the resulting quadratic equation. The cross over point is found to be at

$$\sigma_{\text{co}} = \frac{1}{\sqrt{3}}\sqrt{2\eta r_N^2 + \sqrt{\frac{4}{3}\eta^2 r_N^4 + (0.42\lambda f)^2(1-\eta)}}, \tag{6.64}$$

$$\sigma_{\text{co}} \approx 2\sqrt{\frac{\eta}{3}} r_N, \tag{6.65}$$

which corresponds to a beam standard deviation of under 300 nm for the geometry considered. The assumptions made in section 6.3.3 mean that the absolute value of the beam flux may not be accurate, however the relative scaling between the zone plate and pinhole is expected to be more accurate.

6.5.2 Achievable signal to noise

Although the total beam flux is important, when acquiring images in a real microscope the key parameter is the signal to noise ratio achieved per pixel. SHeM images are usually limited by shot noise, so the signal to noise ratio determines the dwell time per pixel required to reduce the noise sufficiently to observe contrast. Images collected by Fahy et al. show that for a sample with large topographic changes and minimal background contributions, the images appear relatively noise free above a signal to noise ratio of 30 [62]. Non-topographic contrast mechanisms will create smaller changes in the scattered helium distribution and so will require a better signal to noise ratio to observe.

Within a certain dwell time, Δt, the total signal per pixel is determined by the total (optimised) flux incident on the sample, F_{tot}, the proportion of atoms accepted by the detector,

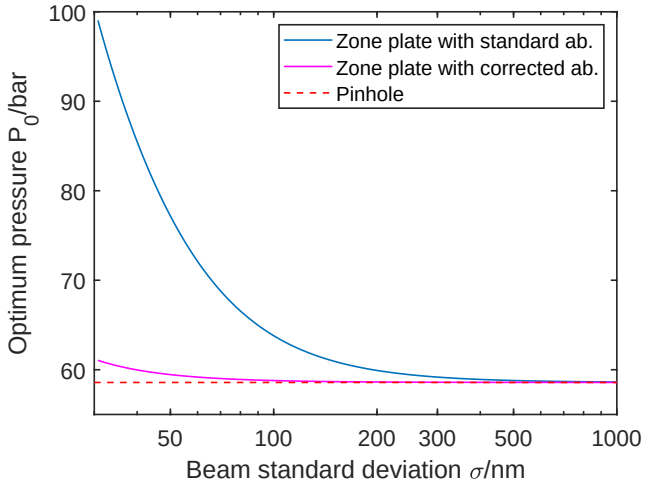


Figure 6.16: Plot of optimum nozzle pressure versus target beam standard deviation for the pinhole geometry (dashed-line) and zone-plate geometry (solid-curves) using the parameters given in the body text. The blue curve is obtained using a chromatic aberration modelled with $\sigma_{chrom} = r_N/(\sqrt{2}S)$, while the purple curve is obtained by using the corrected chromatic aberration term $\sigma'_{chrom} = r_N/(5.066\sqrt{2}S)$. The curves are obtained using the constrained numerical optimisation. At large target beam standard deviations zone-plate chromatic aberrations are insignificant and the optimum pressure corresponds to the brightness maximum in fig. 6.13. At smaller target standard deviations chromatic aberrations become significant and a higher speed ratio is needed, so the optimum pressure shifts from the brightness maximum.

P_{acc}, and the efficiency of the detector, P_{det},

$$N_{\text{tot}} = F_{\text{tot}}\, P_{\text{acc}}\, P_{\text{det}}\, \Delta t. \tag{6.66}$$

In reality, there are a number of experimental complications including instabilities, background gases in the detector, and temperature variations throughout the system, although these effects are neglected here. The probability of atoms scattering into the detector on the current machine in Cambridge can be estimated from the solid angle to be $P_{\text{acc}} \approx 0.5\%$, and a reasonable estimate for next generation detectors gives the detector efficiency as $P_{\text{det}} \approx 0.1\%$.

It should be stressed that in the light of the results presented in chapter 4, it seems unlikely that the background in the detector can be ignored as the flux is reduced. If the efficiency of the detector were $P_{\text{det}} \approx 0.1\%$, then the background would be likely to be of order 10,000 cps which will be very significant when compared to the fluxes being measured with small beam widths. Initially the background will be ignored to focus on the fundamental limits of the machine.

For a pinhole the signal to noise is given by $\sqrt{N_{\text{tot}}}$ since all the flux passing through the pinhole is in the beam. In the absence of any unwanted diffuse beams from insufficient differential pumping, the only helium incident on the surface is in the narrow beam formed by the pinhole.

For a zone-plate, only a small proportion of the total flux contributes to the central spot, while the majority of the signal forms a wide background. The zone plate has multiple focal points with a diminishing fraction of the flux outside of the first order focus, and a

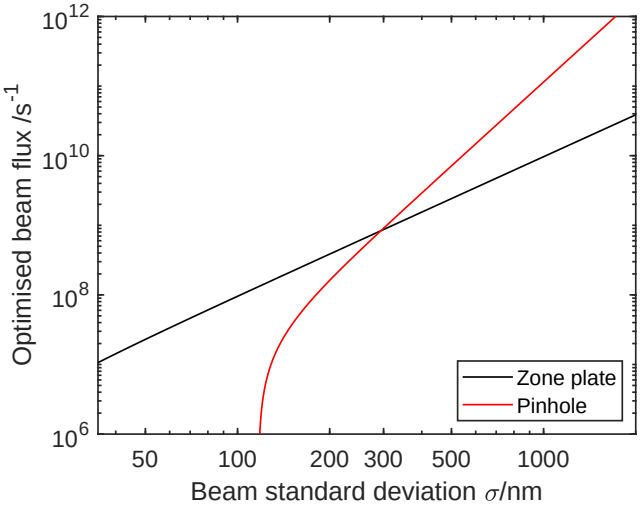


Figure 6.17: Plot of the optimum flux versus target beam standard deviation, obtained from a numerical constrained optimisation of the beam formed, in both zone-plate and pinhole geometries. At larger beam standard deviations the pinhole geometry delivers a greater optimum flux, but as the target standard deviation is reduced, the zone plate becomes the preferred method. The calculation was performed with the parameters described in the body text and a helium wavelength of $\lambda = 0.0563$ nm which corresponds to a room temperature beam.

large component in the zero order of the beam that passes straight through the zone plate. The shot noise that is expected is then given by $\eta N_{\text{tot}}/\sqrt{\alpha N_{\text{tot}}}$ where $\alpha = 0.374$ is the open fraction of the zone plate, and N_{tot} is the number of atoms that would be detected if the zone plate were replaced with a fully open aperture of the same radius. In deriving the expression it is assumed that the zero order beam is not contributing towards imaging and is simply acting as a background. Deconvolution of the diffraction pattern from the images could improve the signal to noise beyond the expression given above.

Figure 6.18a plots the signal to noise ratio expected for both pinhole and zone-plate geometries. For a zone plate without a zero order stop, the signal to noise ratio is reduced and the pinhole becomes the preferred method at lower beam sizes when compared to simply comparing the fluxes.

The performance of zone-plate focussing has been shown to be improved by introducing an aperture after the zone plate, to block the majority of the flux and only allow the central spot through [52]. In the case of a zero order stop, the expected signal to noise ratio is $\sqrt{\eta N_{\text{tot}}}$ since now the only flux of helium incident on the sample is in the main focussed beam.

The black-line in Fig. 6.18a shows the effect of the zero order stop, which increases the signal to noise ratio such that the cross over between the pinhole and zone plate geometry occurs in the same place as the flux comparison.

An approximate expression for the cross over point for the signal to noise ratio between a zone plate with a zero order stop and a pinhole can be derived in the same way as before. Provided the source brightness is still at a maximum and therefore has not changed between the pinhole and zone plate cases (which may not be a good assumption if the chromatic aberrations are significant), the cross over point is given by

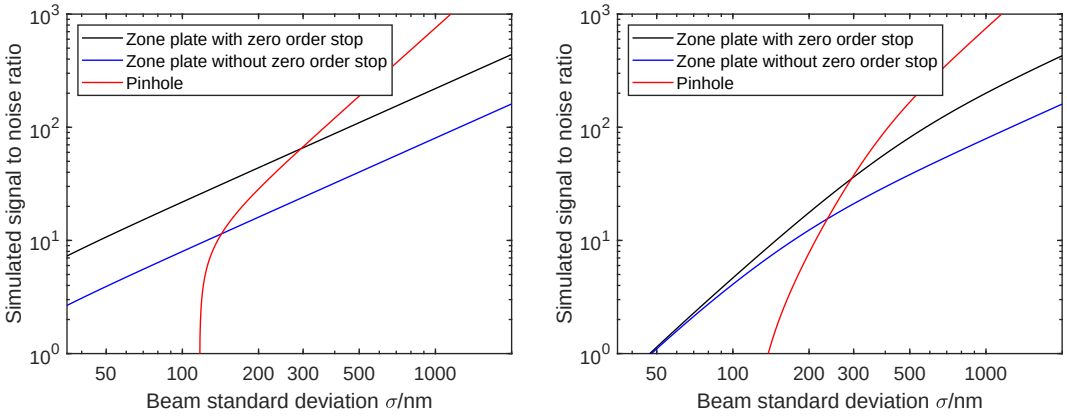


Figure 6.18: (a) Plot of the optimum simulated signal to noise ratio for a dwell time of 1 s as a function of target beam standard deviation for a pinhole and zone plate setup with and without a zero order stop. Without a zero order stop the signal to noise ratio is degraded and so the pinhole remains the preferred choice till smaller beam standard deviations. The calculation was performed with the parameters described in the body text and a helium wavelength of $\lambda = 0.0563\,\text{nm}$. (b) The same calculation as above but with a background count rate of 10,000 cps which significantly degrades the signal to noise ratio at smaller beam widths. Future detectors for use at low beam widths will need to reduce the background or increase the solid angle collection area.

$$\sigma_{\text{co},2} = \frac{1}{\sqrt{3}}\sqrt{2\frac{\eta^2}{\alpha}r_N^2 + \sqrt{\frac{4\eta^4}{3\alpha^2}r_N^4 + (0.42\lambda f)^2(1-\frac{\eta^2}{\alpha})}}, \tag{6.67}$$

$$\sigma_{\text{co},2} \approx \frac{2\eta r_N}{\sqrt{3\alpha}}. \tag{6.68}$$

Figure 6.18b illustrates the signal to noise ratio for the methods described above but with a background of 10,000 cps. The significant degradation in the signal to noise ratio illustrates how a significant background will make it difficult to use a machine with a small beam and therefore a small flux. Therefore if the background that was described in section 4.5.2 cannot be eliminated, the machine will need to be modified to achieve a small beam standard deviation.

