Systems of Transition Metal Dichalcogenides : Controlling Applied Strain and Defect Density With Direct Impact on Material Properties

Drew Edelberg

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Systems of Transition Metal Dichalcogenides : Controlling Applied Strain and Defect Density With Direct Impact on Material Properties

Drew Edelberg

Transition metal dichalcogenides (TMDs) are crystalline layered materials that have significantly impacted the field of condensed matter physics. These materials were the first exfoliatable semiconductors to be discovered after the advent of graphene. The focus of this dissertation is utilizing multiple imaging and characterization techniques to improve and understand the impact of strain and lattice defects in these materials. These inclusions to the lattice, alter the semiconducting performance in controllable ways. A comprehensive study using scanning tunneling spectroscopy (STM), spectroscopy (STS), scanning transmission electron microscopy (STEM), and photoluminescence (PL) in this work will provide a breadth of ways to pinpoint and cross-examine the impact of these factors on these materials. In the first half of this work we focus on the control of lattice defects through two growth processes: chemical vapor transport (CVT) and self-flux. By fine tuning the growth procedure we are both able to determine the intrinsic defects of the material, their electronics, and consistently diminish their density. The second half uses an in-situ strain device to reversibly control and examine the effects of applied strain on transition metal dichalcogenide layers. Utilizing the scanning tunneling microscope to image the lattice, we characterize the change of lattice parameters and observe the formation of strain solitons within the lattice. Measuring these solitons directly we look at the dynamics of a special class of line defects, folds within the top layer of the material, that occur naturally as strain is relieved within the monolayer. With the available imaging techniques and theoretical models we uncover a host of properties of these materials that are only accessible within the high strain regime

Table of Contents

List of Figures

List of Tables

Acknowledgements

My advisor and mentor Abhay Pasupathy of Columbia University once told me – “Science does not occur in a vacuum”. I have undeniable proof, it is only with thanks to the multitude of friends, family and colleagues that this work can be presented before you today. And of the *few* people mentioned here, know that there were countless others that inspired me to walk the path of a physics PhD. To express my gratefulness to all those who have joined me throughout my life’s trajectory, I wish to illuminate the teeniest speck of the experiences that were, and still fuel my inspiration.

Although one could say that an elaborate tapestry of childhood teachers and experiences organically flow into the present. Or that each, is a step along a road that cascades forward and informs future decisions, there are pivotal moments where the seeds of what will transpire takes root. For me that moment was in 2002, a 7th grade math teacher, Mrs. Vicky Decker gave me my first physics book, it was on Supersymmetry a topic far above my head at the time. This gift not only planted a seed, but made me receptive to the teachers who through their own passion for learning made me see physics as a pursuit in itself; and not just a class with a grade. Two teachers furthered this dream, Mr. Carlos Rodriquez and Mrs. Dana Blumfield, who went out of their way to answer and inspire questions that surpassed the typical high school education.

Briefly I would like to also acknowledge my undergraduate advisor, Professor Michael Crommie whose laboratory environment inspired my love of electron microscopy. Here at Berkeley, I was thrilled to see my first atomic surface during the initial discovery of graphene freshman year. It was this experience that propelled me to Columbia University upon graduation in 2012.

To my current advisor Professor Abhay Pasupathy, I am so grateful for all of the days and nights we spent building, designing, measuring and writing the totality of the text to follow. Without your support, this thesis would never have been possible. Through your vision, I was paired with the many researchers who’ve passed through our laboratory. And as an advocate for my success, I want to thank you for integrating me into the Columbia MRSEC community. It is the insightful discussions I’ve had with Professor Jim Hone, Professor Xiaoyang Zhu, and Professor Katayun Barmak, which enabled the

breadth of techniques performed, and introduced me to the graduate students & postdocs who've shared their expertise in the field.

Three postdoctorates that especially directed and guided my path: now Professor Pinshane Huang, Doctor Lior Embon and Doctor Leo McGilly. Each appeared at a different stage of my graduate career and inspired me to be a better researcher – these are the people I look up to, and aspire to be. These are the personal giants whose shoulders I have stood on, each taking me to higher aspirations and wishing me the utmost success.

Of course I must thank the core of the Pasupathy lab and the affiliated researchers that participated in my project: Dr. Ethan Rosenthal, Dr. Chris Gutiérrez (with special thanks for coding expertise), Dr. Ali Dadgar, Dr. Carlos Arguello, Dr. Erick Andrade, Dr. Ayelet Notis, Dr. Sky Cheung, Alex Kerelsky, Minghao Cheng, Andrew Wieteska, Simon Turkel, and Mihir Bhaskar. With external help from a Dr. Daniel Chenet, and Dr. Daniel Rhodes of the Hone lab, for their expertise in crystal growth and extreme patience as I learned to scan their materials.

Again, without the support of my friends which kept me levelheaded on my PhD journey. The rarity through which we encounter steadfast friendship, my roommate Evan Telford that has been there with such camaraderie and commiseration of our shared PhD experiences. For Angelo Esposito, my other roommate - it was such a pleasure to live with like minds. Additionally the philosophical and meaningful discussions with Milica Iličić, and for showing me so much more of New York City.

Lastly, I am so thankful to have the support of my family who have been there through every step of this process, this includes those people who by extension have known me since before I was born. I consider you to be family not by blood, but by shared experience. With a special thanks to the Horrocks family, both Susan for helping me see a world full of positivity, music & light and Grant for discussions of the universe, imagined beyond limitation. I am so grateful to have met your family through piano and art, thank Anna and Grace for all the fond memories, and with best of luck to William Horrocks who will be starting his own PhD journey soon. A special honor to my childhood friends Max Cecil and Claire Horrocks who shaped my world, and shared so many of life's great mysteries.

Most of all my brother Bryce Edelberg, every hour we spend together is the biggest inspiration, and more than I could ever ask for. With love for my parents Robert Edelberg and Meredith Reu, to the journey that is yet to be written, thank you for always accompanying me. As this chapter closes it opens so many more avenues and successes to come.

With many thanks to the lab of Abhay Pasupathy

And my family, Bryce & Robert Edelberg

and Meredith Reu

Chapter 1

Introduction to Scanning Tunneling Microscopy

1.1 Principles of Scanning Tunneling Microscopy

Scanning tunneling microscopy (STM) is a microscopy technique which utilizes the quantum nature of electrons to visualize a surface[1-5]. This technique has many wide reaching applications in condensed matter physics, and has been used to locally map the atomic structure and electronic properties of many metals, semiconductors, and superconductors. With many more materials being discovered each year, this has become a core technique to understand new and emergent physics. Since this probe is low energy in nature, an STM can be built easily at low cost to fit many specifications allowing for custom tailoring to almost any research project [6].

Before getting too far into material studies we first focus on how the scanning probe system works and how it can be tailored to our needs. STM is able to work thanks to a process known as quantum tunneling [7]. This takes place when an electron or other particle with quantum nature crosses a barrier which is forbidden classically. Quantum tunneling is thanks to an overlap of the wave function of the electron and therefore only occurs at very short distances. In order to use this in an application we first envision two metallic surfaces separated by a vacuum barrier. Since the electron has a large swath of states in either metal it can move freely within them. However in vacuum there are no states for the electron, so this acts as an edge which stops it from leaving the material. In the special instance where the two metals become close enough together (roughly the distance of 1 atom) however the electrons in one metal are able to pass through this barrier with a probability that is exponential with distance between the surfaces. This is known as a tunneling current and is very sensitive to small changes in the vacuum barrier.

To utilize quantum tunneling, the STM is a system that pairs an atomically sharp metallic tip with a surface of interest. From this, we setup a feedback loop on the tunneling current. Here the measured current will be proportional to the integrated states in the material [4]. This means that for a typical surface

we can detect variations in electronic density in a controllable way. As a consequence, STM is a way to directly probe the wavefunctions of the material. With this simple basis to start our surface study, the STM can measure individual atoms and the interplay they have within a larger electronic structure.

The typical method by which we can see simple lattices is shown in Figure 1. This shows an STM tip made of a platinum-iridium blend scanning over a surface of gold. We see that in this is very similar to reading brail, our feedback system is able to travel over this surface and map the change in density. STM is governed by utilizing the tunneling equation; a cartoon explaining this equation is shown in Figure 1b.

$$I(d) \propto \mathrm{eV} * e^{-\frac{2d\sqrt{2m\,\Delta\Phi}}{\hbar}}$$

This graphic shows the current I as a function of distance d. We find that the tunneling current is proportional to the exponent of the distance between the surface and tip, and multiplied by the applied bias offset between them V, with other prefactors that are constant (such as the workfunctions $\Delta\Phi$ of the materials and mass of the electron m). This governing equation is central to scanning tunneling microscopy, and as such, all further parts of this chapter deal with how we can realize this in a controllable way.

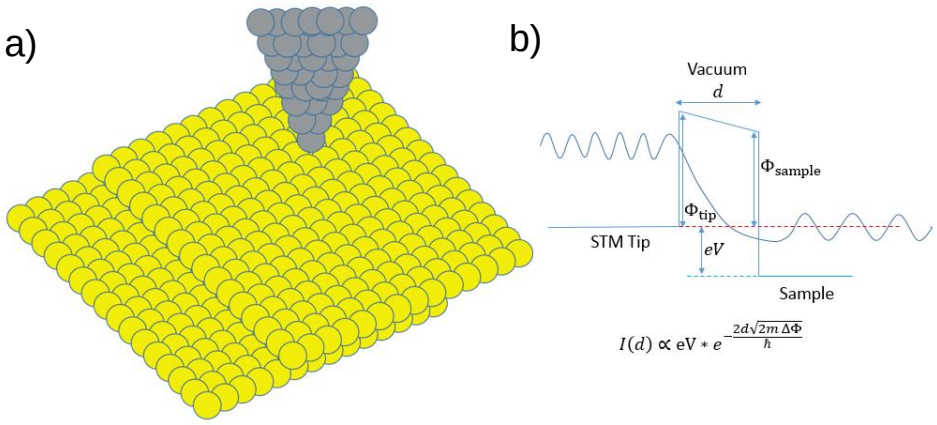

Figure 1.1 : Cartoon Model of Scanning Tunneling Microscopy

a) A cartoon model of the scanning tunneling system b) The tunneling equation which governs STM feedback

1.2 Building the Scanning Probe System

In order to control the distance between the surface and the tip with atomic precision a typical STM is built off of another discovery of condensed matter physics, piezoelectricity [5,8]. This phenomenon links physical motions to applied voltage. To get motions of less than an atom, we use piezoelectric crystals which can expand or contract with the application of an electric field. This occurs because the material is made of linked polar molecules which experience a small distortion in the presence of an electric field. This distortion allows the material to set up a counter propagating electric field to cancel the applied voltage. Since this is an electronic property it can be tuned easily and has a high degree of precision. Shown in Figure 1.2 is a piezoelectric scan tube. This cylindrical piezoelectric has four quadrant electrodes which oppose a central grounded core. When the STM is in operation we can apply separate voltages to each of these quadrants to bend the tube resulting in a raster scan of the surface [9,10]. In the quadrant geometry the application of a voltage to all the electrodes simultaneously results in a full expansion or contraction of the cylinder so in addition to X-Y motion the scanner has access to the Z axis. This simple crystal allots the system to all three degrees of atomic motion. We can see the physical scanner used in our STM construction mounted in Figure 1.2**b**. From this picture the wires used to control the quadrants are visible passing through access holes from the top of the image. Affixed to this scan tube is the electrical contact for the current which we will receive from a mounted platinum-iridium wire [11-13], or STM tip (not pictured).

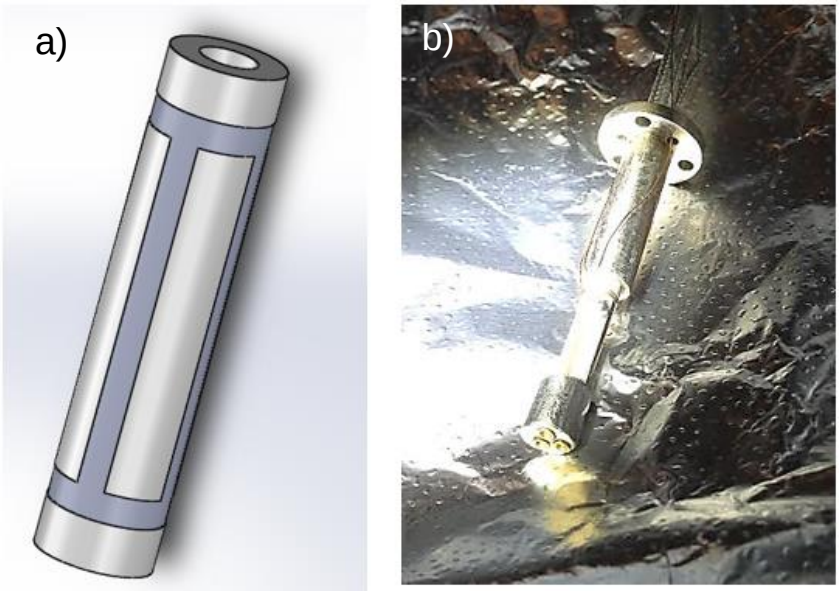

Figure 1.2: Piezoelectric Scan Tube

a) A 3D rendered model of a four quadrant scanning tube for STM b) The STM assembly after wiring before inserting into the coarse system Z mover

With a piezoelectric scan tube we have a method for fine control on the atomic scale, however since this object is stationary we have not yet described how we can manipulate the system to reliably move the scanner both across the crystal of interest and towards and away from a sample plate. Again for the coarse motion we chose to use piezoelectric crystals [14], however we machine a clever geometry which utilizes differences in static versus kinetic friction.

In order to translate over large distances using a voltage it is impossible to use a single stroke, which is limited to a full range voltage pulse from our amplifier, since the piezos maximum response under these conditions is on the order of microns. However we can utilize the difference between a fast motion and a slow one. For a slow motion ramp the piezoelectric that is in contact with the object which we want to push, can move slowly enough that it retains its static friction with the surface. This allows it to move the sample slider unhindered and smoothly. Alternatively for a fast motion the inertia of the slider changes very little because the contact surface experiences a kinetic friction. So instead for a full stroke in a unipolar or bipolar set of slow and fast motions we can engineer the force transfer mechanism to be unidirectional. Pictured in Figure 1.3 is the fully assembled coarse mover for the planar motion of

the sample. Glued inside of identically machined clamshells are four sets of shear piezo stacks each of these can be seen in Figure 1.3. This assembly has wires for electrical contact to each stack and stainless steel tabs which allow for the application of voltage. When working properly these 4 stacks work in tandem to propel the sample stage in a jellyfish like motion, as pictured in Figure 1.3. We see that the four stacks shift the bar over in the direction of the red arrow, and then slip back into place for the next stroke [15,16]. Our full motion for this assembly is 7 mm along both the X and Y axis of the system.

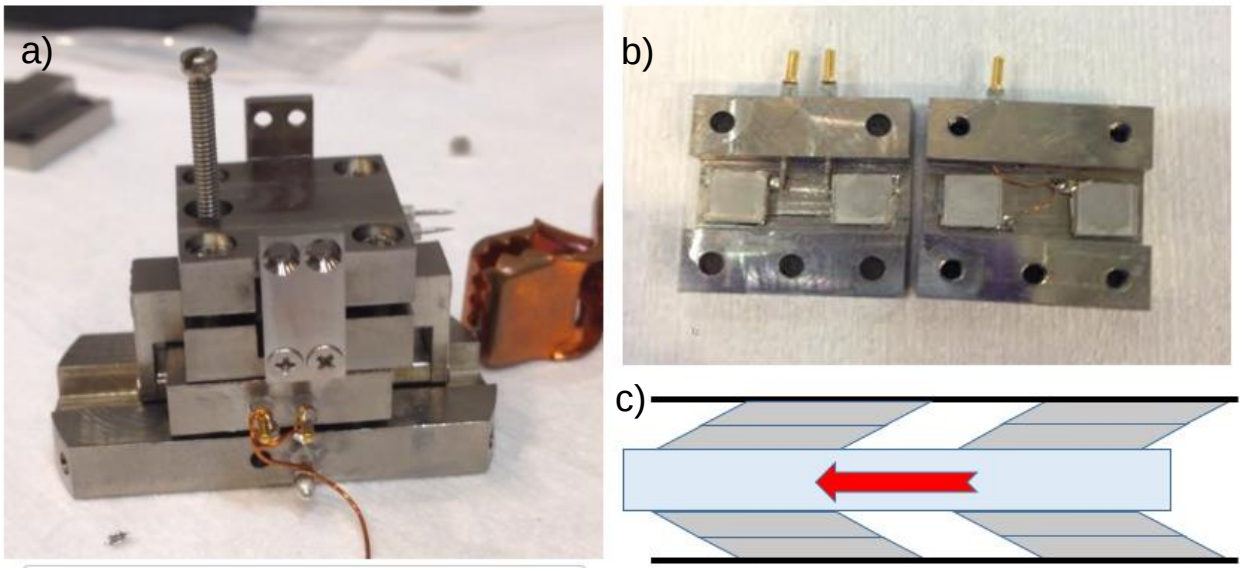

Figure 1.3: The Coarse X-Y Mover

a) The fully assembled X-Y mover for sample translation in the STM b) A single mover clamshell for translating the STM stage c) The working principle of the piezoelectric motion a slider bar is pushed by four sets of shear piezos

For the Z motion we can create a similar geometry, however Z motion is more complicated because we need high precision to combat hysteresis and gravity. To meet these constraints the Z – mover utilizes three sets of piezo stacks which surround a triangular prism core. Pictured in Figure 1.4 is the completed Z mover where the triangular prism can be seen protruding from the top of its titanium frame. A sapphire ball and spring system is used to directly tune the normal force applied to each piezo stack, allowing for mechanical control of the frictional parameters. Additional sapphire is placed on the cap of each stack making the prism run smoothly for its full range. Pictured in Figure 1.4b is the view

inside of the STM system where we can see the Z – mover holding the scanning tip at a constant position over the crystal.

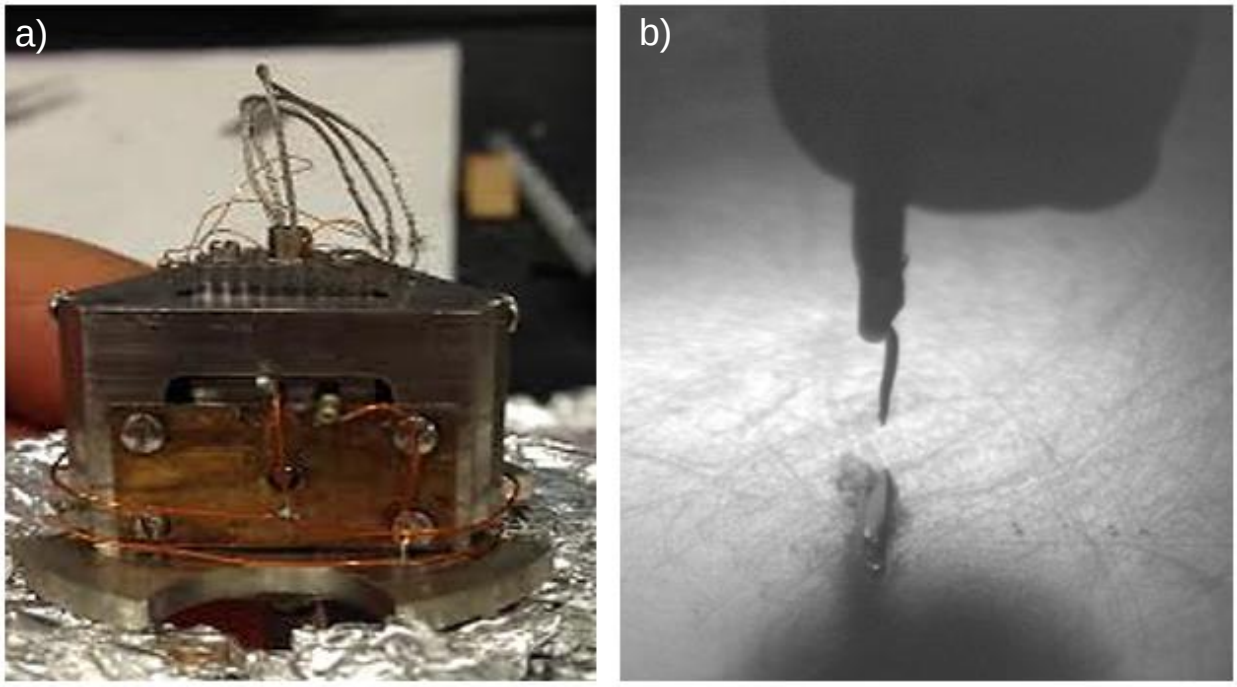

Figure 1.4: The Coarse Z-Mover and STM Tip

a) The fully assembled Z mover with moveable prism and sapphire ball for tuning exact frictional force b) The STM tip over a in-situ cleaved sample as an example of the Z-Mover's working functionality

Piezoelectric crystals along with precision machining, compose a completed method for motion in the STM system. We placed STM into ultra-high vacuum (UHV) where voltage control allows us to scan systems without exposing them to atmosphere. As another benefit the UHV chamber is equipped with a cryostat letting the STM see surfaces with atomic resolution from any temperature in the range of liquid helium to room temperature, 4K to 320K [17].

1.3 Topography Calibration: Preparing the First Surface

Now that we have a fully working scanning probe system we must calibrate it [18,19], to ensure that the tunnel junction is accurately reproducing a known atomic surface and to set the conversion between

the current signal and both height and distance. To do this it is common practice to use a freshly prepped surface of Au (111). A gold crystal is a good choice for calibration because it exhibits two features: atomic steps with well-defined height, and a surface reconstruction known as herringbone with a well-defined length. Most gold however is not perfectly atomic, and is typically amorphous. It is important that we have a method to create the necessary cleaning and calibration crystal in our UHV chamber, which we must form from a cut Au (111) single crystal through cycles of annealing and argon bombardment. A picture of the mounted crystal can be seen in the STM in Figure 1.5, this is the same crystal used in calibration for every experiment after repeated cleaning cycles.

In the process of cleaning gold it is necessary to heat the gold to temperatures in the range of 600 C to 800 C. To reach these temperatures we devised a filament annealer that can be manipulated in UHV using a magnetic transfer arm. This annealer works through radiative heating, since there is no convection in vacuum the only possible methods for heat transfer are either conductive or radiative. As our heat source we use a coiled tungsten rhenium filament which we pass 4.4 A and 30 V through. Effectively this creates a ~130 W light bulb which we contain within a reflective medium made from high heat materials such as molybdenum, tantalum and alumina. Closing this box with a gold sample holder we can easily reach the required temperature for cleaning, which we verify using the radiative color and a pyrometer. The heated gold sample is shown in Figure 1.5b. For comparison we show the cooled shiny gold surface in its sample mount in Figure 1.5c**.**

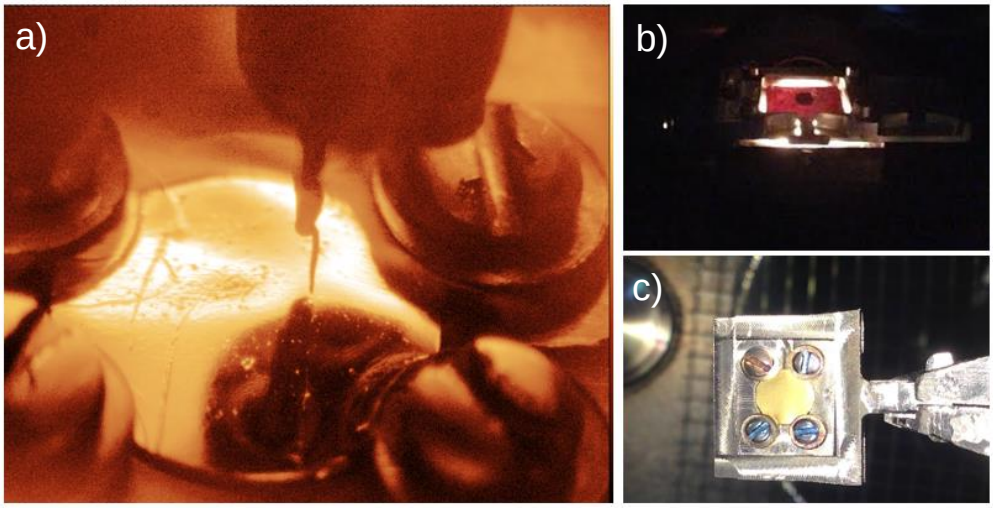

Figure 1.5: Gold Calibration Crystal and Annealer

a) The gold crystal cleaned and in the scanning tunneling microscope during the tip cleaning process b) The annealing system built to heat the gold to 650 C within UHV c) The mounting system used to affix the gold to its sample plate

As promised we look at the surface using STM, looking for our two calibration parameters. The typical surface that is characteristic of Au (111) is shown in Figure 1.6. In this image we can clearly see an atomic step in the top left corner and a rippling structure across the surface. When gold is allowed to relax in a vacuum potential, as is developed through repeated annealing the surface buckles to maximize the contact between gold atoms. The resulting pattern which resembles the textile “herringbone” is well established in literature to have a spacing of 6.3 nm [18,20]. We use this to calibrate the STM by looking at a line profile as is demonstrated in Figure 1.6b**.** One period of the pattern is defined from pair to pair, using this we can exactly map our scanning voltage to our distance calibration. We can do the same thing for a single atomic terrace of gold. A step on gold should be 240 pm [21], we calibrate to this in Figure 1.6c.

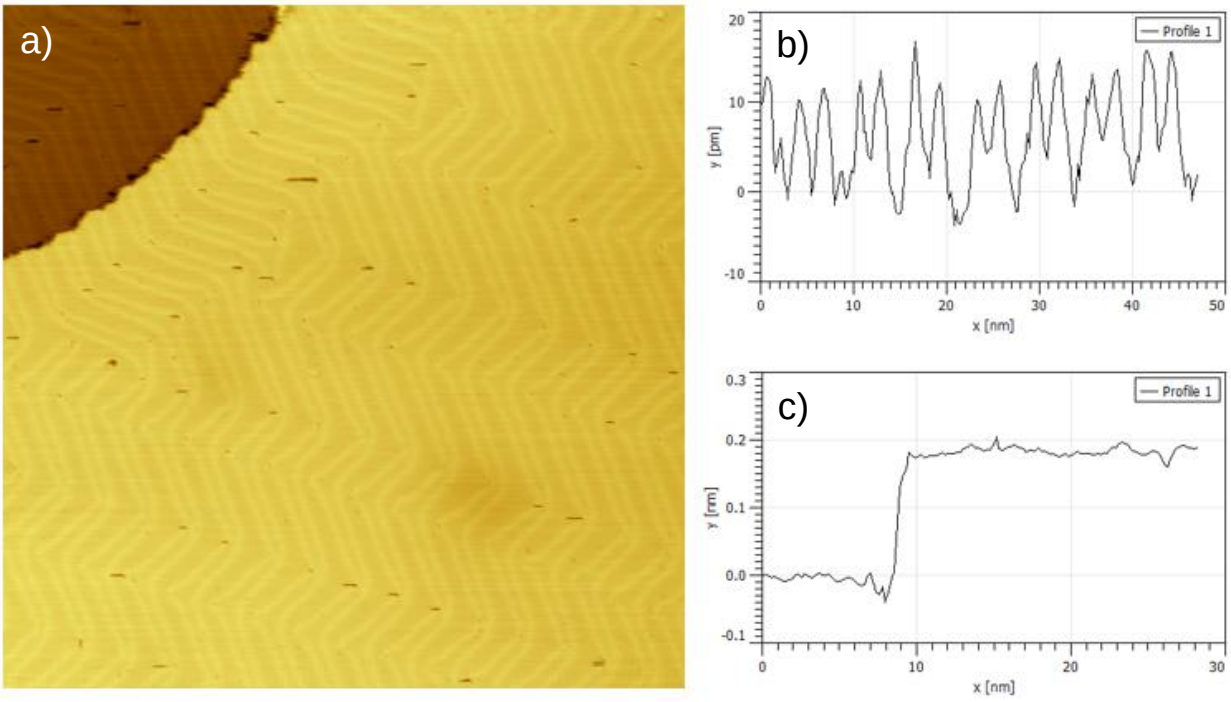

Figure 1.6: Standard Gold Calibration

a) The herringbone and step edge patterns characteristic of a gold surface when properly cleaned b) A cross section of the herringbone surface to calibrate the STM the lateral spacing should be 6.3 nm between periods c) A cross section showing the height profile of a single atomic step. The resulting height should be 240 pm

1.4 Intro to Scanning Tunneling Spectroscopy

The calibration capabilities when scanning gold aren't just topographical, another capability of STM is measurement of the electronics on the surface [22,23]. Since the current is proportional to the integrated density of states we can take the derivative using a lock-in amplifier. This allows us to measure dI/dV, the local density of states on the surface. In a clean metal the band dispersion is simple, we expect parabolic bands – and we can image those using scanning tunneling spectroscopy (STS). For gold we can see a series of slices of this band which are circular in nature, and since these also represent allowed momentum states we effectively measure the surface for different electron wavelengths. In Figure 1.7 we show the scattering of the gold surface band [24], this is a circular ring of momentum states that are present at this energy. This image comes from the Fourier transform of the real space map of the surface at this energy, which is pictured next to it in Figure 1.7b. We can extract the band structure because the

electrons scatter from point defects and step edges. At the step edge in the top of the image we even can see a linear scattering of incoming electrons, which looks quite like water waves hitting a barrier, or light diffracting from a knife edge. Since the available states at this energy come from a momentum ring of length k, the scattering vector which can be detected in the STS is a ring of magnitude 2k [25,26].

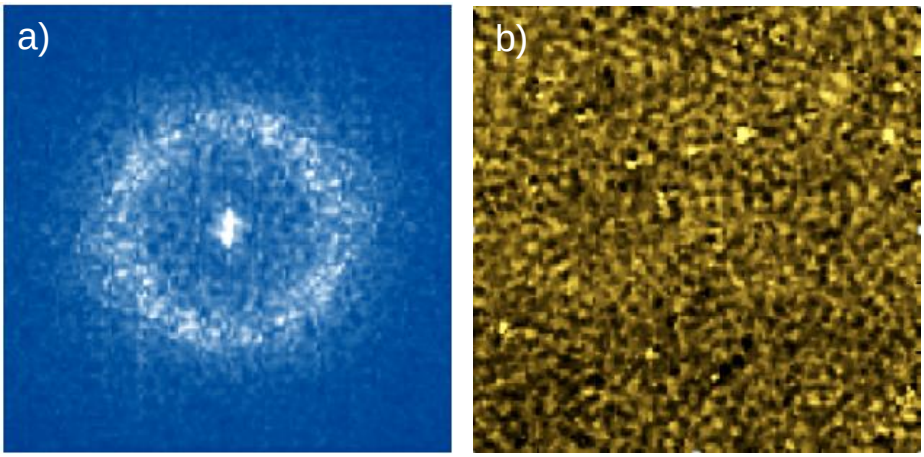

Figure 1.7: The Spectroscopic Properties of Gold

a) FFT of a scanning tunneling spectroscopy map taken on a gold surface. The resulting ring is a slice of scattering caused by the density of states at this energy b) The surface of the crystal at the energy which this dI/dV map was taken, showing that the FFT is a powerful tool in extracting information regarding the electronic structure.

With a working STM, using the techniques demonstrated on gold we now are able to turn our efforts towards systems of interest. While there is a large class of materials that STM can scan, some of the most interesting are layered Van Der Waals structures, specifically the transition metal dichalcognides known as TMDs.

Chapter 2

Introduction to Transition Metal Dichalcogenides

2.1 Backgrounds of Transition Metal Dichalcogenide Experiments

The transition metal dichalcogenides (TMDs) were developed in an effort to make materials that were graphene like, but with a wider range of electronic properties [27,28]. The defining property of such layered materials is the formation of sheets that bind to one another through a weak Van Der Waals interaction. This allows layers of these materials to be separated from one another using exfoliation methods. In the case of graphene the most famous layered material, the Van Der Waals force binds together prefect planes of hexagonally connected carbon. The discovery of graphene opened many new doors because it was a way to confine electrons to two dimensions. As soon as the exfoliation of graphene was realized, many other available Van Der Waals crystals became candidates for the creation of 2D monolayers. One of the first such natural crystals to be exfoliated was a lubricant material MoS_2. This material was found to have a rich range of physics because unlike graphene the parent crystal was a semiconductor. With monolayers of this material the indirect bandgap seen in bulk switched to a direct one, making this material of much interest to the optical community, since strong excitonic resonances require a direct gap.

Developing synthetic materials based on the same physics was the natural course for the physics community. Using MoS_2 as a base to start a materials search the class of TMDs became a much larger crystal family. It was found that the same hexagonal structure could be reproduced using any transition metal M, paired to two oxygen column elements, or chalcogens denoted X. The simplest resulting structure was determined to be a unit cell containing two stacked layers with equivalent hexagonal spacing. A shorthand for this form is 2H – MX_2 where the 2H is the stacking order and the metal constituent can stand for any of M = Ta, Mo, W, Nb, Pt, etc. Similarly X denotes one of the three chalcogens X = S,Se,Te. We plot the top view of the hexagonal plane in Figure 2.1a. Also shown is the unit vectors $\vec{a}$ and $\vec{b}$ for the typical TMD this distance is 3.3 Å. In Figure 2.1b we show the lateral stacking

of these sheets. Since this crystal has a two layer unit cell the vector bridging these sheets is $\vec{c}/2$ which measures roughly 6.5 Å

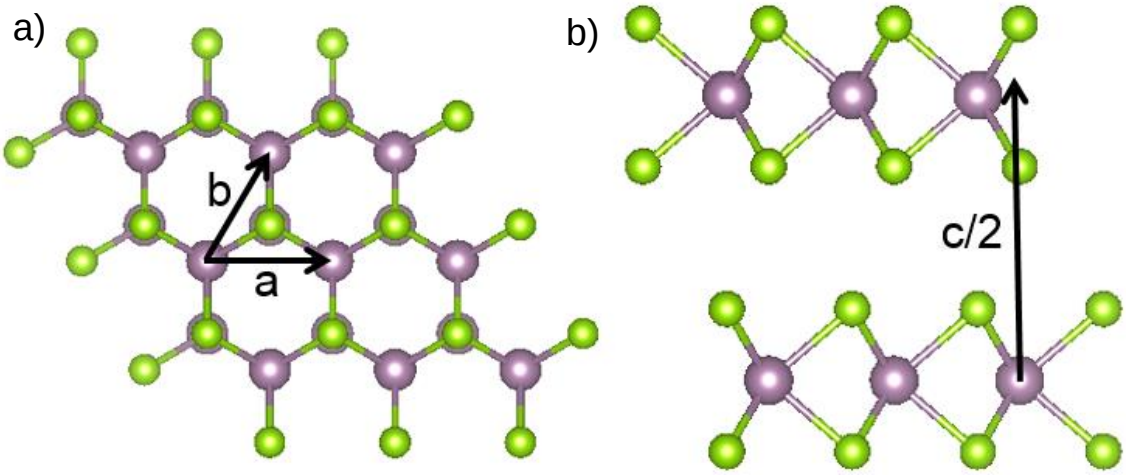

Figure 2.1: Transition Metal Dichalcogenide Lattice

a) The in plane lattice structure of the two dimensional dichalcogenide sheet with two lattice vectors **a** and **b** as shown
b) The lattice spacing between layers and the atomic stacking of the 2H structure

The significant properties of these sheets are highly dependent on the differing elements that are used to construct these materials, they range anywhere from semiconductors, superconductors, and Weyl metals [29]. A subclass that has been heralded for two dimensional circuitry is the TMDs that consist of a metal M= Mo,W and chacogens X = S, Se, Te. The band structure for the bulk material of one of this subfamily is shown in Figure 2a. The optical modes of the uppermost valence band and lowest conduction bands are the A, and B excitons at the K-point and a smallest indirect one from Γ to Γ – K denoted X [30-33]. When exfoliated to a single layer the bands of this material change to make the smallest gap be the A exciton, which can be excited optically. These new bands are shown in Figure 2b**.**

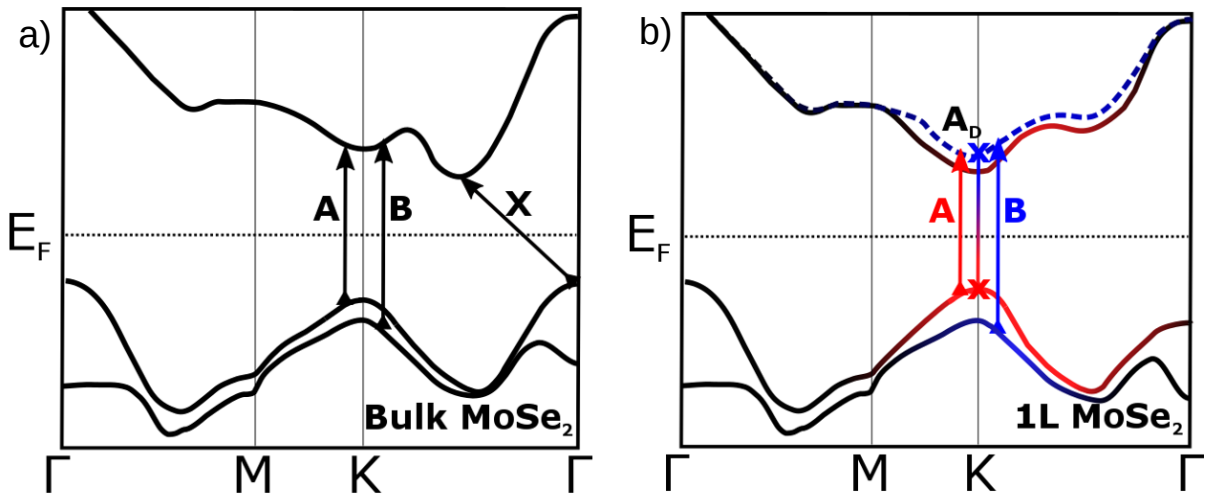

Figure 2.2: The Band Structure of the Semiconducting TMDs

a) The bulk band structure of a typical TMD semiconductor specifically $MoSe_2$. There are three recombination channels the A and B direct excitons and the X indirect exciton pair b) The monolayer structure of the TMD band structure, the A exciton splits into a bright and dark pair with close energies, the B exciton remains.

2.2 Scanning Probe Imaging of TMDs

The STM can be used to see the atomic structure of these materials, since their bandgap is not large enough to limit quantum tunneling, the biggest of which is seen in MoS_2 of 1.5 V is still easily scannable when the bias is within either the conduction or valence bands. Since the STM probes the integrated bands it will always detect the smallest bandgap. For bulk this means that the STM will detect a gap the size of X, roughly .7 V, but in monolayer it will revert to the K-point gap. Since these materials can be very sensitive to air contamination, getting a clean surface into the vacuum system requires that we have a method to exfoliate the material in vacuum.

With Van Der Waals materials the interlayer binding is weak enough that most adhesives are stronger than the out of plane bonds of the material. Therefore to exfoliate the system in vacuum we glue a cleaving post to the top of the surface. This consists of small shaped piece of scrap metal and an epoxy. We then transport the crystal into vacuum where hitting the post with a sharp, fast motion results in a freshly exposed layer.

The STM's imaging capabilities were tested on many candidate TMDs before deciding on one for an in-depth study. Since these materials are mainly semiconducting atomic resolution can be typically found near either band edge since in this state the carriers are localized. An example image of a semiconducting TMD crystal is shown in Figure 2.3a with an inset that shows the atomic placement in Figure 2.3b. This image was taken at a bias of 2V and a current of 150pA.

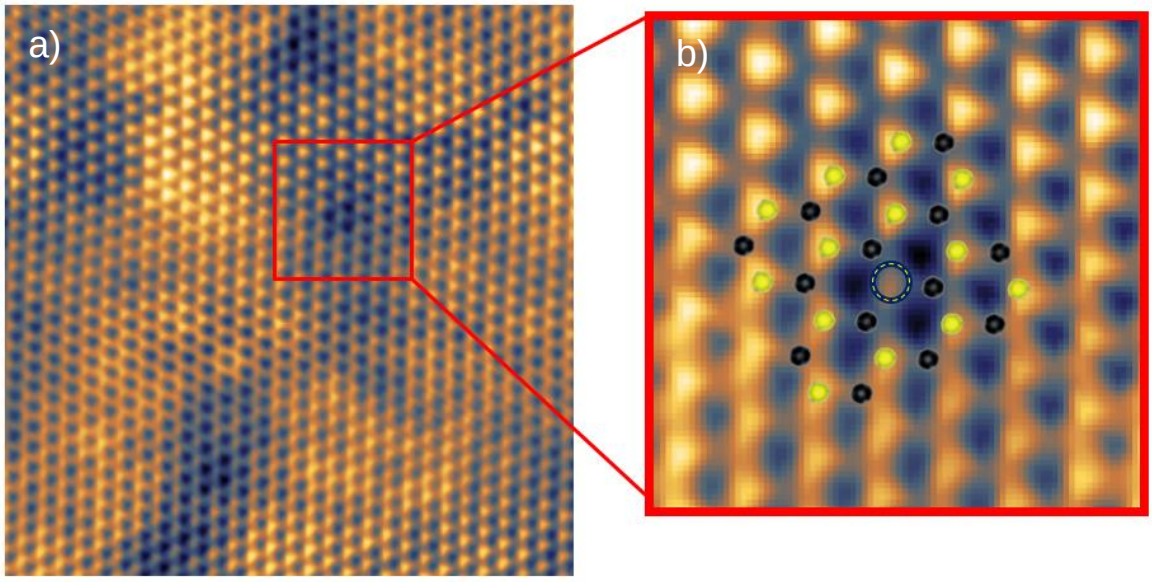

Figure 2.3: Typical 2H TMD Lattice Structure with STM

a) A large area atomic scan of a typical 2H structure b) Zoom in on a lattice vacancy showing clear resolution of the atomic sites and individual defects

Another TMD structure results from a distortion of the hexagonal 2H structure this phase is stable for a couple of the semiconducting class specifically for ones including tellurium (Te). In this distorted structure known as 1T' the resulting crystal is metallic [34,35]. We can again see this system with atomic resolution however it is much harder to visualize atomic positions at reasonable tunneling currents. An image of this TMD is shown in Figure 2.4a with an inset Figure 2.4b that overlays the lattice. In this metallic system to achieve this resolution we use a bias of -1.5 mV and -28 nA of current.