Since only $P_{\text{acc}} \approx 0.5\%$ of the scattered atoms are measured, one method to improve the signal to noise would be to reduce the contrast and increase the collected solid angle, thus moving towards the design by Witham and Sánchez [60]. When increasing the detector aperture, some care must be taken to ensure that the pumping speed of the detector is larger than the pumping speed of the detector aperture so that gas flows into the detector (more details on the concept were detailed in section 4.5.1 and fig. 4.41). However if the working distance were also reduced, the pumping speed of the detector aperture could be maintained while increasing the solid angle of the aperture.

6.5.3 Further improvements

So far, a 'realistic optimum' configuration for a scanning helium microscope has been described, including several parameters that were fixed by experimental constraints. It is use-

ful to discuss these constraints further, and the potential for further improvement.

In the pinhole geometry, the working distance affects the optimised flux at the sample strongly, with one term varying as $F_{\mathrm{o,p}} \propto f^{-2}$ (equation 6.59). By lowering the working distance, larger angular source sizes can be used leading to higher fluxes through the pinhole. The microscope developed by Witham et al. [60] uses working distances typically between 10-$50\,\mu\mathrm{m}$ and thus achieves very high fluxes. The trade off for very small working distances is that the depth of field will be reduced, and sample mounting and positioning becomes much more critical. Nonetheless, large gains can potentially be made from the 1 mm reference working distance used here.

In the zone-plate geometry, the minimum feature size is a critical parameter, with one term varying as $F_{\mathrm{o,zp}} \propto {\Delta r_n}^{-2}$ (equation 6.63). Reducing the minimum feature sizes leads to larger area zone plates that collect and focus a greater amount of helium. Larger zone plates lead to more severe chromatic aberrations, however for most resolutions this is not a limiting condition. Gains through the reduction of Δr_n will be limited by the development and advancement of nano-fabrication technologies.

The working distance in the zone plate geometry is a less important criteria, but should still be minimised. When the chromatic aberrations are negligible, the working distance is unimportant. As the working distance is increased, the decrease in the optimum source size is cancelled out by the increase in collection area. However, when chromatic aberrations are important, the optimum flux increases as the working distance decreases. The optimum flux depends asymptotically on the working distance, approaching a maximum value when the chromatic aberrations are negligible.

6.6 Summary and Conclusions

In order to understand the properties of a Fresnel zone plate, a diffraction calculation was implemented. The diffraction integral was solved through a convolution that allowed rapid calculation of the solution. The dependence of the focal spot size on a finite sized source and non-monochromatic beam was parametrised for use in an optimisation.

A method for optimising the geometry of a scanning helium microscope has been achieved by using a constrained optimisation of the flux arriving at a sample. For a pinhole arrangement the process can be performed analytically. In a zone-plate geometry, chromatic aberrations link the source properties with the zone-plate characteristics, so a numerical routine must be used to optimise the entire system. The method provides insight into both the design and operating conditions of an optimised microscope, and the method of constrained optimisation can be applied to more complex models of both the beam intensity and beam standard deviation, as they become available.

The optimisation process demonstrates that the optimum flux drops very rapidly as the beam standard deviation is reduced, with the fourth and second powers of resolution for a pinhole and zone-plate respectively. To enable high-resolution measurements to be obtained within reasonable measurement times, future machines will require small working distances (relevant to pinholes) or fine feature sizes (relevant to zone-plates), combined with a high brightness source (small nozzle, high pressure) and extremely high efficiency detectors with negligible backgrounds as currently being developed [90].

Chapter 7

Conclusions and outlook

In this thesis I have described the design of a scanning helium microscope and how it is possible to use atoms to form images of the surface of a material without any damage. The sample is mounted into the machine using a kinematic mounting principle to ensure that it can be removed and replaced with a high repeatability. The three positioners that are now installed in the machine allow for careful manipulation of the sample and variation in the outgoing scattering angle of the atoms. The concept of a z scan has been introduced and the factors that determine the variation of the measured flux with distance between the pinhole and the sample have been discussed. The virtual source of the helium atoms was measured and found to be compatible with other measurements. The importance of differential pumping was also highlighted to avoid a large effuse beam forming that swamps the direct narrow beam passing through the pinhole.

Future instruments will need to follow the example illustrated by Fahy et al. [62] and improve the differential pumping directly behind the pinhole to boost the observed contrast in images. While the ratio of the direct beam to effusive beam will scale with the size of the pinhole, it will still consistently smear out any subtle contrast changes in images.

The kinematic mount used to support the sample stage provides a stable base for the three positioners that are currently being used on the machine. In the future it should be possible to install additional positioners to achieve rotation and tilting of the sample. Some care must be taken to ensure that the additional weight of any extra manipulators does not overwhelm the delicate slip-stick drive used for vertical positioning. Further improvements to the sample stage also include temperature control which opens a huge avenue of potential experiments including investigating the effects of adsorbates, samples such as ice that are not stable at room temperature and perhaps even phase transitions.

A new detector has been implemented on the microscope that provides an improvement to the detection efficiency of helium atoms and therefore also a large improvement in the signal to noise ratio in images. Many of the long standing issues with solenoidal ionisers that have been developed in the group, in particular the behaviour of a conversion dynode and electron multiplier, are now far better understood. There are two remaining issues with the detector, one is the complex behaviour of the ionisation efficiency and the other is the observed background in the detector. It seems likely that the complex behaviour of the ionisation efficiency is due to the space charge not being sufficiently controlled in the ioniser. Therefore higher efficiencies seem possible given some changes to the ioniser.

Addressing the background is the most important remaining issue with the detector since it will make the measurement of lower fluxes extremely difficult. The predictions made on the obtainable signal to noise ratio of a detector in chapter 6 assume that there is no background present in the detector. Without removal of the background, it will be difficult to imagine an instrument that can collect images in a reasonable time if the helium flux were further reduced. Since the background seems to depend on the pressure in the ioniser, it will be necessary to reduce the pressure of background gas. However, the detector relies on a low pumping speed at the front to maximise the ionisation efficiency, so the additional pumping will need to be selective in nature so that it removes contaminants and not helium. One potential method would be to coat the liner with a getter material to absorb the background gas.

The contrast mechanisms for rough surfaces have been discussed showing that masking, shadowing and diffuse illumination are important processes. The scattering distribution from rough surfaces is similar to a cosine distribution and therefore variations in the angle that the surface normal makes with the detector also creates contrast in images. These mechanisms are also present even if the structure is smaller than the beam width and cannot be spatially resolved. Given that sub-beam structure is important, the roughness of a sample was shown to be significant which has important consequences for previously reported chemical contrast.

Two forms of novel contrast have been observed and described in chapter 5 for the first time. A mixed beam was used to simultaneously image a surface with two atom beams at different wavelengths and energies simultaneously. A mixed beam of argon and helium was used, however the presence of a background at mass 40 made the measurements difficult. If a mixed beam were to be used in the future, an inert gas will need to be used that has a mass with a low background in vacuum systems. Diffraction contrast was also observed for the first time from a LiF surface by observing structure in a z scan that was not present with rough surfaces.

The measurements of diffraction contrast were obtained with a relatively large detector aperture, meaning that direct evaluation of the diffraction pattern was difficult since many of the diffraction peaks were simultaneously measured. However, if the detector aperture were made smaller and a rotation stage were added so that the azimuthal angle could be varied it would be possible to measure the diffraction pattern from a small area of a sample. These adaptations could lead to an exciting application of the microscope to measure the atomic structure of a polycrystalline material using helium atoms.

Finally, an optimisation of the geometry of a helium microscope with either a pinhole or zone plate was performed using Lagrange multipliers in chapter 6. First the diffraction from a zone plate was computationally solved to obtain a model for the width of the focussed spot produced with a realistic design for the zone plate. Since the zone plate has chromatic aberrations and the source can either have a high brightness or monochromaticity it is necessary to optimise both the microscope geometry and source properties simultaneously. It was found that for small beam widths, using a zone plate will produce the highest flux, however a pinhole is the optimum method for larger beam widths. In general, it was found that a reduction of the distance between the sample and the pinhole increase the flux in the helium spot.