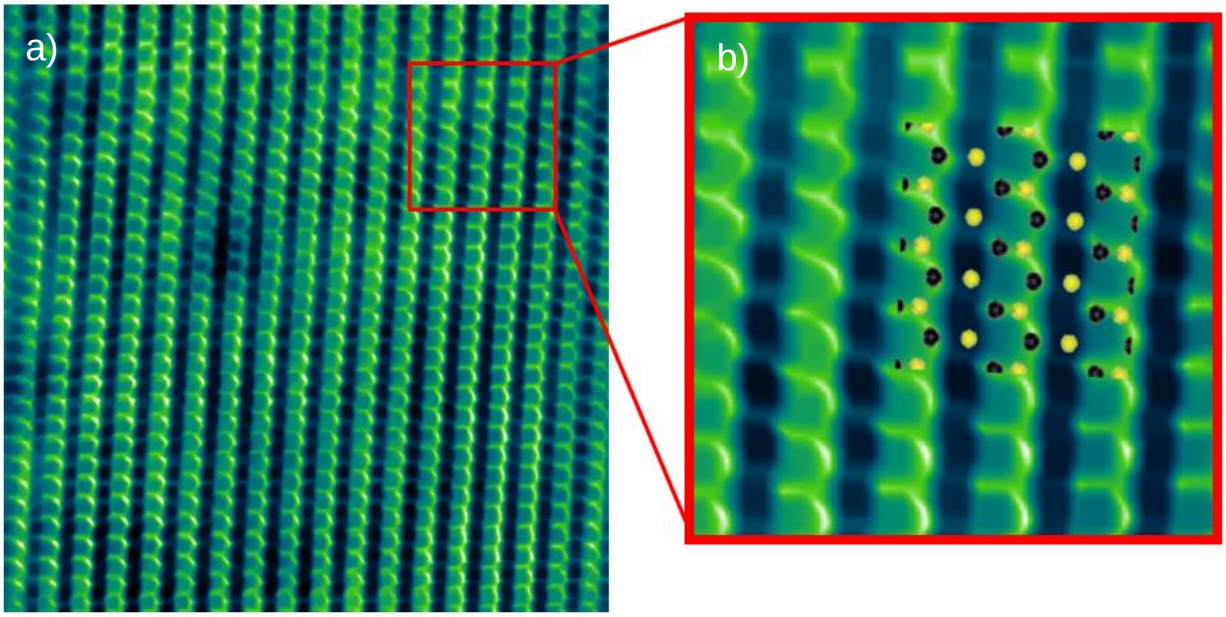

Figure 2.4: Typical Lattice Structure of the 1T' TMD Using STM

a) A large area atomic scan of a typical 1T' structure b) Zoom in of the lattice showing clear resolution of the atomic sites in orthorhombic fashion

2.3 TMD Crystal Growth

To study transition metal dichalcogenides, requires a source for repeated measurements [36] [37]. Since we have seen that the STM is a good probe for seeing the atomic placement and electronic properties of a surface, the best course is to examine the methods for synthetic crystal growth. Bulk TMDs can be grown through two well established methods, with many others that focus on growing the monolayer directly. The two methods that were used to make crystals for our experiments were chemical vapor transport (CVT) and self-flux growths. These methods vary in both total growth time, crystal size and material quality.

When growing crystals through the chemical vapor transport method, constituent elements are placed on one end of a quartz ampoule with an additional transport agent, typically a halogen such as bromine (Br) or iodine (I) [38] [39,40]. This ampoule is then placed in a furnace with a gradual temperature gradient. At the hot end of the tube the transport agent helps the constituent elements to sublime where they recondense as MX_2 crystals in the cold end of the furnace. Since the sublimation point of the

halogen is very low it continues to cycle from the hot to cold ends of the ampoule until all of the reactants have been used. This method is pictured in Figure 2.5a where iodine is being used as the transport agent.

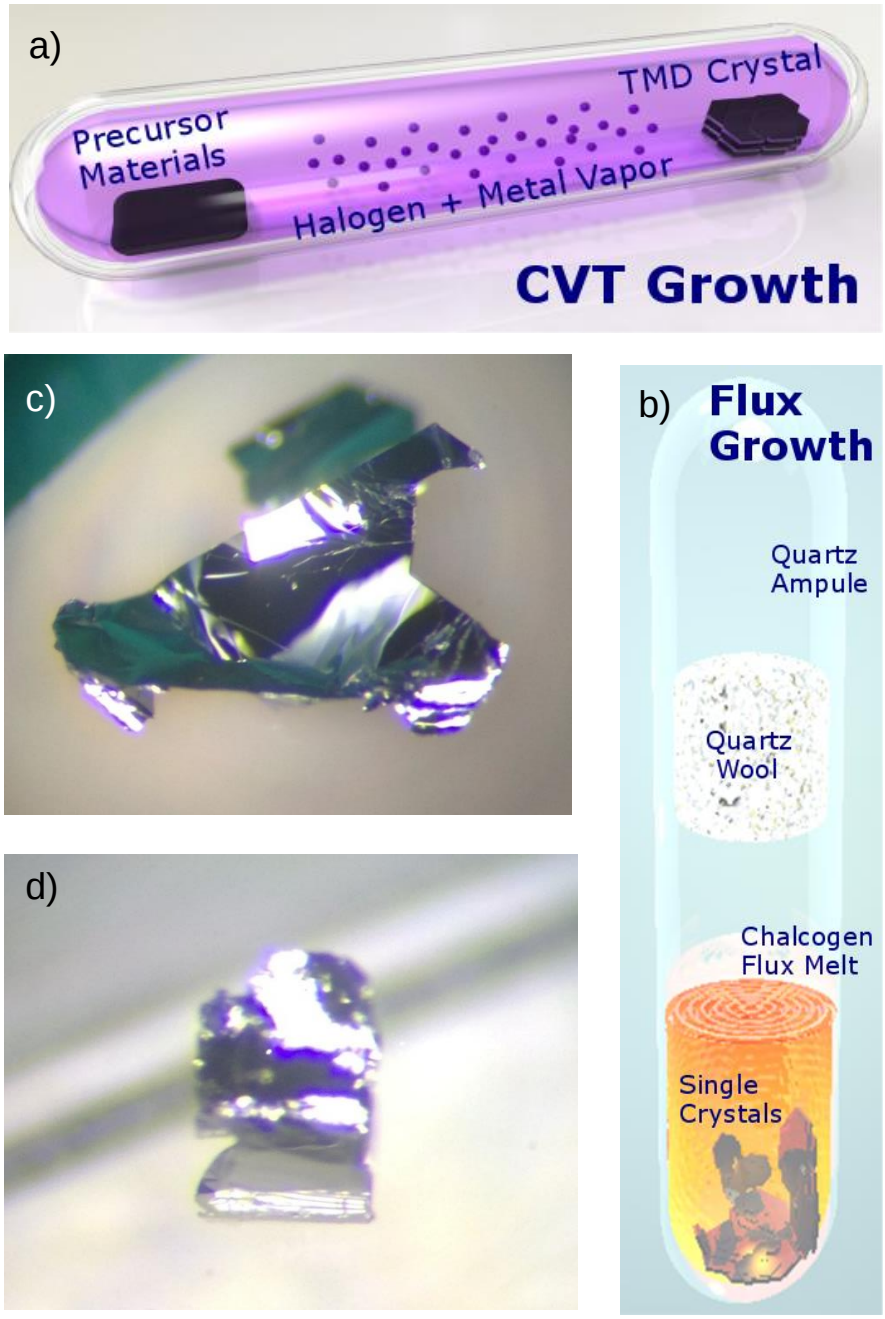

Figure 2.5: Growth Procedures for TMD Crystal Formation

a) A cartoon model of the chemical vapor transport (CVT) growth technique for the formation of TMD crystals b) A similar cartoon showing the growth conditions of self-flux crystals c) A example of a typical CVT crystal after growth d) An example of a typical self-flux crystal after growth

The second method for creation of TMD crystals is the self-flux method [41,42]. Unlike the transport method this uses the chaclogen itself to dissolve the metallic component at high temperatures. Single crystals seed and grow out of this flux-melt over a period of months. The system is then inverted to allow the flux to drain through a barrier of quartz wool. This prevents the crystals from becoming encased in solid chalcogen. An additional anneal cycle afterward allows for the removal of further excess chalcogen since the mixture must contain on the order of 20 X : 1 M. A diagram denoting the growth method is shown in Figure 2.5b.

The crystals that resulted from these two methods typically varied in macroscopic parameters, however some general trends were noted before applying scanning tunneling microscopy techniques. Chemical vapor transport (CVT) methods created larger crystals in a much faster time. Crystals of this growth method are shown in Figure 2.5b. It was found that these batches would occasionally appear with microcrystals of pure selenium, additionally the halogen used had to be dissolved by heating the crystal in isopropanol or acetone. Flux method resulted in much smaller crystals shown in Figure 2.5d. Although much smaller it appeared that the longer growth time resulted in crystals that were consistently monocrystaline. With such varied macroscopic parameters we would expect the microscopic composition to be affected; which is exactly the type of question we can answer through STM.

2.4 Experimental Methods

All STM measurements to follow were performed using a custom built, variable temperature, UHV STM system. Single crystals of $MoSe_2$ and WSe_2 were mounted onto metallic sample holders using a vacuum safe silver paste. Samples were then transferred into the STM chamber and cleaved in-situ, exposing a clean surface. A Pt-Ir STM tip was cleaned and calibrated against a gold (111) single crystal prior to the measurements. Measurements were collected at 82 K and 300 K.

$MoSe_2$ and WSe_2 crystals were synthesized by reacting Mo/W powders, 99.999%, with Se shot, 99.999%, typically in a ratio of 1:20. These materials were first loaded into a quartz ampoule. A piece of quartz wool is then pressed into a cylindrical shape and pushed into the quartz ampoule, approximately 1

cm above the raw elements. The ampoule was then evacuated and sealed at ~10^{-3} Torr. For growth, the ampoule is heated to 1000 °C over 48 hours, held there for 3 days, then cooled at a rate of 1.5 °C down to 400 °C and subsequently flipped and centrifuged. Crystals are then harvested from the quartz wool filter and annealed at a temperature of 250 °C with the empty end of the quartz ampoule held approximately at 100 °C for 48 hours.

$MoSe_2$ crystals were synthesized by reacting Mo powder, 99.999%, with Se shot, 99.999%, in stoichiometric proportions with iodine 99.999% as a transport agent. These materials were first loaded into a quartz ampoule 12 cm in length, 1 cm in diameter, then evacuated and sealed at ~10^{-3} Torr. For growth, the ampoule is heated to 1000 °C over a period of 48 hours, held there for 1 week, then cooled for 3 days to 750 °C and subsequently quenched in air. Crystals are then harvested and rinsed in acetone and isopropanol to remove iodine residue, and left to dry.

Chapter 3

Defect Dynamics of Transition Metal Dichalcogenides

3.1 TMD Semiconductors

Two dimensional (2D) transition-metal dichalcogenide (TMD) semiconductors have been shown to possess many novel optical and electronic properties, with multiple potential engineering applications. This is all the more true as synthetic methods for crystal growth become well characterized. Starting from the macroscopic crystals we quantify their properties on a microscopic scale; specifically using metrics where optical examination only provides qualitative suggestion. We use the STM to characterize the atomic and electronic nature of point defects that are intrinsic to single crystals of these materials synthesized by two different methods - chemical vapor transport and self-flux growth. When combining scanning tunneling microscopy (STM) with scanning transmission electron microscopy (STEM), we show that the two major intrinsic defects in these materials are metal vacancies and chalcogen antisites. With further control of the synthetic conditions, we can reduce the defect concentration from above $10^{13}/cm^2$ to

below 10^{11}/cm^2. Because naturally occuring point defects act as centers for non-radiative recombination of excitons, this improvement in material quality leads to a hundred-fold increase in the radiative recombination efficiency.

The semiconducting transition metal dichalcogenides (TMDs) promise a wide range of applications in electronics and optoelectronics [43-46]. These materials host novel phenomena such as valley physics [47 48], interlayer tunneling[49,50], topological properties[51,52], and exciton superfluidity[47] which are of fundamental interest and may enable new device functionality. However, crystalline disorder obscures intrinsic phenomena and imposes an upper limit on achievable functionality[47,48,53,54]. In particular, point defects[55,56,57,58] strongly impact TMD monolayers: these defects cause carrier scattering and localization[59-61], act as centers for non-radiative recombination[62-65], and give rise to localized emission from excitonic traps[66,67]. Pioneering transport[49,68,69] and STEM studies[55,56,57] have explored the atomic nature and electronic impact of defects arising in monolayer TMDs, and indicate that the quality of these materials remains far behind the classic semiconducting materials such as Si and GaAs. Addressing the quality of these materials is urgently needed to advance their science and engineering applications.

3.2 Intrinsic Defects in Layered Structures

In two-dimensional materials, disorder can arise from both intrinsic sources, such as point defects and grain boundaries in the crystal itself; and extrinsic sources arising from the environment, such as inhomogeneous strain, and charge traps / adsorbates in the substrate[56,70]. In the case of mechanically exfoliated graphene, the intrinsic defect density is extremely low (10^9-10^{10}/cm^2 [71,72]), and reducing extrinsic disorder by encapsulation in hexagonal boron nitride (hBN) has enabled spectacular advances in device performance[73-75]. However, most other 2D materials do not necessarily possess graphene's ultrahigh purity. In the case of semiconducting TMDs, hBN encapsulation also results in improved performance[76,77], but the physical properties of these devices are still far from their theoretical limits, indicating that intrinsic disorder plays an important role. This is consistent with studies showing point

defect densities exceeding $10^{12}/cm^2$ in commonly used TMD materials[78]. Therefore, continued progress in the field necessitates the characterization, quantification, and minimization of defects in TMD materials. Toward this end, a particular challenge is the diversity of material sources, which include natural or synthesized single crystals, and large-area films grown by chemical vapor deposition (CVD)[39,40], metal organic chemical vapor deposition[79], physical vapor deposition[80], and molecular beam epitaxy[81], each of which can give rise to a different density and type of defects. For instance, STEM imaging of MoS_2 reveals that CVD-grown films are dominated by S-vacancies, whereas defects in natural MoS_2 crystals are predominantly Mo vacancies[56].

As mentioned in the introduction to transition metal dichalcogenides and other Van Der Waals materials, we focus on the quality of synthesized TMD single crystals. Currently, single crystals remain the source of the highest-quality TMD monolayers[39,53,82,83], and do not suffer from grain boundaries and phase separation[58,84,85] observed in large-area films. While much initial work on TMDs has utilized naturally occurring minerals[49,86], laboratory-synthesized crystals provide a wider materials selection, and can offer a higher degree of quality control and reproducibility. Toward this end, a number of companies are currently supplying synthesized TMD crystals for laboratory use. However, synthesized TMD crystals have not been well characterized, and virtually no experimental work has examined the correlation between defect density and optoelectronic properties of monolayers derived from these bulk crystals. We use STM and STEM imaging to determine the type and density of intrinsic defects present in single crystals of $MoSe_2$ and WSe_2 synthesized by the chemical vapor transport (CVT) and flux growth techniques.

3.3 Review of Growth Processes

The CVT technique utilizes a transport agent, usually a halogen, to transport starting materials from a hot region into a cooler growth region where they form crystals[38]. CVT provides large crystals in a relatively short growth time at moderate temperatures, and thus has become the prevalent technique for TMD synthesis. The self-flux method in which crystals are grown directly from the molten phase is an

alternative method known to create higher-quality, albeit smaller, crystals[42]. In this work, we characterize three types of crystals: commercially obtained crystals grown by CVT and tested without further annealing (as-grown CVT or ag-CVT); crystals synthesized in our furnaces by CVT then annealed in a temperature gradient (treated CVT or t-CVT); and crystals grown by the self-flux method (flux). Details of the growth procedures, temperatures, and cooling rates are given in the previous chapter.

We first examine defects in bulk crystals through scanning tunneling microscopy (STM), which can provide defect lattice positions, local electronic structure, and defect density. To avoid surface contamination, crystals were cleaved *in situ* under UHV conditions. Scanning tunneling microscopy is a valuable tool to characterize defects because it has access to atomic resolution of the surface layer. Referring to the structure of the 2H phase of these compounds we see that based on the Van Der Waals binding the STM is in direct contact with the top chalcogen plane of the X-M-X structure. The metal sites, since not directly in contact with the STM tip, must further be inferred from the positions of these top chalcogen atoms which form a triangular lattice. Electronic states will also play into our recognition of the lattice, therefore we must observe the surface under many scanning conditions to determine defect lattice sites and composition

3.4 STM Topography of TMD Semiconductors

STM images were taken for a series of scanning bias voltages to compare the height contrast of defects, providing an assay of defect evolution with bias and current. It was found that defects changed their appearance for the opposite sign of scanning bias, which can be attributed to a change of electronic density when switching carriers. Otherwise bias alterations had little effect on the defect contrast. To illustrate this contrast change under bias flip, images taken on the same chalcogen site defect are presented in Figure 3.1a,b taken at positive (a) and negative bias (b) respectively. We additionally observe a central atom is present for this defect type at all biases, as indicated within the circled region.

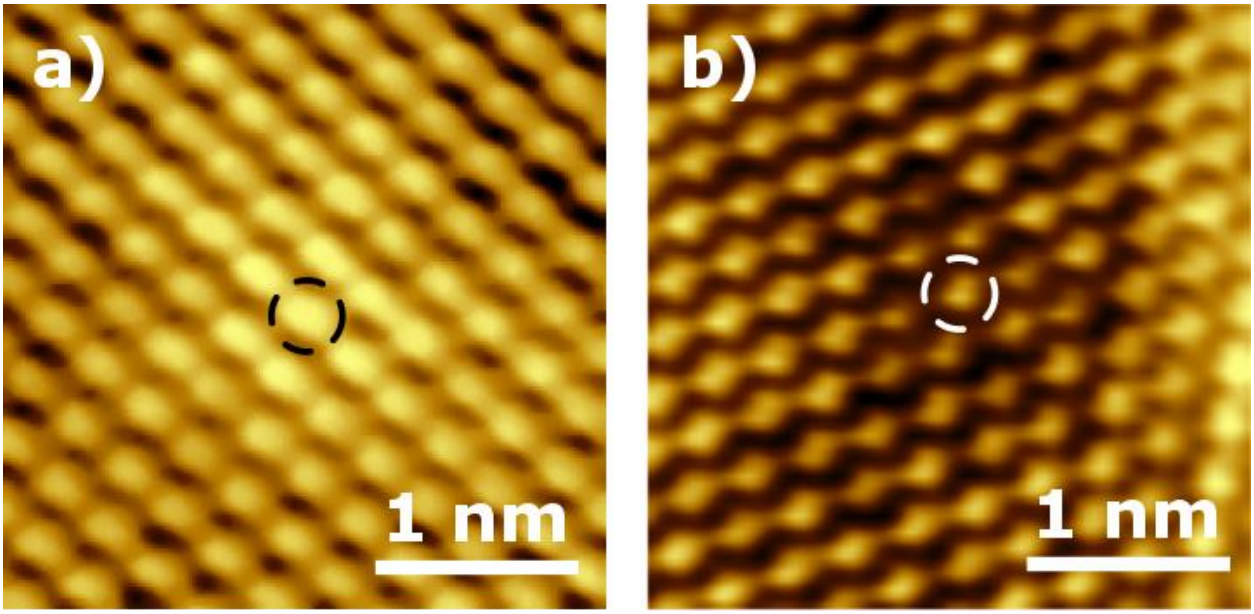

Figure 3.1: Bias Effects on STM Imaging

a) The surface of $MoSe_2$ when imaged at a positive bias of 1V a single defect can be seen identified within the black circle b) Another image of the same region but at a negative bias of -1 V at this voltage the same bright defect appears dark, however all atoms remain visible as is seen in the white circle

With knowledge of how the surface will respond under various scanning conditions we examine a 25 nm square region of $MoSe_2$ (Figure 3.2a), we observe two defect types that can be initially identified by contrast as either “dark” or “bright”. As discussed further below, these two predominant types of defects account for the vast (>99%) majority of defects imaged in both flux and CVT samples. Figure 3.2b and Figure 3.2c show atomic resolution images of these defects. The bright defects, which we denote – X, are located on a selenium site (Figure 3.2b). Since there is no missing atom associated with this defect, it is not a selenium vacancy[78,87], but is rather a substitutional impurity on the chalcogen site, as we saw in Figure 3.1 the atomic site for this defect has an atom at all measured scan conditions. Se vacancies can indeed be observed by STM but are roughly two orders of magnitude less common than the vacancies and substitutional impurities

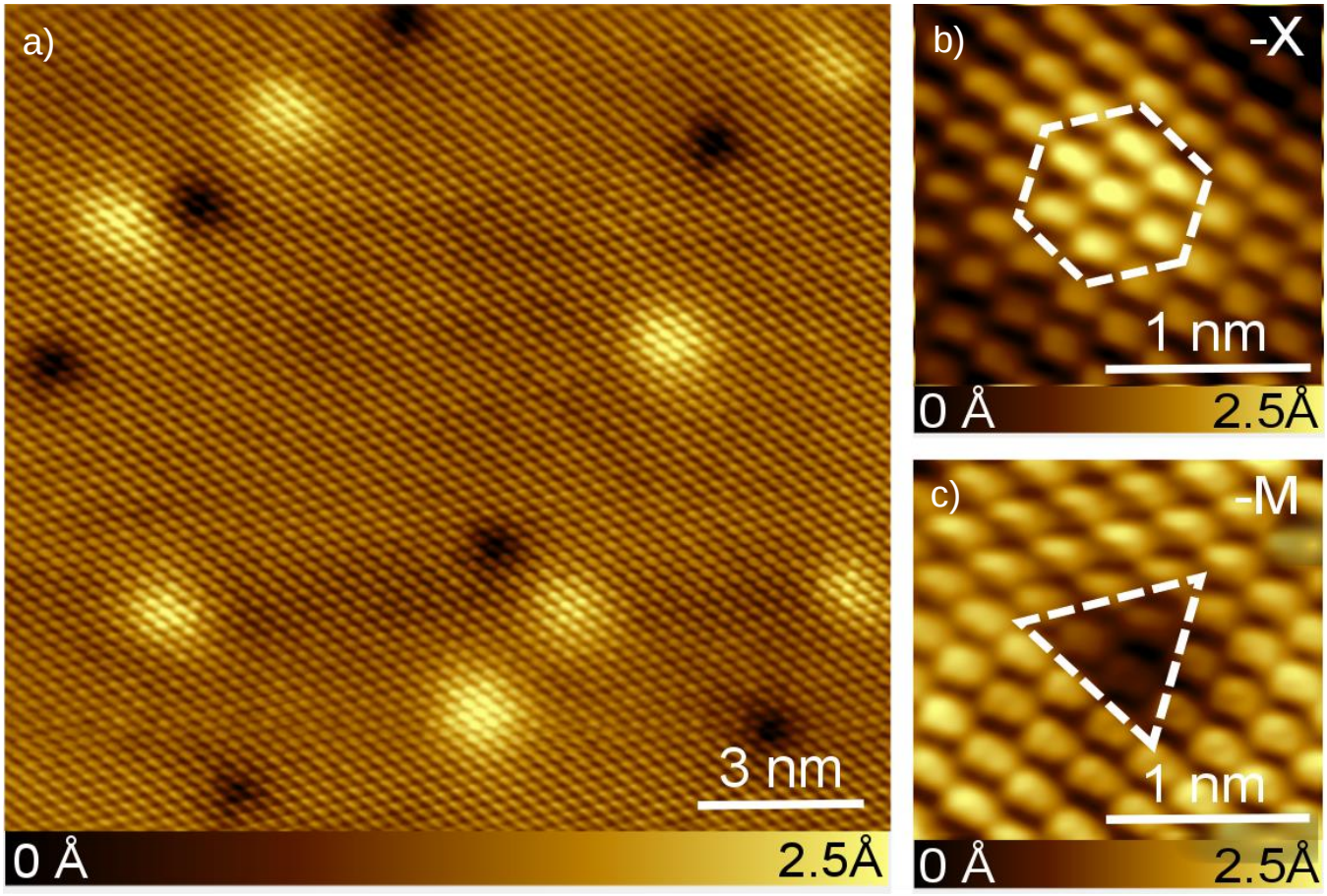

Figure 3.2: STM Topography Showing Two Predominant Defect Types

a) An atomic resolution image of the surface of $MoSe_2$ showing two defect types a "dark" and a "bright" species. b) A 2nm x 2nm zoom in around the bright defect with its characteristic hexagonal shape created by a defect centered on the chalcogen (-X) site. c) A 2nm x 2nm zoom in around the dark defect with its characteristic triangular shape created by a defect centered on the metal (-M) site.

The dark defects observed in STM images, which we denote –M, are aligned with the Mo sites, which are located in the center of a triangle of selenium atoms (Figure 3.2c). Interestingly, whereas flux-grown $MoSe_2$ possesses both –M and –X defects, similarly grown WSe_2 displays predominately –M defects. Atomic resolution measurements were taken on WSe_2 as seen in Figure 3.3a. These images appear with the opposite contrast when compared to those shown previously due to opposite choice of STM bias. As was the case of $MoSe_2$ two defect types were observed, one occurring on the metal site and one occurring on the chalcogen site. We again label the metal site defect as seen in Figure 3.3b as a metal vacancy and the chalcogen site defect seen in Figure 3.3c as a metal antisite. STM scans on both WSe_2 and $MoSe_2$ suggest that these defect types are common to the TMD family.

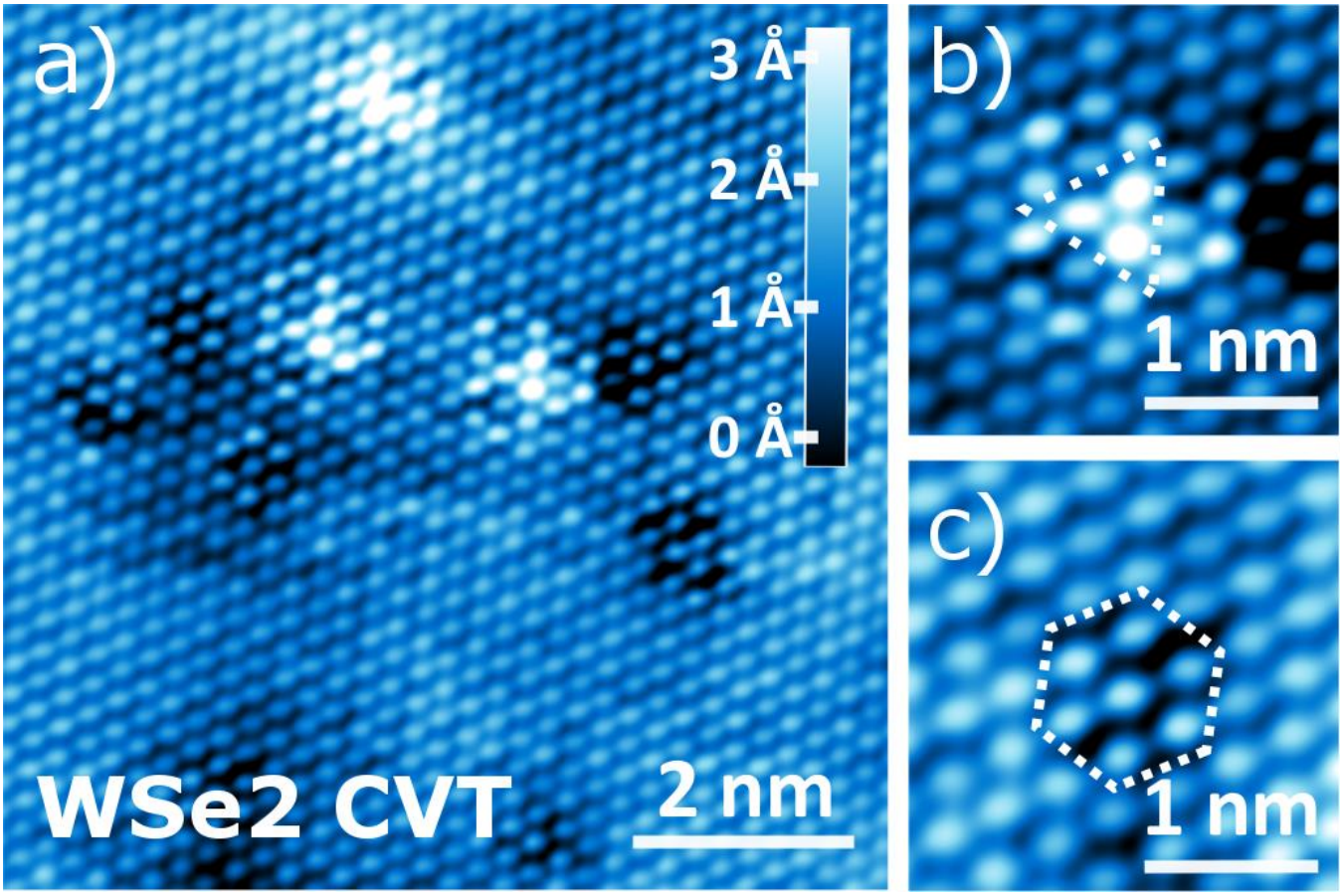

Figure 3.3: STM Topography of WSe_2 Showing Two Predominant Defect Types

a) An atomic resolution image of the surface of WSe_2 showing two defect types a "dark" and a "bright" species due to an inverted bias this image is the opposite contrast of the previous figure. b) A 2nm x 2nm zoom in around the characteristic triangular shape created by a -M defect. c) A 2nm x 2nm zoom in around the characteristic hexagonal shape created by a defect centered on the -X site.

3.5 Scanning Tunneling Spectroscopy of TMD Defects

Further insight into these defect types can be obtained by looking at the electronic structure using scanning tunneling spectroscopy (STS). Figure 3.4a shows a sequence of tunneling spectra measured at varying distances from a single -X defect, out to a distance of 3.5 nm. Directly over the defect, we measure a broad resonance pinned to the edge of the conduction band. The resonance shifts as a function of distance from the defect site, likely due to band bending effects[88]. This indicates that the –X defects behave as n-type dopants. STM itself cannot identify the elemental composition of defects. However, we have observed that the concentration of –X defects can vary widely between crystals

depending on the growth method, even when the same starting raw materials are used in the syntheses. This indicates that the defect is not associated with a foreign substituent, but is most likely an antisite defect, i.e, a Mo atom substituting for a Se atom as suggested by theoretical calculations. The antisites are observed in roughly equal numbers on the top and bottom selenium layers of the MX_2 unit cell as expected as we will see later for large area scans.

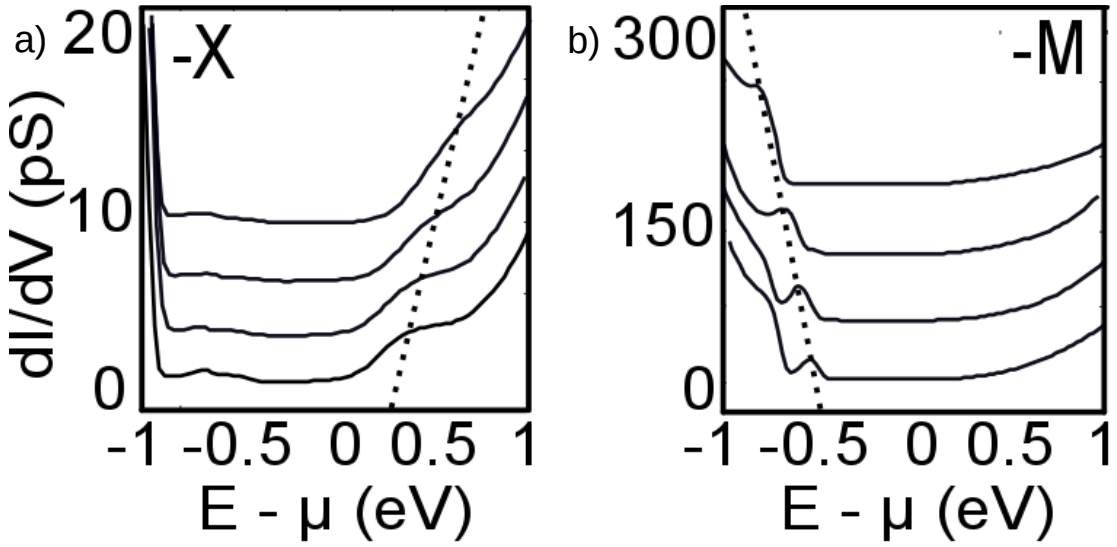

Figure 3.4: Scanning Tunneling Spectroscopy Taken Over Common Defects Types

a) Scanning tunneling spectroscopy (STS) taken over a chalcogen type defect shows a state on the edge of the conduction band that persists even 5nm away from the defect site b) A similar spectra taken over a metal defect shows a gap edge state in the valence band

In contrast to the –X defects, the -M defects show a resonance near the edge of the valence band, as shown in Figure 3.4b. This indicates that the –M defects are acceptor type defects. From a combination of spectra and positioning we can attribute this defect to metal vacancies. Since this is the absence of an atom there is no metal present to supply electrons to the lattice at this location, making the defect act as an electron acceptor. Defect spectra are further compared to spectra taken at 10 nm away shown in Figure 3.5a,b. In this figure we can see the same curves presented in Figure 3.4 presented against a dashed curve which is the corresponding pristine spectra. When looking at the comparison in this manner the defect peak is readily apparent.

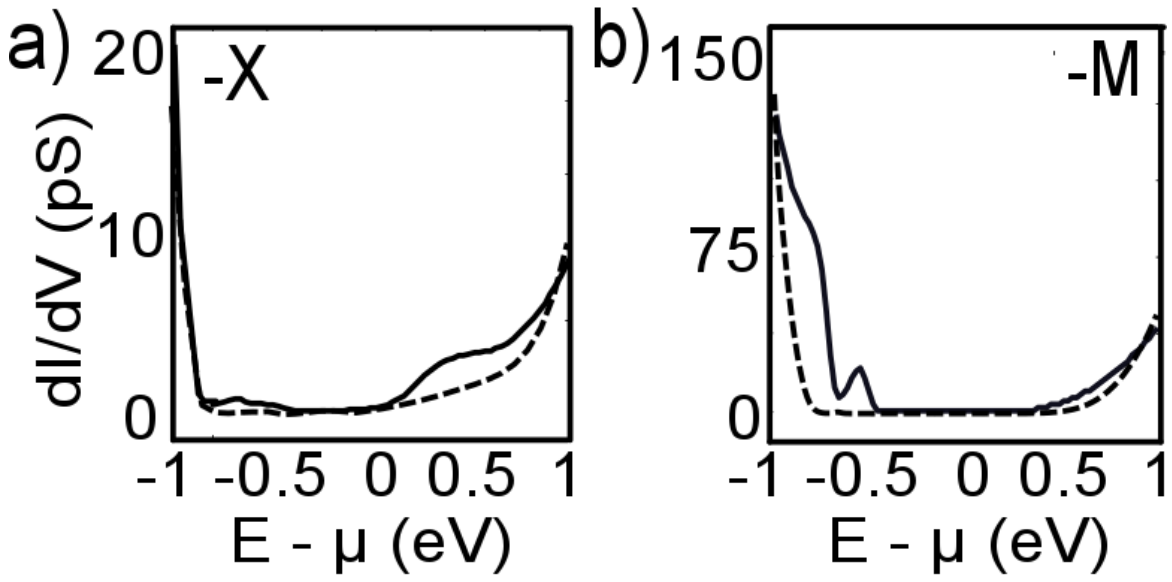

Figure 3.5: Defect Spectra Compared WIth Pristine $MoSe_2$ Spectra

a) A comparison of scanning tunneling spectroscopy (STS) taken over a chalcogen type defect compared with a dashed pristine curve b) A similar spectra taken over a metal defect where gap edge state can clearly be seen above the pristine

3.6 Rare TMD Defect Types

After imaging all possible native defects we found that transition metal vacancies and anti-sites are found to be the most-common defect types, with Se vacancies being much rarer. In our STM imaging experiments, >99% of all of the point defects observed were either the metal vacancy (-M) or chalcogen antisite (-X). We have very occasionally observed missing chalcogen structures in the lattice as shown in Figure 3.6. The vacant atom is observed as a depression in STM topography at all bias voltages, consistent with the missing atom on the surface. The depth of these chalcogen vacancy structures were measured to be 300pm, an order of magnitude deeper than other observed defect types. This defect does not have an observable in-gap state and does not affect the bandgap significantly. Referencing the formation energy from DFT single chalcogen vacancies have a formation energy of 1.81eV, and double vacancies require much higher formation energy of 4.88eV. Therefore we conclude that this vacancy structure represents a single point vacancy. The density of all chalcogen vacancies structures were too low to estimate accurately the area density for the crystals we scanned. Since the density should be

proportional to formation energy, we do not report findings of the double chalcogen vacancy through STM.

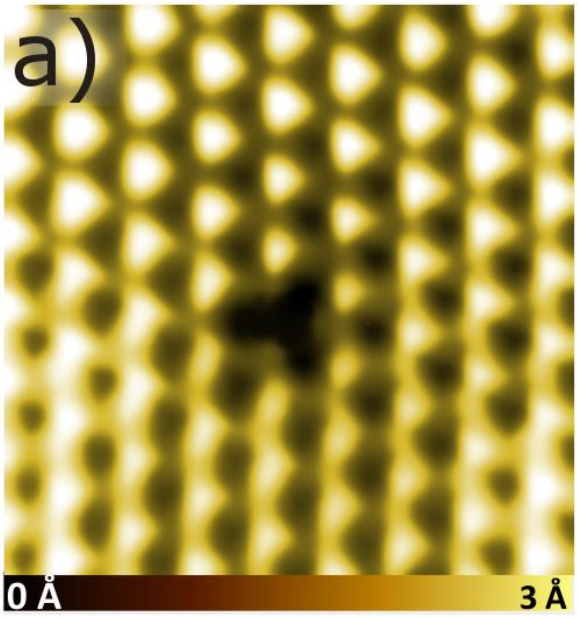

Figure 3.6: Chalcogen Vacancy Defect

a) Scanning tunneling imaging of a chalcogen vacancy a rare defect occurring in pristine bulk crystals

3.7 Large Scale TMD Defect Analysis

Now having fully characterized the defects present on the small scale we now proceed to quantify defect densities on a larger scale. To count defects and obtain a density we switch to large-area STM imaging. Figure 3.7a-c show topographic scans (0.5 µm x 0.5 µm) of the three $MoSe_2$ materials under study. In the ag-CVT sample (Figure 3.7a), the defect density is high enough such that the individual point defects have overlapping electronic signatures, and therefore STM can only provide a lower bound on the defect density of $>10^{13}/cm^{-2}$ (1% of unit cells). This defect density is dramatically reduced, to (2.5 ± 1.5) $x10^{12}$ cm^{-2} (0.2%) in the t-CVT sample (Figure 3.7b). The self-flux crystals display still lower defect density of (1.7 ± 0.5) $x10^{11}$ cm^{-2} (0.01%) (Figure 3.7c).

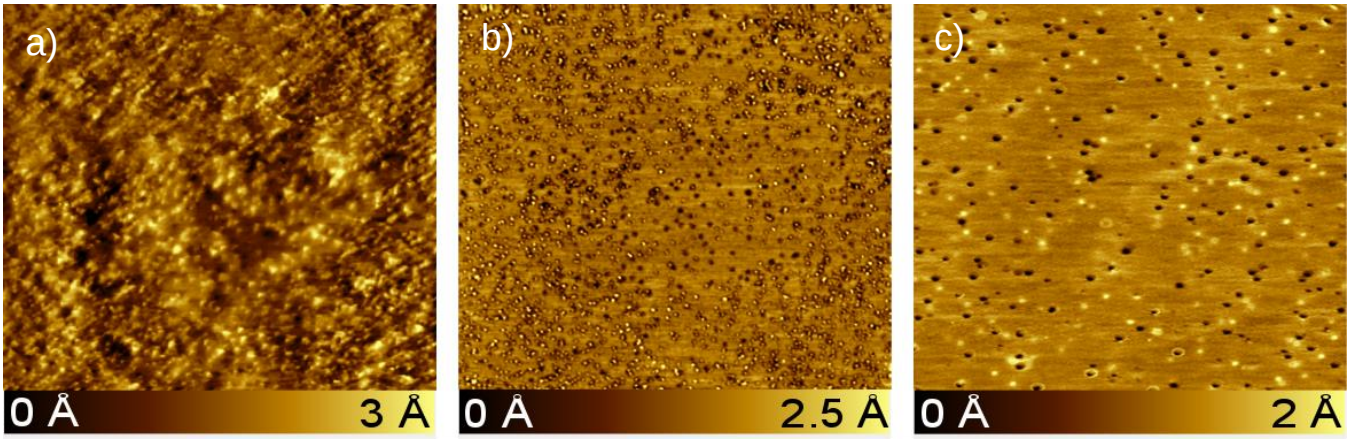

Figure 3.7: $MoSe_2$ Growth Comparison on a 0.5x0.5 μm^2 Scale

STM imaging of a 0.5 x 0.5 μm^2 area in order to count defect density in a) ag-CVT b) t-CVT and c) self-flux grown crystals of $MoSe_2$

Shifting our focus to WSe_2 we perform the same comparison. Commercial ag-CVT WSe_2 exhibits a very high defect density and STM imaging can only provide a lower bound of $>10^{12}$ cm^{-2} (0.1 %) (Figure 3.8a). In the flux-grown WSe_2, the defect density is dramatically smaller, (7.0 ± 2.2) $x10^{10}$ cm^{-2} (0.006%) (Figure 3.8b). This defect density is by far the lowest reported for any TMD semiconductor.

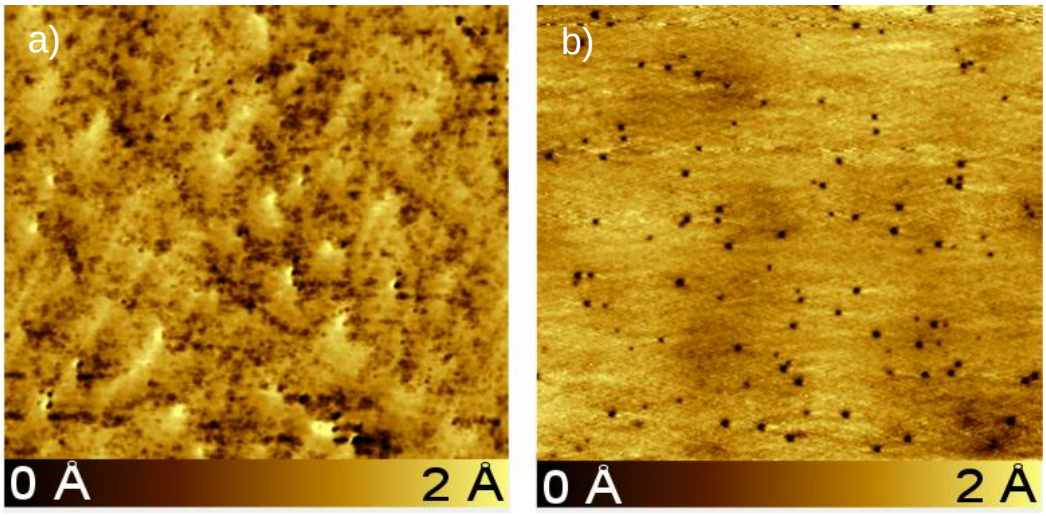

Figure 3.8: WSe_2 Growth Comparison on a 0.5x0.5 μm^2 Scale

STM imaging of a 0.5 x 0.5 μm^2 area in order to count defect density in a) ag-CVT and b) self-flux grown crystals of WSe_2

Since the STM is able to distinguish defect types by color contrast we can process STM images to count individual defect densities. Chalcogen antisite (-X) and metal vacancy (-M) defects are distinctly different from each other at an imaging bias of V=1.25 V and current of I=100 pA. Under these conditions, the –X defect appears as a roughly 10 pm bump in STM topography, as shown in Figure 3.9a, while the –M defect appears as a roughly 10 pm depression (Figure 3.9b). Large scale images can be processed by applying height thresholds to identify directly the –X (red) and –M (blue) defects, Here we specifically demonstrate this procedure on the raw data from Figure 3.7c, this results in the image in Figure 3.9c**.** Labeled defects can now be counted either manually or by edge detection algorithms. After sorting defect counts by type for each growth method, we produce a summary of these results for each growth in Table 3.1.

Table 3.1 : Defect Densities by Growth Methods Based on Defect Species and Crystal Type

Crystal Growth Method	Total Defect Count $Defects/cm^2$	M- Site Defects $Defects/cm^2$	X- Site Defects $Defects/cm^2$
$MoSe_2$ ag-CVT	$> 10^{13}$	$> 10^{13}$	$> 10^{13}$
$MoSe_2$ t-CVT	$(2.5 \pm 1.5) \times 10^{12}$	$(1.3 \pm 0.8) \times 10^{12}$	$(1.12 \pm 0.6) \times 10^{12}$
$MoSe_2$ Self-flux	$(1.7 \pm 0.5) \times 10^{11}$	$(7.0 \pm 0.2) \times 10^{10}$	$(9.91 \pm 0.2) \times 10^{10}$
WSe_2 ag-CVT	$> 10^{12}$	$> 10^{12}$	$> 10^{11}$
WSe_2 Self-flux	$(7.0 \pm 2.2) \times 10^{10}$	$(6.41 \pm 2.0) \times 10^{10}$	$(5.83 \pm 1.8) \times 10^{9}$

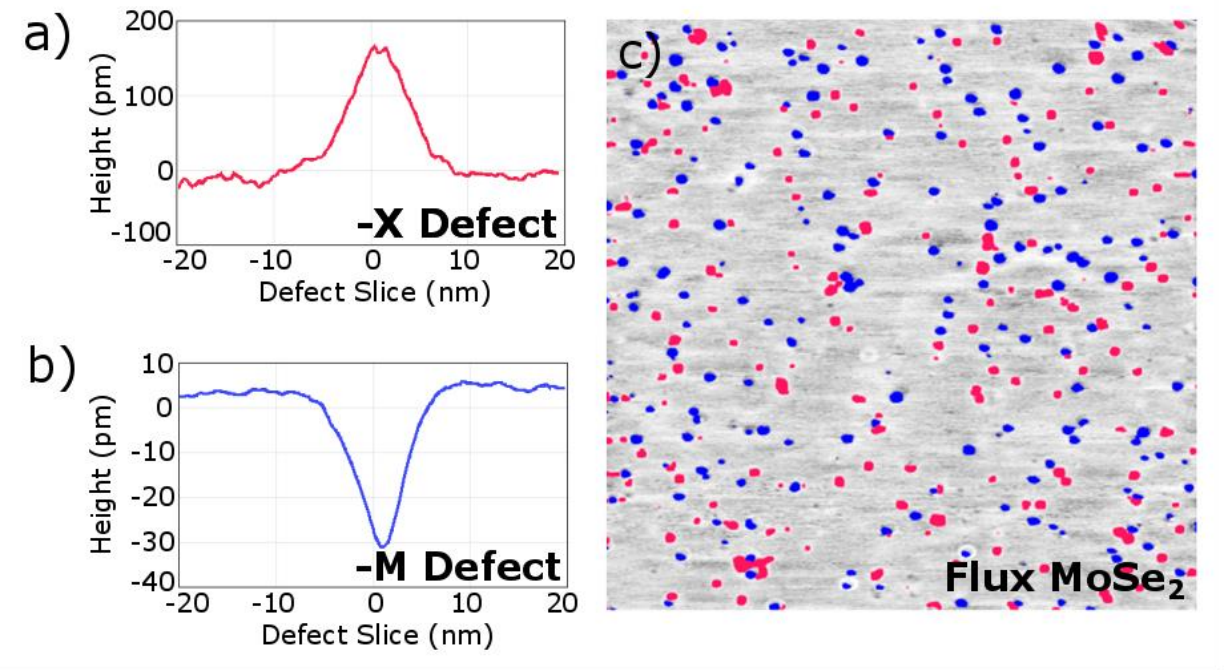

Figure 3.9 : Automated Counting of Defect Species Using a Height Threshold

a) Lattice profile of a chalcogen site defect, where it is visualized as a 200 pm electronic protusion b) A profile of a metal site defect with an apparent electronic depth of 30 to 40 pm c) A false color reproduction of an image taken on the flux crystal to demonstrate the defect counting algorithm. Defect species are labeled by color –X (red) and –M (blue)

3.8 TMD Defect Sorting

In the counting procedure we additionally can sort chalcogen defects by position within a single TMD layer. Since the metal atoms are capped by a top and bottom chalcogen we have two possible orientations for our antisite defects (the red marked ones in Figure 3.9c), either X-M-D or D-M-X, where D represents the lattice defect. Since defect placement is random we expect chalcogen antisite defects (bright) should occur on both sides of this layer with an equal probability. In our STM images, we can distinguish between these two antisite defects from their apparent topographic height. We first demonstrate this procedure on a small area. Using the topographic image of Figure 3.10a,b (reproduced from Figure 3.2) we switch to a three dimensional banded colorscale. From the three dimensional image in Figure 3.10c, we can see that the bright defects (antisite) have two topographic heights (110 and 170 pm). This is consistent with some of them being on the top chalcogen layer while some of them are in the

bottom chalcogen layer. The dark defects on the other hand all have the identical topographic depth of -30 pm. The consistent observation that the dark defects have the same apparent depth while the bright defects have two apparent heights (consistent with the top and bottom chalcogen layer). This convinces us that the point-defect features we see in our images all come from the first MX_2 layer. We do see weak topographic features over much larger length scales (nanometers) that could arise from additional electronic and topographic detection of defects in our large area topographs. We note that these extended corrugations may have multiple origins including defects from layers below[39] or grain defects occurring on a larger scale.

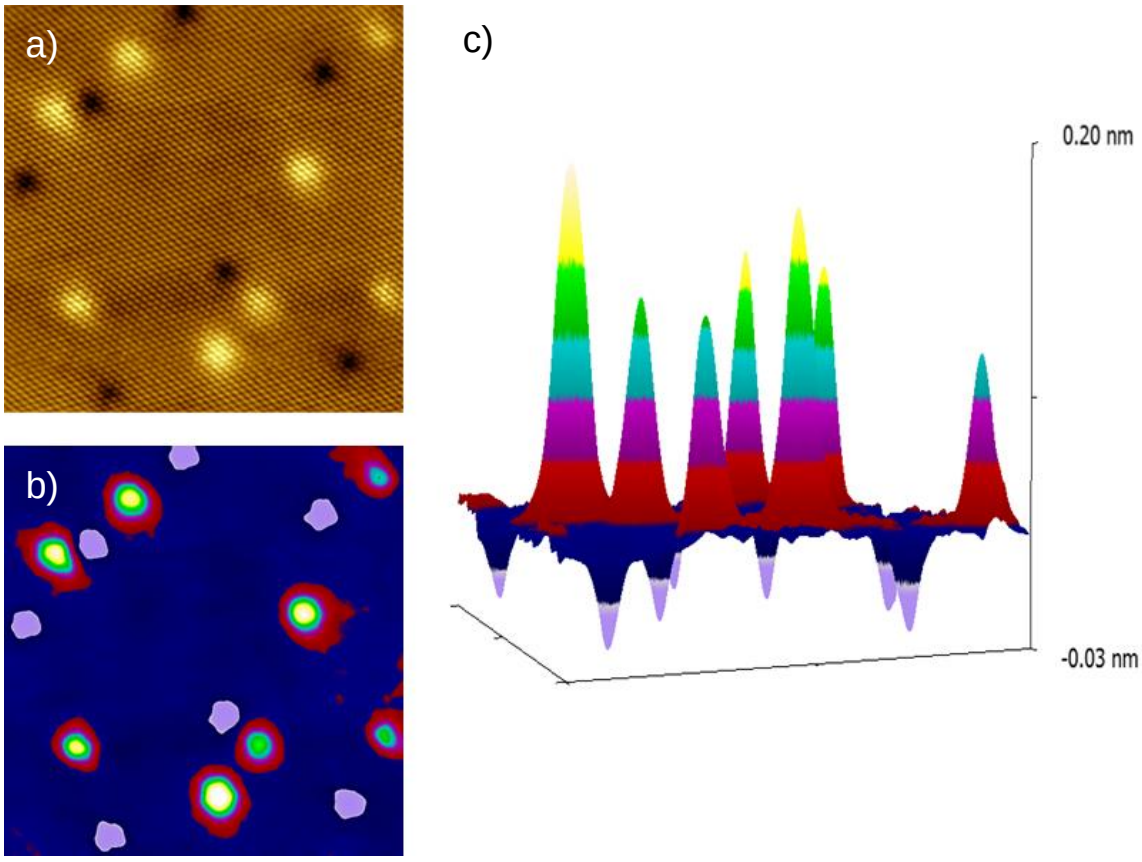

Figure 3.10: Distinguishing -X Defect Lattice Position by Apparent Height

a) Reproduction of the $MoSe_2$ image where we look for two variants of bright defects. b) Labeling these defects using a banded colorscale we can see that there are two apparent heights which should correspond to their lattice position, top or bottom chalcogen site within a single layer c) The same area but plotted in 3D to highlight the consistent variation of height

Shown in Figure 3.11a we use the same procedure to label a large scale STM image (again with nicely separated bright defects). We choose a different color scale that emphasizes the difference in the two types of antisite defects. More quantitatively, we plot a histogram of the topographic heights of all the bright defects extracted from this image in Figure 3.11b. The histogram clearly shows the presence of a bimodal distribution of the antisites. For this particular image, we see 61 antisite defects on the top layer and 79 on the bottom layer. Similar bimodal distributions of antisites are seen in other images.

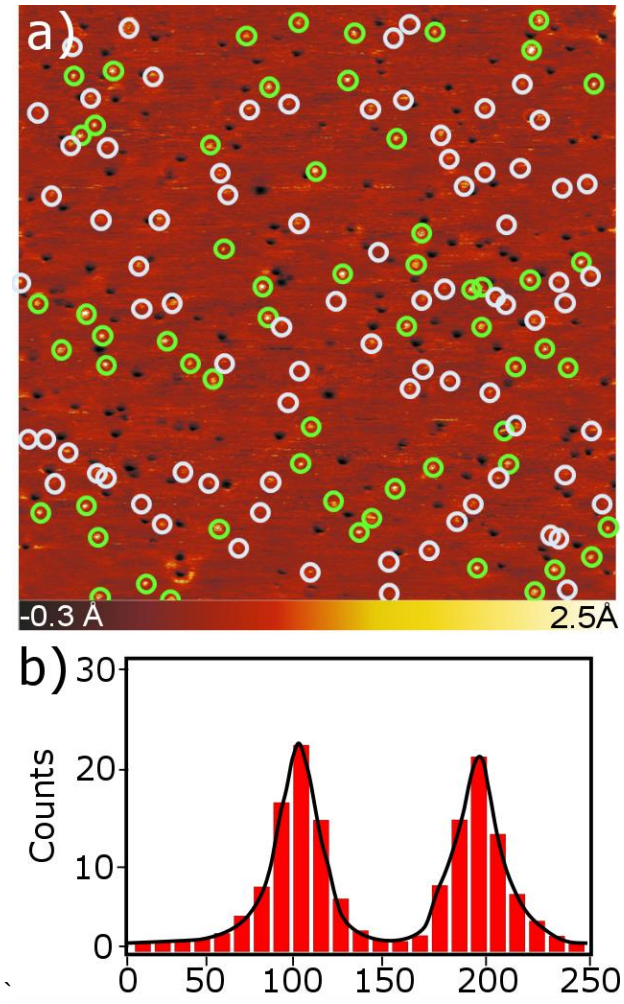

Figure 3.11: Large Scale Delineation of Bright Defects Based on Positions Within the Top Layer

a) We now apply the same fine height filtering to the larger image confirming that two close variants of bright defects can be distinguished based on their electronic heights b) Plotting the height distribution we see that these match nicely with the atomic scale image, and therefore can infer that these are the same defect type

Chapter 4
Scanning Transmission Electron Microscopy and Theoretical Approaches to TMD Defect Identification

4.1 Overview of Scanning Transmission Microscopy

Since the scanning tunneling microscope provides information about defects, without exact chemical composition, another technique is useful in crosschecking the data acquired. For this we use scanning transmission electron microscopy (STEM) which provides data regarding the elemental composition of lattices and their defects.

For the preparation of electron transparent samples for point defect density measurements, the TMD crystals were mechanically exfoliated using Scotch™ tape. The exfoliated flakes were transferred onto oxidized silicon wafer substrates. The monolayer flakes of the exfoliated TMD crystals were identified using light optical microscopy. The monolayers were then transferred onto Quantifoil® holey carbon TEM grids using isopropyl alcohol as a medium. After the alcohol evaporated, the holey carbon grid was attached to the wafer and the monolayer. The wafer sections were then slowly immersed in 1M potassium hydroxide (KOH) solution to etch the very top surface of the oxide and release the TEM grid and the exfoliated crystals attached to it. Distilled water was used to dilute and wash away the KOH solution from the TEM grids. As a last step, TEM grids were immersed in warm (40 ^{o}C) acetone for 10 minutes to dissolve any residue that remained from the exfoliation and sample transfer. The resulting sample is presented in Figure 4.1.

a)

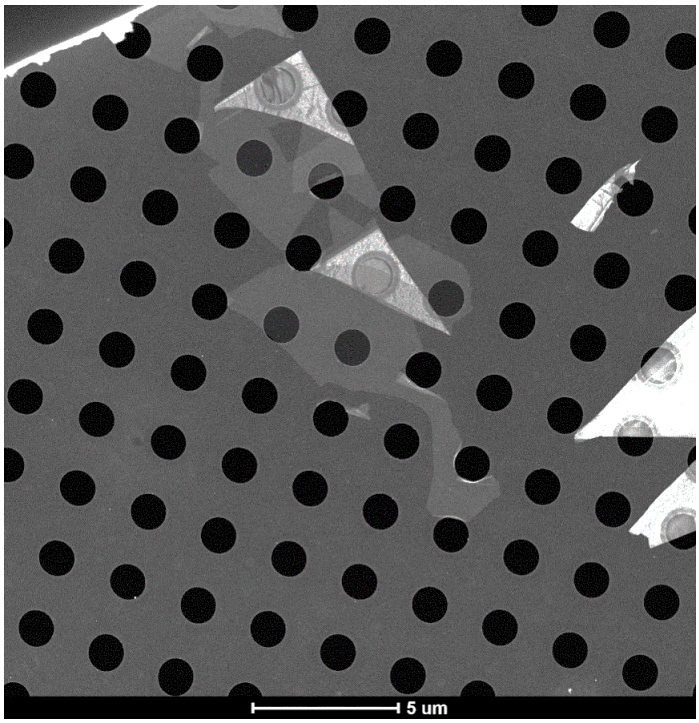

Figure 4.1: WSe_2 monolayer on Scanning Transmission Microscopy Grid

a) A monolayer of WSe_2 on a holey carbon substrate for use in scanning transmission electron microscopy

The STEM imaging of TMD monolayers was carried out in an FEI Talos F200X instrument operated at 200 kV. A low beam current (~60 pA) was used to reduce the amount of beam damage. By using the smallest condenser aperture (50 µm) and beam size 9, the convergence angle of the probe was calculated to be approximately 10 mrad. Images were acquired at a series of times to determine the rate at which metal vacancies formed in the samples when irradiated by the electron beam. For $MoSe_2$, the density of metal vacancies was found to increase linearly with time allowing extrapolation to time zero to determine the initial density. For WSe_2, no new metal vacancies were found to be created up to 40 seconds of imaging. Thus, for these samples, images from multiple regions were collected using a 20 second acquisition time. To improve the contrast and reduce the noise in the images for quantification of point defect density, the Butterworth filter in Gatan Digital Micrograph and the Wiener deconvolution in MATLAB were used. In addition, the simulated structure of the monolayer using the CrystalMaker ~~s~~oftware was overlaid on the processed STEM images to aid in identification of the metal and the chalcogen positions.

4.2 STEM Imaging of WSe_2

Shown in Figure 4.2a is a large-area image of one of these flakes of flux grown WSe_2 having several different thicknesses of TMDs. A part of the flake (false colored in yellow) of monolayer thickness is examined for defects. Shown in Figure 4.2b is a section of this flake with atomic resolution imaging. A single W vacancy is seen in this image. To prove that this defect was not caused by beam damage in these films, defect concentration was measured as function of image acquision time was analyzed. Shown in Figure 4.2c and Figure 4.2d are roughly the same small area of pristine monolayer seen in Figure 4.2b imaged for 20 seconds (Figure 4.2c) and for 40 seconds total (Figure 4.2d). Eight additional selenium vacancies are observed in Figure 4.2c (circled in yellow) when compared to Figure 4.2c, which has three vacancies (circled in black). From a systematic set of such imaging experiments, we have determined the intrinsic concentration of selenium defects is zero (as we see in figure 6b) as well as the knock-off rate. We find that the beam damage is limited to knock-off at the selenium sites (which appear a 180 degree rotation to those at W sites), and that much longer times are needed to remove the W atoms. Therefore our count of W vacancies will be unaffected by the STEM imaging conditions.

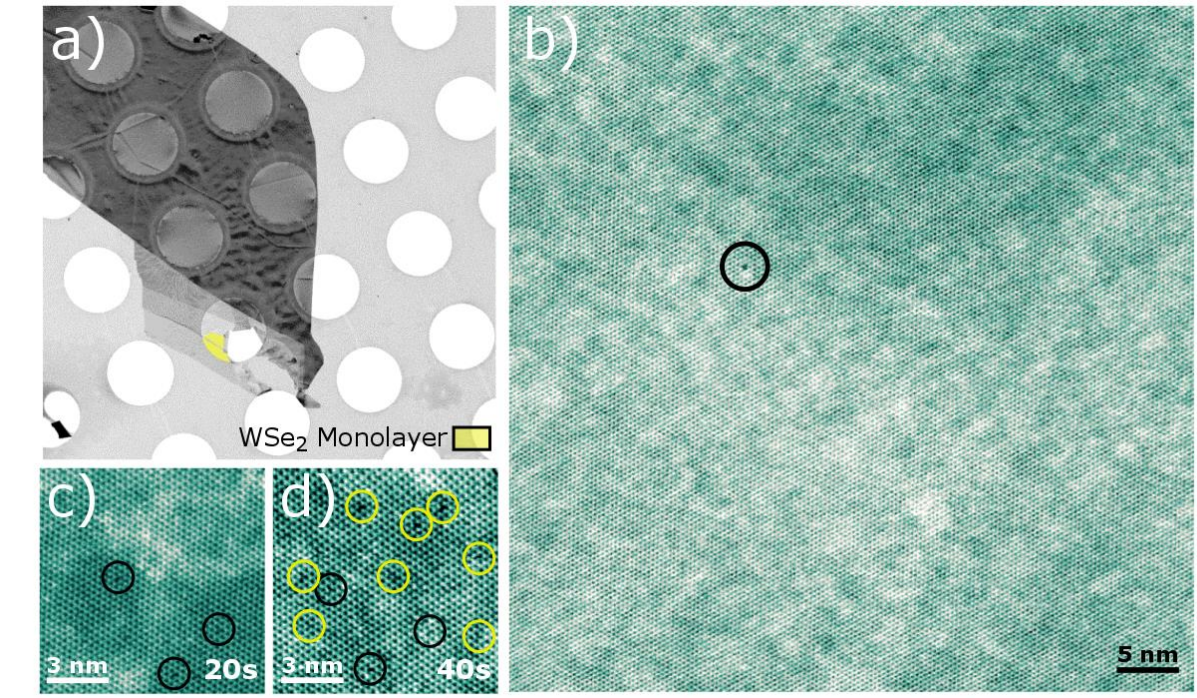

Figure 4.2: STEM of a WSe_2 Monolayer Showing Intrinsic Defects and Calibration of Knock-off Damage

a) Highlighted region of WSe_2 on holey carbon where the following images were acquired b) An area of pristine WSe_2 showing a single metal vacancy circled in black c) A region where continuous imaging was tested for knock-off damage d) the same region after 40s of exposure to the electron beam, created defects are circled in yellow

We can perform the same analysis for the selenium concentrations for multiple areas. In our STEM experiments, when imaging the monolayer using a typical beam voltage of 80kV, we do not see evidence for a significant number of Se vacancies - certainly not at levels matching the concentration of bright defects seen in STM. As additional evidence, we discuss measurements performed at a beam voltage of 200 kV. At this voltage, STEM induced knock off damage occurs, typically due to local increases in temperature, structural instabilities and deformations at higher beam energies. We find a significant number of Se vacancies can be created in this manner (up to 10^{13} /cm^2 at the largest exposure in this sequence of images). Shown in Figure 4.3 is a time series of images that show increasing numbers of Se vacancies that are created as a function of exposure time (Figure 4.3 b-e). For each of these images, we extract the vacancy concentration, and plot this in Figure 4.3a. Extrapolating to zero time indicates that the native Se vacancy concentration is too small to be statistically measurable by STEM imaging.

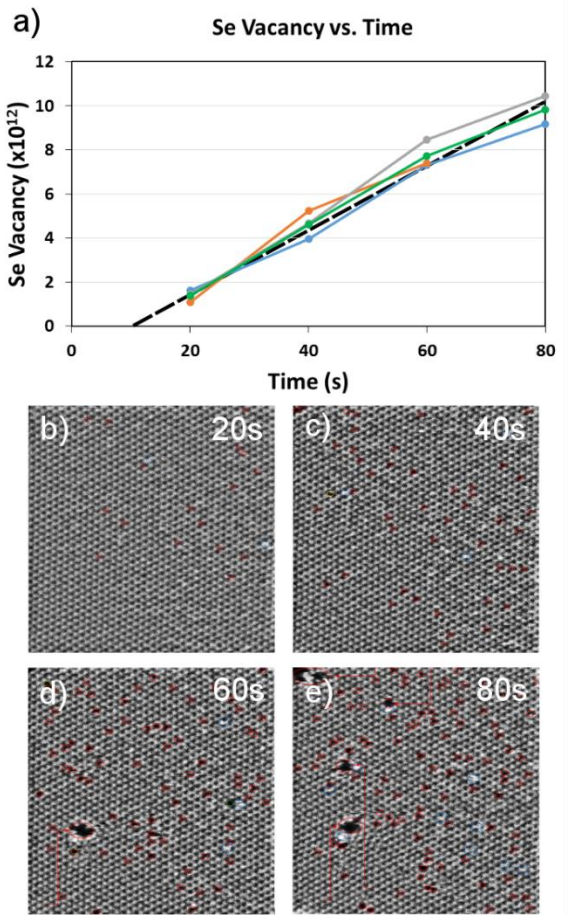

Figure 4.3: STEM Time Series of the Inclusion of Knock-off Defects

a) Experimental extrapolation of knock-off damage in the WSe_2 system, we see that within a short acquisition time (~20s) the monolayer remains pristine b-e) An image series of the same region measuring the knock off defect density as a function of time

Shown in Figure 4.4a is an STEM image that shows the only type of point defect observed in the flux-grown WSe_2 sample. The bright atoms in this image are W due to its high atomic number, indicating that the defect is a missing W atom, i.e., a metal vacancy. This chemical assignment of the –M defect is consistent with the STS observation that they act as electron acceptors.

In our STEM imaging, we do regularly see defects consistent with the antisite defects, which should show a higher contrast at the chalcogen site. One such STEM image is shown in Figure 4.4b. However, we need to be cautious in the interpretation of such images. Since our STEM images are obtained on exfoliated monolayers, we cannot exclude the possibility of adatoms on the surface during the processing required to go from a bulk crystal to a monolayer sample suitable for STEM. To make this

claim accurately would require element specific EELS mapping at the picometer scale together with several controls for the exfoliation process. This is a separate long-term project by itself. So while our STEM data is completely consistent with the presence of antisites at levels comparable to the bright defects seen in STM, we do not wish to over-interpret the STEM data shown in Figure 4.4b.

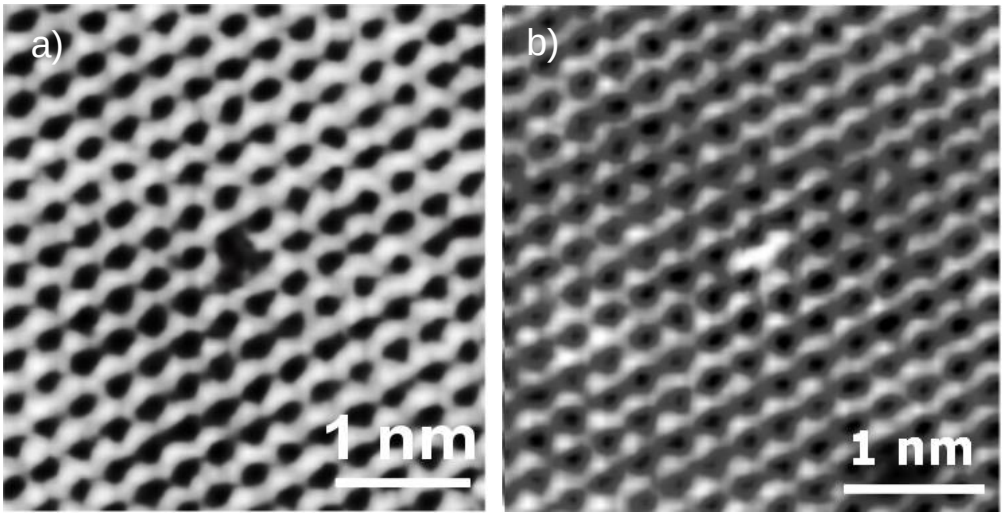

Figure 4.4: Two Defect Types Measured By Scanning Transmission Electron Microscopy

a) A metal vacancy in WSe_2 as imaged through STEM within the window where the lattice remains pristine b) A proposed metal antisite defect seen on the chalcogen site with STEM imaging

4.3 Theoretical Predictions of Lattice Defects

Our STEM, STM and STS measurements can be compared to theoretical expectations of defect formation energies and electronic structure from density functional theory (DFT). We wish to compare the theoretical formations for three defects, a metal vacancy, a metal antisite, and a selenium vacancy. We find that a metal vacancy requires an additional 5.22 eV per defect site for its formation. Using the Kröger-Vink (K-V) formalism to examine the chemistry of this defect gives a charge state of 4^- when referenced to the neutral crystal causing this defect type to act as an electron acceptor (i.e., $\square_M^{4-} + 4h^\bullet$, where $\square_M^{4-}$ denotes the vacancy on the metal site for a compound with chemical formula MX_2 - this defect has a negative charge relative to the filled metal site in the neutral crystal reference state typically used in the K-V formalism, and $h^\bullet$ denotes the requisite holes in the valence band to maintain overall charge

neutrality). DFT predictions for the local density of states match nicely with the observed STS, finding a shallow acceptor state (not pictured). The metal antisite defect has a formation energy of 4.81 eV. These defects exhibit a charge state of 6^+ when referenced to the neutral crystal making them electron donors, consistent with STM and STS (i.e., $M_X^{6+} + 6e'$ where M_X^{6+} denotes the metal antisite and e' represents the electrons required to retain charge neutrality). Apart from the two observed defect types in experiments, we also calculate the formation energy for chalcogen vacancies. We find this energy to be 1.81 eV, which is lower than that of a metal vacancy or antisite. The observed lack of these vacancies in spite of their lower formation energy suggests that kinetics plays a large factor in determining observed defect concentrations. Details of these calculations can be found in the methods and a table of calculated values can be seen in Table 4.1. Additional experiments that control for the kinetics of the reaction (for example by species availability) are required to relate the observed defect concentrations to the formation energies discussed here.

Table 4.1: Defect Formation Energies and Charge States

Defect Type	Kröger-Vink	Formation Energy	Charge State
Mo Vacancy	$\square_M^{4-}$	6.68 eV	4^-, Acceptor
Mo Antisite	M_X^{6+}	5.04 eV	6^+, Donor
Se Vacancy	$\square_X^{2+}$	1.72 eV	2^+, Donor

First-principle calculations for defect formation energies were done using density functional theory (DFT) within the projected augmented wave method [89,90], as implemented in the VASP code [91,92]. The generalized gradient approximation [93] is employed to treat exchange and correlation in DFT. Projected augmented wave method (PAW) was used in the description of the bonding environment for W, Mo, and Se. The structures are fully relaxed until all interatomic forces are smaller than 0.02 eV/Å. The Brillouin zone was sampled with a 5×5×1 k-mesh under the Monkhorst-Pack scheme [94]. Plane-wave energy cut offs of 400 eV and 500 eV are used for structural relaxation and static runs, respectively. The defect formation energies are defined via $E_{form}=E_{defect}-(E_{pristine}+\Sigma n_i\mu_i)$, where E_{defect} stands for the total energy of a

defected monolayer, $E_{pristine}$ is the total energy of a pristine monolayer, n_i is the number of removed (minus sign) or added (plus sign) species i and μ_i is the chemical potential of species i. The size of the supercell was determined by convergence tests, resulting in a 6×6×1 with a 15 Å vacuum space. The chemical potentials of each species are constrained by the relation [94], $\mu_{MoSe2} = \mu_{Mo} + 2\mu_{Se}$, where μ_{Mo} and μ_{Se} are the chemical potentials for Mo and Se, respectively; μ_{MoSe2} is the total energy per formula unit of $MoSe_2$. We determine the range of chemical potentials with two extreme cases: Mo-rich environment and Se-rich environment. For the Mo-rich environment, μ_{Mo} is chosen to be the total energy per atom of Mo in the bcc structure. For the Se-rich environment, μ_{Se} is chosen to be the total energy per atom of Se in the trigonal phase. The defect formation energies as a function of μ_{Mo} are presented in Figure 4.5. These formation energy values are extracted with the chemical potentials following $\mu_i = E_i + E_{bond}$, where E_i = Total energy of bulk metal or crystal of chalcogenide and $E_{bond} = (E_{MX2} - E_M - 2E_X)/3$ [95].

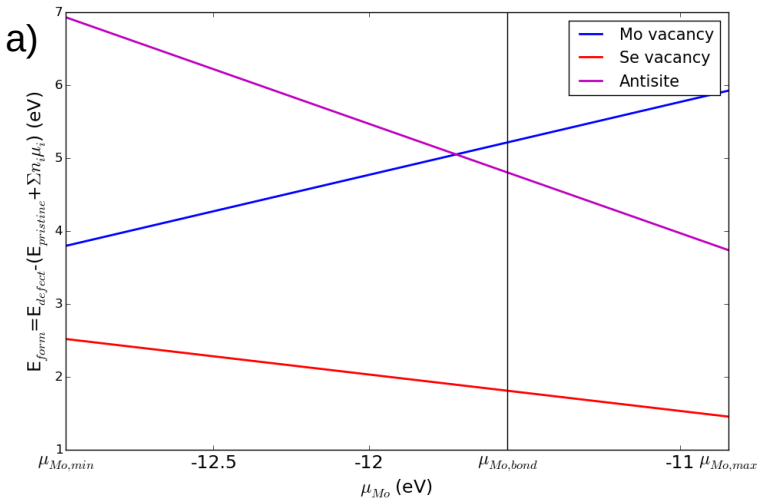

Figure 4.5: Formation Energy as a Funciton of Binding Energy

a) Formation energy of three defect types as a function of the energy required to disassemble the parent compounds in vacuum

4.4 Basic Semiconductor Theory Applied to TMD Defects

From the observed defect densities and binding energies, we can use semiconductor theory to calculate the chemical potential in various samples of $MoSe_2$ as a function of temperature[96]. This model

depends on the concentration of the two defect species as well as the temperature of the system. A four dimensional plot showing these relationships is plotted in Figure 4.6a. If we make this result specific to the defect species we measured for each crystal we arrive at Figure 4.6b for the t-CVT and flux-grown samples. For the CVT material, we expect a relatively constant chemical potential due to large but compensated numbers of donor and acceptor defects. The flux-grown crystal, in contrast, has smaller defect density but a dominance of donors, resulting in a strong chemical potential shift towards the conduction band edge as a function of temperature.

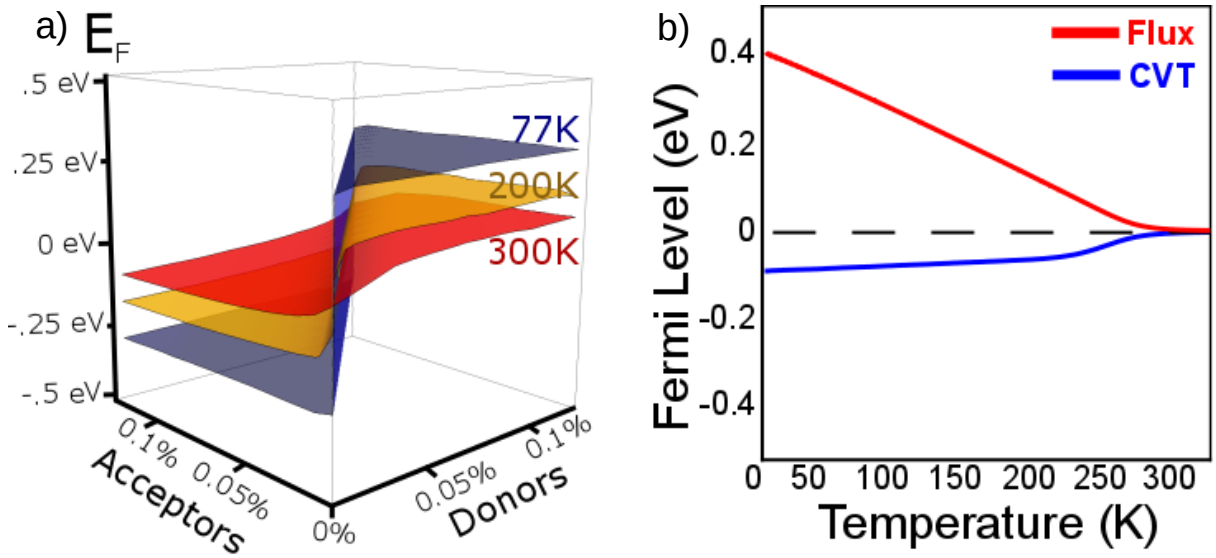

Figure 4.6: Theoretical Semiconducting Bandgap Size as a Function of Defect Concentrations

a) A sampling of three dimensional curves showing the resulting Fermi Energy as a function of temperature and acceptor and donor concentrations b) Predicted Fermi Energy of our two crystal growths as a function of temperature, using the acceptor and donor concentrations measured in experiment

Chapter 5

Scanning Tunneling Spectroscopy and Large Area Defect Electronics

5.1 Overview of Scanning Tunneling Spectroscopy Maps

These considerations indicate that we should measure chemical potential shifts in our STS spectra between different samples. In order to measure this with high spatial resolution, we use STS spectroscopy to measure the local semiconducting gap at every pixel of a 256 x 256 px grid overlaid on a 0.5 x 0.5um^2 area. At each point, we extract a local value of the conduction and valence band edges from the local spectrum. A sum of these two values gives the local value of the semiconducting bandgap.

To obtain bandgap values systematically, gap sizes were found using an automated extraction method for each pixel of our spatially resolved spectroscopy maps. Since the measured dI/dV signal is a convolution of the density of states with the Fermi-Dirac distribution of thermally activated carriers, we use a linear extrapolation to extract the conduction and valence band edges as shown in Figure 5.1a. This figure illustrates the procedure at a location far away from defects. A similar process is employed when spectra are taken over a defect as shown in Figure 5.1b. The band edge is still extracted by linear extrapolation from higher energies as in the defect free case. To extract the energy of the defect state an additional resonance (assumed to be Gaussian in shape) is added near the band edge.

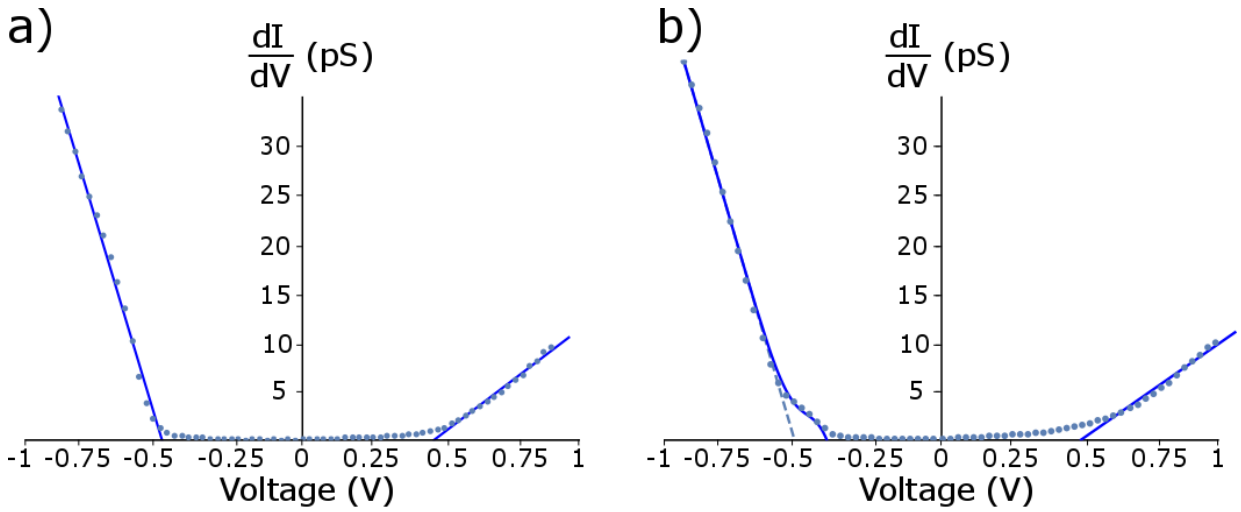

Figure 5.1: Bandgap Extraction from dI/dV Curves

a) Bandgap extrapolation of a typical $MoSe_2$ dI/dV spectra as taken in a pristine region b) The same bandgap extraction using a summation of a numerically fit Gaussian to remove the defect state from the spectra

5.2 Growth Comparison of Defect Electronics

Shown in Figure 5.2a is a color scale image of the measured bandgap variation in a t-CVT crystal. Color variations in this picture represent gap variations in the vicinity of defects. The average bandgap in this crystal is 860 meV, with defect-induced gap variations of order 50 meV. The bulk bandgap is in reasonable agreement with the theoretical gap of 840 meV [97]. The variation seen for the t-CVT crystal is to be contrasted with a similar gap variation image for the flux crystal shown in Figure 5.2b. While the flux crystal also exhibits an average bandgap of 860meV, it displays much smaller variations in space due to the lower concentration of defects.