The method used in chapter 6 does not rely on the models used for the centreline inten-

sity or beam width from the zone plate or pinhole. Instead, the method describes a general procedure for balancing the competing goals of maximising the flux and reducing the beam width. As more experimental data becomes available for sources that are intended to maximise the brightness for microscopy, it will be possible to feed the data into the optimisation method described in chapter 6 to obtain more accurate design criteria for the microscope.

In conclusion, scanning helium microscopy is a rapidly advancing technique which potentially allows the study of a wide range of materials with unique contrast mechanisms. Continued development of detectors will allow the resolution to be improved, allowing the study of materials on length scales unobtainable with a conventional optical microscope. Additionally, as the contrast from conventional mechanisms like topography are better understood, it should become easier to search for novel forms of contrast, unlocking new applications for scanning helium microscopy.

Appendix A

Practical implementation of the detector

Interlock for protecting solenoid from loss of water cooling

The solenoid on the detector dissipates a large amount of energy and requires water cooling to keep the temperature down. Accidental loss of water cooling pressure while a current flows through the solenoid poses a significant risk which could cause serious damage.

A solenoid interlock was designed by Bill Allison to ensure that in the event of a loss of cooling water the solenoid is turned off. The power supply for the solenoid is a TDK power supply which can be set so that it will only provide a current if there is an electrical short between pins 1 and 14 on the J1 port.

A protection box was built that contains both a temperature and flow switch to ensure that cool water is passing through. The flow switch is a Cynergy3, 2 L/min Flow Controller (RS Stock No. 257-076) which ensures that there is a continuous flow of water into the solenoid. The temperature switch is a Honeywell NC 10 A Bi-Metallic Thermostat (RS Stock No. 228-2636) which opens at 50° ensuring that the cooling water does not get too hot. The temperature switch is mounted to the pipe through a metal block to ensure that it has a good thermal contact with the cooling water.

By wiring both the switches in series and connecting them between pins 1 and 14 on the J1 port on the power supply, it is then possible to ensure that the solenoid will only ever have a current flowing through it if the cooling water is flowing and is not too hot.

Flyback diode to prevent damage to the solenoid power supply

The solenoid acts as a large inductor and therefore changes in the current flow will generate a back emf. Rapidly changing the current flow through the solenoid, for example in a loss of power or accidental unplugging of a connector, will create a large back emf that could damage the power supply.

To protect the power supply from being damaged, a flyback diode (1N5401) was installed across the terminals on the power supply. In normal operation no current will flow through the diode and it will not affect the current flow through the solenoid. However, if a large back emf is generated the diode will then allow current flow through it, avoiding any current flow through the power supply that may damage it.

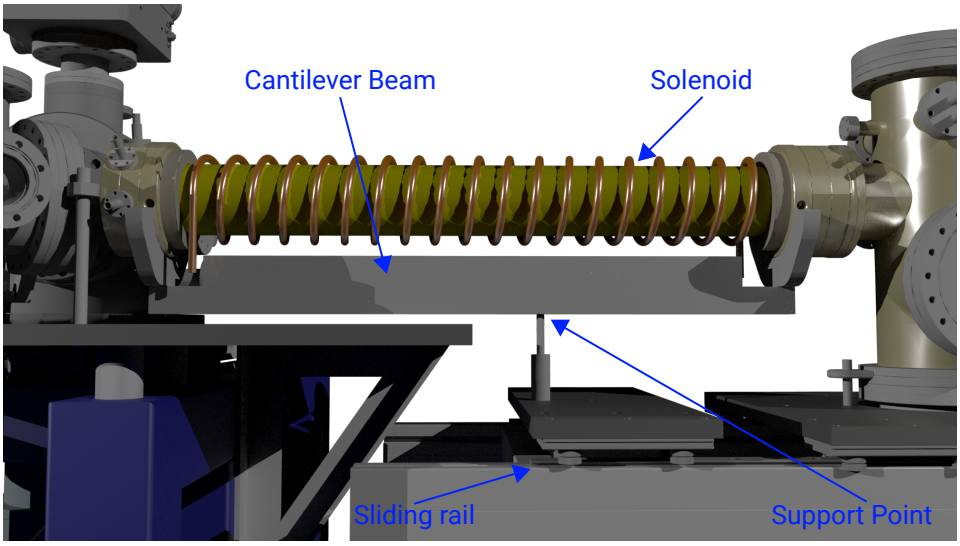


Figure A.1: Render of the detector as installed on the microscope illustrating the metal bar used as a cantilever to support the solenoid.

Mechanical support of the solenoid

Figure A.1 shows how in order to facilitate easy changes to the front of the detector, the detector was installed onto a set of sliding rails to easily remove it from the rest of the machine. The entire support for the detector is built from components from Hepco Motion including both the slides (GV3 size 25 slide) and the actual frame (45x90 profile with Rapilok).

The front of the detector needs to be attached directly to the sample chamber and therefore cannot be supported directly from below and requires a cantilever for support.

Figure A.1 is a render showing how a cantilever beam is used to support the front of the detector. The solenoid rests on an aluminium beam which then is held in position by an M12 bolt that attaches to a sliding plate to allow the detector to be retracted.

Figure A.2a) shows the forces on the whole detector and the support bar. The weight of the solenoid, Mg, is balanced by the support force from the M12 bolt, S which is displaced by a distance x from the centre of the solenoid. However, in order to balance the moments, a reaction force from the detector chamber is also needed which is provided by the weld on the main detector chamber which is a distance x_2 from the end of the solenoid.

The support force from the M12 bolt can be calculated by balancing the forces,

$$S = Mg + C, \tag{A.1}$$

and the moments about the support force,

$$Mg\left(\frac{L}{2} + y\right) = S\left(\frac{L}{2} + x_2 - x\right). \tag{A.2}$$

Solving the simultaneous equations gives the reaction force as,

$$S = Mg\frac{L/2 + x_2}{L/2 + x_2 - x}. \tag{A.3}$$

Assuming the mass of the solenoid is $M = 60\,\mathrm{kg}$, the length of the solenoid is $L = 826\,\mathrm{mm}$,

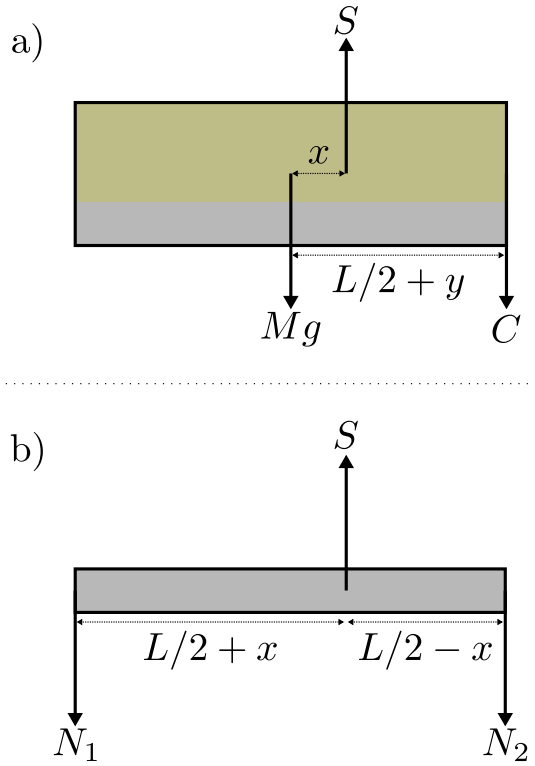


Figure A.2: a) A force diagram of the solenoid and the support bar. b) A force diagram of the support bar

the support is $x = 88$ mm from the centre of mass and the detector weld is $x_2 = 128.5$ mm from the end of the solenoid then the reaction force is given by $S = (71.6)g$. Therefore, there is an 11.6 kg force on the weld on the detector chamber which is acceptable.

Figure A.2b) then shows the forces acting on just the support bar. The large force on the end of the bar may cause it to deflect, so the bending of the bar needs to be analysed. The force on the support bar at the front of the detector N_1, can be calculated by balancing the linear forces,

$$N_1 + N_2 = S, \tag{A.4}$$

and then balancing the moments about the support force,

$$N_1 \left(\frac{L}{2} + x \right) = N_2 \left(\frac{L}{2} - x \right). \tag{A.5}$$

Solving the two simultaneous equations gives the force on the end of the support bar as,

$$N_1 = S \left(\frac{1}{2} - \frac{x}{L} \right). \tag{A.6}$$

For the parameters given earlier, the force is given by $N_1 = (28.2)g$.

The deflection of the bar, y, can then be calculated by solving the bending equation,

$$EIy'' = B(z), \tag{A.7}$$

where E is the Young's modulus of the bar, I is the area moment of inertia of the bar and

$B(z)$ is the bending moment at a position, z, down the length of the bar. The area moment of inertia for a rectangular bar is given by $I = bh^3/12$, where b is the width of the bar and h is the height in the dimension that the bar deflects in. Solving the resulting differential equation with the condition that at the support point the deflection and the first differential are zero gives the deflection at the end of the bar, $y_d = y(0)$, as,

$$y_d = -\frac{(L/2 + x)^3 N_1}{EI}. \tag{A.8}$$

The first iteration of the support bar was made from a 1" x 1" aluminium bar, which visibly deflected under the weight of the solenoid. Using the above equations, the expected deflection is given by $y_d = 15\,\text{mm}$ which agrees with the clear visible deflection that was observed.

Since the moment of inertia scales as h^3, the height of the bar was increased to ensure that the deflection is minimal. By instead using a 3" x 1" support bar, the deflection is expected to only be $y_d = 0.5\,\text{mm}$ which is acceptable. Using a stainless steel was an option, but the Young's modulus is only about two times higher and it makes machining much more difficult so the bar was kept as aluminium.