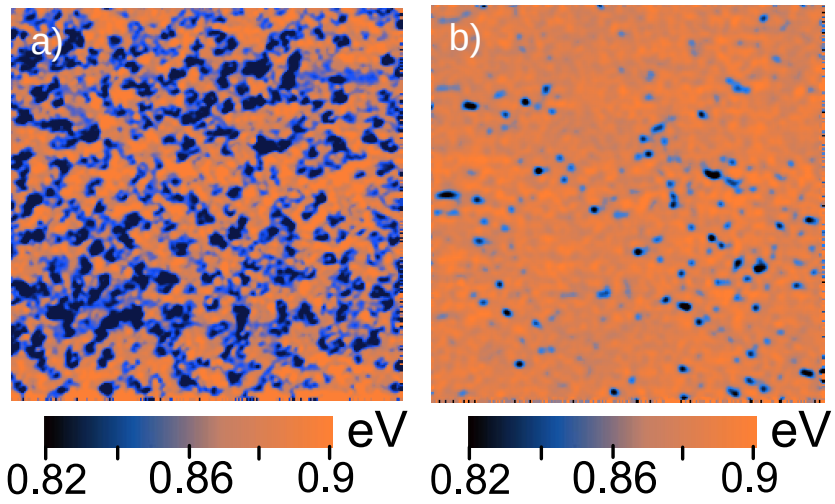

Figure 5.2: Bandgap Maps over a 0.4 x 0.4 μm² Area

a) Bandgap map taken on a CVT crystal showing a large number of defect states infringing on the gap edge b) Bandgap map taken on a self-flux sample. These crystals show much more bandgap continuity

To visualize these differences, histograms detailing the spread of gap sizes for t-CVT and flux are plotted in Figure 5.3a and Figure 5.3b respectively. The observed gap variation in the t-CVT (3σ) is 50 meV, while the flux crystal shows a gap variation of 20 meV. As part of our measurement, STS was able to extract the valence and conduction band edges separately. Therefore, we further the analysis by examining the impact defects have on the valence and conduction band edges. In Figure 5.3c**,**d we plot the valence band onset distributions for t-CVT (centered at –370meV) and flux (centered at –760meV) respectively. From these plots we see that most of the gap variation arises from defect states on the valence edge, which we found earlier correspond to metal vacancies. A similar analysis of the conduction band edge is plotted for t-CVT in Figure 5.3e (centered at 490meV) and flux in Figure 5.3f (centered at 100meV). Here we see almost no variation, especially in the case of the flux crystal.

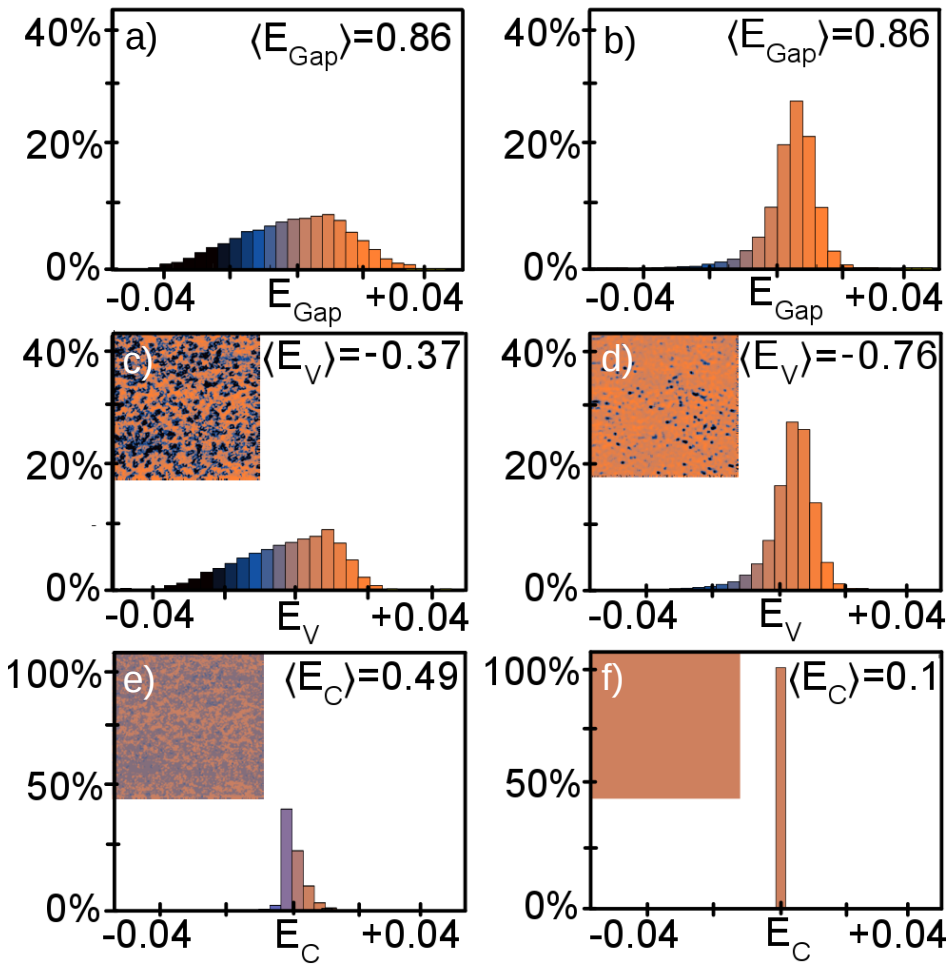

Figure 5.3: Gap and Band Edge Histograms

a-b) Bandgap distribution for the CVT and self-flux crystals respectively c-d) valence band distibutions of the maps pictured in inset for CVT and flux, we see that most of the variation of the bandgap occurs on this edge e-f) The same treatment for the conduction band, we see that the distribution on this edge is much sharper

Our detailed gap maps can also be used to estimate the position of the chemical potential relative to the gap midpoint from $\mu = -\frac{<E_C>+<E_V>}{2}$. For the t-CVT sample, this gives a chemical potential 60 meV below the gap center, indicating slight p-type doping. The self-flux crystals have a chemical potential of 330 meV above the gap center, making them n-type. At 77K, we expect from Figure 4.6b that the chemical potential should be 50 meV below the gap center for t-CVT and 320meV above the gap center for self-flux. This agreement with STS mapping over a 500x500nm^2 region indicates that we have properly accounted for all of the dopants in the semiconductor. Additionally, the observed behavior explains the commonly observed p-type ambipolar FET devices[98] that have been made from t-CVT crystals. To

visualize this further average dI/dV spectra for each bandgap map are presented on the same plot in Figure 5.4.

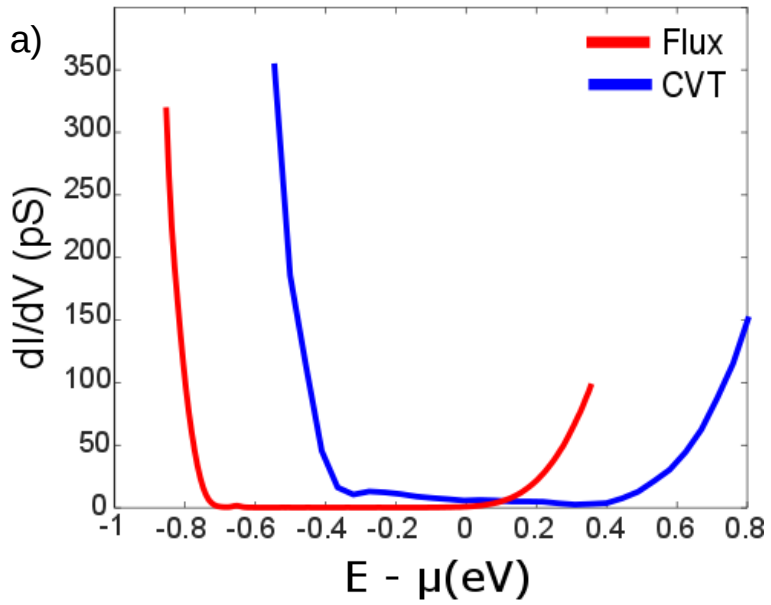

Figure 5.4: Average dI/dV curves for CVT and Self-Flux

a) The average spectra for both the CVT and flux crystals as averaged across the full gap map. We see that the Fermi energy is very different at 77K as we expected from semiconducting band theory. In this case the CVT has a compensated number of defects while the self-flux is pinned to the conduction edge due to a higher concentration of donor states

Flux growth achieves 1-2 orders of magnitude lower defect density than CVT does, and defect densities lower than $10^{11}/cm^2$ from our flux-grown sample are by far the lowest reported for TMD materials. This improvement is reflected in reduction of band-edge disorder measured by scanning tunneling spectroscopy, therefore we expect orders of magnitude increase in photoluminescence quantum yield.

Chapter 6

Optical Analysis of TMD Crystal Growth Utilizing Photoluminescence Measurements

6.1 Overview of Photoluminescence Setup

To connect the large disparity of crystal imperfections versus growth observed across different crystals in STM to the monolayer limit, we carry out photoluminescence (PL) measurements as a simple way to measure the radiative response versus its quality. We isolate single layers through mechanical exfoliation from bulk single crystals[39,82,83]. Monolayers were simultaneously exfoliated from ag-CVT, t-CVT and flux grown crystals and each sample was subsequently handled under identically conditions to eliminate extrinsic factors.

Optical stacks of BN/TMD/BN were fabricated using the polypropylene carbonate (PPC) method as described in [99] and placed on passivated SiO_2[76]. $MoSe_2$ samples were measured using a closed-cycle He cryostat (Attocube Attodry 1100) and an excitation wavelength of 532 nm using a cw diode laser with an approximate power of 2.0 μW. For WSe_2, samples were loaded into a cryostat with a sapphire window which is combined with a homemade photoluminescence setup with an excitation wavelength of 532 nm using a cw diode laser and a power of 80 μW. For cooling, either helium-4 or liquid nitrogen were continuously flowed through the cryostat chamber, immersing the sample while temperature was modulated with a stage heater.

After ensuring monolayer was encapsulated in BN and placed on a passivated SiO_2 surface[76]. The resulting stacks were measured under the same conditions (laser excitation power, spot size, and acquisition time). Additional data was taken to ensure that these curves were taken within the linear regime of the laser, as is seen in Figure 6.1.

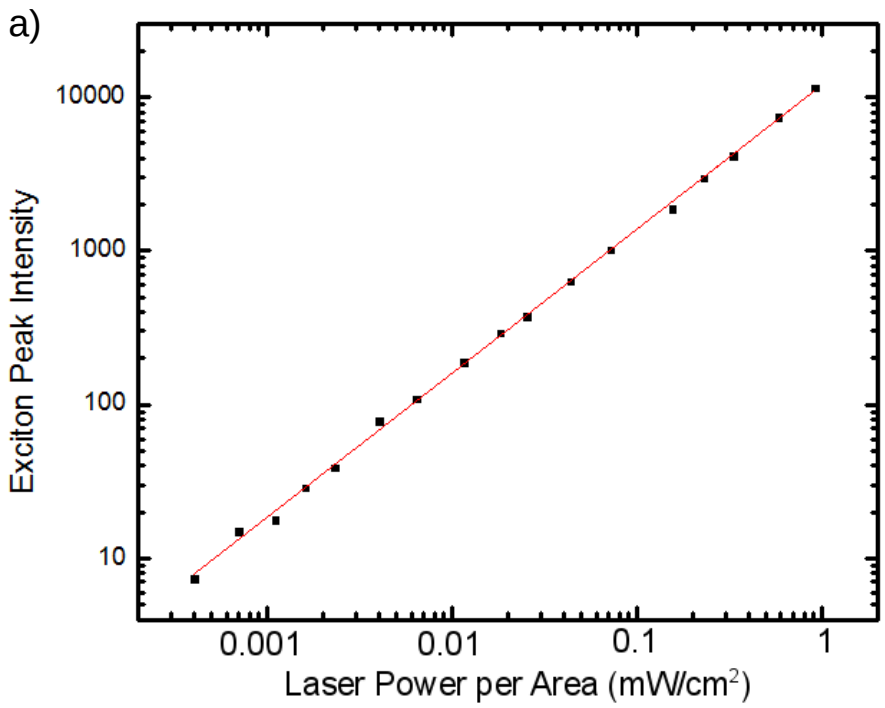

Figure 6.1: Laser Power as a Function of Laser Spot Size

a) The power regime of the laser used in the photoluminescence experiments to follow

6.2 Photoluminescence Comparison

The raw PL data for $MoSe_2$ at 4K are plotted on a log scale in Figure 6.2a. While the peak position shifts very slightly as crystalline quality is enhanced, the linewidth (FWHM) shows an obvious decrease with improved crystal quality, from 4 meV in ag-CVT to 3 meV in t-CVT to 2 meV in self-flux sample. The decrease in FWHM is consistent with the improvement in homogeneity of monolayer sample as defect density decreases. A dramatic effect is seen in the total light emission intensity (proportional to the quantum yield), with the self-flux sample having a 10-fold increase in light intensity over the t-CVT monolayer, and a 100-fold increase over the ag-CVT monolayer. We note that improvements in PL yield can also be achieved by various surface treatments[100,101] that can affect the exciton lifetime or light absorption. In our PL experiments, we take care to keep the environment of the films as identical as possible across samples via encapsulation in BN. The large suppression of excitons with little change to peak position suggests that defects provide non-radiative pathways for the recombination of excitons, via

exciton localization[102] or defect enhanced Auger processes[62,64,65], lowering the PL yield from the intrinsic limit of a pristine sample.

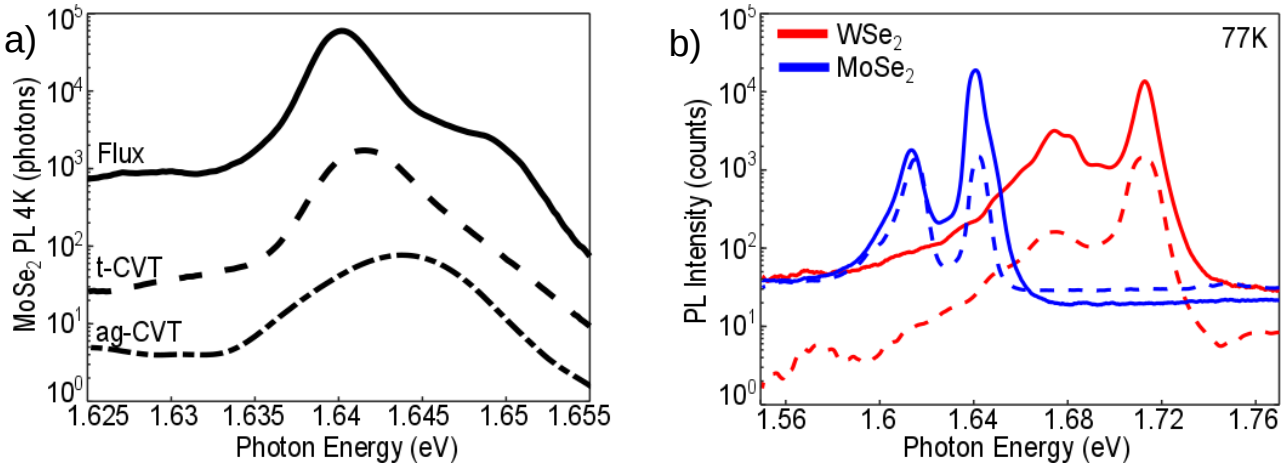

Figure 6.2: Raw Photoluminescence Comparison of Growth Methods

a) Photoluminescence curves taken at 4K showing the large power disparity between the three growth methods, each is roughly separated by a factor of x10 b) Spectra taken at 77K comparing CVT (dashed) and self-flux (blue) for both $MoSe_2$ and WSe_2

To quantify the impact of defects on the excitonic properties, we extend the PL measurements described above to various temperatures on t-CVT and flux monolayers (ag-CVT is omitted due to a lack of intensity at high temperatures). We plot the data taken for both $MoSe_2$ (red) and WSe_2 (blue) on a log scale at 77K in Figure 6.2b where t-CVT is dashed and flux is solid.

6.3 Integrated Photoluminescence

Since the major observed difference from the crystals is the overall PL intensity, we plot the integrated PL intensity as seen for $MoSe_2$ in Figure 6.3. To model the shape of the integrated PL signal we must account for the unique band structure of monolayer TMDs. In the TMD materials, both the valence and conduction bands are spin split due to spin-orbit coupling. The magnitude of the splitting in the valence band is roughly an order of magnitude larger than the conduction-band splitting (~300 meV versus ~30 meV). Due to the spin splitting in the conduction band, one of the two transitions from the conduction band to the upper valence band (A exciton) is dark, while the other is optically bright. In $MoSe_2$ the lower of the two transitions is bright, while the situation is reversed in WSe_2.

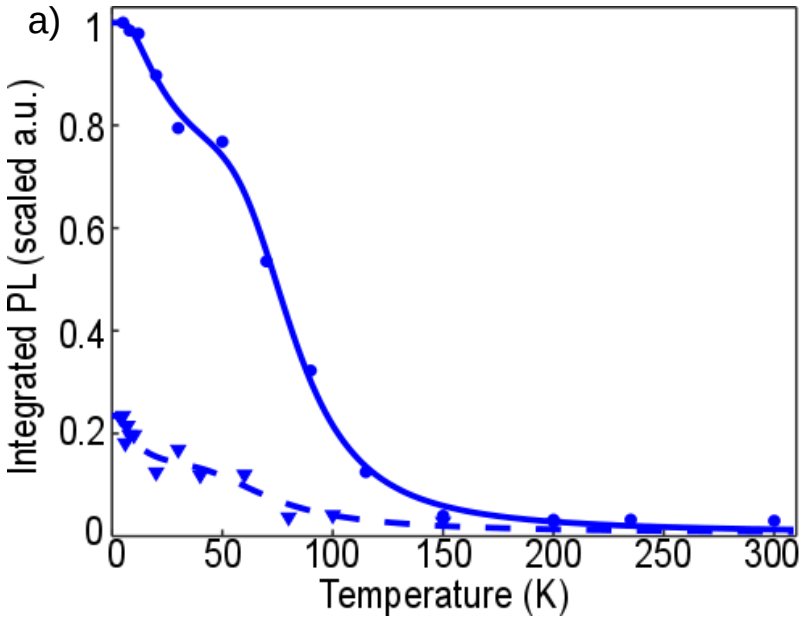

Figure 6.3: Integrated Photoluminescence for $MoSe_2$

a) A comparison of the integrated photoluminescence signal for CVT (dashed) and self-flux (solid). We see that there is a kink in the self-flux signal which corresponds to the energy at which the dark exciton is screened out by temperature. This does not occur in the CVT crystal because defects help the excitons decay non-radiatively

To model the integrated photoluminescence a modified Arrhenius like formula was used. This formula accounts for the presence of dark excitons as well as other non-radiative factors. To accurately represent this PL signature we first look at the band structure of a single layer. Shown in Figure 6.4a is a simplified band structure of monolayer MoX_2 compounds. Due to orbital hybridization of the $d_{x^2-y^2}$ and d_{xy} orbitals a spin splitting of (100 – 500 meV) occurs in the valence band. A smaller splitting of (10-100 meV) exists in the conduction band due to strong spin orbit coupling (SOC) of the transition metal orbitals. In Figure 6.4b a three dimensional rendering of the A exciton bands can be seen. The two lowest energy excitons therefore corresponds to a hole in the top of the valence band and an electron in one of the two spin split conduction bands. The lowest energy exciton is bright while the other is dark in MoX_2 due to spin selection rules (Figure 6.4c). The SOC splitting of the conduction band is opposite in the case of WX_2 causing the lowest energy exciton to be dark.

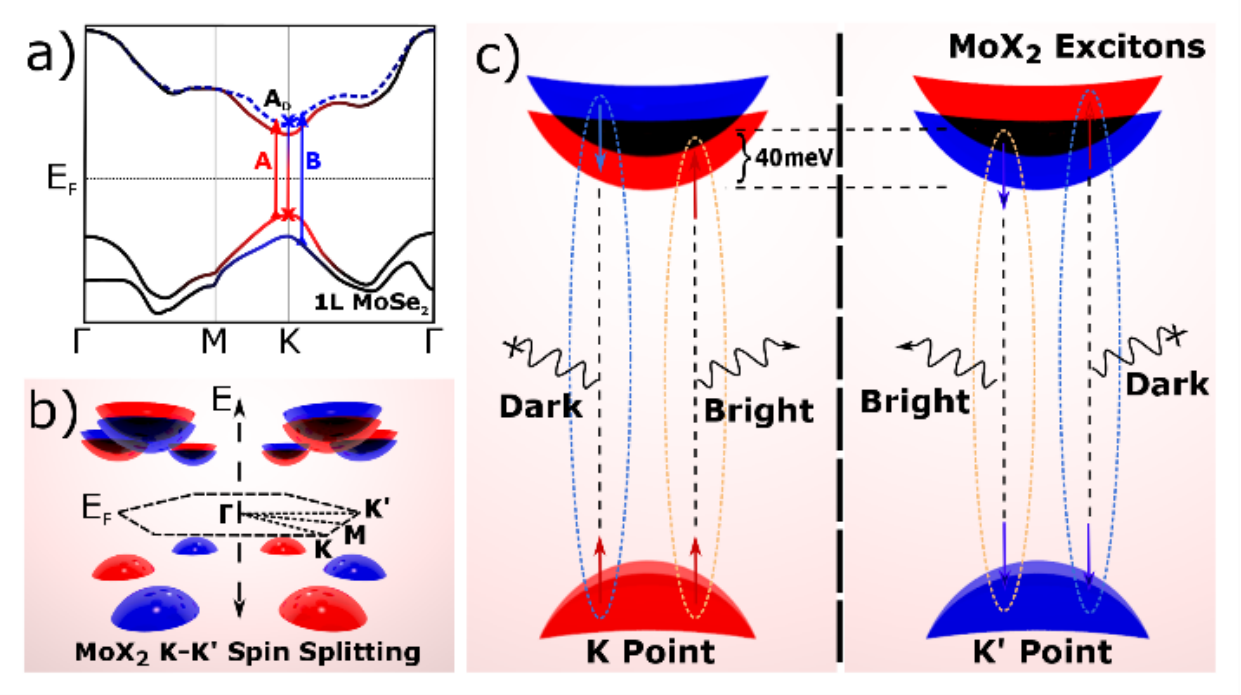

Figure 6.4: Model of Exciton Recombination

a) A reproduction of the TMD band structure for $MoSe_2$ showing the difference between the dark and bright A exciton b) Extending this model to three dimensions we can see the conduction band splitting at both the K and K' points c) This spin splitting informs the allowed pairings of the excitonic pairs. We see that like spins are the only possible channels for recombination since the electron must be able to fill the place of the hole in the valence band.

For every photon that is absorbed there are three recombination channels[103] - radiative recombination via the bright A exciton, non-radiative recombination from the dark A exciton and defect/phonon assisted recombination. We can write the rates of each of these three processes (assuming them to be independent) in terms of their energies and rate constants [104]:

$$W_{PL} = C_{Bright} * e^{-\frac{E_{Bright}}{k_B T}},\ W_{Dark} = C_{Dark} * e^{-\frac{E_{Dark}}{k_B T}},\ W_{NR} = C_{NR} * e^{-\frac{E_{NR}}{k_B T}}$$

We can put these together to extract the temperature dependent integrated PL from the ratio of incoming to outgoing photons

$$MoX_2: \frac{I_{PL}}{I_{Tot}} = \frac{C_{Bright} * e^{-\frac{E_{Bright}}{k_B T}}}{C_{Dark} * e^{-\frac{E_{Dark}}{k_B T}} + C_{NR} * e^{-\frac{E_{NR}}{k_B T}} + C_{Bright} * e^{-\frac{E_{Bright}}{k_B T}}} = \frac{1}{C_1 * e^{-\frac{\Delta E_{SOC}}{k_B T}} + C_2 * e^{-\frac{\Delta E_{NR}}{k_B T}} + 1}$$

A similar analysis can be performed for WSe_2 where the only difference is the sign on the SOC term.

$$WX_2: \frac{I_{PL}}{I_{Tot}} = \frac{1}{C_1 * e^{\frac{\Delta E_{SOC}}{k_B T}} + C_2 * e^{-\frac{\Delta E_{NR}}{k_B T}} + 1}$$

We fit the integrated PL intensity to an empirical Arrhenius equation as described above, which accounts for both the spin split exciton[104], and non-radiative processes which average multiple recombination rates. We find that the integrated PL emission can be described as follows for $MoSe_2$:

$$I_{Tot} = \frac{1}{1 + C_1 * e^{-\frac{\Delta E_{NR}}{k_B T}} + C_2 * e^{-\frac{\Delta E_{Dark}}{k_B T}}}$$

Here the non-radiative term arises both from defects and phonon scattering at higher temperatures, and other term arises from thermal equilibrium between the dark and bright excitons. Looking back at Figure 6.3 a sudden drop can be seen in the PL signal of the self-flux $MoSe_2$ monolayer above roughly 60K. This can be attributed to the Boltzmann distribution of electrons able to access the dark exciton state. As per our fit prescribed above we extract the dark exciton energy to be 40 meV above the bright exciton binding energy. This matches well with reports utilizing a backgate to directly measure the dark exciton at 30 meV above the bright one[105].

Like $MoSe_2$, the PL resonances of WSe_2 have roughly the same peak position across growth methods. The temperature dependent integrated intensity shows distinct behaviors for the flux versus CVT crystals as shown in Figure 6.5. The self-flux crystal shows an initial increase in the PL intensity with decreasing temperature down to about 150 K, below which the intensity drops sharply. The CVT crystal however shows a decreasing intensity with decreasing temperature starting from room temperature. This difference in the case of WSe_2 arises due to the fact that the lowest energy transition is dark[106]. This implies that at sufficiently low temperature only the dark state is populated, exponentially suppressing the PL intensity. At high enough temperature, on the other hand, the PL intensity is suppressed with increasing temperature due to phonon scattering. We can model both effects together using the same formula as $MoSe_2$ with a sign change:

$$I_{Tot} = \frac{1}{1 + C_1 * e^{-\frac{\Delta E_{NR}}{k_B T}} + C_2 * e^{\frac{\Delta E_{Dark}}{k_B T}}}$$

This change of sign indicates that the dark exciton is now at a lower energy than the bright. Thus, PL shows a maximum at some intermediate temperature which is determined by the interplay between defect and phonon recombination versus dark exciton coupling. Using the fitting for the integrated intensity from the flux crystal, we extract a conduction-band splitting of 43 meV. Our measurement of this splitting is within experimental error of the value directly measured from magnetic field measurements (47 meV)[107,108]. We additionally find that CVT crystals exhibit stronger defect mediated coupling to the dark exciton state diminishing the PL signal out to much higher temperatures.

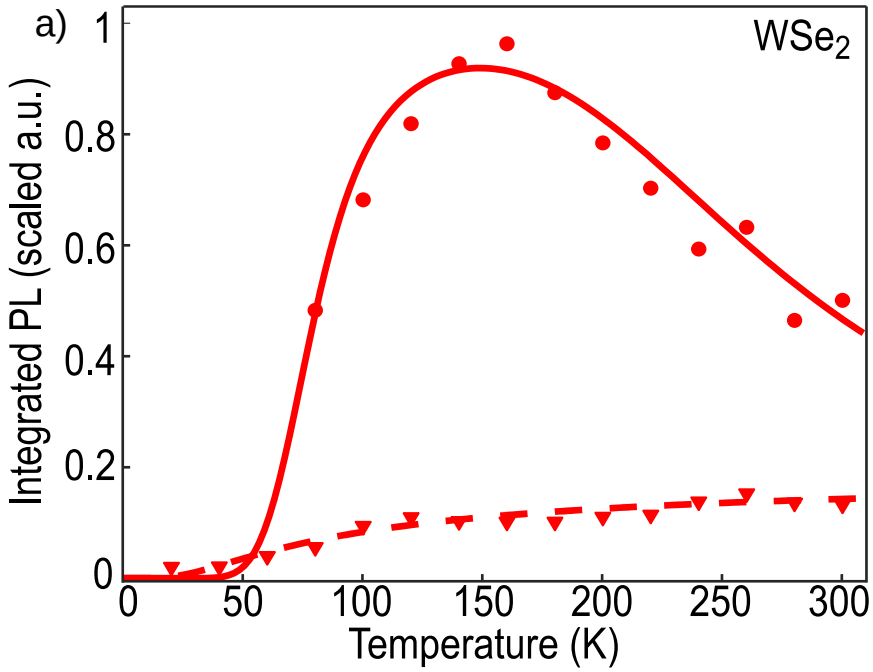

Figure 6.5: Integrated Photoluminescence for WSe_2

a) Comparison of the integrated photoluminescence signal for CVT (dashed) and self-flux (solid). The dark exciton coupling strongly suppresses the photoluminescence signal at low temperatures. However since the self-flux crystal is so clean we are able to extract a dark exciton energy of ~50meV using the Arrhenius equation

6.4 Conclusions of Defect Studies

Our studies of defects in transition-metal dichalcogenides show definitively that there is a direct link between intrinsic point defect concentration in bulk crystals and the optoelectronic properties of

exfoliated monolayers. The improvements in synthesis presented here have led to a lowering of the defect concentration by two orders of magnitude when compared to the current state of the art. Such improvements are a necessary step towards achieving many of the predicted optical phenomena that require high exciton concentration as well as transport phenomena that require long scattering times. We note in conclusion that while our synthetic achievements set a new benchmark for TMD semiconductors, the lowest bulk defect concentrations achieved here (~10^{18} /cm^3) are still significantly higher than those achieved in the best III-V semiconductor films, indicating that there is still room for refining synthetic processes to achieve higher quality TMD materials.

Chapter 7

Strain Measurements of TMD Semiconductors

7.1 Strain Experiments in Two Dimensional Systems

As we demonstrated in the previous chapters the transition metal dichalcogenides share many of the control parameters that are common to traditional semiconductors such as silicon and the III-V compounds. These include the ability to chemically tune the bandgap, excitonic response, and single crystal growth. However there is one principle difference between the transition metal dichalcogenides (TMDs) and conventional semiconductors, a new parameter, two dimensionality[109].

Layered materials research, has seen the creation of two dimensional analogs for traditional silicon based electronics, however these materials have garnered much more interest due an intersection of nano-scalability, flexibility[110] and desirable electronic properties[28,43,46,111]. Development has focused on a myriad of Van Der Waals materials including conductors and insulators such as graphene and boron nitride, and semiconductors such as the transition metal dichalcogenides, and black phosphorous. As a consequence of their two dimensional nature, these compounds can easily be integrated into electronic circuits with unparalleled thinness, which unlike traditional circuitry can further be made to conform to non-Euclidean geometries[112-114]. To understand how these materials react within such an environment, the resulting pursuit has been to understand performance under applied strain[115-122].

As with many layered materials the first studies of strain due to geometric constraints were performed on graphene. While most methods of induced strain were initially due to variations in mechanical scratching techniques using scotch tape[123], controlling strain in layered structures is a promising new parameter available to two dimensional materials. Strain in the two dimensional limit has unique access to a regime unavailable to conventional bulk systems due to the extreme in-plane strength of monolayer bonds. Monolayers therefore can experience much higher strain percentages without

deformation or breaking. Initial experiments regarding strain, found however that for any system of more than one layer, the material was especially susceptible to strain relief due to weak coupling in between the Van Der Waals layers. This is especially true for graphene, in this system strain applied to fixed layers is relieved through the formation of atomic slips along principle directions. These can occur in any free layer not in contact with the strain substrate. The resulting phenomenon is akin to edge dislocations in bulk materials, and has been coined the strain soliton. These solitons form naturally through energy relaxation, and play a major role in stretched or twisted systems of 2D materials.

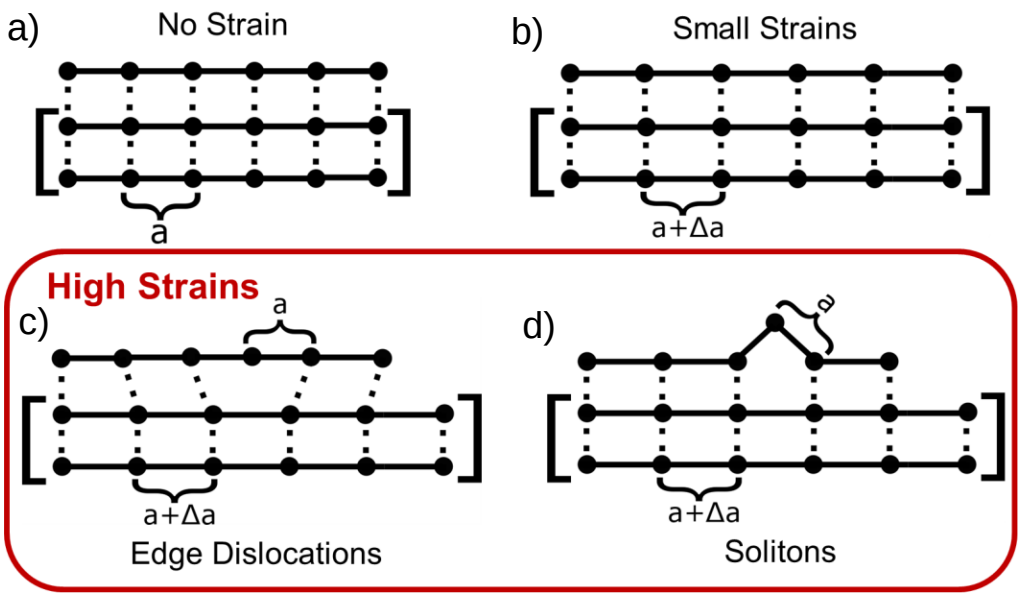

Figure 7.1: Strain Regimes and Interlayer Relief Mechanisms

Cartoon model of strain regimes in the presence of a) no strain b) low strains c) edge dislocations in the high strain regime d) Strain solitons occurring as an instance of an edge dislocation whose formation is exclusive to two dimensional systems

To demonstrate the process by which strain solitons are formed we need to focus on two regimes, the low strain and high strain limits. Shown in Figure 7.1a is a simple representation of a crystal without any strain. The dashed bonds in this figure align perfectly with the layer above and represent the Van Der Waals interlayer binding of atoms. We see that the lattice constant additionally has an unstrained length of $\boldsymbol{a}$. If we apply strain to the bound layers represented by the brackets we arrive at Figure 7.1b. Since only a small strain has been applied all layers remain commensurate with a new strained length of $\boldsymbol{a + \Delta a}$. This crystal still retains its Van Der Waals binding between layers and the strain is uniformly

distributed throughout the crystal. The typical descriptor for this is a strain percentage $\boldsymbol{\epsilon}$, which is an equivalent representation of our strain field $\boldsymbol{a} + \boldsymbol{\epsilon} * \boldsymbol{a}$. If we continue to apply strain we enter the high strain regime as denoted by Figure 7.1c,d**.** In this case layers become too strained to remain in contact, and energetically prefer to slip by a lattice constant. This relaxation of the in-plane elastic constraints lowers the strain, creating either an edge dislocation[124,125] (Figure 7.1c), or a strain soliton (Figure 7.1d). The main difference between the two high strain systems in the 2D limit is that a soliton has no constraints from above, and therefore can further restore its Van Der Waals bonds by lifting from the surface. Therefore we can see this as a special type of edge dislocation only available to 2D crystals. Again, both of these line defects relax the local region around the dislocation to its original length $\boldsymbol{a}$, thereby reducing lattice strain in the defective layers. We also can envision additional strain geometries that form from a full lattice mismatches with the strain substrate, as in the case of BN and graphene[126-129]. In this special case the lattice must relieve through triaxial strain, so instead energetics cause formations such as moiré mismatch[130,131] and pyramidal nanobubbles[126]. In graphene these have been shown to exhibit strong pseudomagnetic fields[126,132] and confining potentials[133] due to orbital overlap in the regions of deformation.

Strain in transition metal dichalcogenides (TMDs) has been harder to access than that of graphene, but has much wider application due to the semiconducting nature of these compounds with optical bandgap[129,134]. This is mainly due to much stronger coupling between sheets of the MX_2 hexagonal lattice, where constituents can be any of the transition metals, {M=Mo,W,Re} and chalcogens, {X = S,Se,Te}. While many other combinations are possible we choose to focus on the semiconducting class of these compounds as studied in previous chapters. We note that further research opportunities lie in the application of strain to other forms of TMDs including layered superconductors such as $NbSe_2$, and Weyl metals like WTe_2. This work however will address high strain in TMDs through controllable strain soliton formation occurring within sheets of $MoSe_2$, a typical TMD semiconductor.

The impact of high strain within the TMD lattice has been a long withstanding goal due to predicted change of phase from semiconducting to metallic[121]. Pursuits in this direction have measured the shift of the bandgap through optical spectra for strains as high as 6%[115-117,135]. For bilayer and greater, however strains of this magnitude unfortunately have not been achievable due to delamination of layers from one another[136]. This makes strain soliton regime of interest, because it is an intermediary region that exhibits partial detachment. To begin to breach this topic we first turn our attention to past works, specifically graphene, where a much weaker Van Der Waals force makes this intermediate region much more accessible.

Studies on few layered graphene provide the mechanisms behind layer delamination under application of strain. Although many groups found strain solitons occurring naturally, the first atomic identification of these features were measured through transmission electron microscopy (TEM). Since this technique can see atoms from every layer on thin substrates these microscopy measurements detected that each soliton contained an extra atomic row (versus the bulk lattice) across its length. This extra row allows for a relaxation of the top lattice to its unstrained length an experimental confirmation of Figure 7.1c,d. The induced relaxation occurs along the atomic direction that is parallel to the short axis of the fold. As a fold or ripple forms in the sheet, its geometry forces a local detachment of the layer through breaking of the Van Der Waals force. Further inspection of soliton theoretically, found that these formations exist due to a competition of interlayer binding force and input strain energy. These calculations were based on the shear lag model[137,138], which both predicts the critical value of strain required for soliton formation, and lattice dynamics which cause an increasing soliton density for further applied strain. Predictions for graphene place the critical soliton formation value at 0.6%, which is easily achievable, as was seen experimentally[139]. For the TMDs however their much higher binding force causes this value to hover near 2%.

7.2 Achieving High Strains in Experiment

In practice achieving high strain has been an experimental challenge. Multiple apparatuses have been produced to control the internal strain applied to two dimensional materials, most using a combination of polymers, piezoelectrics and micromanipulators. Since strain solitons are a local phenomenon which should exhibit electronic features, scanning tunneling microscopy (STM) and spectroscopy (STS) was chosen as the best local probe. To strain in-situ within an STM system, many groups have used piezoelectrics because they provide electronic manipulation in ultra-high vacuum. This method typically requires gluing a crystal to a piezoelectric medium, then using the intrinsic expansion and contraction of the piezo as the strain field for the whole crystal. The drawback to this method is in strain percentage, since each piezo has a limited range, the crystal can only experience the maximum strain of the piezo. Strain percentage therefore is solely dependent on piezo dimensions and material. Given the constraints of our STM environment the best choice of piezo only results in a maximal strain of 0.2%. Since the soliton regime requires strain magnitude a factor of x10 larger, we circumvent these limitations by instead deploying a similar scheme using shear piezos. The apparatus used is pictured in Figure 7.2 is the resulting development for much higher strains.

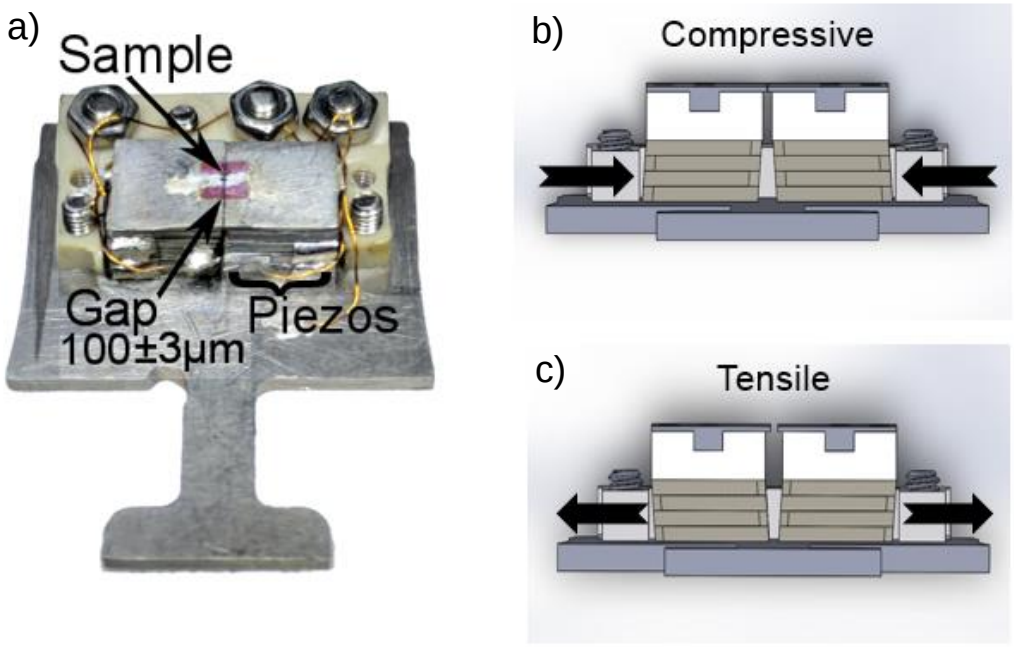

Figure 7.2: Strain Device and Principles of Operation

a) Experimental strain device showing sample mount, strain gap and shear piezo stacks b) The device in its b) compressive and c) tensile modes of operation

Again using the definition of strain, we bind our crystals over a length of L. Therefore the applied strain is given as the change in length ΔL, and so for our strained lattice $L \pm \Delta L$. If we control the length over which strain is applied we have an easy way to tune the strain up to arbitrarily high percentages. To do this the device pictured in Figure 7.2a is formed from two towers of shear piezos which tilt in opposite directions upon the application of a strain voltage as is shown in Figure 7.2b,c. This geometry is beneficial since it allows for two types of strain application compressive (Figure 7.2b) and tensile (Figure 7.2c). Sitting atop these piezo towers are metallic plates for the application of STM bias voltage. To avoid shear, plates have been sanded to level to within a few micron. To demonstrate the constraints applied to the crystal, Figure 7.3a shows the experimental setup this is the way in by which the crystal was attached. Sapphire plates were affixed to ensure a uniform gap and the crystal ends were encased in epoxy to apply a uniform strain within the exposed region. Before and after gluing a crystal to these piezoelectrics, they were tested for the control of the gap separating them as shown in Figure 7.3b. For the $100\ \mu m$

separation they were found to move by $\pm 3\mu m$ at liquid nitrogen temperatures, resulting in both 3% compressive and tensile strain for anything bound across the gap.