Appendix B

Circuit diagram for measuring the emission current

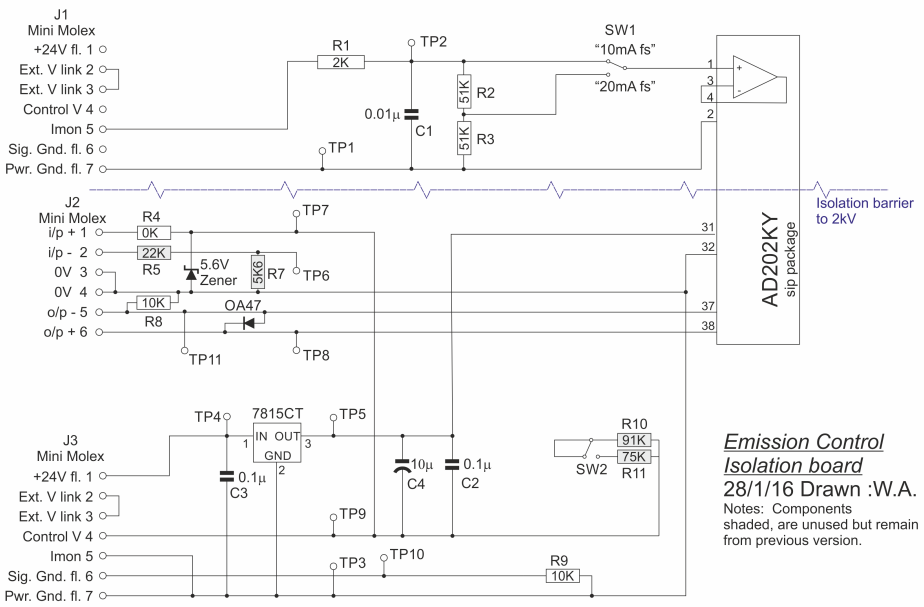


Figure B.1: Circuit diagram for the modified printed circuit board used to measure the emission current with the detector in Cambridge. To maximise the emission current that can be measured the potential divider at the top of the diagram to achieve a full scale of about 20 mA. To achieve filament currents larger than 2.5 A, an op amp is used as an x2 amplifier between control V and ground on the bottom molex connector leading to the filament power supply. Diagram drawn by Bill Allison.

Appendix C

Time response of the pressures at the front of the detector

The pressures at the front of the detector are determined by the differential equations in equations 4.15 and 4.16. The steady state properties can be extracted by setting the time differentials to zero and solving the resulting simultaneous equations. The pressure p_1 in the pipe to the detector is given by,

$$\begin{aligned} \lim_{t\to\infty} p_1(t) &= \frac{Q}{S_{\text{eff}}} \\ &= \frac{Q}{S_s + (1/C_p + 1/S_i)^{-1}}. \end{aligned} \tag{C.1}$$

The steady state gas flux that flows down the detector to the pump Q_{out} is given by,

$$\begin{aligned} \lim_{t\to\infty} Q_{\text{out}}(t) &= Q\,\frac{(1/C_p + 1/S_i)^{-1}}{S_s + (1/C_p + 1/S_i)^{-1}} \\ &= Q\,\frac{1}{S_s\ (1/C_p + 1/S_i) + 1}. \end{aligned} \tag{C.2}$$

Finally, the steady state pressure at the front of the ioniser p_2 is given by,

$$\begin{aligned} \lim_{t\to\infty} p_2(t) &= \frac{Q_{\text{out}}}{S_i} \\ &= Q\,\frac{1}{S_s\ (S_i/C_p + 1) + S_i}. \end{aligned} \tag{C.3}$$

The full time response of the pressures is plotted in fig. C.1, showing that there is a difference in the pressure evolution just after the gas flux is changed. The steady state pressures are also plotted and provide evidence that the time evolution predicted by equations 4.15 and 4.16 is correct. A single exponential can be fitted to the pressure variation in the front of the ioniser after 45 ms that gives a time constant of 256.7 ± 0.2 ms.

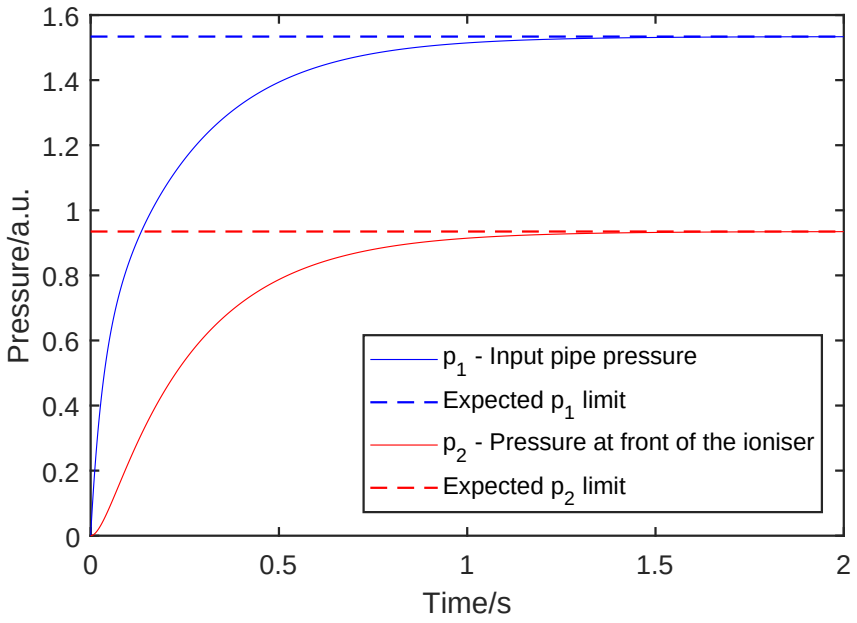


Figure C.1: Simulated time variation of the pressure in the input pipe to the detector and the pressure at the front of the detector. Just after the flux is changed there is a significant difference in the time response due to the gas taking time to diffuse into the detector. However at larger times the time evolution both seem to follow exponential curves.

Appendix D

Connection between the Sikora model and a virtual source model

The Sikora model was simultaneously dervied by Anderson and Sikora in his PhD thesis [145] in 1973, and describes the on axis intensity from a free jet expansion that incorporates the physical size of the source and skimmer. Figure D.1 illustrates the main features of the model, that due to the source not being a point like object, the addition of a collimator (or skimmer as it is typically referred to now) will reduce the total intensity that would reach an on-axis detector. In the Sikora model, we consider the atoms to transition into molecular flow at the quitting surface, and the velocity distribution of the atoms can be determined from the speed ratio. Sikora then describes a series of complex approximations to solve the integrals and obtains an expression for the on axis intensity I as,

$$\left(\frac{I}{I_0}\right) \approx \left[1 - \exp\left(-S^2\left(\frac{r}{x_1}\right)^2\left(\frac{x_D}{x_D - x_C}\right)^2\right)\right], \tag{D.1}$$

where I_0 is the on axis intensity without any collimation, S is the speed ratio, r is the radius of the collimating aperture (skimmer), x_1 is the radius of the quitting surface, x_D is the distance from the nozzle to the detector and x_C is the distance from the nozzle to the collimator.

The Sikora model can be extended by including that the velocity distribution at the quitting surface depends on the direction. In the ellipsoidal velocity distribution model, the velocity distributions are different when considering the velocity either parallel or perpendicular to the streamlines. More recently, the method of tracing particles from a quitting surface through the skimmer has also been investigated numerically by Hustler-Wraight in his thesis [154].

While this model is useful, the quitting surface is a theoretical concept that cannot be experimentally measured, and reformulating the model into different terms can be more useful. The virtual source model can be experimentally measured and is therefore a more useful quantity to base the model on. A model for the intensity after collimation is easier to derive with this simpler model, and as we shall see the equation has the same form as the Sikora model.

The brightness of the source, B, is defined as the on-axis intensity as measured from

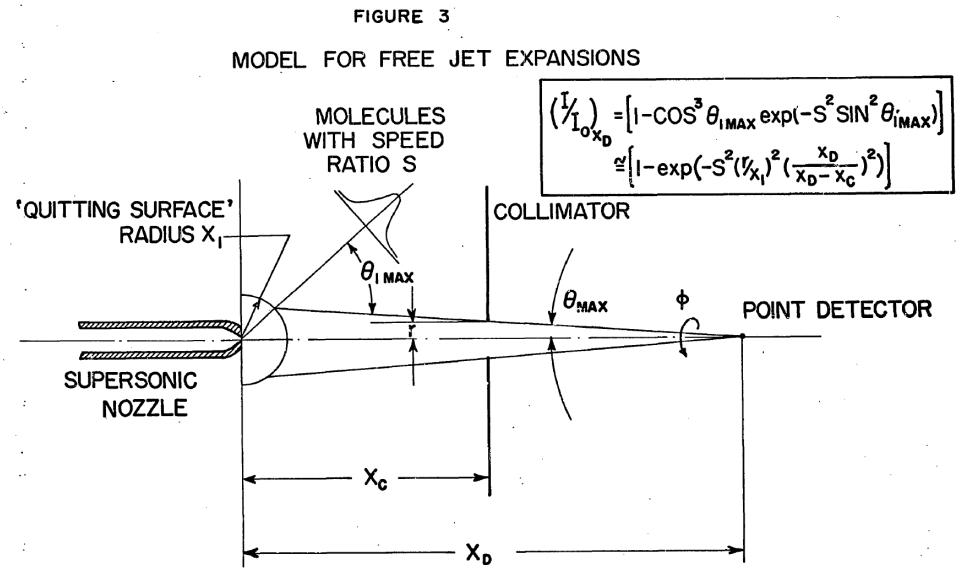


Figure D.1: Illustration of the main features in Sikora model for the intensity, taken from fig. 3 in Sikora's PhD thesis [145].