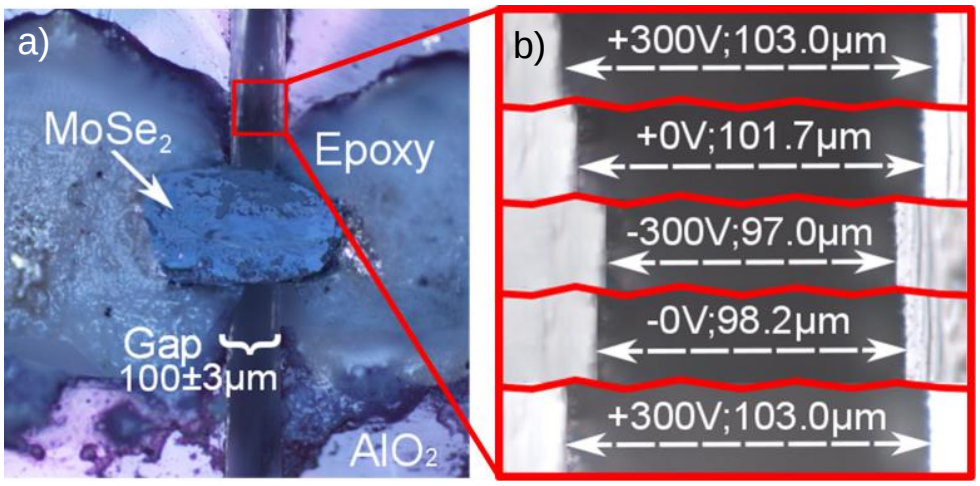

Figure 7.3: Experimental Verification of Gap Motion

a) $MoSe_2$ crystal used in the strain experiment mounted across the strain gap b) Proof of shear piezo motion with bound crystal calibrated to LN_2 temperatures

With this strain setup, achieving strain solitons is theoretically possible within the strain range. To visualize these as part of the experimental setup a cartoon model of their formation is shown in Figure 7.4. In this figure we see that the top (free) layer is not in contact with the constraining epoxy, while the rest of the bulk crystal is. When strain is applied to this system the top layer is free to relax by folding upward creating several strain solitons.

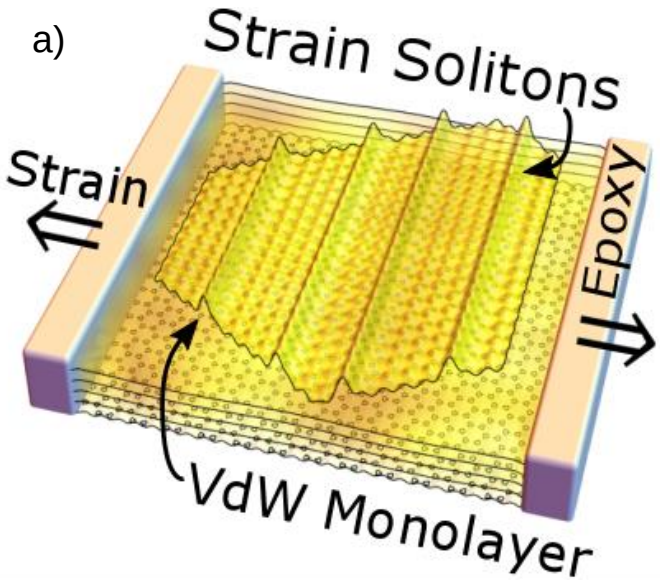

Figure 7.4: Cartoon model of a Free Monolayer with Soliton Formation

a) Linear solitons occurring in the top monolayer of a material. Here the bulk crystal is bound by epoxy and experiences the full strain field. The free monolayer begins to form soliton folds in a competition between strain energy and the Van Der Waals force

The last piece of this puzzle is how to achieve a free monolayer, such as the one pictured in the cartoon. Many groups have noted the existence of unbound monolayers due to the cleaving procedure required for UHV-STM. For a typical cleave we therefore expect that some portions of the freshly exposed bulk will not be in direct contact with our constraint epoxy, but instead be floating monolayers. In these regions under the application of strain we predict the formation of strain solitons to relax the high strain of the layers below creating a linear folding of the unbound layer in many discrete solitons.

Voltage control of the strain system has an additional benefit, we can choose a single large scan area to use during the experiment, and since the strain is applied gradually the STM can remain in feedback with the region throughout. Presented in Figure 7.5a is the area chosen before the application of any strain voltage. Here we only see slight variations due to lattice defects, the multitude of which have been categorized in previous chapters. Using these as landmarks, we see that they persist, providing further proof the STM is remaining in stable position. After application of tensile strain we see that this area develops long range linear features, whose density are strain dependent. Figure 7.5b,c are the resulting area when under 1% and 2% strain respectively. From the macroscopic perspective the

formation of line defects under strain fit the prescription of strain solitons, which similarly occur only under high strain magnitudes. Looking closer, we see that the network of linear features are connected at three way –Y like junctions, additionally suggesting a preferred directionality.

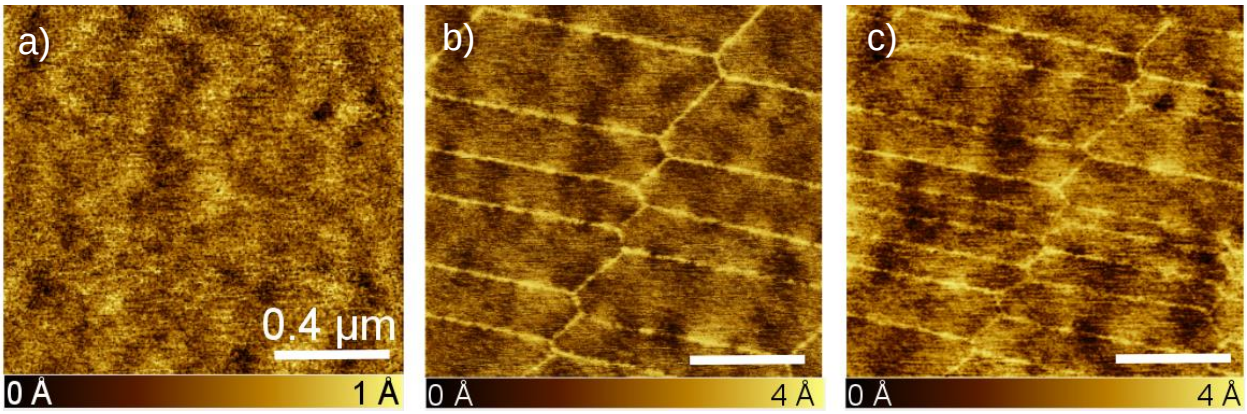

Figure 7.5: $MoSe_2$ Soliton Formation at 0, 1, and 2% Strains

a) A 1.5 x 1.5 µm^2 area of pristine $MoSe_2$ at 0% strain. Since the strain is applied in-situ we image the same region (using defects as markers) as we apply b) 1% and c) 2% strain. The resulting line defects fit with the prescription of strain solitons, we see their density increase in order to relieve higher strains.

To more accurately represent the data we update our cartoon model to show strains that don't conform to a crystallographic axis. The resulting three dimensional structure is shown in Figure 7.6a. Here since the strain axis does not have a single component the lattice must relax in some linear combination of existing high symmetry vectors. This enables the creation of solitons along all three crystallographic axes. From this cartoon it is much easier to see the terminations of these folds on the edges of the atomic sheet. Further proof that the features seen in Figure 7.5 are indeed solitons is to find the monolayer edge experimentally. We image the step edge where our features stop in Figure 7.6b. Through this image we see that our linear defects terminate at the boundary of our free monolayer, and do not propagate into bulk. We additionally find that the solitons can be reversibly tuned through compressive cycling allowing us to control the fold density in this free standing layer. With this system we should both be able to scan and measure transmitted strain effects, both for the bulk material and folded sheet.

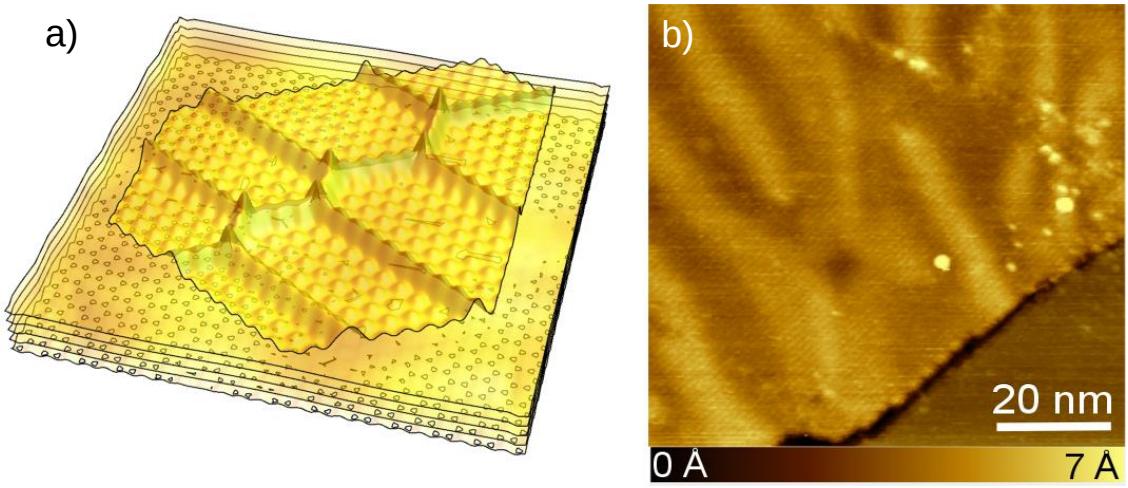

Figure 7.6: Soliton Formation Pattern at the Monolayer Edge

a) We update our soliton model to include components along all three zigzag axes, where we see the expected fold structure should start at the monolayer edge b) Experimental observation of the monolayer edge itself. We see no evidence of solitons occurring in the bulk crystal

7.3 Atomic Scale Confirmation of Strain Solitons

So far we have only discussed strain of the $MoSe_2$ lattice with a large field of view. Without atomic resolution this only provides a topical nature of these features, to achieve an accurate identification we must look at the atoms themselves. Because these features are folds of the lattice, the most striking feature of a strain soliton is the addition of an extra row of atoms along its length. Since we expect this addition to occur smoothly, we should be able to distinguish a soliton from a tear or point lattice defect. Folds cannot remove or break the atomic periodicity. For conclusive proof of a soliton, we must devise a way to detect a change in atomic frequency over the length of the feature. The easiest way to accomplish this is to take atomic resolution images using STM. Figure 7.7a is a 5x5 nm2 square zoom in on a single strain soliton, notable dimensions are a 3 Å height and roughly a 3nm width.

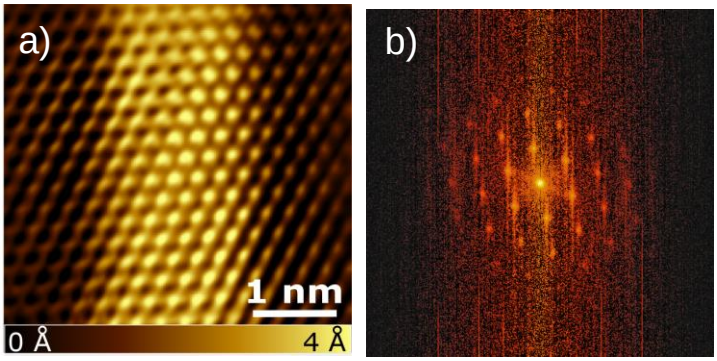

Figure 7.7: Atomic Resolution of a Single Soliton and its FFT

a) An atomic resolution image across a single soliton, notable dimensions are a 3Å height and a 3nm width b) The FFT of such features. We can see atomic peaks however we are unable to discern a clear difference in peak position (lattice compression) due to broadening of the FFT.

With lattice visibility, STM images also determine that every soliton is a contraction along the zigzag direction (short axis), and that this is perpendicular to the direction of propagation along armchair (long axis). Over the feature we see no evidence of atomic distortions, mirror twin or other grain boundaries allowing us to rule out these types of lattice defects. With a soliton distance of 3.3 nm there are roughly 9-10 atoms across the soliton.

Measurements of lattice spacing are useful, but do not prove beyond doubt that this feature has contracted under strain. A more accurate way is to detect an extra row of atoms is direct comparison with the strained lattice. Typically to detect an abrupt change in the lattice frequency we would use the Fourier space of an image, however here signal strength is proportional to the number of atoms that carry that specific frequency. Since the addition of a single atom in a smooth manner only results in small changes, detecting this slightly altered atomic placement would require extremely high resolution in Fourier space, whereas for our data this simply appears as a blurring effect, which is exemplified in Figure 7.7b. Here the atomic peaks and their harmonics are clearly visible however the resolution on the central atomic peak is not high enough to detect a change in lattice constant. This is in part due to the infrequency of solitions compared to the perfect lattice, diminishing any available signal. Instead we find it much easier to detect the extra atomic row in real space.

Since the soliton is a three dimensional feature, we can use a high pixel density image to extract a height profile and local atomic spacing. We present this data in Figure 7.8a, where we can clearly see

each atom individually as a ripple occurring atop a Gaussian-like profile. With this information we can extract the lateral profile by determining the distance between local maxima. If we do this for height slices taken on both a strained off (blue) and on (red) soliton region the result is Figure 7.8b. In this figure, we plot the lateral spacing $\vec{x}$ versus distance detecting exact compression of the soliton along its length. This allows us to measure the total distance spanned by the soliton and subtract it from a pristine region. The resulting difference is equivalent to a strained lattice constant, specifically 3.4 Å, indicating that both sides of the strain soliton are commensurate. More useful information can be extracted from the strained region's lattice spacing of 3.4 Å, this strain value (roughly 3%) matches the maximum strain of our device.

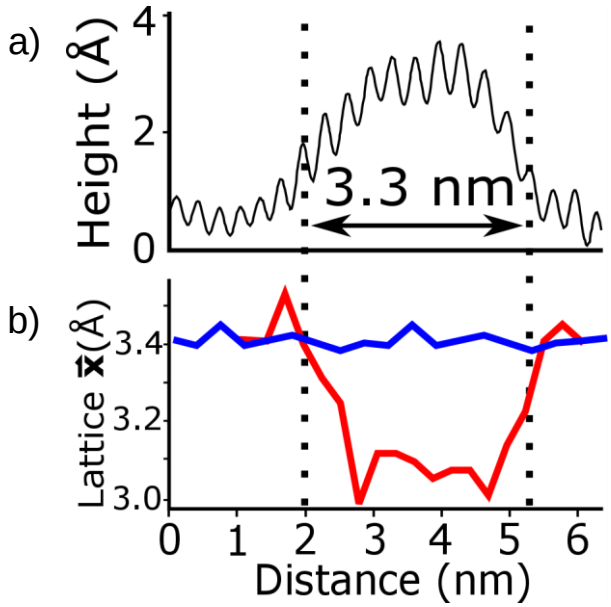

Figure 7.8: Real Space Topographic Slice and Lateral Atomic Spacing

a) The height profile of a single soliton, we see that it is 9-10 atomic lengths and roughly 3Å high b) An extracted lateral spacing on (red) and off the soliton (blue). We see that the soliton is indeed a lattice compression, however when adding in the vertical component we regain the original lattice spacing of 3.3 Å. The blue curve corresponds to 3% applied strain

To further this analysis we extend this same concept into two dimensions. Since we are now attempting to compare two regions locally it will be easiest to visualize this as a beat frequency implicit from the change of atomic placement. Mismatching atoms on the soliton will deviate from a pristine area, over the length of the soliton and return to matching if the two sides of the soliton are commensurate. To combat experimental drift or other distortions due to acquisition time we need to choose our rows along the fast axis of the scan. Put simply, we need to acquire the both lattices simultaneously for comparable

results. We present the two regions chosen for this analysis in Figure 7.9a. Our scanning direction is oriented so that each image line contains half of the soliton region and half the pristine. On the left side of the image is an incoming soliton from which we will extract the (red) region, the middle a junction where the soliton splits out of frame, and the right side a fully pristine lattice from which we will extract the (blue) region. Stitching the left and right side of the experimental data together we arrive at Figure 7.9b. To easily detect a beat frequency we label each armchair row separated by ~1.8 Å on the pristine side. At the top of the image it is seen that same row 1 matches on both sides, but on the bottom the pristine (blue row) 18 now matches with row 19 (red). A real space analysis is conclusive, crossing the soliton we gain a single atomic row, additionally we find that the length is again 3.3 nm which is roughly what we expected from simply measuring its distance in Figure 7.8a,b.

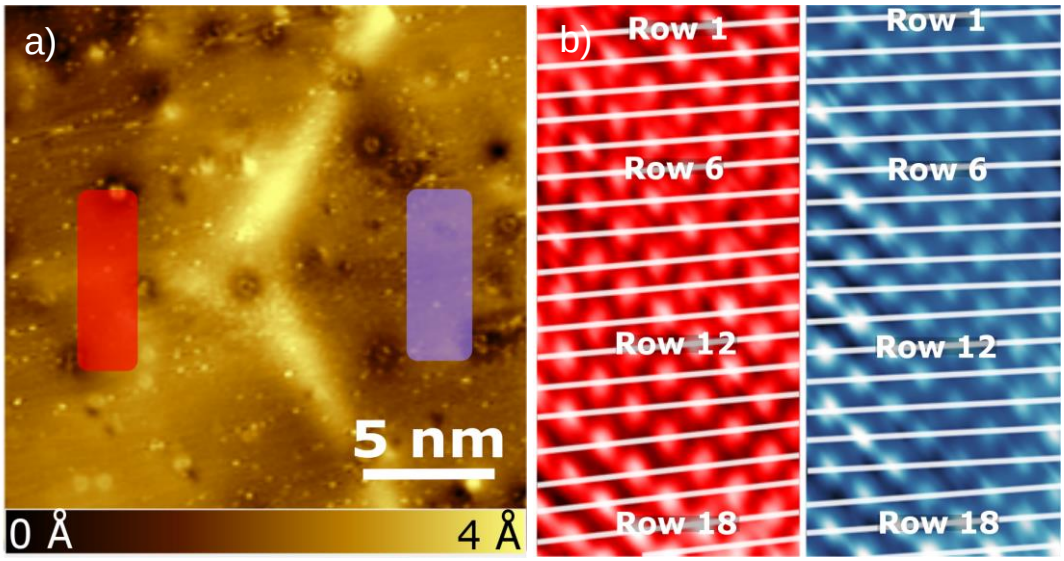

Figure 7.9: Comparison Pristine Lattice with Soliton

a) Comparison regions for simultaneous acquisition of a strain soliton (red) and a pristine area (blue) b) The atomic resolution of the two regions for comparison. We see while the first row matches, after crossing the soliton row 19 matches with row 18 this means that across the length we have gained an extra row of atoms

7.4 Theoretical Models of Soliton Direction

With confirmation of lattice contractions along an axis, we have experimentally verified the presence of strain solitons. We now address why these contractions were only seen along the zigzag direction. This contrasts solitons found in graphene, which naturally occur along any direction of high

symmetry and can be manipulated to take almost any shape. Therefore we suspect terms in the Van Der Waals binding must break the symmetries present in graphene. The main difference for TMDs is the three atom unit cell, here unlike graphene the bonds out of plane are much stronger, dictating which folds will be energetically unfavorable due to an increase of Van Der Waals coupling. To account for this, we directly calculate the Van Der Waals binding energy between layers using a two layer model. In this calculation, we measure the binding of one sheet as the other is incrementally moved through all possible translation vectors. The resulting two dimensional energy landscape for the two sheets of MoSe2 is given in Figure 7.10a. Looking at the high symmetry directions we see that the lowest energy configuration for contraction occurs along the zigzag orientation and the highest occurs along armchair. Taking this further we use this result in a shear-lag model[137,138] to calculate soliton formation for strains applied along the principle axes. For strains applied to zigzag, our model predicts parallel lines along the armchair direction, shown in Figure 7.10b. The inclusion of solitons into the sheet propagate inward from the monolayer edge. Performing a similar analysis along armchair we find that the top layer solitons form in a linear combination of zigzag directions, forming –X like soliton junctions as seen in Figure 7.10c. Empirically the experimental data supports these formation tendencies. The switch from linear to –X junctions followed closely by –Y junctions is seen in an image series taken for increasing strain, results are presented in Figure 7.11.

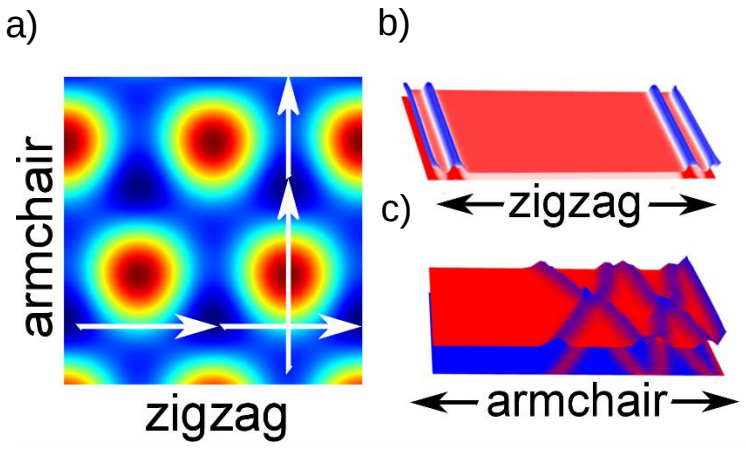

Figure 7.10: Van Der Waals Binding Energy and Shear Lag Soliton Formation

a) Van Der Walls energy as a function of layer stacking b) Soliton formation when straining along the zigzag direction using a shear lag model. We see that the solitons fold using zigzag as their short axis c) the same calculation when straining along armchair similarly solitons relieve the lattice using only zigzag components

7.5 Observed Soliton Formation and Modeling

In an effort to understand why the lattice prefers to form the –Y junction we reversibly watch the formation of the soliton features in the same region. To create this image series we take STM topographies as we increment the applied strain. In Figure 7.11a we detect the first inclusion of a strain soliton propagating inward from the left hand corner of the image. Further applying strain to this region we arrive at Figure 7.11b where we see formation of the second fold creating an –X junction. This junction persists for some time, however when we reach strains upwards of 1% the junction begins to split, as we see in Figure 7.11c. This image is specifically zoomed in to see the initial split of this feature. As we continue to apply strain the two junctions begin to separate from one another resulting in two –Y junctions, as were seen for the entirety of most of the sample, Figure 7.11d.

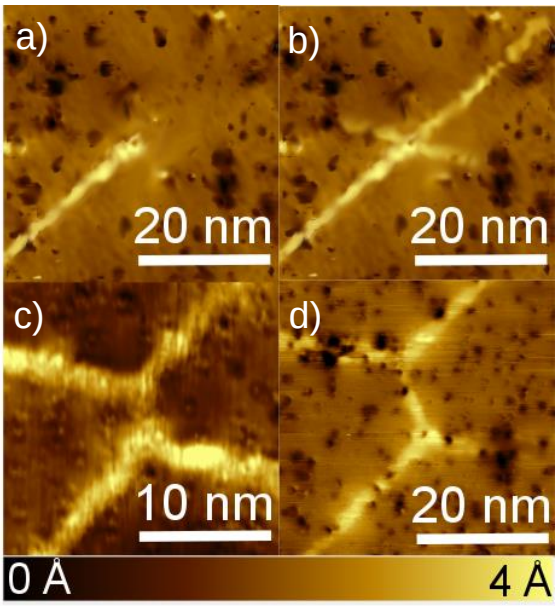

Figure 7.11: Soliton Formaiton Pattern

a) The first soliton enters the monolayer propagating upward from the bottom left hand corner b) After the first linear soliton has fully formed a second one seeds relieving the next highest strain direction c) The completed –X junction begins to separate creating a third soliton d) The completed soliton network with two –Y juncitons

What we have described above is a three step process, solitons seed, –X junctions form, then those transform into –Y junctions. With simple mathematics we can describe the processes through which these soliton building blocks incorporate themselves into the layer. Starting with a description of the

soliton seed, we need to see how a strained and unstrained lattice combine to form the proto-soliton. We do this simply by noting the formation of linear moiré patterns that form from the application of a strain field. It is these moiré that are the initial points that energetically favor the soliton fold formation. We can see these quite distinctly in Figure 7.12, where we plot the free monolayer versus its entirely strained partner, the bulk crystal.

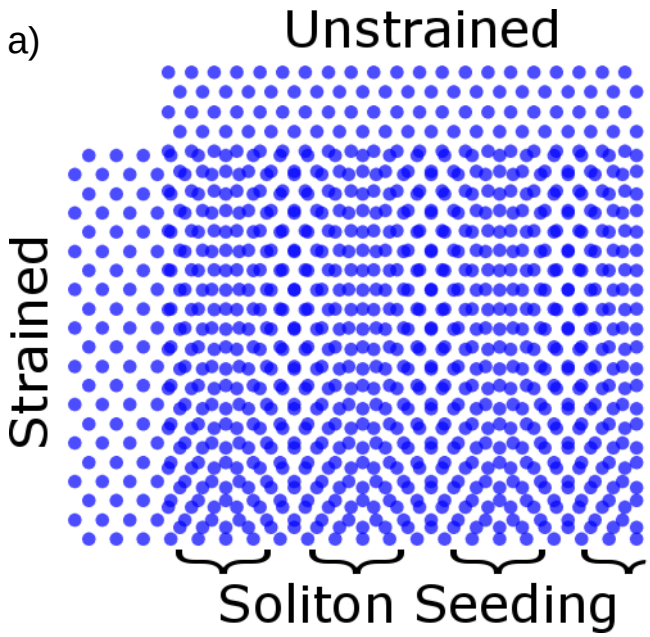

Figure 7.12: Demonstration of a Strain Moiré and Its Effect on Soliton Seeding

a) A linear moiré lattice formed at the intersection of a strained and unstrained lattice. This formation is likely influential in propagating the first strain solitons throughout the lattice

In this figure we can see at the overlap of our two lattices the resulting moiré is exactly the necessary bunching needed to start a soliton superlattice[140]. We also note that due to defects and strain direction this will become more complicated and less periodic than under the perfect conditions we apply to the layers of Figure 7.12.

Calculations above provide good theoretical match to observed soliton direction, and from the models presented we can understand how strain relief drives soliton formation in a linear strain regime; but how do we use these to describe the –X and –Y junctions. To answer this we construct the theoretical formation of the experimental data in Figure 7.11 from two postulates. First, that the strain vector will take

any arbitrary direction in experiment. And second, soliton components can only relieve strain along the three zigzag directions. With these simple rules the solitons created through relaxation of the lattice will match with our observations of –X intersections followed closely by –Y junctions.

Intuitively we know that any arbitrary vector can be described using a two component basis. Therefore, the necessary inclusion of three vector components seems contradictory to basic linear algebra. Fortunately, this is expected from mathematical formalisms, as strain vectors affect both lattice directions simultaneously through matrix multiplications. So instead of needing to describe a simple position vector, we need a basis that spans the possible two dimensional 2x2 strain matrices. With an arbitrary strain direction, the inverse matrix which denotes strain relief additionally must be compositely formed from transforms along all three zigzag directions (from our second postulate), since solitons can only form along these high symmetry vectors. Our new objective is to find a suitable matrix decomposition for the inverse of applied strain, using only these three principle directions. As mentioned earlier, a simple vector analysis is insufficient; the inverse matrix for an arbitrary direction that contains two out of three zigzag components will necessarily result in persistent off diagonal terms (barring cases where the applied strain shares a high symmetry axis).

To begin the decomposition process, we first write out the matrix for application of an arbitrary strain. Since we wish to apply strain at any choice of angle θ, by definition we need to project vectors that compose our lattice basis onto the strain direction. After projecting, we extend the parallel components, while leaving those perpendicular unchanged. The resulting vector is then transformed back to its original angle, and will compose our newly strained lattice basis. The matrix representation of this process is:

$$S_\theta(\epsilon) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \epsilon + 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Since each strain matrix contains an angle and a magnitude we shorten the description to $S_\theta(\epsilon)$ making it easier to perform operations without writing so many rotation matrices. We now need a procedure to write the inverse matrix for the above strain matrix using only three components, namely the zigzag directions of our lattice.

$$v_1 = a_0\binom{1}{0}, v_2 = a_0\begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}, v_3 = a_0\begin{pmatrix} 1/2 \\ -\sqrt{3}/2 \end{pmatrix}$$

These principle directions occur at 0°,60° and –60° respectively. Concisely, we use these to write the following decomposition for our arbitrary strain matrix:

$$S_\theta^{-1}(\epsilon) = S_{-60}(\lambda_3) * S_{60}(\lambda_2) * S_0(\lambda_1)$$

Here the λs must be determined numerically based on the applied strain magnitude ϵ and direction θ. This can be done for any given direction or magnitude, and becomes a simple set of four equations with four unknowns, namely $\theta, \lambda_1, \lambda_2, \lambda_3$ (through direct comparison of the individual entries of both resulting 2x2 matrices). For demonstration, we visualize the application of these matrices one by one to a strained lattice showing that all three components are required to reverse the strain. We start by applying an arbitrary strain to a lattice, pictured in Figure 7.13a and our objective is to return to an unstrained lattice as seen in Figure 7.13b. The first step of this process is to apply a reverse strain along the x-axis (v_1) this results in Figure 7.13c**,** where we can see that the bottom corner atom is now almost corrected in x but with a vertical offset. Since strain changes both vectors at once this additionally places the upper atom of the triangle closer to its starting position, but with an offset at roughly 60°. Next we apply the second matrix which compresses along the 60°, or the v_2 axis. Again both atomic directions are affected since both vectors have a component along this direction. Therefore the best we can do is get both atoms almost to their starting position. If we move one perfectly back the other overshoots, and vice-versa. The best option is to align them along the third axis resulting in Figure 7.13d. Now since both lie along the v_3 direction we simply apply a small correction along this axis restoring the atoms to their original positions as in Figure 7.13b. Since this process is occurring in our soliton system it is useful in explain their formation pattern, first solitons relieve the highest magnitude of strain. This is chosen to be the direction that is closest aligned to the applied strain. Then additional solitons form to relieve the next largest strain component. This creates multiple –X junctions throughout the monolayer. Finally with very high strains the –Xs are required to break into –Y junctions to return the lattice to a perfect match utilizing a slight component along v_3 as in this scenario.

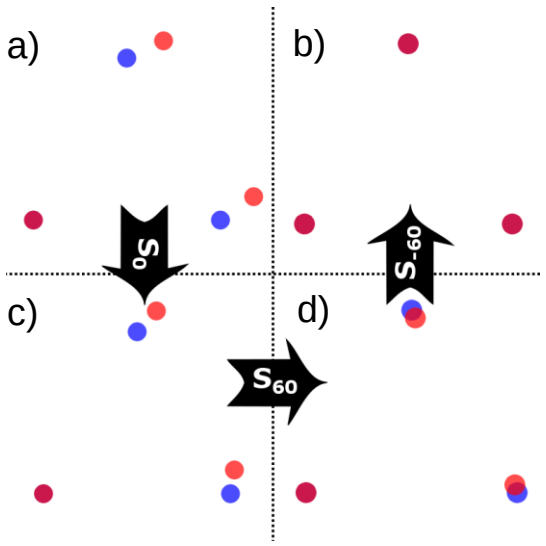

Figure 7.13: Strain Reversal Using a Three Zigzag Component Inverse Matrix

a) The unit cell of $MoSe_2$ with an arbitrary strain applied to it b) The resulting lattice after all three zigzag components have been enacted c) The lattice shown in a after only the x-axis inverse is applied d) The lattice shown in a after having applied an inverse along both principle lattice vectors. Here we see that a third component is necessary since strain application is matrix multiplication not simply vector addition

We can additionally plot the magnitudes of the λs as a function of strain angle; since the system has multiple symmetries we expect that these will look the same for each vector direction with an applied rotation of 120°. Similarly strain applied along 0° is the same as that along 180° therefore we expect the plot to have mirror symmetry along its principle direction. The resulting λs are plotted in polar fashion versus strain angle θ in Figure 7.14a,b,c where each panel corresponds with the magnitude of $S_0(\lambda_1)$, $S_{60}(\lambda_2)$ and $S_{-60}(\lambda_3)$ respectively. These represent the components required to compose the $S_\theta^{-1}(\epsilon)$ matrix.

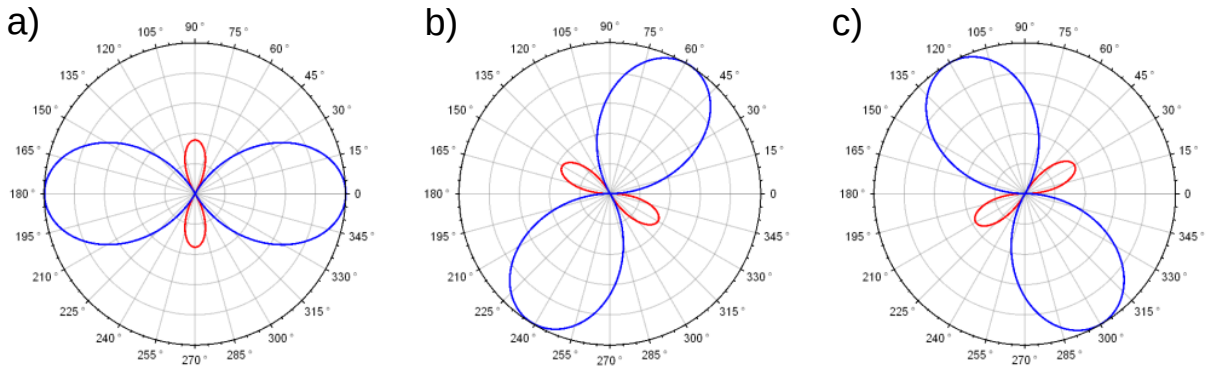

Figure 7.14: Matrix Coefficients for Each Zigzag Direction as a Function of Strain Angle

a) The necessary relaxation coefficient λ_1 for inverting stain applied along the given angle b) λ_2 c) λ_3 We note that along high symmetry directions only one (or two) components are needed to invert the strain field

7.6 Soliton –Y Junction and its Consequences

So far, our experimental analysis has been focused on data that confirms the presence of a single soliton and how that drives the formation of the soliton network. We now shift focus to the physics of the –Y junction. Notably, these are unlike the –X junctions or linear solitons; these are created by breaking parallel folds into intersecting soliton line segments. Therefore implicit to their creation we expect many consequences that must retain the original topology of the parent system. To begin our analysis let us first look at this atomistically, as we did earlier for the linear soliton. Imaging the junction using STM, a typical image is shown in Figure 7.15a. Parsing this image we see that there are three pristine regions separated by the solitons, single point defects in the lattice and elevated rows that correspond to the three solitons. Taking a closer look at the soliton center we can see that each of the three extra rows enter into what appears to be a lattice defect. To see this intersection at a much closer scale, we zoom in using Figure 7.15b to image the soliton center, arrows denoting the incoming rows from each soliton. From this picture we see that the center is not actually a simple lattice defect but one that has generated itself at a node in the soliton network. With three incoming rows the intersection therefore looks like a zero dimensional “point” soliton. Like a soliton, this point contains an extra atom, however this cannot be its own inclusion since for the –X junction does not require such a node. Therefore we must determine how this extra atom has entered the lattice.

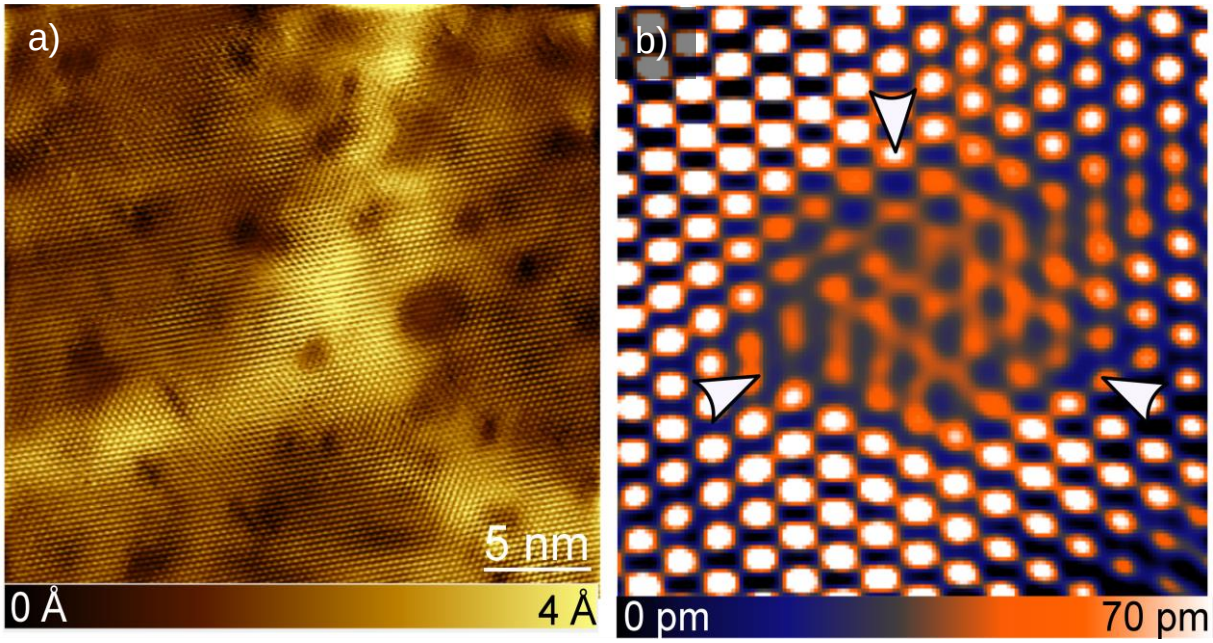

Figure 7.15: -Y Junction and Zoom in Around Atomic Center

a) Atomic resolution around a single –Y junction showing three incoming solitons and an apparent lattice defect in the center b) A zoom in around this intersection point we see that the node where all three solitons meet is additionally the center of this electronic defect

To start by analogy, this three way intersection is an atomic zipper. We can use the properties of our linear solition, most importantly the inclusion of an extra unit, to understand why the extra atom is present, where it came from, and how the two –Y junctions generated are intrinsically linked.

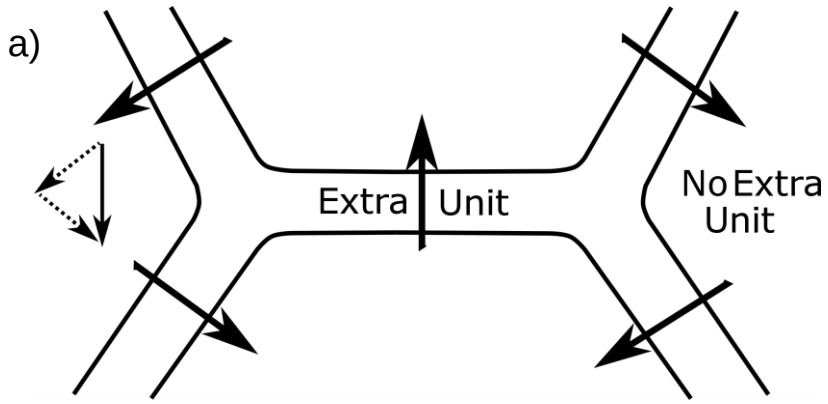

Figure 7.16: Vectorization of the Soliton Intersection

a) A vectorization of the soliton lattice where each vector is the defined lattice shift accumulated crossing the soliton. We see that as soon as a basis is chosen the remaining vectors are decided since it is necessary to cancel the extra unit on both the left and right sides. A cartoon of this cancelation is seen on the left

Momentarily, let us return to Figure 7.9, the image taken on both sides of a single –Y junction. Since crossing the soliton gives us an extra unit we can choose to represent this in a vectorized way – we define a unit vector that points along the short axis of the soliton. Whenever we compare regions separated by the soliton we add the denoted extra lattice constant, which occurs normally when crossing any soliton boundary. Of course we can choose this vector one of two ways, however once we have decided how one vector is oriented, it defines the whole system. To elaborate our reasoning, two choices must exist because is there is no explicit directionality when crossing a soliton, only that one unit must be gained. We find therefore that only two choices are possible, the chosen vector or a 180° flip where its head becomes its tail.

Once we have labeled the incoming soliton (the red region of Figure 7.9a) we can create an equivalent vectorized cartoon; the final result of which is the vector assignment of Figure 7.16. Here we can extract the vectors for the remaining solitions because we know that the blue region of Figure 7.9a does not contain an extra unit. Therefore the remaining vectors must reverse (or unzip) the incoming soliton by additively canceling. We can see that this is true for regions on both sides of the –Y junctions, where a vectorized addition is shown on the lefthand side of Figure 7.16. This resulting vector is equivalent to the negative of extra unit gained from the central soliton thereby halting its propagation through the lattice.

From this simple cartoon, we find something that has much bigger physical implications. Looking closely at each –Y junction we see that when traveling around a single junction there is a preferred handedness. Again using Figure 7.16 we see that the junction on the left is a counterclockwise set of shifts, while on the right a clockwise set. We illustrate this better in Figure 7.17, showing that each junction has a defined parity. As mentioned earlier, the choice of vectorization is arbitrary, but this is only relevant to which junction has which parity. So although we can't distinguish the handedness of an individual junction in experiment, we should be able to tell that they have a different parity from one another. This means that we can look for differences between each –Y, and that those differences will hold for every –Y of its kind globally. Effectively when we break the –X junction we create a system of –Ys with equal and opposite parity, and this in turn conserves the parity of the entire system.