an infinitesimal part of the skimmer, so that $dI = B\,dA$ where dA is an infinitesimal area of the virtual source that is visible through the skimmer. The brightness in a supersonic expansion is usually described with a 2D Gaussian function,

$$B = Ke^{-\frac{\rho^2}{2\sigma^2}}, \tag{D.2}$$

where ρ is the radial distance from the centre of the nozzle, σ is the virtual source size as measured as the standard deviation of a Gaussian function and K is a normalisation factor for the function which is given by

$$K = \frac{I_0}{2\pi\sigma^2}. \tag{D.3}$$

Following the nomenclature used in Sikora's derivation, if a skimmer of radius r is used, then by considering similar triangles from the detector, the visible part of the virtual source has a size,

$$r_{\text{pro}} = \frac{x_D}{x_D - x_C} r. \tag{D.4}$$

The on-axis intensity can then be calculated by integrating the brightness,

$$I = \int B\,dA, \tag{D.5}$$

$$I = \int^{r_{\text{pro}}} Ke^{-\frac{\rho^2}{2\sigma^2}} \rho\,d\rho\,d\theta. \tag{D.6}$$

Performing a substitution of variables and computing the integral gives,

$$I = 2\pi K\sigma^2 \left[1 - \exp\left(-\frac{r^2}{2\sigma^2}\left(\frac{x_D}{x_D - x_C}\right)^2\right)\right], \tag{D.7}$$

$$I = I_0 \left[1 - \exp\left(-\frac{r^2}{2\sigma^2} \left(\frac{x_D}{x_D - x_C} \right)^2 \right) \right]. \tag{D.8}$$

We can then relate this back to a quitting surface model using the expressions dervied by Beijerinck et al. [72]. The authors show in eq. (18), that for a quitting surface with a velocity component $v_\perp(z)$ perpendicular to the streamlines, characterised by $\alpha_\perp(z) = (2kT_\perp(z)/m)^{1/2}$, and assuming that all the molecules have the same parallel velocity $v_\parallel = u_\infty$ we find that the virtual source radius is given by,

$$R \equiv \sqrt{2}\sigma = \frac{z\,\alpha_\perp(z)}{u_\infty}. \tag{D.9}$$

Where z is the distance downstream from the nozzle. Beijerinck et al. also show in eq. (25) that the speed ratio is given by,

$$S = \frac{u_\infty}{\alpha(x_1)}. \tag{D.10}$$

However since the Sikora model only uses one temperature to model the behaviour at the quitting surface, we can then substitute into eq. D.9 and obtain,

$$R \equiv \sqrt{2}\sigma = \frac{x_1}{S}. \tag{D.11}$$

Therefore, the intensity expression derived in eq. D.8 is equivalent to the Sikora model in eq. D.1.

Additionally, the full expression for a finite size aperture is relatively complex and the exponential term means that it is difficult to obtain analytic expressions for the optimum conditions. Instead, consider the limit of a small skimmer and $x_D >> x_C$, so that after Taylor expanding the exponential function the intensity is approximately given by,

$$I \approx I_0 \frac{r^2}{2\sigma^2}, \tag{D.12}$$

which after some rearrangement with a factor of π gives the product of the normalisation constant K of the brightness and the size of the skimmer. Therefore it is possible to vastly reduce the complexity of the Sikora model in the limit of small skimmers since the brightness is approximately constant in the centre of the expansion.

Appendix E

Advantage of using Lagrange multipliers for constrained optimisation

In general, we will construct models for both the total flux passing through an atom optics element, $F = f(\underline{x})$, and the size of the beam produced by the element, $\sigma = \phi(\underline{x})$. Typically, larger beam sizes will have higher fluxes so these functions will oppose each other. Kaltenbacher [146] attempts to maximise the beam flux while simultaneously minimising the beam size, however these opposing objectives will only produce Pareto-optimal solutions.

Instead, let us attempt to maximise the total flux passing through the atom optics element for a fixed beam width. So we are attempting to solve

$$\begin{aligned} &\text{maximise} \quad F = f(\underline{x}), \\ &\text{subject to} \quad \sigma = \phi(\underline{x}). \end{aligned} \tag{E.1}$$

Palau et al. [143, 144] perform the constrained optimisation using the substitution method. If the constraint is a relatively simple function, then it may be possible to parametrise one of the variables in terms of the constraint so that instead of $\sigma = \phi(x_1, x_2, ..., x_n)$ we obtain,

$$x_1 = h(\sigma, x_2, x_3, ..., x_n). \tag{E.2}$$

We can then substitute back into our equation for the flux, and now solve the new unconstrained problem

$$\text{maximise} \quad F = f(h(\sigma, x_2, x_3, ..., x_n), x_2, x_3, ...x_n). \tag{E.3}$$

The substitution approach is only possible if it is possible to rearrange the original constraint into a new form in terms of one of the variables. Some care does need to be taken with this approach, you cannot simply eliminate the original constraint without considering if x_1 can take any value or does also carry a constraint (e.g. $x_1 > 0$).

A more elegant, general solution is to use Lagrange multipliers to solve the constrained

optimisation problem. We create a Lagrange function, $\mathcal{L}$, that is defined as,

$$\mathcal{L} = f(\underline{x}) - \Lambda\left(\phi(\underline{x}) - \sigma\right). \tag{E.4}$$

Now, we look to find the stationary points of the Lagrange function by solving the set of equations given by,

$$\frac{\partial \mathcal{L}}{\partial \underline{x}} = 0. \tag{E.5}$$

By solving this set of equations, we can find the constrained optimum of the flux, but the constraint is easily incorporated and we can avoid explicit parametrisation in terms of the constraints which may not be possible in a complex model that could be produced from experimental data.

Bibliography

[1] R. Hooke, *Micrographia.* (Royal Society, London, 1667).

[2] J. Zhang, P. Chen, B. Yuan, W. Ji, Z. Cheng, and X. Qiu, "Real-Space Identification of Intermolecular Bonding with Atomic Force Microscopy," Science **342**, 611–614 (2013).

[3] G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, "Surface Studies by Scanning Tunneling Microscopy," Physical Review Letters **49**, 57–61 (1982).

[4] G. Binnig, C. F. Quate, and C. Gerber, "Atomic Force Microscope," Physical Review Letters **56**, 930–933 (1986).

[5] U. Dürig, D. W. Pohl, and F. Rohner, "Near-field optical-scanning microscopy," Journal of Applied Physics **59**, 3318–3327 (1986).

[6] D. Attwood, *Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*, 1st ed. (Cambridge University Press, Cambridge, 2007).

[7] D. B. Williams and C. B. Carter, *Transmission electron microscopy: a textbook for materials science / David B. Williams and C. Barry Carter.* (Plenum, New York, New York ; London, 1996).

[8] A. M. Donald, "The use of environmental scanning electron microscopy for imaging wet and insulating materials," Nature Materials **2**, 511–516 (2003).

[9] N. Kardjilov, I. Manke, A. Hilger, M. Strobl, and J. Banhart, "Neutron imaging in materials science," Materials Today **14**, 248 – 256 (2011).

[10] J. G. King and W. R. Bigas, "Molecular Scanner," Nature **222**, 261–263 (1969).

[11] Y. Harada, S. Yamamoto, M. Aoki, S. Masuda, T. Ichinokawa, M. Kato, and Y. Sakai, "Surface spectroscopy with high spatial resolution using metastable atoms," Nature **372**, 657–659 (1994).

[12] S. Yamamoto, S. Masuda, H. Yasufuku, N. Ueno, Y. Harada, T. Ichinokawa, M. Kato, and Y. Sakai, "Study of solid surfaces by metastable electron emission microscopy: Energy-filtered images and local electron spectra at the outermost surface layer of silicon oxide on Si(100)," Journal of Applied Physics **82**, 2954–2960 (1997).

[13] D. Farias and K.-H. Rieder, "Atomic beam diffraction from solid surfaces," Reports on Progress in Physics **61**, 1575 (1998).

[14] A. Tamtögl, E. Bahn, J. Zhu, P. Fouquet, J. Ellis, and W. Allison, "Graphene on Ni(111): Electronic Corrugation and Dynamics from Helium Atom Scattering," The Journal of Physical Chemistry C **119**, 25983–25990 (2015).

[15] G. Benedek and J. P. Toennies, "Helium atom scattering spectroscopy of surface phonons: genesis and achievements," Surface Science **299–300**, 587–611 (1994).

[16] A. P. Jardine, J. Ellis, and W. Allison, "Quasi-elastic helium-atom scattering from surfaces: experiment and interpretation," Journal of Physics: Condensed Matter **14**, 6173 (2002).

[17] A. P. Jardine, H. Hedgeland, G. Alexandrowicz, W. Allison, and J. Ellis, "Helium-3 spin-echo: Principles and application to dynamics at surfaces," Progress in Surface Science **84**, 323–379 (2009).

[18] D. J. Riley, M. Mann, D. A. MacLaren, P. C. Dastoor, W. Allison, K. B. K. Teo, G. A. J. Amaratunga, and W. Milne, "Helium Detection via Field Ionization from Carbon Nanotubes," Nano Letters **3**, 1455–1458 (2003).