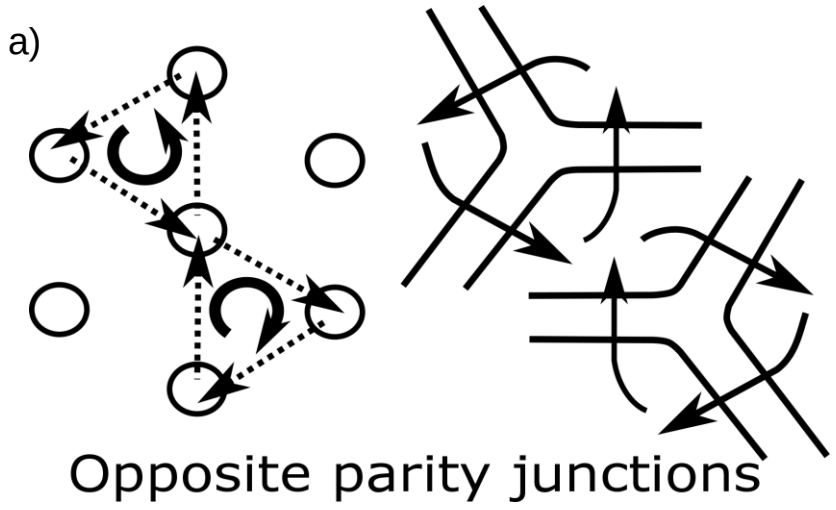

Figure 7.17: Two Parity System as Defined by -Y Junctions

a) The two possible parities that can be taken by a –Y junction based on the two ways we can add two vectors. Both of these parities exist in our lattice we now write these as –Y and –ʎ junctions.

With all of this analysis in mind we plot a close up image of a pair of –Y junctions. This is shown in Figure 7.18. In this image we have a high enough resolution to see both central intersections. This results in a stark distinction for each –Y. We see that the partner junction to Figure 7.15a,b has a center that is opposite in electronic contrast. Therefore we can extrapolate that instead of gaining a unit cell from nowhere, the system has split an existing unit, placing a metal (M) at the center of one junction and a chalcogen (X) at the other. With this determination, our soliton network is actually composed of multiple pairs of Frenkel defects whose charge is dictated by the intersection parity.

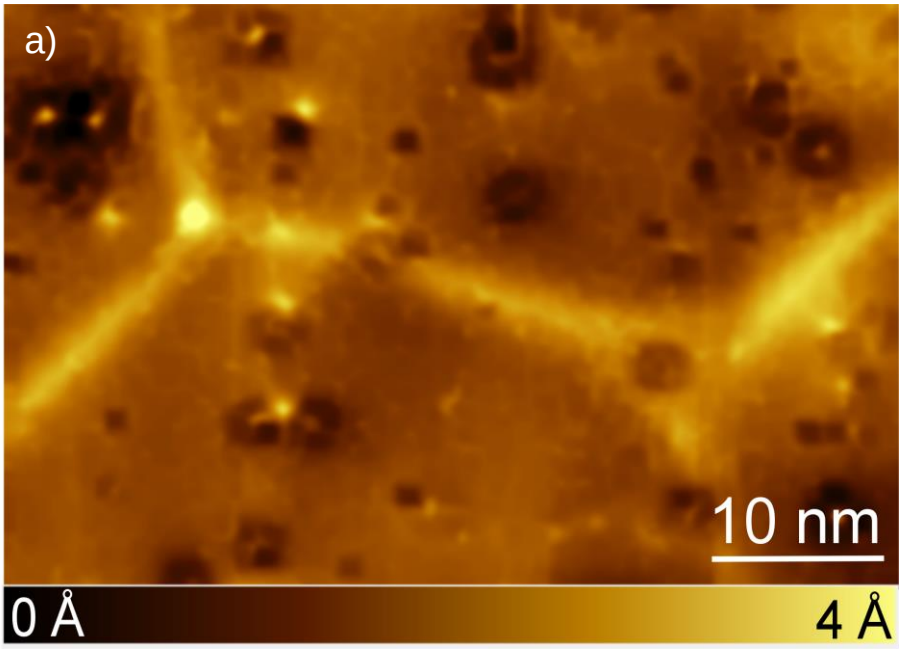

Figure 7.18: Distinguishing Junction Parity Experimentally

a) The –Y and –λ junctions, we see the existence of two defect types one at each center

7.7 Macroscopic Extension of Soliton Dynamics

The vector representation of solitons is useful in extracting macroscopic properties. Since we know that each soliton contains a detectable lattice shift we can use a two dimensional correlation to measure the lattice distortion[141,142]. This is a direct extension of our analysis for the linear soliton, since with the atomic resolution of Figure 7.15 we can exactly measure the lattice positions of the soliton –Y junction. We compare these positions with respect to the perfectly generated lattice, which was found using the FFT and calibrated to one side of the junction. This setup lets us correlate every atom in the image with its predicted positioning. We expect two sides of the image correlate perfectly to the projected lattice, by returning a value of unity, proving that the region is also commensurate. An analysis of this kind will determine when two lattices have both the same k vectors and offset (phase). Any region that is perfectly correlated therefore can only differ by a phase of 2π, or a full lattice period. Shown in Figure 7.19 is the result of a correlation of this kind. For all three pristine sides of the image we return values that are consistently greater than .9 with only slight deviations occurring around point defects. Thus we can

conclude that all three sides that are separated by solitons are matched in phase and k. As expected, the same analysis performed on the soliton itself return values slightly less than zero. We find that the soliton lattice differs along one k vector, and hence will never perfectly correlate with the pristine lattice. A soliton also does not anti-correlate, since other k vectors are retained. Hence the comparison returns values that indicate a null correlation, no connection between the lattices whatsoever.

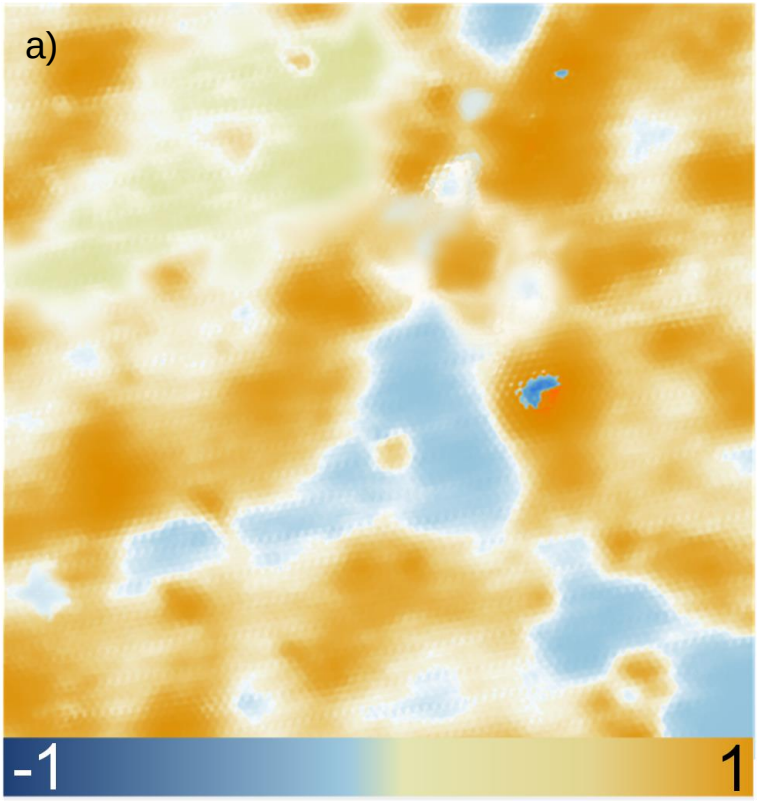

Figure 7.19: Distortion Analysis of Soliton Junction

a) A lattice correlation showing commensurate regions of the soliton lattice with the pristine

We can extend this result by performing a line integration along the strain direction for a full image where we label each region that is commensurate in the same process as we did for Figure 7.19. For an integration of this type crossing any soliton is cumulative, each crossing adds an additional lattice constant. Every pristine region therefore takes a constant value corresponding to its net shift versus a region which is defined to be at the origin. Again due to the vectorized choice this will show strain direction but only be able to reveal information relative to our origin. Our analysis results in Figure 7.20, in

this integration the bottom of the image was defined to be zero lattice shifts, resulting in seven shifts across the full image. We see that traveling along the strain direction results in the fastest ascent, top to bottom, from the point defined to be at zero. Traveling perpendicular to this we see that there is negligible change in phase, or roughly a band of constant valued slips. We can conclude, that the lattice does not need to relieve strain along this direction.

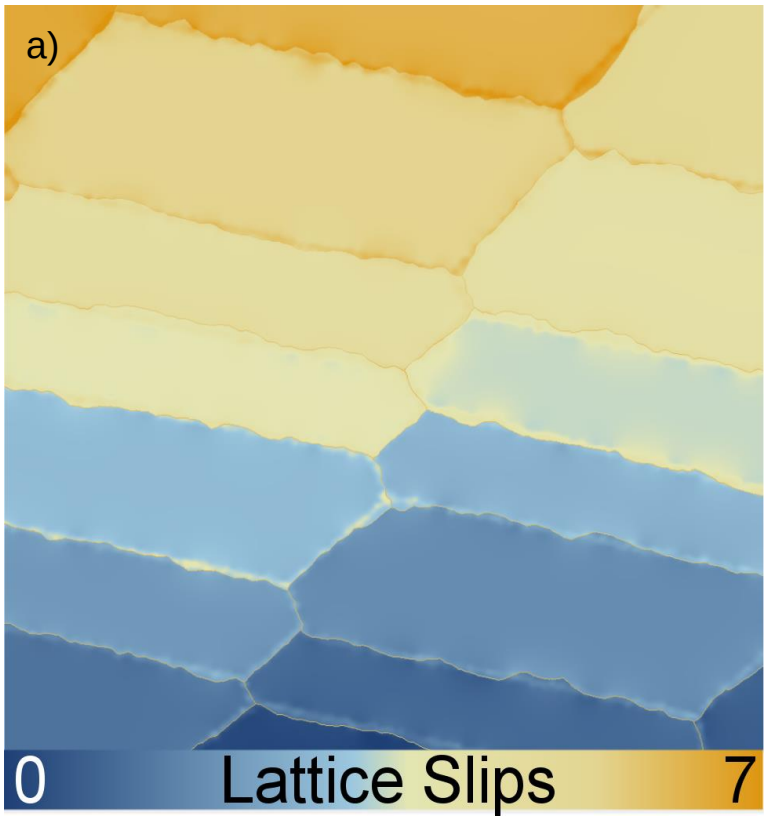

Figure 7.20: Lattice Slips Along the Strain Direction

a) A measure of the number of lattice slips along the strain direction. We see that traveling top to bottom the most strain has been relieved, while left to right there is a small difference since this is perpendicular to the strain direction

Our vector labels have allowed us to extract much more information from the raw data, and additionally predict pairs of Frenkel defects[143] throughout the soliton network. We further our lattice analysis by looking at the –Y junctions electronically using scanning tunneling spectroscopy (STS). We expect that given the parity difference, the electronic signatures of each junction will be easily distinguishable. To extract the interactions around the –Y junctions however we first need to controllably understand the lattice itself when exposed to large strains.

7.8 Electronic Properties of Strained TMDs

As mentioned in previous chapters, when taking spectra the STM will only be able to discern the smallest gap value. Therefore we will effectively measure the bulk indirect bandgap as a function of strain. Presented in Figure 7.21a we use STS to take spectra up to 3% strain and track the gap size using a set of dashed lines. We find that the average gap size for different strain percentages follows a roughly linear trend which we present in Table 7.1

Table 7.1: A Measure of the Indirect Gap Size versus Strain Percentage

Strain %	Lattice Constant	Gap Size
0 %	3.30 Å	0.86 V
1 %	3.33 Å	0.73 V
2 %	3.36 Å	0.6 V
3 %	3.40 Å	0.47 V

This spectra was taken in a region that was entirely under strain, we can use this as an additional calibration for strain percentage, without needing to measure the lattice constant directly. From here we can look at our solitons at a given strain, for the remaining experiments we chose to stay at a value of 2%. Since we expect the different –Y junctions to have differing electronics let us take spectra on them each individually and compare them with a pristine area within the vicinity. The resulting spectra presented as a waterfall plot is shown in Figure 7.21b**.** From the top down, we see the pristine region with its strained gap size of 0.6 V, an inverted –ʎ junction (to distinguish parity) and an upright –Y junction. When we look at both junctions we see that they both have the same average bandgap size as the pristine, however they also have additional features on the band edges. And in the case of the upright –Y junction there are two in gap states that appear around 100 meV and 175meV respectively.

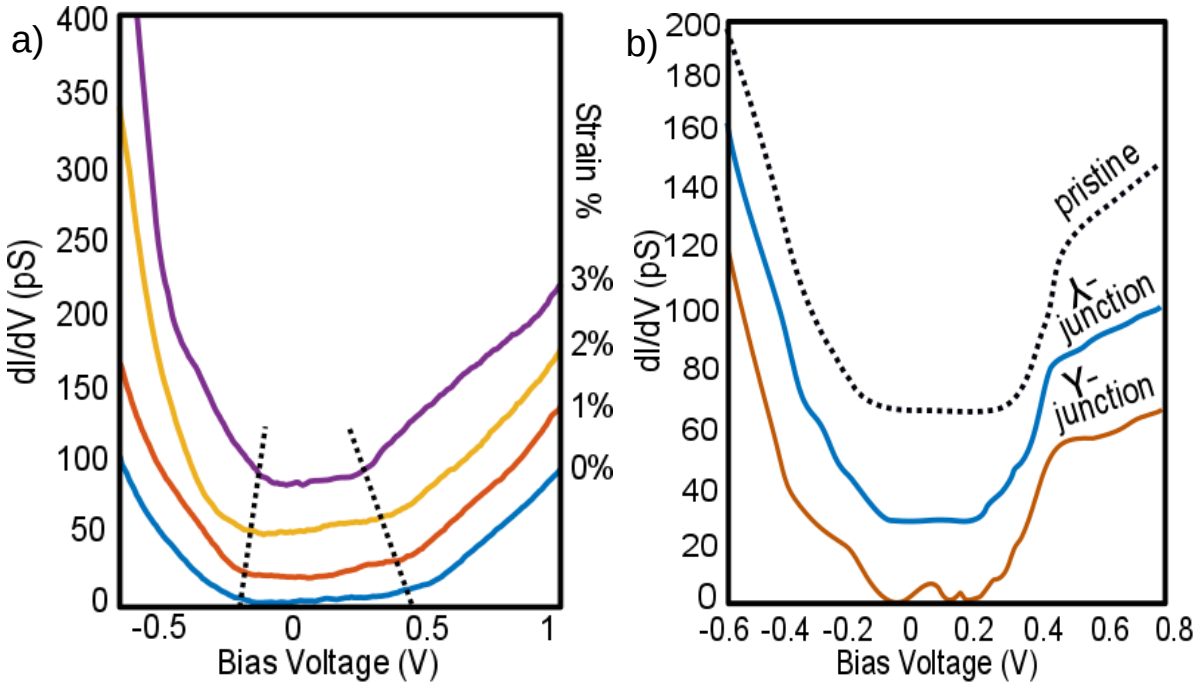

Figure 7.21: dI/dV Spectra as a Function of Strain and Position on Junction

a) A measure of the indirect gap as a function of strain percentage b) Spectra taken at 2% strain for a pristine region, a –λ junction and a –Y junction. In gap states are seen in one distinguishing them by parity

7.9 Electronic Signatures of the Soliton Network

While the STS can resolve electronic features locally using point spectra, it is a much more powerful when combined with scanning capabilities allowing us to create spatial maps of the electronic structure. Shown in Figure 7.22 is the density of states difference between regions on the soliton and in the pristine lattice for a slice at 325 meV. In this image we see that the soliton network has full resonances that begin immediately as the pristine band edge of the material opens, and we note that additional resonances occur at voltages higher in the conduction band. Again in this spatial map we can already clearly detect an electronic difference between our two junctions by looking at the size and magnitude of these states.

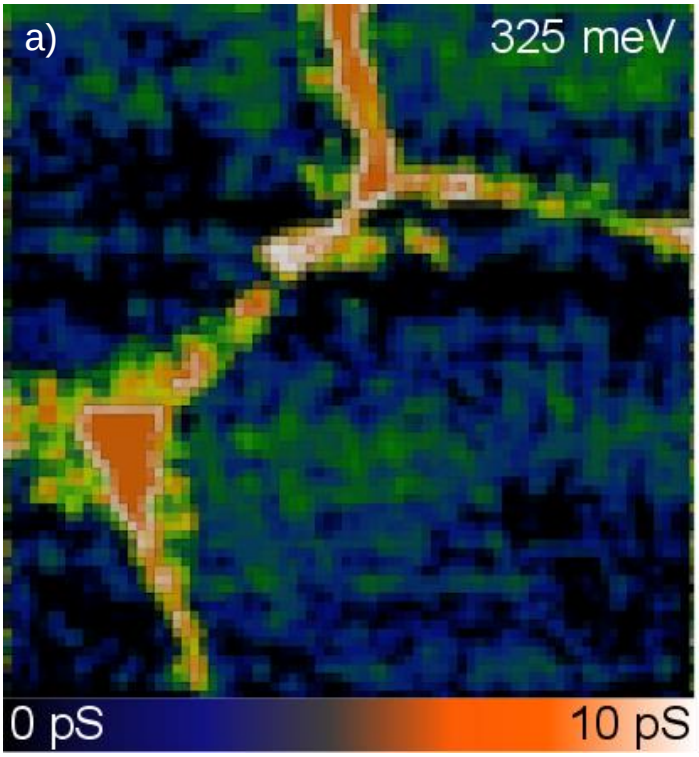

Figure 7.22: Energy Slice Showing Density of States at 325 meV

a) dI/dV map slice at 325meV highlighting difference between the two junctions at the conduction band edge. Here the states on the solitons are higher than the surrounding band, we note the difference in electronic size of the two junctions

The electronic differences between both junctions were readily apparent from both the peaks in Figure 7.21b and the gap slice of Figure 7.22. Now to fully explore these differences, we compare band slices at every energy of the dI/dV spectral map. The easiest way to visualize this is to plot the radial average around the –Y and –λ junctions. This will provide a profile of states in the junction vicinity as well as give us better intuition as to what causes these states themselves. To generate these spatially averaged spectra slices we take the full dI/dV map (from which we extracted Figure 7.22) and average the spectra found in radial rings surrounding the center of each junction. The resultant Figure 7.23a,b compares the available states for a –λ (Figure 7.23a) and –Y **(**Figure 7.23b**)** junction respectively.

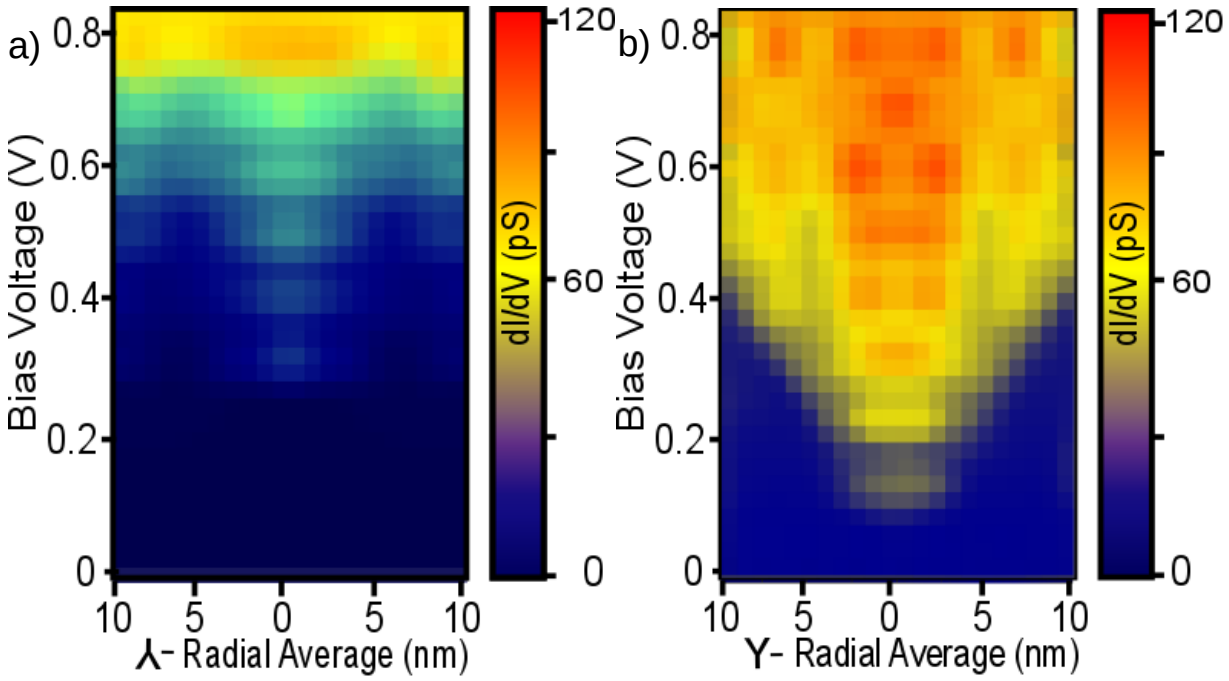

Figure 7.23: Radial Average of Density of States around Each Junction

a) Spectra taken over a –λ junction on the band edge we can see the presence of confinement states due to the local geometry b) States over a –Y junction in addition to the confinement states we see two defect states in the gap and overall electronic enhancement

Within these pictures we can see that there are two effects occurring around the band edge and in the gap. And that the junctions behave fairly differently within the same parameters. To start with the description of the simpler one we describe the –λ junction of Figure 7.23a. In this radial average we see that there are states that are evenly spaced beginning at the conduction band edge. With such an even dispersion of approximately of 75meV we can conclude that these states must arise from a harmonic oscillator like potential. As for the origins of this potential we should expect quantum confinement which arises from the altered curvature of the soliton network[120,122]. And indeed these states can be seen occurring in the other junction as well starting at roughly the same energy. However in the other junction Figure 7.23b we see an additional effect. Below the energy at which the quantum confinement takes effect we see two almost equally sized in-gap states. Since these additional states occur within the bandgap, their origin is due to the electron donor nature of this half of the Frenkel defect. This defect state may also be responsible for the enhancement of the confinement levels since we see that they are of much higher intensity within the same colorscale. Most importantly, this defect is how we can distinguish

the two –Y junctions by parity, and is a measureable consequence of the soliton network. Since we can tune the system reversibly using strain we can control the distance between these junctions, making this state of further interest especially because of its topological origin.

7.10 Quantum States at –Y Junction Centers

Although further analysis of the transport and optical effects of this soliton topology is not available to the STM we can plot the exact wavefunction magnitudes for both the defect states and the quantum confinement occurring in the band. We plot the first four quantum confinement states in Figure 7.24a-d, where we can see a clear node – antinode alternation between states. We would expect such states due to the density of a typical two dimensional harmonic oscillator or similar potential landscape. Even within the linear channel we can detect harmonic oscillator states, but these are a much weaker effect and occur at a much closer energy spacing that is hard to capture within our resolution.

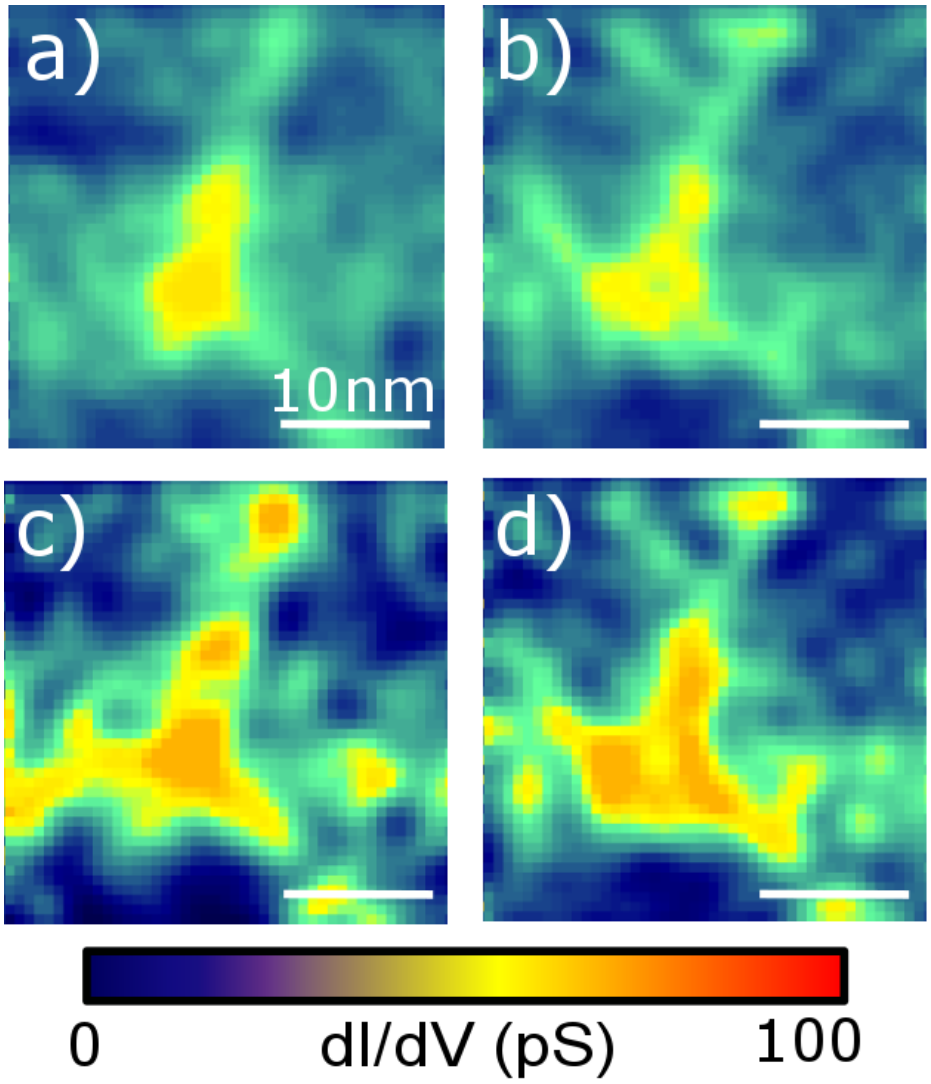

Figure 7.24: Measure of the First Four Confinement State Wavefunction Magnitudes

The quantum confinement states seen in the vicinity of both junctions at the following energies; a) 325 meV b) 400 meV c) 475 meV and d) 550 meV

Exact reasons to expect this behavior make this system promising for applications and further research. Since the soliton is a local effect, any model describing the addition of electronic states must arise from some brand of quantum confinement. The simplest of these as presented, must result in the introduction of a quantum wire system. To create this conducting channel two candidate theories are likely or some combination thereof. Strain in TMDs causes a change in the tight binding hopping parameters. Analogous to nanobubbles in graphene, these new parameters create trapping potentials for electrons, hence cause confinement within the soliton channels. Alternatively confinement can arise from

a local change from a 2H to 3R stacking configuration. This has been shown to create topological edge states, it will induce a change in the band structure of the TMD locally. Further understanding confinements of this nature can be critical to moiré and twisted angle systems and their impact on systems comprised of TMD monolayers.

To complete our analysis of the strain system we plot the in-gap defect state which is available to only the –Y junction. The two defect states are plotted in Figure 7.25a,b. From this dI/dV map we see that the defect state is roughly 5 nm across. While the defects look similar to the confinement states we note that the second state is the only state of the system that is not radially symmetric. Looking at Figure 7.25b we notice that the asymmetric state is similar in quality to the triplet state or of a hydrogenic system. Since the TMDs have a strong spin dependence this may be due to the conduction band splitting. Additionally the axis of asymmetry is pointed to the closest partnering –λ. Alternatively, this may be due to a polarization caused by the formation of Frenkel defects in a multiple charge system. Therefore we may be measuring an induced a dipole state, the exact dynamics of which may be of interest for further study.

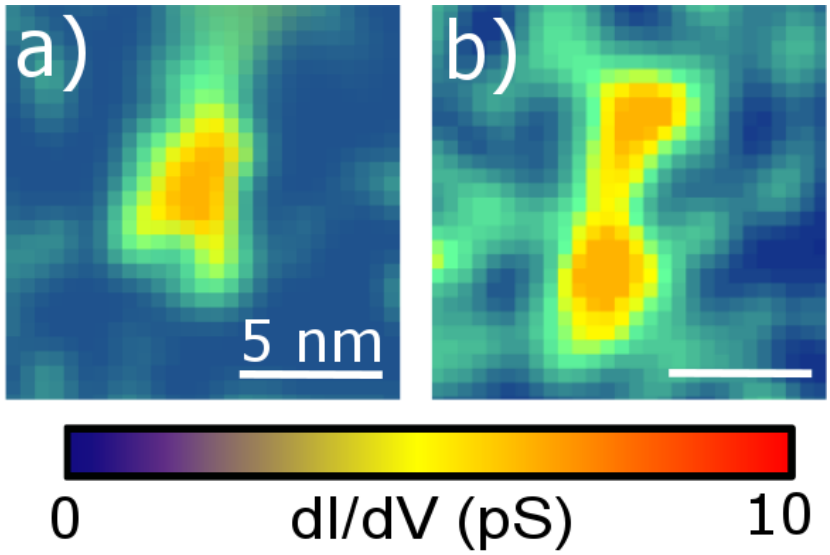

Figure 7.25: Conduction Band Defect States Occurring Exclusively at the -Y Junction

For states occurring in the gap over the –Y junction we plot their wavefunciton magnitude at energies of a) 100 meV and b) 225 meV. As a defect "orbital" this state is not spherically symmetric as was the case for the confinement states, and therefore may reflect the nature of the junction formation and its parity.

7.11 Conclusions of TMD Strain

Throughout the text we have explored the impact of defects and strain on the TMD system. And as with any research we have outlined the methods and designs that can be built upon to take any of the ideas presented further. There are many more avenues and innovations that can be made from the systems and devices that were produced at Columbia University. Within this thesis and accompanying technical documents it is my hope that future generations of graduate students will find and extract some of the knowledge I've put to paper; may it assist them on their own journey toward a PhD.

References

1 Binnig, G. & Rohrer, H. Scanning tunneling microscopy. *IBM Journal of research and development* **44**, 279 (2000).

2 Hansma, P. K. & Tersoff, J. Scanning tunneling microscopy. *Journal of Applied Physics* **61**, R1-R24 (1987).

3 Binnig, G. & Rohrer, H. Scanning tunneling microscopy. *Surface science* **126**, 236-244 (1983).

4 Tersoff, J. & Hamann, D. Theory of the scanning tunneling microscope. *Physical Review B* **31**, 805 (1985).

5 Chen, C. J. *Introduction to scanning tunneling microscopy*. Vol. 4 (Oxford University Press on Demand, 1993).

6 Mamin, H., Guethner, P. & Rugar, D. Atomic emission from a gold scanning-tunneling-microscope tip. *Physical review letters* **65**, 2418 (1990).

7 Tersoff, J. & Hamann, D. Theory and application for the scanning tunneling microscope. *Physical review letters* **50**, 1998 (1983).

8 Cady, W. G. *Piezoelectricity: Volume Two*. (Courier Dover Publications, 2018).

9 Binnig, G. & Rohrer, H. (Google Patents, 1982).

10 Besocke, K. An easily operable scanning tunneling microscope. *Surface Science* **181**, 145-153 (1987).

11 Musselman, I. H. & Russell, P. E. Platinum/iridium tips with controlled geometry for scanning tunneling microscopy. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* **8**, 3558-3562 (1990).

12 Libioulle, L., Houbion, Y. & Gilles, J. M. Very sharp platinum tips for scanning tunneling microscopy. *Review of Scientific Instruments* **66**, 97-100 (1995).

13 Kerfriden, S., Nahle, A., Campbell, S., Walsh, F. & Smith, J. The electrochemical etching of tungsten STM tips. *Electrochimica Acta* **43**, 1939-1944 (1998).

14 Kuk, Y. & Silverman, P. Scanning tunneling microscope instrumentation. *Review of scientific instruments* **60**, 165-180 (1989).

15 Chatterjee, K., Boyer, M., Wise, W. & Hudson, E. An auxiliary capacitor based ultrafast drive circuit for shear piezoelectric motors. *Review of Scientific Instruments* **80**, 095110 (2009).

16 Gupta, A. K. & Ng, K.-W. Compact coarse approach mechanism for scanning tunneling microscope. *Review of Scientific Instruments* **72**, 3552-3555 (2001).

17 Swartzentruber, B. & Schacht, M. Kinetics of atomic-scale fluctuations of steps on Si (001) measured with variable-temperature STM. *Surface science* **322**, 83-89 (1995).

18 Chen, W., Madhavan, V., Jamneala, T. & Crommie, M. Scanning tunneling microscopy observation of an electronic superlattice at the surface of clean gold. *Physical Review Letters* **80**, 1469 (1998).

19 Chidsey, C. E., Loiacono, D. N., Sleator, T. & Nakahara, S. STM study of the surface morphology of gold on mica. *Surface Science* **200**, 45-66 (1988).

20 Reinert, F. & Nicolay, G. Influence of the herringbone reconstruction on the surface electronic structure of Au (111). *Applied Physics A* **78**, 817-821 (2004).

21 Hasegawa, Y. & Avouris, P. Direct observation of standing wave formation at surface steps using scanning tunneling spectroscopy. *Physical Review Letters* **71**, 1071 (1993).

22 Crommie, M. Observing electronic scattering in atomic-scale structures on metals. *Journal of Electron Spectroscopy and Related Phenomena* **109**, 1-17 (2000).

23 Lang, N. Spectroscopy of single atoms in the scanning tunneling microscope. *Physical Review B* **34**, 5947 (1986).

24 Fujita, D. *et al.* Anisotropic standing-wave formation on an Au (111)-(23×√ 3) reconstructed surface. *Physical review letters* **78**, 3904 (1997).
25 Avouris, P., Lyo, I.-W. & Molinas-Mata, P. STM studies of the interaction of surface state electrons on metals with steps and adsorbates. *Chemical physics letters* **240**, 423-428 (1995).
26 Heller, E., Crommie, M., Lutz, C. & Eigler, D. Scattering and absorption of surface electron waves in quantum corrals. *Nature* **369**, 464 (1994).
27 Wilson, J. A. & Yoffe, A. The transition metal dichalcogenides discussion and interpretation of the observed optical, electrical and structural properties. *Advances in Physics* **18**, 193-335 (1969).
28 Wang, Q. H., Kalantar-Zadeh, K., Kis, A., Coleman, J. N. & Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nature nanotechnology* **7**, 699 (2012).
29 Chhowalla, M. *et al.* The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets. *Nature chemistry* **5**, 263 (2013).
30 Böker, T. *et al.* Band structure of ${\mathrm{MoS}}_{2},$ ${\mathrm{MoSe}}_{2},$ and $\ensuremath{\alpha}-{\mathrm{MoTe}}_{2}:$ Angle-resolved photoelectron spectroscopy and \textit{ab initio} calculations. *Physical Review B* **64**, 235305 (2001).
31 Arora, A., Nogajewski, K., Molas, M., Koperski, M. & Potemski, M. Exciton band structure in layered MoSe2: from a monolayer to the bulk limit. *Nanoscale* **7**, 20769-20775, doi:10.1039/C5NR06782K (2015).
32 Mann, J. *et al.* 2-Dimensional Transition Metal Dichalcogenides with Tunable Direct Band Gaps: MoS2(1–x)Se2x Monolayers. *Advanced Materials* **26**, 1399-1404, doi:10.1002/adma.201304389 (2014).
33 Mattheiss, L. Band structures of transition-metal-dichalcogenide layer compounds. *Physical Review B* **8**, 3719 (1973).
34 Wang, Z. *et al.* ${\mathrm{MoTe}}_{2}$: A Type-II Weyl Topological Metal. *Physical Review Letters* **117**, 056805, doi:10.1103/PhysRevLett.117.056805 (2016).
35 Zhang, C. *et al.* Charge Mediated Reversible Metal–Insulator Transition in Monolayer MoTe2 and W x Mo1–x Te2 Alloy. *ACS nano* **10**, 7370-7375 (2016).
36 Jariwala, D., Sangwan, V. K., Lauhon, L. J., Marks, T. J. & Hersam, M. C. Emerging Device Applications for Semiconducting Two-Dimensional Transition Metal Dichalcogenides. *ACS Nano* **8**, 1102-1120, doi:10.1021/nn500064s (2014).
37 Wang, X. *et al.* Chemical Vapor Deposition Growth of Crystalline Monolayer MoSe2. *ACS Nano* **8**, 5125-5131, doi:10.1021/nn501175k (2014).
38 Ubaldini, A., Jacimovic, J., Ubrig, N. & Giannini, E. Chloride-Driven Chemical Vapor Transport Method for Crystal Growth of Transition Metal Dichalcogenides. *Crystal Growth & Design* **13**, 4453-4459, doi:10.1021/cg400953e (2013).
39 Lee, Y. H. *et al.* Synthesis of Large-Area MoS2 Atomic Layers with Chemical Vapor Deposition. *Advanced Materials* **24**, 2320-2325 (2012).
40 Chen, J. *et al.* Chemical Vapor Deposition of Large-size Monolayer MoSe2 Crystals on Molten Glass. *Journal of the American Chemical Society* (2017).
41 Han, G. H. *et al.* Seeded growth of highly crystalline molybdenum disulphide monolayers at controlled locations. *Nature communications* **6** (2015).
42 Zhang, X. *et al.* Flux method growth of bulk MoS2 single crystals and their application as a saturable absorber. *CrystEngComm* **17**, 4026-4032, doi:10.1039/C5CE00484E (2015).
43 Mak, K. F. & Shan, J. Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides. *Nat Photon* **10**, 216-226, doi:10.1038/nphoton.2015.282 (2016).

44 Lee, J., Mak, K. F. & Shan, J. Electrical control of the valley Hall effect in bilayer MoS2 transistors. *Nat Nano* **11**, 421-425, doi:10.1038/nnano.2015.337

http://www.nature.com/nnano/journal/v11/n5/abs/nnano.2015.337.html#supplementary-information (2016).

45 Chen, X., Yan, T., Zhu, B., Yang, S. & Cui, X. Optical Control of Spin Polarization in Monolayer Transition Metal Dichalcogenides. *ACS nano* (2017).

46 Xu, X., Yao, W., Xiao, D. & Heinz, T. F. Spin and pseudospins in layered transition metal dichalcogenides. *Nat Phys* **10**, 343-350, doi:10.1038/nphys2942 (2014).

47 Fogler, M., Butov, L. & Novoselov, K. High-temperature superfluidity with indirect excitons in van der Waals heterostructures. *arXiv preprint arXiv:1404.1418* (2014).

48 Sie, E. J., Frenzel, A. J., Lee, Y.-H., Kong, J. & Gedik, N. Intervalley biexcitons and many-body effects in monolayer ${\mathrm{MoS}}_{2}$. *Physical Review B* **92**, 125417 (2015).

49 RadisavljevicB, RadenovicA, BrivioJ, GiacomettiV & KisA. Single-layer MoS2 transistors. *Nat Nano* **6**, 147-150, doi:http://www.nature.com/nnano/journal/v6/n3/abs/nnano.2010.279.html#supplementary-information (2011).

50 Ma, N. & Jena, D. Interband tunneling in two-dimensional crystal semiconductors. *Applied Physics Letters* **102**, 132102 (2013).

51 Soluyanov, A. A. *et al.* Type-II weyl semimetals. *Nature* **527**, 495-498 (2015).

52 Bruno, F. Y. *et al.* Observation of large topologically trivial Fermi arcs in the candidate type-II Weyl semimetal WT e 2. *Physical Review B* **94**, 121112 (2016).

53 Choi, W. *et al.* Recent development of two-dimensional transition metal dichalcogenides and their applications. *Materials Today* (2017).

54 Zeng, Q. & Liu, Z. Novel Optoelectronic Devices: Transition-Metal-Dichalcogenide-Based 2D Heterostructures. *Advanced Electronic Materials* (2018).

55 Lin, Z. *et al.* Defect engineering of two-dimensional transition metal dichalcogenides. *2D Materials* **3**, 022002 (2016).

56 Hong, J. *et al.* Exploring atomic defects in molybdenum disulphide monolayers. *Nat Commun* **6**, doi:10.1038/ncomms7293 (2015).

57 Zhou, W. *et al.* Intrinsic structural defects in monolayer molybdenum disulfide. *Nano letters* **13**, 2615-2622 (2013).

58 Yankowitz, M., McKenzie, D. & LeRoy, B. J. Local Spectroscopic Characterization of Spin and Layer Polarization in ${\mathrm{WSe}}_{2}$. *Physical Review Letters* **115**, 136803 (2015).

59 Walukiewicz, W. Carrier scattering by native defects in heavily doped semiconductors. *Physical Review B* **41**, 10218 (1990).

60 El-Mahalawy, S. & Evans, B. Temperature dependence of the electrical conductivity and hall coefficient in 2H-MoS2, MoSe2, WSe2, and MoTe2. *physica status solidi (b)* **79**, 713-722 (1977).

61 Fivaz, R. & Mooser, E. Mobility of charge carriers in semiconducting layer structures. *Physical Review* **163**, 743 (1967).

62 Wang, H., Zhang, C. & Rana, F. Ultrafast dynamics of defect-assisted electron–hole recombination in monolayer MoS2. *Nano letters* **15**, 339-345 (2014).

63 Palummo, M., Bernardi, M. & Grossman, J. C. Exciton radiative lifetimes in two-dimensional transition metal dichalcogenides. *Nano letters* **15**, 2794-2800 (2015).

64 Moody, G., Schaibley, J. & Xu, X. Exciton dynamics in monolayer transition metal dichalcogenides. *JOSA B* **33**, C39-C49 (2016).

65 Wang, H. *et al.* Fast exciton annihilation by capture of electrons or holes by defects via Auger scattering in monolayer metal dichalcogenides. *Physical Review B* **91**, 165411 (2015).
66 Bastard, G., Delalande, C., Meynadier, M., Frijlink, P. & Voos, M. Low-temperature exciton trapping on interface defects in semiconductor quantum wells. *Physical Review B* **29**, 7042 (1984).
67 Ky, N. H. & Reinhart, F. Amphoteric native defect reactions in Si-doped GaAs. *Journal of applied physics* **83**, 718-724 (1998).
68 Ovchinnikov, D., Allain, A., Huang, Y.-S., Dumcenco, D. & Kis, A. Electrical transport properties of single-layer WS2. *ACS nano* **8**, 8174-8181 (2014).
69 Baugher, B. W. H., Churchill, H. O. H., Yang, Y. & Jarillo-Herrero, P. Intrinsic Electronic Transport Properties of High-Quality Monolayer and Bilayer MoS2. *Nano Letters* **13**, 4212-4216, doi:10.1021/nl401916s (2013).
70 Raja, A. *et al.* Coulomb engineering of the bandgap and excitons in two-dimensional materials. *Nature Communications* **8**, 15251 (2017).
71 Liu, L., Qing, M., Wang, Y. & Chen, S. Defects in graphene: generation, healing, and their effects on the properties of graphene: a review. *Journal of Materials Science & Technology* **31**, 599-606 (2015).
72 Zhong, J.-H. *et al.* Quantitative correlation between defect density and heterogeneous electron transfer rate of single layer graphene. *Journal of the American Chemical Society* **136**, 16609-16617 (2014).
73 Dean, C. R. *et al.* Boron nitride substrates for high-quality graphene electronics. *Nature nanotechnology* **5**, 722-726 (2010).
74 Mayorov, A. S. *et al.* Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano letters* **11**, 2396-2399 (2011).
75 Britnell, L. *et al.* Electron tunneling through ultrathin boron nitride crystalline barriers. *Nano letters* **12**, 1707-1710 (2012).
76 Ajayi, O. *et al.* Approaching the Intrinsic Photoluminescence Linewidth in Transition Metal Dichalcogenide Monolayers. *2D Materials* (2017).
77 Wang, J. I.-J. *et al.* Electronic transport of encapsulated graphene and WSe2 devices fabricated by pick-up of prepatterned hBN. *Nano letters* **15**, 1898-1903 (2015).
78 McDonnell, S., Addou, R., Buie, C., Wallace, R. M. & Hinkle, C. L. Defect-dominated doping and contact resistance in MoS2. *ACS nano* **8**, 2880-2888 (2014).
79 Eichfeld, S. M. *et al.* Highly scalable, atomically thin WSe2 grown via metal–organic chemical vapor deposition. *ACS nano* **9**, 2080-2087 (2015).
80 Muratore, C. *et al.* Continuous ultra-thin MoS2 films grown by low-temperature physical vapor deposition. *Applied Physics Letters* **104**, 261604 (2014).
81 Zhang, Y. *et al.* Direct observation of the transition from indirect to direct bandgap in atomically thin epitaxial MoSe2. *Nature nanotechnology* **9**, 111-115 (2014).
82 Van Der Zande, A. M. *et al.* Grains and grain boundaries in highly crystalline monolayer molybdenum disulphide. *Nature materials* **12**, 554-561 (2013).
83 Roy, A. *et al.* Structural and electrical properties of MoTe2 and MoSe2 grown by molecular beam epitaxy. *ACS applied materials & interfaces* **8**, 7396-7402 (2016).
84 Kerelsky, A. *et al.* Absence of a Band Gap at the Interface of a Metal and Highly Doped Monolayer MoS2. *Nano Letters* (2017).
85 Kang, K. *et al.* High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity. *Nature* **520**, 656 (2015).

86 Manzeli, S., Ovchinnikov, D., Pasquier, D., Yazyev, O. V. & Kis, A. 2D transition metal dichalcogenides. **2**, 17033, doi:10.1038/natrevmats.2017.33 (2017).
87 Voiry, D., Mohite, A. & Chhowalla, M. Phase engineering of transition metal dichalcogenides. *Chemical Society Reviews* **44**, 2702-2712 (2015).
88 Zhang, S. *et al.* Defect Structure of Localized Excitons in a WSe 2 Monolayer. *Physical review letters* **119**, 046101 (2017).
89 Kresse, G. & Hafner, J. Ab initio molecular dynamics for open-shell transition metals. *Physical Review B* **48**, 13115 (1993).
90 Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B* **54**, 11169 (1996).
91 Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Physical review letters* **77**, 3865 (1996).
92 Blöchl, P. E. Projector augmented-wave method. *Physical review B* **50**, 17953 (1994).
93 Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B* **59**, 1758 (1999).
94 Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Physical review B* **13**, 5188 (1976).
95 Haldar, S., Vovusha, H., Yadav, M. K., Eriksson, O. & Sanyal, B. Systematic study of structural, electronic, and optical properties of atomic-scale defects in the two-dimensional transition metal dichalcogenides M X 2 (M= Mo, W; X= S, Se, Te). *Physical Review B* **92**, 235408 (2015).
96 Jin, Z., Li, X., Mullen, J. T. & Kim, K. W. Intrinsic transport properties of electrons and holes in monolayer transition-metal dichalcogenides. *Physical Review B* **90**, 045422 (2014).
97 Tongay, S. *et al.* Thermally Driven Crossover from Indirect toward Direct Bandgap in 2D Semiconductors: MoSe2 versus MoS2. *Nano Letters* **12**, 5576-5580, doi:10.1021/nl302584w (2012).
98 Pradhan, N. R. *et al.* Ambipolar molybdenum diselenide field-effect transistors: field-effect and hall mobilities. *ACS nano* **8**, 7923-7929 (2014).
99 Wang, L. *et al.* One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614-617 (2013).
100 Amani, M. *et al.* Near-unity photoluminescence quantum yield in MoS2. *Science* **350**, 1065-1068 (2015).
101 Amani, M. *et al.* High luminescence efficiency in MoS2 grown by chemical vapor deposition. *ACS nano* **10**, 6535-6541 (2016).
102 Godde, T. *et al.* Exciton and trion dynamics in atomically thin MoSe 2 and WSe 2: Effect of localization. *Physical Review B* **94**, 165301 (2016).
103 Fang, Y. *et al.* Investigation of temperature-dependent photoluminescence in multi-quantum wells. *Scientific reports* **5** (2015).
104 Baranowski, M. *et al.* Dark excitons and the elusive valley polarization in transition metal dichalcogenides. *2D Materials* **4**, 025016 (2017).
105 Quereda, J., Ghiasi, T. S., van Zwol, F. A., van der Wal, C. H. & van Wees, B. J. Observation of bright and dark exciton transitions in monolayer MoSe 2 by photocurrent spectroscopy. *2D Materials* (2017).
106 Dery, H. & Song, Y. Polarization analysis of excitons in monolayer and bilayer transition-metal dichalcogenides. *Physical Review B* **92**, 125431 (2015).
107 Zhang, X.-X., You, Y., Zhao, S. Y. F. & Heinz, T. F. Experimental evidence for dark excitons in monolayer WSe 2. *Physical review letters* **115**, 257403 (2015).

108 Wang, G. *et al.* Magneto-optics in transition metal diselenide monolayers. *2D Materials* **2**, 034002 (2015).
109 Chhowalla, M., Liu, Z. & Zhang, H. Two-dimensional transition metal dichalcogenide (TMD) nanosheets. *Chemical Society Reviews* **44**, 2584-2586 (2015).
110 Akinwande, D., Petrone, N. & Hone, J. Two-dimensional flexible nanoelectronics. *Nature communications* **5**, 5678 (2014).
111 Yun, W. S., Han, S., Hong, S. C., Kim, I. G. & Lee, J. Thickness and strain effects on electronic structures of transition metal dichalcogenides: 2H-M X 2 semiconductors (M= Mo, W; X= S, Se, Te). *Physical Review B* **85**, 033305 (2012).
112 Johari, P. & Shenoy, V. B. Tuning the electronic properties of semiconducting transition metal dichalcogenides by applying mechanical strains. *ACS nano* **6**, 5449-5456 (2012).
113 Bhattacharyya, S. & Singh, A. K. Semiconductor-metal transition in semiconducting bilayer sheets of transition-metal dichalcogenides. *Physical Review B* **86**, 075454 (2012).
114 Roldán, R. *et al.* Theory of 2D crystals: graphene and beyond. *Chemical Society Reviews* **46**, 4387-4399 (2017).
115 Wang, Y. *et al.* Strain-induced direct–indirect bandgap transition and phonon modulation in monolayer WS2. *Nano Research* **8**, 2562-2572 (2015).
116 Mennel, L. *et al.* Optical imaging of strain in two-dimensional crystals. *Nature communications* **9**, 516 (2018).
117 Zhu, C. *et al.* Strain tuning of optical emission energy and polarization in monolayer and bilayer MoS 2. *Physical Review B* **88**, 121301 (2013).
118 Luo, X. *et al.* Effects of lower symmetry and dimensionality on Raman spectra in two-dimensional WSe 2. *Physical Review B* **88**, 195313 (2013).
119 Wang, Y., Cong, C., Qiu, C. & Yu, T. Raman spectroscopy study of lattice vibration and crystallographic orientation of monolayer MoS2 under uniaxial strain. *Small* **9**, 2857-2861 (2013).
120 He, K., Poole, C., Mak, K. F. & Shan, J. Experimental Demonstration of Continuous Electronic Structure Tuning via Strain in Atomically Thin MoS2. *Nano Letters* **13**, 2931-2936, doi:10.1021/nl4013166 (2013).
121 Duerloo, K.-A. N., Li, Y. & Reed, E. J. Structural phase transitions in two-dimensional Mo-and W-dichalcogenide monolayers. *Nature communications* **5** (2014).
122 Ghorbani-Asl, M., Borini, S., Kuc, A. & Heine, T. Strain-dependent modulation of conductivity in single-layer transition-metal dichalcogenides. *Physical Review B* **87**, 235434 (2013).
123 Soumyanarayanan, A. *et al.* Quantum phase transition from triangular to stripe charge order in NbSe2. *Proceedings of the National Academy of Sciences* **110**, 1623-1627 (2013).
124 Cockayne, D., Ray, I. & Whelan, M. Investigations of dislocation strain fields using weak beams. *Philosophical Magazine* **20**, 1265-1270 (1969).
125 Reed-Hill, R. E., Abbaschian, R. & Abbaschian, R. Physical metallurgy principles. (1973).
126 Levy, N. *et al.* Strain-induced pseudo–magnetic fields greater than 300 tesla in graphene nanobubbles. *Science* **329**, 544-547 (2010).
127 Liu, K. *et al.* Elastic properties of chemical-vapor-deposited monolayer MoS2, WS2, and their bilayer heterostructures. *Nano letters* **14**, 5097-5103 (2014).
128 Davies, A. *et al.* Lattice-matched epitaxial graphene grown on boron nitride. *Nano letters* **18**, 498-504 (2017).
129 Johnson, A. D., Cheng, F., Tsai, Y. & Shih, C.-K. Giant enhancement of defect-bound exciton luminescence and suppression of band-edge luminescence in monolayer WSe2–Ag plasmonic hybrid structures. *Nano letters* **17**, 4317-4322 (2017).

130 Yankowitz, M. *et al.* Electric field control of soliton motion and stacking in trilayer graphene. *Nature materials* **13**, 786 (2014).
131 Lin, J. *et al.* AC/AB stacking boundaries in bilayer graphene. *Nano letters* **13**, 3262-3268 (2013).
132 Guinea, F., Geim, A., Katsnelson, M. & Novoselov, K. Generating quantizing pseudomagnetic fields by bending graphene ribbons. *Physical Review B* **81**, 035408 (2010).
133 Jiang, L. *et al.* Soliton-dependent plasmon reflection at bilayer graphene domain walls. *Nature materials* **15**, 840 (2016).
134 Mitchell, N. P., Koning, V., Vitelli, V. & Irvine, W. T. Fracture in sheets draped on curved surfaces. *Nature materials* **16**, 89 (2017).
135 Lu, P., Wu, X., Guo, W. & Zeng, X. C. Strain-dependent electronic and magnetic properties of MoS 2 monolayer, bilayer, nanoribbons and nanotubes. *Physical Chemistry Chemical Physics* **14**, 13035-13040 (2012).
136 Butz, B. *et al.* Dislocations in bilayer graphene. *Nature* **505**, 533 (2014).
137 Kumar, H., Dong, L. & Shenoy, V. B. Limits of Coherency and Strain Transfer in Flexible 2D van der Waals Heterostructures: Formation of Strain Solitons and Interlayer Debonding. *Scientific reports* **5** (2016).
138 Cox, H. The elasticity and strength of paper and other fibrous materials. *British journal of applied physics* **3**, 72 (1952).
139 Alden, J. S. *et al.* Strain solitons and topological defects in bilayer graphene. *Proceedings of the National Academy of Sciences* **110**, 11256-11260 (2013).
140 Woods, C. *et al.* Commensurate–incommensurate transition in graphene on hexagonal boron nitride. *Nature physics* **10**, 451 (2014).
141 Yothers, M. P., Browder, A. E. & Bumm, L. A. Real-space post-processing correction of thermal drift and piezoelectric actuator nonlinearities in scanning tunneling microscope images. *Review of Scientific Instruments* **88**, 013708 (2017).
142 Fujita, K. *et al.* Direct phase-sensitive identification of a d-form factor density wave in underdoped cuprates. *Proceedings of the National Academy of Sciences* **111**, E3026-E3032 (2014).
143 Li, W. *Point Defects in 2D and 3D Nanomaterials: A Density Functional Theory Exploration*, Utrecht University, (2017).

Appendix A1

Landau Physics

A1.1 Introduction to Landau Physics

This appendix details the inner workings of Landau physics, as is applicable to a general strain field. This was a possible description of the linear soliton system and a circularly symmetric system. This section of work was attached to appendix because it was tangentially relevant to our discussion of strain given the current models we have for the system, but may be useful in describing others. This deals with the derivations of the typical landau physics and how they can be applied for any system, as opposed to the specific case of strain.

For charged particles traveling within a magnetic field quantum mechanics dictates the nature of the classical cyclotron orbits. The quantum states exhibited by this system were first characterized by Lev Landau and have come to be known as Landau levels. In finding the quantum analog, we borrow the techniques of classical physics with operator formalism. Therefore the simplest starting point is a quantum analog to the Lagrangian. We write this as the following, additionally noting the Hamiltonian:

$$\boldsymbol{\mathcal{L}} \equiv \sum_i \frac{1}{2} m \dot{\boldsymbol{x}}_{\boldsymbol{\imath}}^{\,2} + q \dot{\boldsymbol{x}}_{\boldsymbol{\imath}} \widehat{\boldsymbol{A}}_{\boldsymbol{i}} \;, \widehat{\boldsymbol{H}} \equiv \sum_i \frac{1}{2} m \dot{\boldsymbol{x}}_{\boldsymbol{\imath}}^{\,2}$$

Of course $\dot{\hat{x}}_\imath$ is not a well-defined quantum operator, however we can obtain its form by noting that $\widehat{\boldsymbol{p}}_i = \partial \boldsymbol{\mathcal{L}} / \partial \dot{\hat{x}}_\imath$ where $\widehat{\boldsymbol{p}}_i$ is a quantum operator defined in the position representation as $\widehat{\boldsymbol{p}}_i \equiv -\mathbb{i}\hbar \frac{\partial}{\partial x_i}$. Enacting this partial derivative we obtain:

$$\widehat{\boldsymbol{p}}_i = \frac{\partial \boldsymbol{\mathcal{L}}}{\partial \dot{\hat{\boldsymbol{x}}}_\imath} \Rightarrow \widehat{\boldsymbol{p}}_i = m \dot{\boldsymbol{x}}_{\boldsymbol{\imath}} + q \widehat{\boldsymbol{A}}_{\boldsymbol{i}} \Rightarrow \dot{\boldsymbol{x}}_{\boldsymbol{\imath}} = \frac{\widehat{\boldsymbol{p}}_i - q \widehat{\boldsymbol{A}}_{\boldsymbol{i}}}{m}$$

We now choose to define a new operator $m\hat{\dot{\boldsymbol{x}}}_{\boldsymbol{t}}$ as the canonical momentum $\widehat{\boldsymbol{\pi}}_i$ this results in the following Hamiltonian:

$$\widehat{\boldsymbol{\pi}}_i \equiv \left(\widehat{\boldsymbol{p}}_i - q\widehat{\boldsymbol{A}}_{\boldsymbol{i}}\right), \widehat{\boldsymbol{H}} = \frac{1}{2m}\vec{\boldsymbol{\pi}} \cdot \vec{\boldsymbol{\pi}} = \frac{1}{2m}\sum_i \left(\widehat{\boldsymbol{p}} - q\widehat{\boldsymbol{A}}\right)_i^{\,2}$$

We should note the Hamiltonian's similarity to that of a free particle $\widehat{\boldsymbol{p}}^2/2m$ where $\widehat{\boldsymbol{p}}$ has been replaced with the canonical momentum vector operator $\vec{\boldsymbol{\pi}}$ which includes the effects of the magnetic field as represented through choice of a gauge potential $\widehat{\boldsymbol{A}}$. We also note that although displayed for clarity the summation is typically dropped for ease of notation.

So far we have presented the material generally, however for demonstration we shall now work in a simplified case where we choose $\widehat{\boldsymbol{A}}$ as based on a uniform $\vec{\boldsymbol{B}}$ field in the z direction. This arbitrary choice is justified as any constant field direction can be recovered by simple rotation.

$$\vec{B} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix} \Rightarrow \nabla \mathrm{x}\vec{A} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}$$

It should be said that the gauge field is not uniquely defined, as it is not a physical observable, and therefore can be chosen without altering the physics of the system. This choice of gauge will however impact the shape of degenerate eigenstates, because any configuration can be a sum of these wavefunctions with differing weights. We therefore wish to view our gauge choice as instead altering the commutivity of momentum operators with the overall Hamiltonian. Given that the gauge choice will alter the shape of the resulting solutions, we can use this to determine the best solutions to use in perturbation calculations where the B-field decays within a characteristic magnetic length.

To begin with the simplest gauge representation we choose the Landau gauge which is linear in nature. In this gauge we have that:

$$\vec{A} = \begin{pmatrix} 0 \\ B_0 x \\ 0 \end{pmatrix}, \nabla \mathrm{x}\vec{A} = \; det\begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ 0 & B_0 x & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}$$

From here we can write out the Hamiltonian:

$$\widehat{\boldsymbol{H}} = \frac{1}{2m}\left(\widehat{\boldsymbol{p}} - q\widehat{\boldsymbol{A}}\right)^2 = \frac{1}{2m}\left(\widehat{\boldsymbol{p}}_x^2 + \left(\widehat{\boldsymbol{p}}_y - qB_0\widehat{\boldsymbol{x}}\right)^2\right)$$

A1.2 Translation Operators

In order to simplify our Hamiltonian further we now introduce another operator, the translation operator $\widehat{\boldsymbol{T}}(\vec{a})$ this operator shifts the overall wavefunction by the vector $\vec{a}$. Depending on our choice of geometry we will later find it useful to choose this vector along principal axes where we expect conservation. First though we work out relations involving this operator $\widehat{\boldsymbol{T}}(\vec{a})$ and its action on a state $\psi(\vec{x})$.

$$\widehat{\boldsymbol{T}}(\vec{a})|\vec{x}\rangle \equiv |\vec{x} + \vec{a}\rangle$$

Based on this we can define multiple properties that must hold:

$$\widehat{\boldsymbol{T}}(\vec{0}) = \hat{\mathbb{I}} \; ; \; \widehat{\boldsymbol{T}}(\vec{a} + \vec{b}) = \; \widehat{\boldsymbol{T}}(\vec{a})\widehat{\boldsymbol{T}}(\vec{b}) \; ; \; \widehat{\boldsymbol{T}}^{-1}(\vec{a}) = \widehat{\boldsymbol{T}}(-\vec{a}) \; ; \; \widehat{\boldsymbol{T}}^{\dagger}(\vec{a}) = \widehat{\boldsymbol{T}}^{-1}(\vec{a})$$

Unpacking these relations we first note that a zero translation must be equal to the identity operator and return the original wavefunction since $|\vec{x} + \vec{0}\rangle = \; |\vec{x}\rangle$. We also have that two consecutive translations should be the same as applying their summation by the associative property of vectors; $|\vec{x} + (\vec{a} + \vec{b})\rangle = \; |\vec{x} + \vec{a} + \vec{b}\rangle$. From this we immediately can derive the inverse as $\widehat{\boldsymbol{T}}(-\vec{a})$ since the summation will return the zero vector. The last property comes from noting that $\widehat{\boldsymbol{T}}$ is a unitary operator, meaning it has no impact on the magnitude of a wavefunction, hence we can write:

$$\widehat{\boldsymbol{T}}^{\dagger}(\vec{a}) = \widehat{\boldsymbol{T}}(-\vec{a}) \Rightarrow \langle\vec{x}|\widehat{\boldsymbol{T}}(\vec{a})|\psi\rangle = \langle\vec{x} - \vec{a}|\psi\rangle$$

$$\widehat{\boldsymbol{T}}(\vec{a})\psi(\vec{x}) \equiv \psi(\vec{x} - \vec{a})$$

We see that based on our simple definition the translation operator shifts the spatial electronic density in a fairly straightforward way. However we would prefer to have a formalization for this operator in terms of the operators available in the postulates of quantum mechanics. To remedy this we expand $\widehat{\boldsymbol{T}}$ in a series of small translations $\overrightarrow{\delta a}$:

$$\widehat{\boldsymbol{T}}(\delta a) = \widehat{\boldsymbol{T}}(0) + \delta a \frac{d\widehat{\boldsymbol{T}}}{da}(0) + \mathbb{O}(a^2)$$

We will write the first term of this differential series as an operator in its own right $\hat{k} = \frac{d\widehat{\boldsymbol{T}}}{da}(0)$ since this has no positional arguments we can write $\hat{k}$ without any a dependence. We separately use the definition of the derivative acting on $\widehat{\boldsymbol{T}}$ at an arbitrary position a:

$$\frac{d\widehat{\boldsymbol{T}}}{da}(a) = \lim_{\epsilon \to 0} \frac{\widehat{\boldsymbol{T}}(a+\epsilon) - \widehat{\boldsymbol{T}}(a)}{\epsilon} = \lim_{\epsilon \to 0} \left(\frac{\widehat{\boldsymbol{T}}(\epsilon) - \hat{\mathbb{I}}}{\epsilon} \right) \widehat{\boldsymbol{T}}(a) = \lim_{\epsilon \to 0} \left(\frac{\widehat{\boldsymbol{T}}(\epsilon + 0) - \widehat{\boldsymbol{T}}(0)}{\epsilon} \right) \widehat{\boldsymbol{T}}(a) = \widehat{\boldsymbol{k}} * \widehat{\boldsymbol{T}}(a)$$

This provides us with simple differential equation for $\widehat{\boldsymbol{T}}$ with exponential solutions:

$$\frac{d\widehat{\boldsymbol{T}}}{da}(a) = \widehat{\boldsymbol{k}} * \widehat{\boldsymbol{T}}(a) \implies \widehat{\boldsymbol{T}}(a) = e^{a\widehat{\boldsymbol{k}}}$$

However we don't know the exact nature of the $\hat{k}$ operator although we can see that this will be the generator of translations. To elucidate the action of $\hat{k}$ further we look at its effect on an arbitrary spatial wavefunction with the addition of clever zero to rewrite the derivative on a using the product rule:

$$\hat{k}\psi(x) = \frac{d\widehat{\boldsymbol{T}}}{da}(a)\psi(x)\Big|_{a\to 0} = \frac{d\widehat{\boldsymbol{T}}}{da}(a)\psi(x) + \widehat{\boldsymbol{T}}(a)\frac{d}{da}\big(\psi(x)\big)\Big|_{a\to 0} = \frac{d}{da}\Big(\widehat{\boldsymbol{T}}(a)\psi(x)\Big)\Big|_{a\to 0}$$

Since we now have the translation operator acting directly (as opposed to its derivative) we can enact the operator on $\psi(x)$:

$$\frac{d}{da}\Big(\widehat{\boldsymbol{T}}(a)\psi(x)\Big)\Big|_{a\to 0} = \frac{d}{da}\big(\psi(x+a)\big)\Big|_{a\to 0} = \psi'(x)$$

We've taken the spatial derivative along a but we just as easily could have taken this along x as a rephrasing of the terms in the definition of the derivative:

$$\frac{d}{da}\big(\psi(x+a)\big) = \lim_{\epsilon\to 0}\frac{\psi(x+a+\epsilon) - \psi(x+a)}{\epsilon} = \frac{d}{dx}\big(\psi(x+a)\big)$$

Hence we have that the operator $\widehat{\boldsymbol{k}}$ is related to the (canonical) momentum operator, which also takes the form of a derivative in a spatial basis:

$$\hat{\boldsymbol{k}} = \frac{\mathbb{i}}{\hbar}\hat{\boldsymbol{p}} \Rightarrow \hat{\boldsymbol{T}}(\vec{a}) = e^{\frac{\mathbb{i}}{\hbar}(\vec{a}\cdot\hat{\boldsymbol{p}})}$$

We have now vectorized this operator noting that $\vec{a}$ is dotted into the momentum vector operator.

$$\vec{\boldsymbol{p}} = \{\hat{\boldsymbol{p}}_{\boldsymbol{x}}, \hat{\boldsymbol{p}}_{\boldsymbol{y}}, \hat{\boldsymbol{p}}_{\boldsymbol{z}}\}$$

We note that because of our derivative d/dx actually referring to $\hat{\boldsymbol{\pi}}$ we have a more general form of $\hat{\boldsymbol{k}}$ where we must add back the gauge term. (This will be useful in the symmetric case when the gauge is non-zero in both x and y)

$$\hat{\boldsymbol{k}} = \frac{\mathbb{i}}{\hbar}(\hat{\boldsymbol{p}} + q\hat{\boldsymbol{A}})$$

Our transgression with translation operators is nearly complete, lastly we note the effect it has on functions of spatial operators. We have that:

$$\left[\hat{\boldsymbol{x}}, \hat{\boldsymbol{T}}(a)\right]|x\rangle = (\hat{\boldsymbol{x}}\,\hat{\boldsymbol{T}}(a) - \hat{\boldsymbol{T}}(a)\,\hat{\boldsymbol{x}})\,|x\rangle = (x + a - x)\,|x + a\rangle = a\hat{\boldsymbol{T}}(a)|x\rangle$$

Utilizing this commutation we have that:

$$\hat{\boldsymbol{T}}^{\dagger}(\vec{a})\hat{\boldsymbol{x}}\hat{\boldsymbol{T}}(\vec{a}) = \hat{\boldsymbol{T}}^{\dagger}(\vec{a})\left(\hat{\boldsymbol{T}}(a)\hat{\boldsymbol{x}} + a\hat{\boldsymbol{T}}(a)\right) = \hat{\boldsymbol{x}} + a\hat{\mathbb{I}}$$

As a more general statement we can now write out for any function of $\hat{\boldsymbol{x}}$ (such as our Hamiltonian):

$$\hat{\boldsymbol{T}}^{\dagger}(\vec{a})f(\hat{\boldsymbol{x}})\hat{\boldsymbol{T}}(\vec{a}) = f(\hat{\boldsymbol{x}} + a\hat{\mathbb{I}})$$

A1.3 Linear Solution

Let us return to the magnetic field problem at hand, before our interlude we had that:

$$\hat{\boldsymbol{H}} = \frac{1}{2m}\left(\hat{\boldsymbol{p}} - q\hat{\boldsymbol{A}}\right)^2 = \frac{1}{2m}\left(\hat{\boldsymbol{p}}_x^2 + \left(\hat{\boldsymbol{p}}_y - qB_0\hat{\boldsymbol{x}}\right)^2\right)$$

We see that this Hamiltonian commutes with the momentum operator $\hat{\boldsymbol{p}}_y$ because there are no instances of $\hat{\boldsymbol{y}}$. Therefore we know that eigenstates of this Hamiltonian are also eigenstates of $\hat{\boldsymbol{p}}_y$. Therefore we can conclude that this wavefunction will have a measurable y momenta and our wavefunction will be

separable along the axes. As an added benefit this means that k_y, the eigenstate of momentum, will be a good quantum number for our energy state wavefunctions.

$$[\hat{\boldsymbol{p}}_y, \hat{\boldsymbol{H}}] = 0 \;\Rightarrow\; [\hat{\boldsymbol{T}}_y(a), \hat{\boldsymbol{H}}] = 0 \;\; \hat{\boldsymbol{p}}_y|\psi_E\rangle = \hbar k_y|\psi_E\rangle$$

Using the simultaneous diagonizability of $\hat{\boldsymbol{H}}$ and $\hat{\boldsymbol{p}}_y$ we write the wavefunction as a joint state of both energy and y-axis momentum, $\left|\psi_{E,k_y}\right\rangle$. In this representation all $\hat{\boldsymbol{p}}_y$ operators simply act on the wavefunction to replace all values of $\hat{\boldsymbol{p}}_y$ with their eigenvalue $\hbar k_y$ further simplifying our Hamiltonian:

$$\hat{\boldsymbol{H}} = \frac{1}{2m}\left(\hat{\boldsymbol{p}}_x^2 + \left(\hbar k_y - qB_0\hat{\boldsymbol{x}}\right)^2\right)$$

Since k_y is a constant (of degeneracy) we have effectively reduced this to an offset one dimensional harmonic oscillator (H.O.) equation. To condense equation into a typical H.O. form this we write out the angular frequency as $\omega_C = \frac{qB}{m}$, which unsurprisingly is just the classical cyclotron frequency.

$$\hat{\boldsymbol{H}} = \left(\frac{\hat{\boldsymbol{p}}_x^2}{2m} + \frac{1}{2}m\omega_C^2\left(\hat{\boldsymbol{x}} - \frac{\hbar k_y}{m\omega_C}\right)^2\right)$$

Given that we have separated y from our Hamiltonian we can use a convenient conversion of coordinates using translation operators. We start from the set of harmonic oscillator wave functions centered around the origin:

$$\hat{\boldsymbol{H}}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}_x) = \left(\frac{\hat{\boldsymbol{p}}_x^2}{2m} + \frac{1}{2}m\omega_C^2(\hat{\boldsymbol{x}})^2\right) \;;\; \phi_{E_n}(x) = \frac{\pi^{-\frac{1}{4}}}{\sqrt{2^n n!}} e^{-\frac{x^2}{2}} H_n(x)$$

With $H_n(x)$ being given by the physicists Hermite polynomials. These fit into an energy eigen-equation, where the expectation of the energy is given as:

$$E_n = \langle\phi_n(x)|\hat{\boldsymbol{H}}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}_x)|\phi_n(x)\rangle$$

Now inserting a set of specially chosen translation operators (and their inverses) we have that:

$$E_n = \left\langle\phi_n(x)\left|\hat{\boldsymbol{T}}\left(-\frac{\hbar k_y}{m\omega_C}\right)\hat{\boldsymbol{T}}^{\dagger}\left(-\frac{\hbar k_y}{m\omega_C}\right)\; \hat{\boldsymbol{H}}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}_x)\, \hat{\boldsymbol{T}}\left(-\frac{\hbar k_y}{m\omega_C}\right)\; \hat{\boldsymbol{T}}^{\dagger}\left(-\frac{\hbar k_y}{m\omega_C}\right)\right|\phi_n(x)\right\rangle$$

$$E_n = \left\langle \phi_n(x + \frac{\hbar k_y}{m\omega_C}) \right| \widehat{\boldsymbol{H}}(\widehat{\boldsymbol{x}} - \frac{\hbar k_y}{m\omega_C}\hat{\mathbb{I}}, \widehat{\boldsymbol{p}}_{\boldsymbol{x}}) \left| \phi_n(x + \frac{\hbar k_y}{m\omega_C}) \right\rangle$$

This is the exact form of our magnetic Hamiltonian. So we can extract the same energy eigenvalues and wavefunctions with a quick change of basis. Given that $\widehat{\boldsymbol{\xi}} = \widehat{\boldsymbol{x}} - \frac{\hbar k_y}{m\omega_C}\hat{\mathbb{I}}$, with eigenkets of $\xi = x + \frac{\hbar k_y}{m\omega_C}$ we can recover eigenvalues of $\widehat{\boldsymbol{x}}$. We also note that translation operators leave momentum unchanged such that $\widehat{\boldsymbol{p}}_{\boldsymbol{\xi}} = \widehat{\boldsymbol{p}}_{\boldsymbol{x}}$. Another minor point is that shifts along x due to k_y have the opposite sign within the Hamiltonian to that of their eigenvalue, this ensures the correct position for the translated density. A fact that becomes relevant later when relating to symmetric gauge (or from an orbit "perspective" this line density is just an infinitely elongated ellipse).

$$E = \hbar\omega_C\left(n + \frac{1}{2}\right), \psi_{E_n,k_y}\left(x + \frac{\hbar k_y}{m\omega_C}, y\right) = \psi_{k_y}(y) * \phi_{E_n}(\xi) = \frac{\pi^{-\frac{1}{4}}}{\sqrt{2^n n!}} e^{-\frac{\xi^2}{2} + \mathbb{i}(k_y)y} H_n(\xi)$$

Here the x-axis is governed by the physicists Hermite polynomials H_n and the y axis is a free particle with well-defined momentum. We plot the first few wavefunctions noting that the degeneracy only changes the position along the x axis (as expected through choice of k_y):

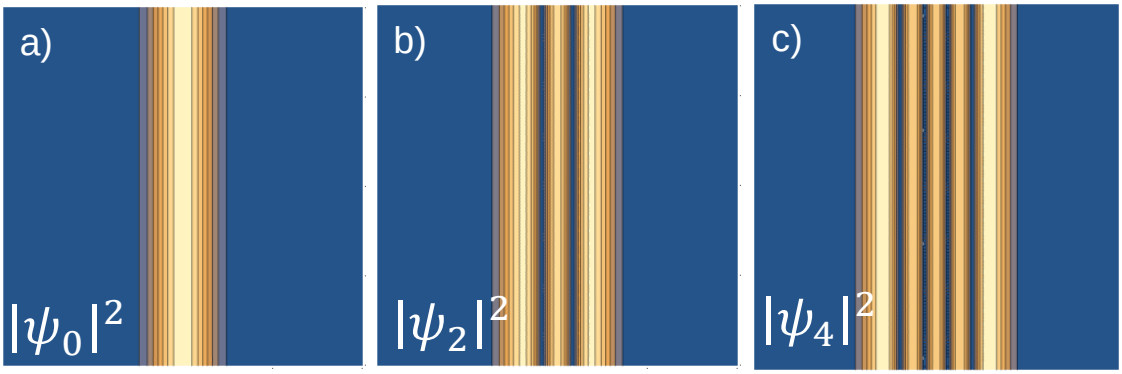

Figure A1.1: The First Few Linear Landau Wavefunctions

A few of the wavefuncitons that occur for a Landau system in linear gauge a) 0th state b) 2nd state c) 4th state

When represented through the Landau gauge, solutions lie parallel with a linear axis. Therefore this gauge should be chosen for perturbation calculations where the B-field takes constant values along a line and variations only occur along a perpendicular direction. We find this geometry to match well with a linear strain profile due to fringing of the field on the sides of a parallel channel.

A1.4 Symmetric Gauge

We now choose to represent our B-field in a symmetric gauge. In this gauge we have that A varies along both the x and y axes as follows:

$$\vec{A} = \frac{1}{2}\begin{pmatrix} -B_0 y \\ B_0 x \\ 0 \end{pmatrix}, \nabla \mathrm{x} \vec{A} = \det\begin{pmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ \partial_x & \partial_y & \partial_z \\ \frac{-B_0 y}{2} & \frac{B_0 x}{2} & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B_0 \end{pmatrix}$$

Using this gauge we need a more complicated Hamiltonian due to the presence of components along both axes:

$$\widehat{\boldsymbol{H}} = \frac{1}{2m}\left(\widehat{\boldsymbol{p}} - q\widehat{\boldsymbol{A}}\right)^2 = \frac{1}{2m}\left(\left(\widehat{\boldsymbol{p}}_x + \frac{qB_0}{2}\widehat{\boldsymbol{y}}\right)^2 + \left(\widehat{\boldsymbol{p}}_y - \frac{qB_0}{2}\widehat{\boldsymbol{x}}\right)^2\right)$$

However with a clever commutator trick we can factorize this as a harmonic oscillator just as we saw in the linear case. To start to tackle this we write out the general form of the commutator $[\widehat{\boldsymbol{\pi}}_x, \widehat{\boldsymbol{\pi}}_y]$

$$\left[\widehat{\boldsymbol{\pi}}_x, \widehat{\boldsymbol{\pi}}_y\right] = \widehat{\boldsymbol{\pi}}_x\widehat{\boldsymbol{\pi}}_y - \widehat{\boldsymbol{\pi}}_y\widehat{\boldsymbol{\pi}}_x = \left[\left(\widehat{\boldsymbol{p}}_x - qA_x(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})\right), \left(\widehat{\boldsymbol{p}}_y - qA_y(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})\right)\right]$$

Simplifing this further using commutator relations we must now calculate the following commutations:

$$\left[\widehat{\boldsymbol{\pi}}_x, \widehat{\boldsymbol{\pi}}_y\right] = \cancel{\left[\widehat{\boldsymbol{p}}_x, \widehat{\boldsymbol{p}}_y\right]} - q\left[\widehat{\boldsymbol{p}}_x, A_y(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})\right] + q\left[\widehat{\boldsymbol{p}}_y, A_x(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})\right] + q^2\cancel{\left[A_x(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}), A_y(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})\right]}$$

We see that the first and last terms vanish because the spatial axes are independent of one another. The middle terms however do not commute with one another in general because each A field can contain position operators along all axes. We have additionally reversed the order of the $\left[A_x(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}}), \widehat{\boldsymbol{p}}_y\right]$ commutator so that it takes the same form as the x-axis commutation with an additional negative sign.

Due to the formalism of momentum operators in the position representation $\widehat{\boldsymbol{p}}_x = -\mathbb{i}\hbar\partial_x$ it is easy enough to prove that the commutator of momentum $\widehat{\boldsymbol{p}}_x$ with a general function of position operators $f(\widehat{\boldsymbol{x}}, \widehat{\boldsymbol{y}})$. Our result will simply return the application of a partial derivative. This is best understood by using the chain rule through application to an arbitrary wavefunction:

$$[\hat{\boldsymbol{p}}_x, f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})]\,\psi = (\hat{\boldsymbol{p}}_x\, f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) - f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\,\hat{\boldsymbol{p}}_x)\,\psi = -\,\mathbb{i}\hbar\left(\frac{\partial}{\partial x}(f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}) * \psi) - f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\frac{\partial}{\partial x}(\psi)\right)$$

Working out the derivatives we obtain the following:

$$[\hat{\boldsymbol{p}}_x, f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})]\,\psi = -\,\mathbb{i}\hbar\left(\frac{\partial}{\partial x}\big(f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\big) * \psi + \cancel{f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\frac{\partial}{\partial x}(\psi)} - \cancel{f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\frac{\partial}{\partial x}(\psi)}\right)$$

Dropping ψ we have that:

$$[\hat{\boldsymbol{p}}_x, f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})] = -\,\mathbb{i}\hbar\frac{\partial}{\partial x}\big(f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\big)$$

Hence we can now write out the commutator of canonical momentum $[\hat{\boldsymbol{\pi}}_x, \hat{\boldsymbol{\pi}}_y]$ as:

$$[\hat{\boldsymbol{\pi}}_x, \hat{\boldsymbol{\pi}}_y] = [\hat{\boldsymbol{p}}_x, A_y(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})] - [\hat{\boldsymbol{p}}_y, A_x(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})] = \mathbb{i}\hbar q\left(\frac{\partial}{\partial x}\big(A_y(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\big) - \frac{\partial}{\partial y}\big(A_x(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\big)\right) = \mathbb{i}\hbar q\big(\nabla \mathrm{x}\vec{\boldsymbol{A}}\big)_z$$

Or more concisely using the definition of the B-field

$$[\hat{\boldsymbol{\pi}}_x, \hat{\boldsymbol{\pi}}_y] = \mathbb{i}\hbar q B_z(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$$

We have calculated this generally because it will become useful for perturbation calculations, however in the case of a uniform symmetric gauge this reduces further to the following:

$$[\hat{\boldsymbol{\pi}}_x, \hat{\boldsymbol{\pi}}_y] = \mathbb{i}\hbar q B_0 = \mathbb{i}\hbar m\omega$$

With knowledge of the final solution we again condense the B field into an angular frequency $= \frac{qB_0}{m}$, based on the classical cyclotron orbit. From here we are able to factorize the Hamiltonian to collect terms in the form of a harmonic oscillator.

$$\hat{\boldsymbol{H}} = \frac{1}{2m}\big(\hat{\boldsymbol{\pi}}_x^2 + \hat{\boldsymbol{\pi}}_y^2\big) = \frac{1}{2m}\Big(\big(\hat{\boldsymbol{\pi}}_x - \mathbb{i}\hat{\boldsymbol{\pi}}_y\big)\big(\hat{\boldsymbol{\pi}}_x + \mathbb{i}\hat{\boldsymbol{\pi}}_y\big) - \mathbb{i}[\hat{\boldsymbol{\pi}}_x, \hat{\boldsymbol{\pi}}_y]\Big)$$

Upon factorization we pick up an additional commutator term which is necessary to cancel the cross terms arising from the multiplication of the canonical momentums. We plug in the definition of this commutator as calculated above.

$$\widehat{\boldsymbol{H}} = \frac{1}{2m}\left(\widehat{\boldsymbol{\pi}}_x^2 + \widehat{\boldsymbol{\pi}}_y^2\right) = \frac{1}{2m}\left(\left(\widehat{\boldsymbol{\pi}}_x - \mathbb{i}\widehat{\boldsymbol{\pi}}_y\right)\left(\widehat{\boldsymbol{\pi}}_x + \mathbb{i}\widehat{\boldsymbol{\pi}}_y\right) + \hbar m\omega\right)$$

Proceeding we look at the commutation relation of our factorized operators

$$\left[\left(\widehat{\boldsymbol{\pi}}_x + \mathbb{i}\widehat{\boldsymbol{\pi}}_y\right), \left(\widehat{\boldsymbol{\pi}}_x - \mathbb{i}\widehat{\boldsymbol{\pi}}_y\right)\right] = \cancel{\left[\widehat{\boldsymbol{\pi}}_y, \widehat{\boldsymbol{\pi}}_y\right]} - \mathbb{i}\left[\widehat{\boldsymbol{\pi}}_x, \widehat{\boldsymbol{\pi}}_y\right] + \mathbb{i}\left[\widehat{\boldsymbol{\pi}}_y, \widehat{\boldsymbol{\pi}}_x\right] + \cancel{\left[\widehat{\boldsymbol{\pi}}_x, \widehat{\boldsymbol{\pi}}_x\right]} = -2\mathbb{i}\left[\widehat{\boldsymbol{\pi}}_x, \widehat{\boldsymbol{\pi}}_y\right] = 2\hbar m\omega$$

Because this commutator results in a constant we can redefine the operators in a characteristic raising/lowering operator form, as is common to all quantum harmonic oscillators:

$$\widehat{\boldsymbol{b}}^\dagger \equiv \sqrt{\frac{1}{2\hbar m\omega}}\left(\widehat{\boldsymbol{\pi}}_x - \mathbb{i}\widehat{\boldsymbol{\pi}}_y\right), \widehat{\boldsymbol{b}} \equiv \sqrt{\frac{1}{2\hbar m\omega}}\left(\widehat{\boldsymbol{\pi}}_x + \mathbb{i}\widehat{\boldsymbol{\pi}}_y\right)$$

Where now we have that $[\widehat{\boldsymbol{b}}^\dagger, \widehat{\boldsymbol{b}}] = 1$, rearranging constants we have converted our symmetric gauge to a harmonic form.

$$\widehat{\boldsymbol{H}} = \frac{1}{2m}\left(2\hbar m\omega * \widehat{\boldsymbol{b}}^\dagger\widehat{\boldsymbol{b}} + \hbar m\omega\right) = \hbar\omega\left(\widehat{\boldsymbol{b}}^\dagger\widehat{\boldsymbol{b}} + \frac{1}{2}\right)$$

A1.5 Angular Momentum and Degeneracy

Earlier we noted that our choice of gauge field $\widehat{\boldsymbol{A}}$ will effectively choose the quantum number of degeneracy. Since the Hamiltonian has both $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ dependence we must look for a momentum that treats them symmetrically to fit our choice of gauge. Therefore following further analysis we will find that the good quantum number will be eigenstates of the $\hat{\boldsymbol{L}}_{\boldsymbol{z}}$ operator where $\hat{\boldsymbol{L}}_{\boldsymbol{z}} \equiv \hat{\boldsymbol{x}}\,\hat{\boldsymbol{p}}_y - \hat{\boldsymbol{y}}\,\hat{\boldsymbol{p}}_x$.