[19] D. A. Maclaren, *Development of a single crystal mirror for scanning helium microscopy*, Ph.D., University of Cambridge (2002).

[20] D. A. MacLaren, B. Holst, D. J. Riley, and W. Allison, "Focusing Elements and Design Considerations for a Scanning Helium Microscope (SHeM)," Surface Review and Letters **10**, 249–255 (2003).

[21] C. S. Adams, M. Sigel, and J. Mlynek, "Atom optics," Physics Reports **240**, 143–210 (1994).

[22] C. J. Foot, *Atomic physics / C.J. Foot.*, Oxford master series in physics (Oxford University Press, Oxford ; New York, 2005).

[23] M. Kohler, H. Daerr, P. Sahling, C. Sieveke, N. Jerschabek, M. B. Kalinowski, C. Becker, and K. Sengstock, "All-optical production and trapping of metastable noble-gas atoms down to the single-atom regime," EPL (Europhysics Letters) **108**, 13001 (2014).

[24] J.-W. Ahn, D. Craig, G. Fiksel, D. J. Den Hartog, J. K. Anderson, and M. G. O'Mullane, "Emission intensities and line ratios from a fast neutral helium beam," Physics of Plasmas **14**, 083301 (2007).

[25] W. Kaiser and C. G. B. Garrett, "Two-photon excitation in CaF_2: Eu^{2+}," Physical Review Letters **7**, 229–231 (1961).

[26] A. L'Huillier, L. A. Lompre, G. Mainfray, and C. Manus, "Multiply charged ions induced by multiphoton absorption processes in rare-gas atoms at 1.064μm," Journal of Physics B: Atomic and Molecular Physics **16**, 1363 (1983).

[27] G. S. Hurst, M. G. Payne, M. H. Nayfeh, J. P. Judish, and E. B. Wagner, "Saturated two-photon resonance ionization of $He(2^1S)$," Physical Review Letters **35**, 82–85 (1975).

[28] S. S. Hodgman, R. G. Dall, L. J. Byron, K. G. H. Baldwin, S. J. Buckman, and A. G. Truscott, "Metastable Helium: A New Determination of the Longest Atomic Excited-State Lifetime," Physical Review Letters **103**, 053002 (2009).

[29] T. Sleator, T. Pfau, V. Balykin, and J. Mlynek, "Imaging and focusing of an atomic beam with a large period standing light wave," Applied Physics B **54**, 375–379 (1992).

[30] H.-R. Noh, K. Shimizu, and F. Shimizu, "Imaging of an atomic beam with electrostatic lenses," Physical Review A **61**, 041601 (2000).

[31] W. G. Kaenders, F. Lison, A. Richter, R. Wynands, and D. Meschede, "Imaging with an atomic beam," Nature **375**, 214–216 (1995).

[32] W. G. Kaenders, F. Lison, I. Müller, A. Richter, R. Wynands, and D. Meschede, "Refractive components for magnetic atom optics," Physical Review A **54**, 5067–5075 (1996).

[33] R. Chaustowski, V. Leung, and K. Baldwin, "Magnetic hexapole lens focusing of a metastable helium atomic beam for UV-free lithography," Applied Physics B **86**, 491–496 (2007).

[34] A. P. Jardine, P. Fouquet, J. Ellis, and W. Allison, "Hexapole magnet system for thermal energy 3he atom manipulation," Review of Scientific Instruments **72**, 3834–3841 (2001).

[35] R. Castillo-Garza, J. Gardner, S. Zisman, and M. G. Raizen, "Nanoscale Imaging of Neutral Atoms with a Pulsed Magnetic Lens," ACS Nano **7**, 4378–4383 (2013).

[36] J. R. Gardner, E. M. Anciaux, and M. G. Raizen, "Communication: Neutral atom imaging using a pulsed electromagnetic lens," The Journal of Chemical Physics **146**, 081102 (2017).

[37] C. G. Aminoff, A. M. Steane, P. Bouyer, P. Desbiolles, J. Dalibard, and C. Cohen-Tannoudji, "Cesium atoms bouncing in a stable gravitational cavity," Physical Review Letters **71**, 3083–3086 (1993).

[38] I. G. Hughes, P. A. Barton, T. M. Roach, M. G. Boshier, and E. A. Hinds, "Atom optics with magnetic surfaces: I. Storage of cold atoms in a curved 'floppy disk'," Journal of Physics B: Atomic, Molecular and Optical Physics **30**, 647 (1997).

[39] H. C. Schewe, B. S. Zhao, G. Meijer, and W. Schöllkopf, "Focusing a helium atom beam using a quantum-reflection mirror," New Journal of Physics **11**, 113030 (2009).

[40] T. A. Pasquini, Y. Shin, C. Sanner, M. Saba, A. Schirotzek, D. E. Pritchard, and W. Ketterle, "Quantum Reflection from a Solid Surface at Normal Incidence," Physical Review Letters **93**, 223201 (2004).

[41] J. Lapujoulade, Y. Lejay, and G. Armand, "The thermal attenuation of coherent elastic scattering of noble gas from metal surfaces," Surface Science **95**, 107 – 130 (1980).

[42] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light / by Max Born and Emil Wolf with contributions by A.B. Bhatia ... [et al.]*., 6th ed. (Pergamon Press, Pergamon, Oxford ; New York, Oxford, 1980).

[43] B. Holst and W. Allison, "An atom-focusing mirror," Nature **390**, 244–244 (1997).

[44] K. Fladischer, H. Reingruber, T. Reisinger, V. Mayrhofer, W. E. Ernst, A. E. Ross, D. A. MacLaren, W. Allison, D Litwin, J. Galas, S. Sitarek, P. Nieto, D. Barredo, D. Farías, R. Miranda, B. Surma, A. Miros, B. Piatkowski, E Søndergård, and B. Holst, "An ellipsoidal mirror for focusing neutral atomic and molecular beams," New Journal of Physics **12**, 033018 (2010).

[45] D. Barredo, F. Calleja, A. E. Weeks, P. Nieto, J. J. Hinarejos, G. Laurent, A. L. Vazquez de Parga, D. A. MacLaren, D. Farías, W. Allison, and R. Miranda, "Si(111)–H(1×1): A mirror for atoms characterized by AFM, STM, He and H2 diffraction," Surface Science **601**, 24–29 (2007).

[46] G. Anemone, A. A. Taleb, S. D. Eder, B. Holst, and D. Farías, "Flexible thin metal crystals as focusing mirrors for neutral atomic beams," Physical Review B **95**, 205428 (2017).

[47] P. Sutter, M. Minniti, P. Albrecht, D. Farías, R. Miranda, and E. Sutter, "A high-reflectivity, ambient-stable graphene mirror for neutral atomic and molecular beams," Applied Physics Letters **99**, 211907 (2011).

[48] E. Hecht, *Optics*, 4th ed. (Addison-Wesley, Reading, Mass, 2001).

[49] O. Carnal, M. Sigel, T. Sleator, H. Takuma, and J. Mlynek, "Imaging and focusing of atoms by a fresnel zone plate," Physical Review Letters **67**, 3231–3234 (1991).

[50] R. B. Doak, R. E. Grisenti, S. Rehbein, G. Schmahl, J. P. Toennies, and C. Wöll, "Towards Realization of an Atomic de Broglie Microscope: Helium Atom Focusing Using Fresnel Zone Plates," Physical Review Letters **83**, 4229–4232 (1999).

[51] S. D. Eder, T. Reisinger, M. M. Greve, G. Bracco, and B. Holst, "Focusing of a neutral helium beam below one micron," New Journal of Physics **14**, 073014 (2012).

[52] S. D. Eder, A. K. Ravn, B. Samelin, G. Bracco, A. S. Palau, T. Reisinger, E. B. Knudsen, K. Lefmann, and B. Holst, "Zero-order filter for diffractive focusing of de Broglie matter waves," Physical Review A **95**, 023618 (2017).

[53] I. Mohacsi, P. Karvinen, I. Vartiainen, V. A. Guzenko, A. Somogyi, C. M. Kewish, P. Mercere, and C. David, "High-efficiency zone-plate optics for multi-keV X-ray focusing," Journal of Synchrotron Radiation **21**, 497–501 (2014).

[54] L. Kipp, M. Skibowski, R. L. Johnson, R. Berndt, R. Adelung, S. Harm, and R. Seemann, "Sharper images by focusing soft X-rays with photon sieves," Nature **414**, 184–188 (2001).

[55] S. D. Eder, X. Guo, T. Kaltenbacher, M. M. Greve, M. Kalläne, L. Kipp, and B. Holst, "Focusing of a neutral helium beam with a photon-sieve structure," Physical Review A **91**, 043608 (2015).

[56] R. Flatabø, M. M. Greve, S. D. Eder, M. Kalläne, A. S. Palau, K. K. Berggren, and B. Holst, "Atom sieve for nanometer resolution neutral helium microscopy," Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena **35**, 06G502 (2017).

[57] M. M. Greve, A. M. Vial, J. J. Stamnes, and B. Holst, "The Beynon Gabor zone plate: a new tool for de Broglie matter waves and hard X-rays? An off axis and focus intensity investigation," Optics Express **21**, 28483–28495 (2013).

[58] M. Koch, S. Rehbein, G. Schmahl, T. Reisinger, G. Bracco, W. E. Ernst, and B. Holst, "Imaging with neutral atoms—a new matter-wave microscope," Journal of Microscopy **229**, 1–5 (2008).

[59] P. Witham and E. Sánchez, "A simple approach to neutral atom microscopy," Review of Scientific Instruments **82**, 103705 (2011).

[60] P. Witham and E. Sánchez, "Exploring neutral atom microscopy," Crystal Research and Technology **49**, 690–698 (2014).

[61] M. Barr, A. Fahy, A. Jardine, J. Ellis, D. Ward, D. A. MacLaren, W. Allison, and P. C. Dastoor, "A design for a pinhole scanning helium microscope," Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 20th International Workshop on Inelastic Ion-Surface Collisions (IISC-20), **340**, 76–80 (2014).

[62] A. Fahy, M. Barr, J. Martens, and P. C. Dastoor, "A highly contrasting scanning helium microscope," Review of Scientific Instruments **86**, 023704 (2015).

[63] A. Fahy, S. D. Eder, M. Barr, J. Martens, T. A. Myles, and P. C. Dastoor, "Image formation in the scanning helium microscope," Ultramicroscopy **192**, 7–13 (2018).

[64] M. Barr, A. Fahy, J. Martens, A. P. Jardine, D. J. Ward, J. Ellis, W. Allison, and P. C. Dastoor, "Unlocking new contrast in a scanning helium microscope," Nature Communications **7**, 10189 (2016).

[65] G. Scoles, D. Bassi, U. Buck, and D. C. Laine, eds., *Atomic and Molecular Beam Methods: Volume 1* (Oxford University Press, New York, 1988).

[66] M. Barr, K. M. O'Donnell, A. Fahy, W. Allison, and P. C. Dastoor, "A desktop supersonic free-jet beam source for a scanning helium microscope (SHeM)," Measurement Science and Technology **23**, 105901 (2012).

[67] H. Hedgeland, A. P. Jardine, W. Allison, and J. Ellis, "Anomalous attenuation at low temperatures in high-intensity helium beam sources," Review of Scientific Instruments **76**, 123111 (2005).

[68] S. Lambrick, *Contrast mechanisms and image formation in helium atom microscopy*, First year report, Cavendish Laboratory, University of Cambridge (2018).

[69] L. C. Hale and A. H. Slocum, "Optimal design techniques for kinematic couplings," Precision Engineering **25**, 114–127 (2001).

[70] S. D. Eder, G. Bracco, T. Kaltenbacher, and B. Holst, "Two Dimensional Imaging of the Virtual Source of a Supersonic Beam: Helium at 125 K," The Journal of Physical Chemistry A **118**, 4–12 (2013).

[71] K. B. Burke, E. J. Luber, N. P. Holmes, A. J. Murray, W. J. Belcher, X. Zhou, D. Mitlin, and P. C. Dastoor, "A knife-edge measurement of the beam profile of STXM 5.3.2.2 using a focussed ion beam milled metallic glass," Journal of Electron Spectroscopy and Related Phenomena **185**, 453–457 (2012).

[72] H. C. W. Beijerinck and N. F. Verster, "Absolute intensities and perpendicular temperatures of supersonic beams of polyatomic gases," Physica B+C **111**, 327–352 (1981).

[73] D. P. DePonte, S. D. Kevan, and F. S. Patton, "Brightness of micronozzle helium source," Review of Scientific Instruments **77**, 055107 (2006).

[74] T. Reisinger, G. Bracco, S. Rehbein, G. Schmahl, W. E. Ernst, and B. Holst, "Direct Images of the Virtual Source in a Supersonic Expansion," The Journal of Physical Chemistry A **111**, 12620–12628 (2007).

[75] E. Keppler, V. V. Afonin, C. C. Curtis, A. V. Dyachkov, J. E. Jr, C. Y. Fan, K. C. Hsieh, D. M. Hunten, W.-H. Ip, A. K. Richter, A. J. Somogyi, and G. Umlauft, "Neutral gas measurements of comet Halley from Vega 1," Nature **321**, 273–274 (1986).

[76] A. Frattolillo and A. D. Ninno, "A powerful tool to quantitatively detect tiny amounts of 4he in a deuterium rich background for fusion research," in *2007 IEEE 22nd Symposium on Fusion Engineering* (2007) pp. 1–4.

[77] D. M. Chisnall, *A high sensitivity detector for helium atom scattering*, Ph.D., University of Cambridge (2013).

[78] H. Hedgeland, *The development of quasi-elastic helium-3 spin-echo spectroscopy as a tool for the study of surface dynamics*, Ph.D., University of Cambridge (2007).

[79] K. M. O'Donnell, A. Fahy, M. Barr, W. Allison, and P. C. Dastoor, "Field ionization detection of helium using a planar array of carbon nanotubes," Physical Review B **85**, 113404 (2012).

[80] S. L Chin and P. Lambropoulos, *Multiphoton ionization of atoms / edited by S.L. Chin, P. Lambropoulos.* (Academic, Toronto ; London, 1984).

[81] K. M. Martini, W. Franzen, and M. El-Batanouny, "Compact He-atom detector for high-resolution surface phonon measurements," Review of Scientific Instruments **58**, 1027–1037 (1987).

[82] A. R. Alderwick, *Instrumental and analysis tools for atom scattering from surfaces*, Ph.D., University of Cambridge (2010).

[83] H. G. Bullman, *Development of a position sensitive atom detector for helium-surface scattering*, Ph.D., University of Cambridge (1998).

[84] J. R. B. Knowling, *Helium atom scattering : experiment and the interpretation of experiment*, Ph.D., University of Cambridge (2000).

[85] S. Dworski, *Atom optical methods for surface studies*, Ph.D., University of Cambridge (2004).

[86] M. DeKieviet, D. Dubbers, M. Klein, U. Pieles, and C. Schmidt, "Design and performance of a highly efficient mass spectrometer for molecular beams," Review of Scientific Instruments **71**, 2015–2018 (2000).

[87] A. V. Kalinin, L. Y. Rusin, and J. P. Toennies, "Ion source with longitudinal ionization of a molecular beam by an electron beam in a magnetic field," Instruments and Experimental Techniques **49**, 709–713 (2006).

[88] H. Deutsch, K. Becker, G. Senn, S. Matt, and T. D. Märkcd, "Calculation of cross sections and rate coefficients for the electron impact multiple ionization of beryllium, boron, carbon, and oxygen atoms," International Journal of Mass Spectrometry **192**, 1–8 (1999).

[89] G. Zschornacka, M. Schmidt, and A. Thorn, "Electron Beam Ion Sources," arXiv:1410.8014 [physics] (2014), 10.5170/CERN-2013-007.165, arXiv: 1410.8014.

[90] A. R. Alderwick, A. P. Jardine, H. Hedgeland, D. A. MacLaren, W. Allison, and J. Ellis, "Simulation and analysis of solenoidal ion sources," The Review of Scientific Instruments **79**, 123301 (2008).

[91] J. Martens, A. Fahy, M. Barr, A. Jardine, W. Allison, and P. C. Dastoor, "Development of a permanent magnet alternative for a solenoidal ion source," Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 20th International Workshop on Inelastic Ion-Surface Collisions (IISC-20), **340**, 85–89 (2014).

[92] G. Zschornack, J. König, M. Schmidt, and A. Thorn, "A compact, versatile low-energy electron beam ion source," Review of Scientific Instruments **85**, 02B703 (2013).

[93] M. Walt, "Introduction to Geomagnetically Trapped Radiation by Martin Walt," (1994).

[94] W. Allison, D. J. Ward, and M. Bergin, *Cambridge NEMI Report*, Private communication (2016).

[95] D. J. Douglas, "Linear quadrupoles in mass spectrometry," Mass Spectrometry Reviews **28**, 937–960.

[96] W. Chen, B. A. Collings, and D. J. Douglas, "High-Resolution Mass Spectrometry with a Quadrupole Operated in the Fourth Stability Region," Analytical Chemistry **72**, 540–545 (2000).

[97] S. Penner, "Calculations of Properties of Magnetic Deflection Systems," Review of Scientific Instruments **32**, 150–160 (1961).

[98] E. Harting and F. H. Read, *Electrostatic lenses* (1976).

[99] L. A. Dietz and J. C. Sheffield, "Secondary electron emission induced by 5–30-keV monatomic ions striking thin oxide films," Journal of Applied Physics **46**, 4361–4370 (1975).

[100] "Channeltron Electron Multiplier Hanbook for Mass Spectrometry Applications," .

[101] S. Weidmann and R. Zenobi, "High-Mass MALDI-MS Using Ion Conversion Dynode Detectors: Influence of the Conversion Voltage on Sensitivity and Spectral Quality," Journal of The American Society for Mass Spectrometry **25**, 950–954 (2014).

[102] K. Ohya, T. Yamanaka, K. Inai, and T. Ishitani, "Comparison of secondary electron emission in helium ion microscope with gallium ion and electron microscopes," Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms Proceedings of the 17th International Workshop on Inelastic Ion-Surface Collisions, **267**, 584–589 (2009).

[103] T. Ishitani, T. Yamanaka, K. Inai, and K. Ohya, "Secondary electron emission in scanning Ga ion, He ion and electron microscopes," Vacuum Selected papers from the 19th conference on Ion-Surface Interactions, **84**, 1018–1024 (2010).

[104] P. Varga and H. Winter, "Slow particle-induced electron emission from solid surfaces," in *Particle Induced Electron Emission II*, Springer Tracts in Modern Physics No. 123, edited by D. D. Hasselkamp, D. H. Rothard, P. D. K.-O. Groeneveld, D. J. Kemmler, D. P. Varga, and P. D. H. Winter (Springer Berlin Heidelberg, 1992) pp. 149–214.

[105] R. A. Baragiola, E. V. Alonso, J. Ferron, and A. Oliva-Florio, "Ion-induced electron emission from clean metals," Surface Science **90**, 240–255 (1979).

[106] P. Riccardi, R. A. Baragiola, and C. A. Dukes, "Electron emission and electronic stopping in the interaction of slow helium ions with aluminum," Physical Review B **92**, 045425 (2015).

[107] I. E. Software, *IES LORENTZ User Guide*, Tech. Rep.

[108] J. Greenwood, "The correct and incorrect generation of a cosine distribution of scattered particles for Monte-Carlo modelling of vacuum systems," Vacuum **67**, 217–222 (2002).

[109] R. Hawley, "Solid insulators in vacuum: A review (Invited paper)," Vacuum **18**, 383–390 (1968).

[110] D. H. Crandall, R. A. Phaneuf, and P. O. Taylor, "Electron-impact ionization of C 3 + and N 4 +," Physical Review A **18**, 1911–1924 (1978).

[111] S. Lambrick, M. Bergin, A. Jardine, and D. Ward, "A ray tracing method for predicting contrast in neutral atom beam imaging," Micron **113**, 61–68 (2018).

[112] R. Kersevan and J.-L. Pons, "Introduction to MOLFLOW+: New graphical processing unit-based Monte Carlo code for simulating molecular flows and for calculating angular coefficients in the compute unified device architecture environment," Journal of Vacuum Science & Technology A **27**, 1017–1023 (2009).

[113] J. M Lafferty, *Foundations of vacuum science and technology / edited by James M. Lafferty.* (Wiley, New York ; Chichester, 1998).

[114] H. Tawara and V. P. Shevelko, "Multiple ionization of negative and positive ions, neutral atoms, and molecules under electron impact: data and databases," International Journal of Mass Spectrometry **192**, 75–85 (1999).

[115] A. Kramida, Yu. Ralchenko, J. Reader, and and NIST ASD Team, *NIST Atomic Spectra Database (ver. 5.5.6)* (2018) published: NIST Atomic Spectra Database (ver. 5.5.6), [Online]. Available: https://physics.nist.gov/asd [2018, August 27]. National Institute of Standards and Technology, Gaithersburg, MD.

[116] R. E. Ellefson, W. E. Moddeman, and H. F. Dylla, "Hydrogen isotope analysis by quadrupole mass spectrometry," Journal of Vacuum Science and Technology **18**, 1062–1067 (1981).

[117] C. Day, "The use of a high-resolution quadrupole gas mass spectrometer system for selective detection of helium and deuterium," Vacuum **51**, 21–30 (1998).

[118] S. Davies, J. A. Rees, and D. L. Seymour, "Threshold ionisation mass spectrometry (TIMS); a complementary quantitative technique to conventional mass resolved mass spectrometry," Vacuum Including rapid communications, original articles and a special section: Measurement Characteristics and Use of Quadrupole Mass Spectrometers for Vacuum Applications, **101**, 416–422 (2014).

[119] J. R. de Laeter, J. K. Böhlke, B. P. De, H. Hidaka, H. S. Peiser, K. J. R. Rosman, and P. D. P. Taylor, "Atomic weights of the elements. Review 2000 (IUPAC Technical Report)," Pure and Applied Chemistry **75**, 683–800 (2003).

[120] M. C. Cowen, W. Allison, and J. H. Batey, "Electron space charge effects in ion sources for residual gas analysis," Measurement Science and Technology **4**, 72 (1993).

[121] P. S. M. Townsend, *Diffusion of light adsorbates on transition metal surfaces*, Ph.D., University of Cambridge (2018).

[122] M. Knudsen, *The kinetic theory of gases: some modern aspects. (3rd ed)*. (London, London, 1950).

[123] Frank O. Goodman, *Dynamics of gas-surface scattering / Frank O. Goodman, Harold Y. Wachman.* (Academic Press, New York, 1976).

[124] John Keith Roberts, *Heat and thermodynamics / by J.K. Roberts ...*, 3rd ed., The Student's physics ; [Vol. IV] (Blackie & son limited, London ; Glasgow, 1945).

[125] G. Comsa, "Angular Distribution of Scattered and Desorbed Atoms from Specular Surfaces," The Journal of Chemical Physics **48**, 3235–3240 (1968).

[126] W. Steckelmacher, "Knudsen flow 75 years on: the current state of the art for flow of rarefied gases in tubes and systems," Reports on Progress in Physics **49**, 1083 (1986).

[127] W. Steckelmacher, "A review of the molecular flow conductance for systems of tubes and components and the measurement of pumping speed," Vacuum **16**, 561 – 584 (1966).

[128] W. Steckelmacher, "The Flow of Rarefied Gases in Vacum Systems and Problems of Standardization of Measuring Techniques," Japanese Journal of Applied Physics **13**, 117 (1974).

[129] F. Celestini and F. Mortessagne, "Cosine law at the atomic scale: Toward realistic simulations of Knudsen diffusion," Physical Review E **77**, 021202 (2008).

[130] R. Feres and G. Yablonsky, "Knudsen's cosine law and random billiards," Chemical Engineering Science **59**, 1541–1556 (2004).

[131] D. R. O'Keefe and R. L. Palmer, "Atomic and molecular beam scattering from macroscopically rough surfaces," Journal of Vacuum Science and Technology **8**, 27–30 (1971).

[132] Pedrotti F. L, *Introduction to optics.* (Prentice Hall, Englewood Cliffs, NJ, 1987).

[133] J. T. Conway, "Analytical solution for the solid angle subtended at any point by an ellipse via a point source radiation vector potential," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **614**, 17–27 (2010).

[134] C. Wade, *Contrast Mechanisms in Scanning Helium Microscopy*, Part III Project (2012).

[135] K. A. O'Donnell and E. R. Mendez, "Experimental study of scattering from characterized random surfaces," JOSA A **4**, 1194–1205 (1987).

[136] T. S. Hegge, T. Nesse, A. A. Maradudin, and I. Simonsen, "The scattering of a scalar beam from isotropic and anisotropic two-dimensional randomly rough Dirichlet or Neumann surfaces: The full angular intensity distributions," Wave Motion **82**, 30–50 (2018).

[137] B. v. Ginneken, M. Stavridi, and J. J. Koenderink, "Diffuse and Specular Reflectance from Rough Surfaces," Applied Optics **37**, 130–139 (1998).

[138] J. P. Toennies and K. Winkelmann, "Theoretical studies of highly expanded free jets: Influence of quantum effects and a realistic intermolecular potential," The Journal of Chemical Physics **66**, 3965–3979 (1977).

[139] J. Braun, P. K. Day, J. P. Toennies, G. Witte, and E. Neher, "Micrometer-sized nozzles and skimmers for the production of supersonic He atom beams," Review of Scientific Instruments **68**, 3001–3009 (1997).

[140] I. Estermann and O. Stern, "Beugung von Molekularstrahlen," Zeitschrift für Physik **61**, 95–125 (1930).

[141] G. Boato, P. Cantini, and L. Mattera, "A study of the (001)LiF surface at 80 K by means of diffractive scattering of He and Ne atoms at thermal energies," Surface Science **55**, 141–178 (1976).

[142] J. R. Bledsoe and S. S. Fisher, "Scattering of helium nozzle beams from LiF(001) and NaCl(001) crystal surfaces: I. Elastic and inelastic transitions," Surface Science **46**, 129–156 (1974).

[143] A. Salvador Palau, G. Bracco, and B. Holst, "Theoretical model of the helium zone plate microscope," Physical Review A **95**, 013611 (2017).

[144] A. S. Palau, G. Bracco, and B. Holst, "Theoretical model of the helium pinhole microscope," Physical Review A **94**, 063624 (2016).

[145] G. S. Sikora, "Analysis of asymptotic behavior of free jets: prediction of molecular beam intensity and velocity distributions." (1973).

[146] T. Kaltenbacher, "Optimization of a constrained linear monochromator design for neutral atom beams," Ultramicroscopy **163**, 62–68 (2016).

[147] J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts & Company Publishers, Englewood, Colo, 2005).

[148] Arnold Sommerfeld, *Optics / by Arnold Sommerfeld ; translated by Otto Laporte and Peter A. Moldauer.*, 1st ed., Sommerfeld, Arnold, 1868-1951. Vorlesungen über theoretische Physik. English ; v. 4 (Academic Press, New York ; London, 1964).

[149] T. Reisinger, S. Eder, M. M. Greve, H. I. Smith, and B. Holst, "Free-standing silicon-nitride zoneplates for neutral-helium microscopy," Microelectronic Engineering The 35th International Conference on Micro- and Nano-Engineering (MNE), **87**, 1011–1014 (2010).

[150] Alan G. Michette, *Optical systems for soft X rays* (Plenum Press, New York, 1986).

[151] J. M. v. Houten and M. H. A. J. Herben, "The elliptical Fresnel-zone plate antenna," in *1995 Ninth International Conference on Antennas and Propagation, ICAP '95 (Conf. Publ. No. 407)*, Vol. 1 (1995) pp. 97–101 vol.1.

[152] R. Flatabø, S. D. Eder, A. K. Ravn, B. Samelin, M. M. Greve, T. Reisinger, and B. Holst, "Fast resolution change in neutral helium atom microscopy," Review of Scientific Instruments **89**, 053702 (2018).

[153] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," Applied Optics **46**, 1819–1829 (2007).

[154] P. T. Hustler-Wraight, *Aspects of atom-surface interactions : considerations for microscopy*, Ph.D., University of Cambridge (2008).