To understand this more generally however we must return to the Lagrangian as represented by quantum operators and calculate the operator of total angular momentum $\hat{\boldsymbol{J}}$. As is apparent from classical mechanics we have that the total angular momentum is $\hat{\boldsymbol{J}} \equiv \sum_i \frac{\partial \mathcal{L}}{\partial \dot{\theta}_i}$. Since we are working in cylindrical coordinates, we only have one angular derivative to enact namely:

$$\frac{\partial}{\partial \dot{\theta}} \equiv -y\frac{\partial}{\partial \dot{x}} + x\frac{\partial}{\partial \dot{y}}$$

$$\mathcal{L} \equiv \sum_i \frac{1}{2} m \dot{\hat{x}}_i^{\,2} + q \dot{\hat{x}}_i \hat{A}_i \;\Rightarrow \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = -\hat{y}\left(m\dot{\hat{x}} + q\hat{A}_x\right) + \hat{x}\left(m\dot{\hat{y}} + q\hat{A}_y\right)$$

More succinctly we have that:

$$\hat{J} = \hat{x}\,\hat{\pi}_y - \hat{y}\,\hat{\pi}_x + q\hat{x}\hat{A}_y - q\hat{y}\hat{A}_x$$

Although it will be simpler to use the canonical form for $\hat{J}$ as given above we can use our definition of $\hat{\pi}_i$ to recover some physical intuition for this system:

$$\hat{J} = \hat{x}\,\hat{p}_y - \hat{y}\,\hat{p}_x - \cancel{q\hat{x}\hat{A}_y} + \cancel{q\hat{y}\hat{A}_x} + \cancel{q\hat{x}\hat{A}_y} - \cancel{q\hat{y}\hat{A}_x} = \hat{L}_z$$

Working cylindrical coordinates does imposes a caveat, but as long as the particle is confined to a plane we see that the canonical form will always reduce to $\hat{L}_z$ regardless of our choice of gauge field. We expect this type of dependence because $\hat{J}$ is a physical observable and therefore cannot depend on $\hat{A}$. We note that more generally, without choice of coordinates, we expect $\hat{J}$ to be an arbitrary function of all $\hat{L}_i$ operators. Since we have chosen a gauge earlier in the symmetric problem we insert that into the canonical $\hat{J}$ now:

$$\hat{J} = \hat{x}\,\hat{\pi}_y - \hat{y}\,\hat{\pi}_x + \frac{qB_0}{2}(\hat{x}^2 + \hat{y}^2)$$

With the ability to measure the angular momentum (knowing that it is conserved) we also need to devise a way to change the angular state. Looking back to the linear problem for inspiration we can make an educated choice by looking at the generators of spatial translation. These come to us in the form of magnetic translation operators. These translations must depend on the canonical momentum $\hat{\pi}$ as opposed to $\hat{p}$ directly (where previously our gauge reduced $\hat{k}_x$ to $\hat{p}_x$), a side effect of their derivation. We have that these generators are instead (without prefactors):

$$\hat{k} = \hat{p} + \hat{A}(\hat{x}, \hat{y})$$

In symmetric gauge these become:

$$\hat{k}_x = \hat{p}_x - \frac{qB_0}{2}\hat{y}\,, \hat{k}_y = \hat{p}_y + \frac{qB_0}{2}\hat{x}$$

We look for a simple relation between these translation generators and the Hamiltonian. Since these look like the canonical momenta but with the opposite sign of B_0 we write out a "Hamiltonian-like" equation:

$$\frac{\hat{\boldsymbol{k}}_x^2 + \hat{\boldsymbol{k}}_y^2}{2m} = \frac{1}{2m}\left(\left(\hat{\boldsymbol{p}}_x - \frac{qB_0}{2}\hat{\boldsymbol{y}}\right)^2 + \left(\hat{\boldsymbol{p}}_y + \frac{qB_0}{2}\hat{\boldsymbol{x}}\right)^2\right)$$

Substituting in $\hat{\boldsymbol{p}}_x = \hat{\boldsymbol{\pi}}_x - \frac{qB_0}{2}\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{p}}_y = \hat{\boldsymbol{\pi}}_y + \frac{qB_0}{2}\hat{\boldsymbol{y}}$ we can further write this as:

$$\frac{\hat{\boldsymbol{k}}_x^2 + \hat{\boldsymbol{k}}_y^2}{2m} = \frac{1}{2m}\left((\hat{\boldsymbol{\pi}}_x - qB_0\hat{\boldsymbol{y}})^2 + \left(\hat{\boldsymbol{\pi}}_y + qB_0\hat{\boldsymbol{x}}\right)^2\right)$$

$$= \frac{1}{2m}\left(\hat{\boldsymbol{\pi}}_x^2 + \hat{\boldsymbol{\pi}}_y^2 + 2qB_0\left(\hat{\boldsymbol{\pi}}_y\hat{\boldsymbol{x}} - \hat{\boldsymbol{\pi}}_x\hat{\boldsymbol{y}} + \frac{qB_0}{2}(\hat{\boldsymbol{x}}^2 + \hat{\boldsymbol{y}}^2)\right)\right)$$

But using our definition for the angular momentum in this gauge we see this reduces to:

$$\frac{\hat{\boldsymbol{k}}_x^2 + \hat{\boldsymbol{k}}_y^2}{2m} = \frac{1}{2m}\left(\hat{\boldsymbol{\pi}}_x^2 + \hat{\boldsymbol{\pi}}_y^2\right) + \omega\hat{\boldsymbol{J}} = \hat{\boldsymbol{H}} + \omega\hat{\boldsymbol{J}}$$

Our choice of starting equation has naturally developed an interesting relationship between our translation generators, energy and the angular momentum. Following a similar treatment to our Hamiltonian let us look for a raising/lowering form. First let us calculate the commutators of $[\hat{\boldsymbol{k}}_x, \hat{\boldsymbol{k}}_y]$:

$$\left[\hat{\boldsymbol{k}}_x, \hat{\boldsymbol{k}}_y\right] = \left[\hat{\boldsymbol{p}}_x + q\hat{\boldsymbol{A}}_x, \hat{\boldsymbol{p}}_y + q\hat{\boldsymbol{A}}_y\right] = \cancel{\left[\hat{\boldsymbol{p}}_x, \hat{\boldsymbol{p}}_y\right]} + \left[\hat{\boldsymbol{p}}_x, q\hat{\boldsymbol{A}}_y\right] - \left[\hat{\boldsymbol{p}}_y, q\hat{\boldsymbol{A}}_x\right] + \cancel{\left[q\hat{\boldsymbol{A}}_x, q\hat{\boldsymbol{A}}_y\right]}$$

$$\left[\hat{\boldsymbol{p}}_x, q\hat{\boldsymbol{A}}_y\right] - \left[\hat{\boldsymbol{p}}_y, q\hat{\boldsymbol{A}}_x\right] = -\mathbb{i}\hbar m\omega$$

We end up seeing that these operators form their own harmonic oscillator equation and are simply the angular momentum raising and lowering operators:

$$\hat{\boldsymbol{a}}^\dagger \equiv \sqrt{\frac{1}{2\hbar m\omega}}\left(\hat{\boldsymbol{k}}_x - \mathbb{i}\hat{\boldsymbol{k}}_y\right), \hat{\boldsymbol{a}} \equiv \sqrt{\frac{1}{2\hbar m\omega}}\left(\hat{\boldsymbol{k}}_x + \mathbb{i}\hat{\boldsymbol{k}}_y\right)$$

A1.6 Symmetric Gauge Solutions

We now have a new formulation for energy using $\hat{\boldsymbol{J}}$ and the angular momentum raising and lowering operators:

$$\hat{\boldsymbol{H}} = \hbar\omega\left(\hat{\boldsymbol{a}}^{\dagger}\hat{\boldsymbol{a}} + \frac{1}{2}\right) - \omega\hat{\boldsymbol{J}}$$

And since $\hat{\boldsymbol{a}}^{\dagger}$commutes with this Hamiltonian we see that changing the angular momentum has no effect on the energy. Therefore in this gauge our Landau levels look like the following:

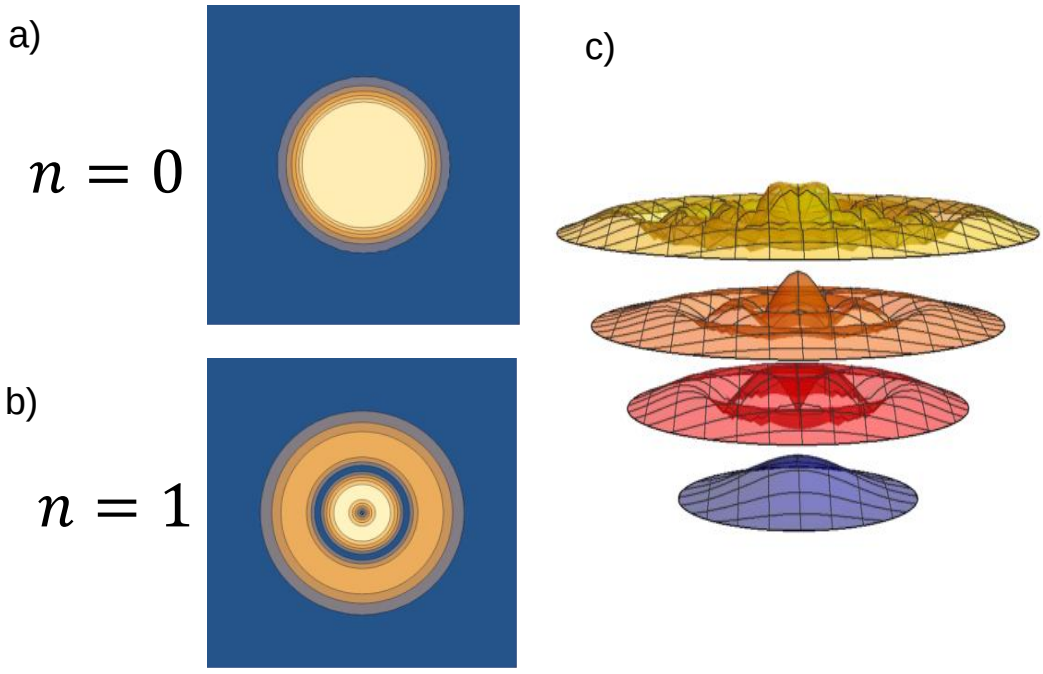

Figure A1.2: Landau Levels in Symmetric Gauge

a) The 0th state of the Landau system in symmetric gauge b) The 1st energetic state of the symmetric gauge c) A three dimensional plot of the first four landau confinement levels

As a result of this appendix, we can create a method for dealing with strain fields, the most apparent connection being a strong correlation with the harmonic oscillators which are observed in experiment. Results this direction are as of yet inconclusive however the background understanding will help press future works regarding the connections between strain, tight binding and Landau levels in TMD materials.

Appendix A2

Distortion Detection Algorithm

A2.1 Phase Matching a Lattice in 1D

We will show an easy way to define the phase match of the system with a single variable, M[ϕ]=

$$\int_{\phi-\frac{\pi}{k}}^{\phi+\frac{\pi}{k}} F[\theta] * G[\theta]\, d\theta$$

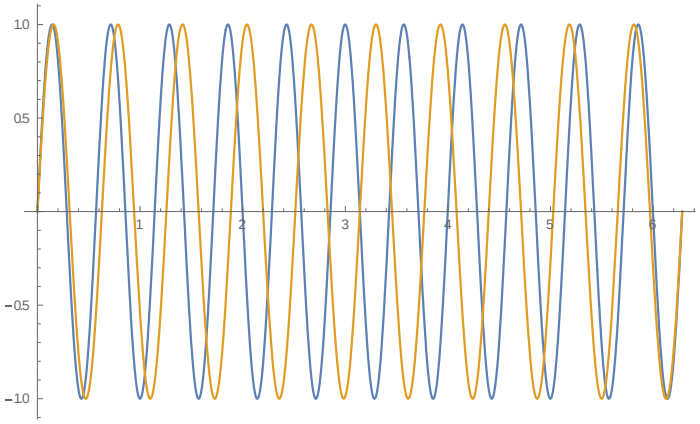

Figure A2.1: Beat Frequency of Two Cosine Waves

The phase difference between the two waves above can be represented using the function M[ϕ] noting how this this function behaves under varying phase conditions. We first comment that M[ϕ] is a sliding integrand, testing the phase matching for a single period of a wave at a specified position. Therefore if we understand how a single period behaves it will give us an understanding for all of space. Here we present some special cases:

$$\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} F[\theta]^2\, d\theta = 1, \quad \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} -F[\theta]^2\, d\theta = -1, \quad \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} F[\theta] * F\left[\theta \pm \frac{\pi}{2}\right] d\theta = 0$$

Here F[θ] is a basic sine or cosine with period (2Pi)/k. We have that when F[θ] is in perfect phase our integrand M[ϕ] is identically 1, out of phase -1, and for a 90 degree shift M = 0.

Therefore for the wave above we simply compute the functional using M[ϕ] as our metric for the corresponding phase shifts as a function of space.

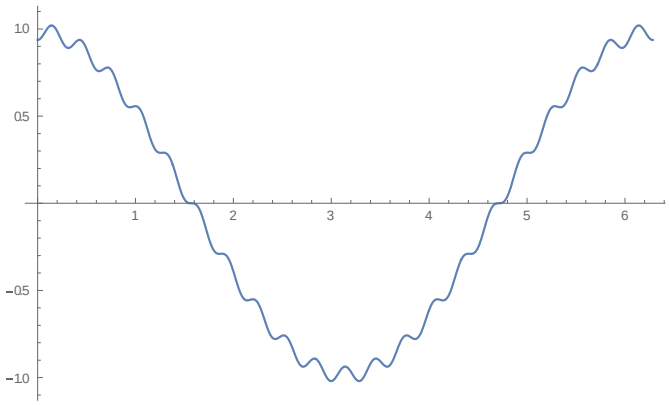

Figure A2.2: Phase Difference between Two Cosines with a Beat Frequency

This returns exactly what we expect when the waves are synchronized we return 1 (at roughly 0 and 2π) and when perfectly mismatched -1 (at roughly π). Because this represents the phase mismatching of two waves occurring with a beat frequency we also extract that the region plotted is a single period of the resulting beat between our competing waves. Additionally we note that the smaller ripples seen in our phase match arise from the fact that our two waves never perfectly match one another beacuse there is no point over a full period where F[k*x]=F[(k+1)*x] indentically.

A2.2 Phase Matching a 1D Soliton

Now we calculate the consequences for the phase difference arising from a 1D soliton. Consider a waveform that changes in a piecewise fashion such that it's position and derivatives match at the stitching point. When we suture these waveforms together we form a function as follows, and we apply our phase match procedure:

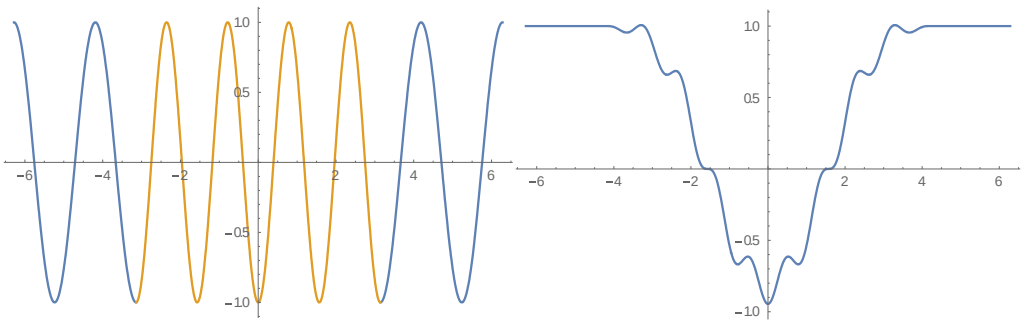

Figure A2.3: Phase Signature of a 1D Soliton

We see exactly now that over the soliton length scale the phase mismatches, then returns to perfectly phase matched. We additionally can create a partial solition (creating a 2π/(k+1) phase difference) and see the effects:

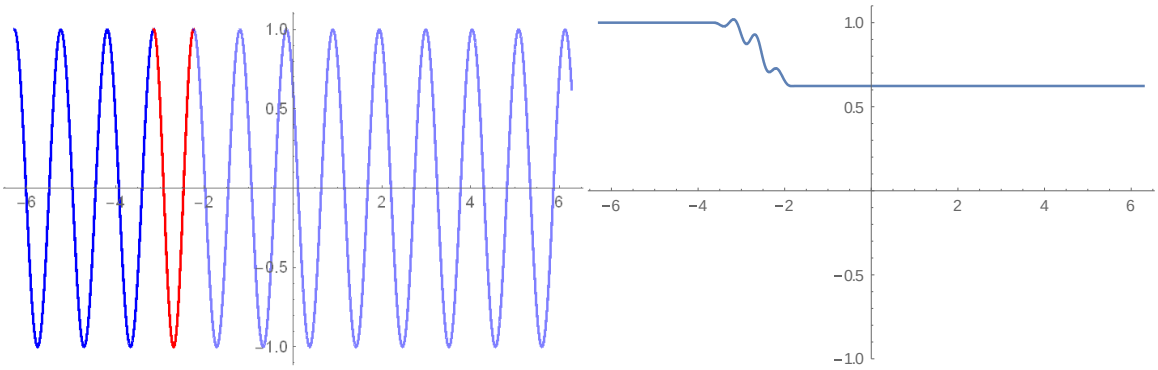

Figure A2.4: Phase Signature of a Partial Soliton

Because after the soliton we return to a wave with a phase mismatch from the original, the phase metric we've created M[ϕ] detects the change by returning a value other than 1. This soliton has caused a lattice mismatch of 2π/(k+1) which we graphically indicate below:

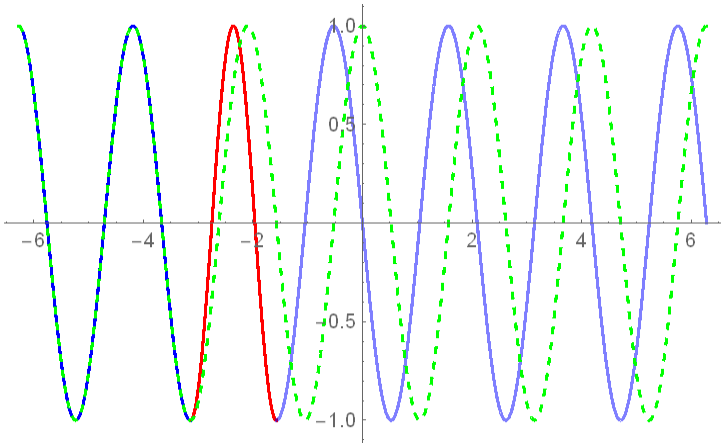

Figure A2.5: Graphical Representation of Phase Mismatch after Partial Soliton

As an additional note we observe Gibbs phenomenon due the discontinuity we introduced into our waveform. This an expected result arising in any Fourier technique, hence our phase matching metric will overshoot if it encounters a feature that causes a discrete change in periodicity.

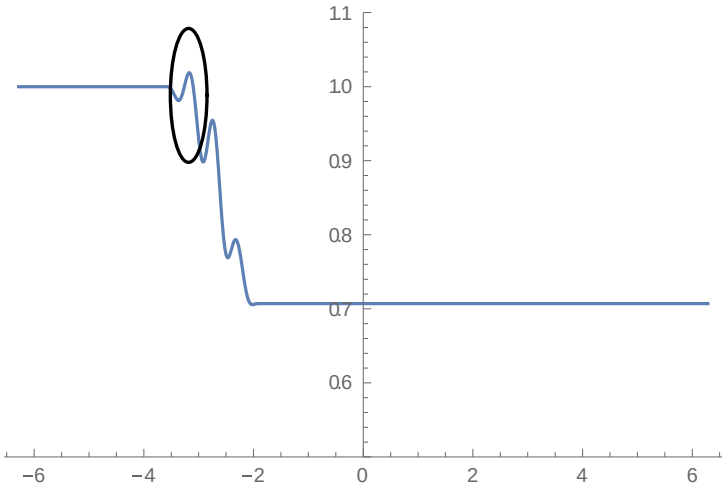

Figure A2.6: Gibb's Phenomenon Occurring in Phase Signature

A2.3 Phase Matching in 2D Using a Square Lattice

Now we need a way to generalize this procedure into higher dimensions. Unlike the 1D case we need a spatial wave in two dimensions. This requires us to create a periodic lattice with a simple wave representation of two components {x,y}. Therefore as is typical with lattices, we can write out principle vectors, which in the ideal case are seperable. Separating variables however requires our vectors to be orthogonal confining us to rectrangular geometries. For a square lattice we additionally know that the principle vectors should also have the same magnitude hence, we write vectors **a** = {1,0} , **b** = {0,1} which are orthonormal. This lets us fomulate our lattice as the sum of trigonometric functions, with additional parameter k for periodicity and ϕ for lattice rotation:

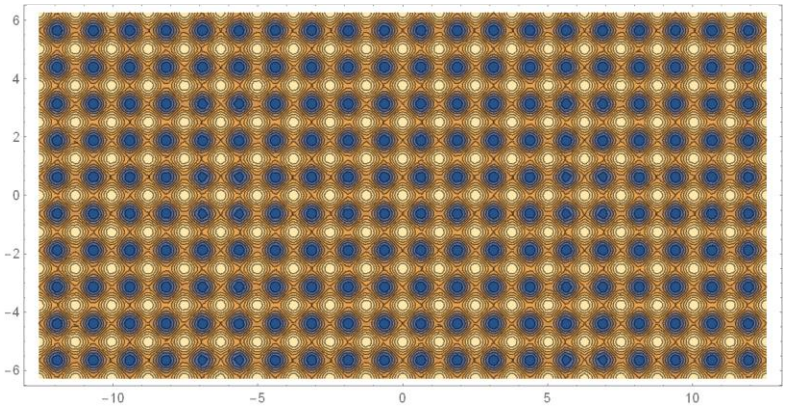

Figure A2.7: A Perfect 2D Square Lattice

Our integrand now must become an area integral over the region we wish to phase match, M[x,y] = ∯ F[ξ,ψ] G[ξ,ψ] dξ dψ . However this doesn't come without additional complications we must ensure that the area we integrate over contains a single periodicity in each direction. Therefore we must search for the smallest possible region which satisfies this constraint:

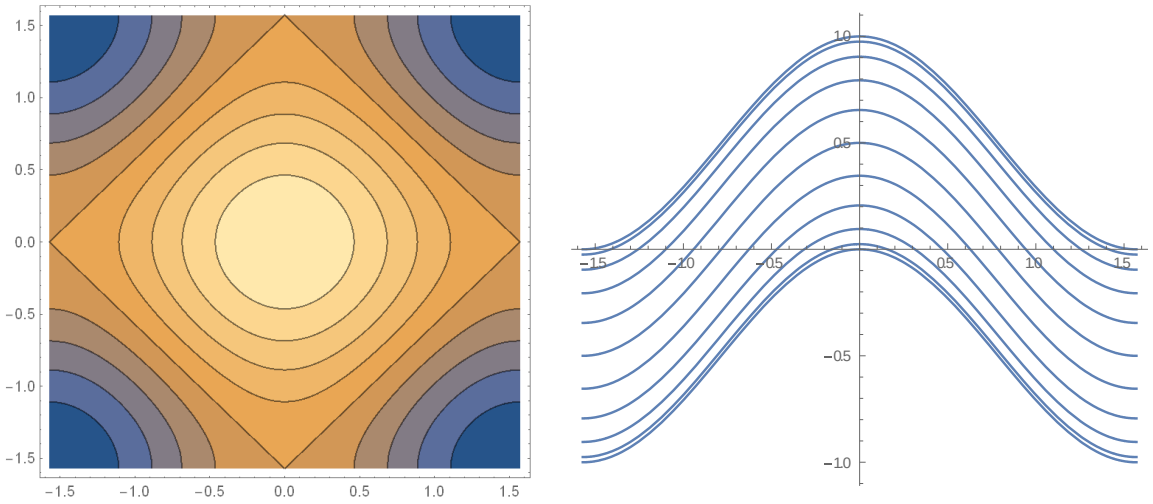

Figure A2.8: Integration Region with a Complete Period of the 2D Cosine Lattice

Here a slice along the x and y directions shows that we contain a single full period of a cosine wave, A+B*Cos[kx] hence this is the smallest acceptable integration region.

We find that the the smallest possible area in the case of our square lattice is simply a square which goes from -Pi/k to Pi/k along both axes, hence we have M[x,y] = $\int_{x-\mathrm{Pi}/k}^{x+\mathrm{Pi}/k} \int_{y-\mathrm{Pi}/k}^{y+\mathrm{Pi}/k} F[\xi, \psi]\, G[\xi, \psi]\, d\xi\, d\psi$

This inegration area comes with a few stipulations:

1. The integration remains consistent for all translations of the lattice represented by F[x,y]
2. Rotations require an additional rotation of the integration area
3. For latticies with different periodicty along the x and y axes the periodicity k must be scaled with the given direction {kx, ky}

Now we demonstrate proof of the statements given above:

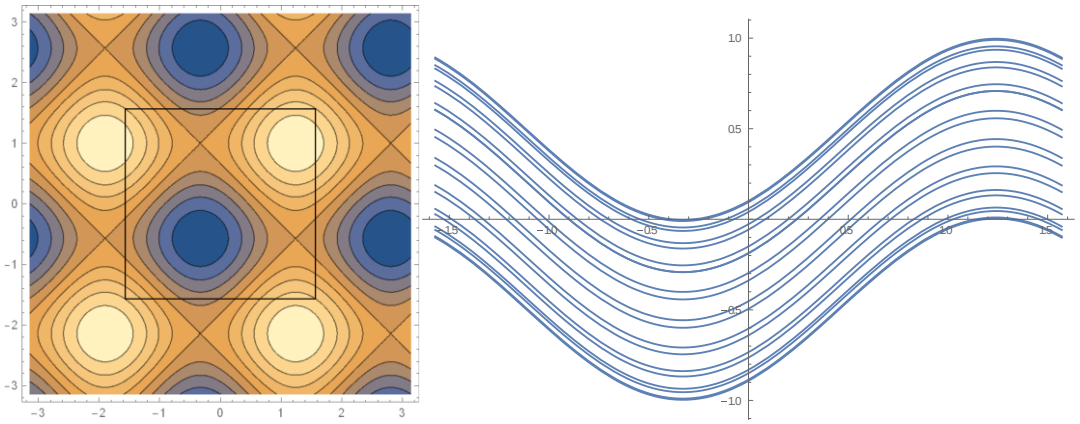

Figure A2.9: Incomplete Period for Un-rotated Integration Region

For a translation, we see that over the same region presented we still contain a single waveform, although it is not in an exact sine or cosine phase. Within our region however this has no impact on the resulting integral. Hence we have that for any arbitrary translation of the lattice our integration region remains accurate. As a corolary, this is the same as noting that F[x,y] matches itself for all of space.

$$M[x,y] = \int_{x-Pi/k}^{x+Pi/k} \int_{y-Pi/k}^{y+Pi/k} F[\xi, \psi]\, F[\xi, \psi]\, d\xi\, d\psi = 1 \text{ , for any choice of } \{x,y\}$$

We now use this translational property to show that for a rotated lattice the same integration region is not adequate:

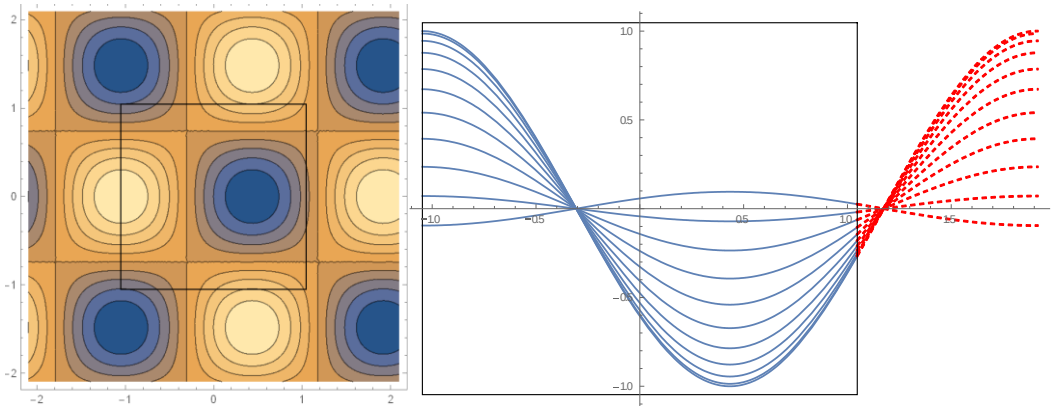

Figure A2.10: Demonstration of a New Complete Integration Region

Above we present the previous integration region for a π/4 (45 degree) shifted lattice. We can see that along the x and y axis the periodicity now matches the diagonal of our original lattice. Hence we must choose a different region for integration to contain a full period. While it is acceptable to increase the size of the region to match the rotated size (extending to the red dashed curves) it will not the smallest

indivisible region of choice. Therefore we perform a rotation on our integration region as follows. Now we see that for any slice along our new integration variables, as represented by the sliding red line, we have a single period of the waveform as in the un-rotated case. Therefore in the original Cartesian coordinates we find that our second integrand now is dependent on the limits defined for the first. For the π/4 rotation our integrand is now:

$$M[x,y] = \int_{x-\sqrt{2}*\frac{Pi}{k}}^{x}\int_{y-\xi+\sqrt{2}*\frac{Pi}{k}}^{y+\xi+\sqrt{2}*\frac{Pi}{k}} F[\xi,\psi]\,F[\xi,\psi]\,d\xi\,d\psi + \int_{x}^{x+\sqrt{2}*\frac{Pi}{k}}\int_{y+\xi-\sqrt{2}*\frac{Pi}{k}}^{y-\xi+\sqrt{2}*\frac{Pi}{k}} F[\xi,\psi]\,F[\xi,\psi]\,d\xi\,d\psi =$$

1

For a general rotation therefore we must describe our region with 3 piecewise integrations. Lastly we turn our attention to the third point, looking at a rectangular region with two periodicities For the rectangular boxed region we find that our phase match integrand is correct, along each integration axis we contain a full period of the lattice, despite kx and ky being different. The only major change is in the integration limits and the normalization factor. We now must normalize by the boxed region (kx/Pi)*(ky/Pi)

A2.4 Phase Matching a 2D Square Soliton

To prove the effectiveness of our phase match metric M[x,y] we create an artificial soliton for detection:

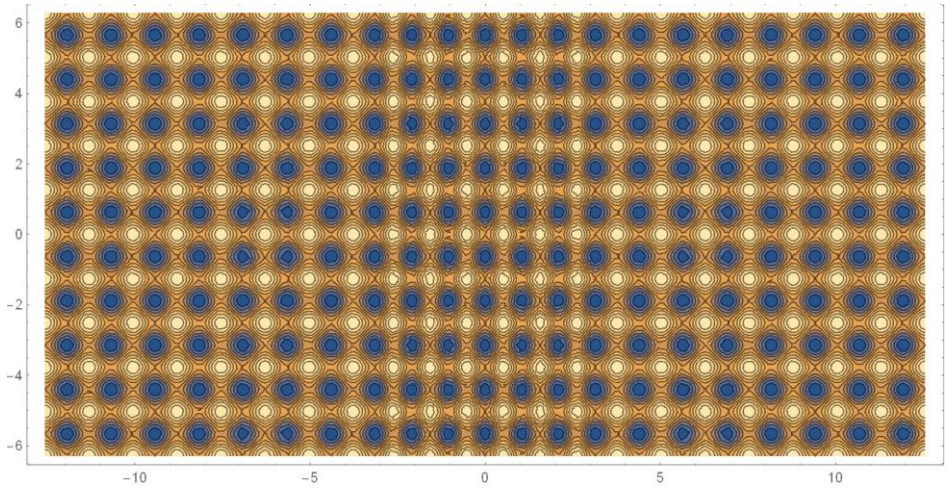

Figure A2.11: Square Lattice with Inclusion of Soliton

Using our phase metric we determine that the soliton creates a phase slip along a single axis, this

corresponds to a 180 phase shift along the x-axis while perfectly matched along y. Hence we see that our metric averages these correponding to M[x,y] = 0. If we instead had a slip along both axes this would correspond with M[x,y] = -1.

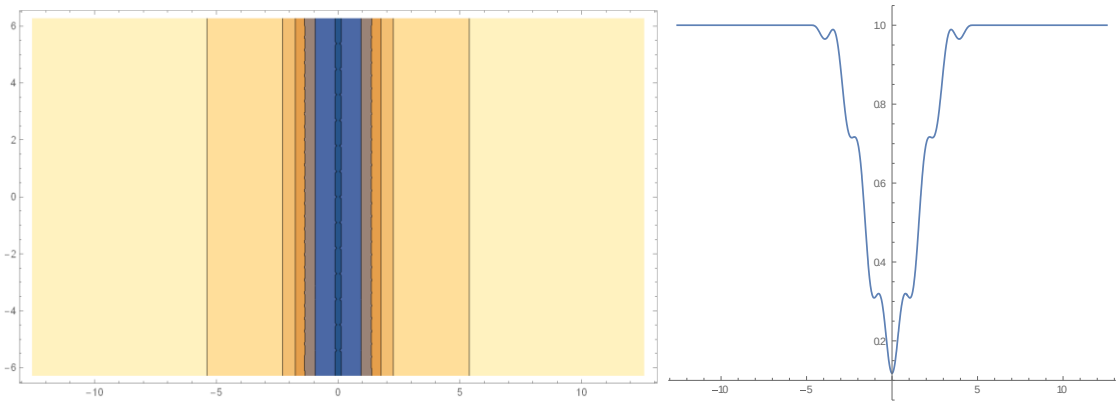

Figure A2.12: Lattice Integration Used to Detect Distortion

A2.5 Phase Matching a 2D Hexagonal Lattice

To exactly match the experiment we must now write a phase match procedure for a hexagonal lattice. The implicit complication lies in separating the three vectors that describe our space; now the third lattice vector is linearly dependent on the principle two we use to define our axes. Dealing with this problem we write instead three principle directions and attempt to classify them instead using their symmetries to our advantage.

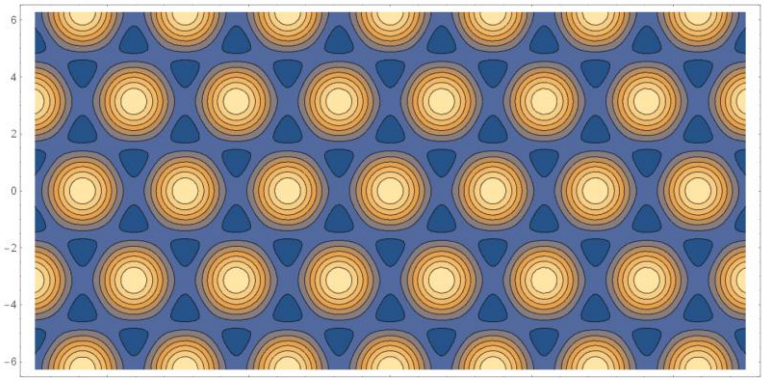

Figure A2.13: A Pristine Hexagonal Lattice

Now that we have a lattice what is the smallest indivisible area that we can phase match, while keeping all directions equivalent?

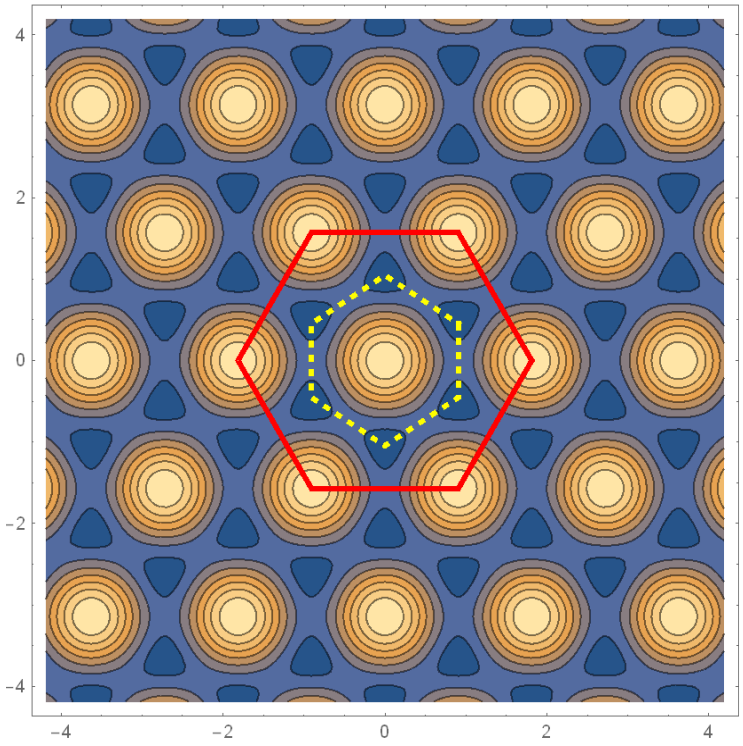

Figure A2.14: Intuitive Choices for Integration Region

The area to integrate over looks slightly counter intuitive and is represented by the solid red region. Perhaps we could take a smaller slice just around the atom (in dashed yellow)? This however is not a complete phase of the lattice waveform. We demonstrate this by taking a slice along the y axis, where our choice becomes much more apparent:

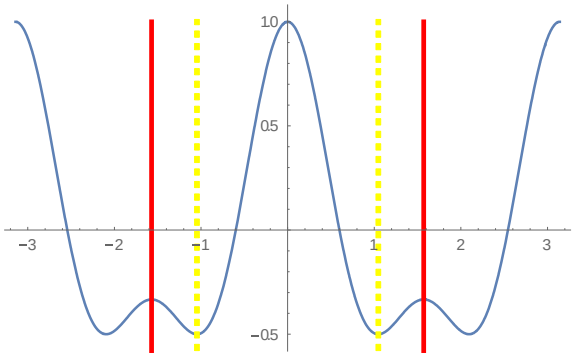

Figure A2.15: Demonstration of Correct Integration for Full Period

If we chose the yellow region only we would end up neglecting contributions from the bridge between neighboring atoms. This causes complications when translating the integration area. Our red region provides the desired result, equivalence between three lattice directions and an integration region

which is invariant under translation. As is apparent by symmetry each direction is composed of an identical set of waveforms. We also note that there are regions which are only accessible to a single lattice vector, these were the regions that were omitted in our smaller integration region. From the perspective where we omit the third vector (as dependent) these are only accessible to the difference vector between the two remaining lattice directions. Before moving on to the hexagonal soliton we simply note that the three proofs we gave earlier still hold for this lattice:

1) Invariance to Translation

2) Lattice Rotation = Integration Region Rotation

3) Deformation = Integration Region Deformation

A2.6 Phase Matching a 2D Hexagonal Soliton

Let's create a strained lattice by deforming along the x-axis, now we'll stitch this into a normal lattice to create a soliton structure

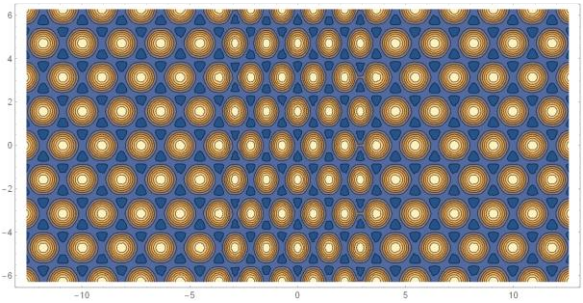

Figure A2.16: Generated Hexagonal Linear Soliton

Completing our calculation, we now apply our phase match metric M[x,y] to our hexagonal soliton.

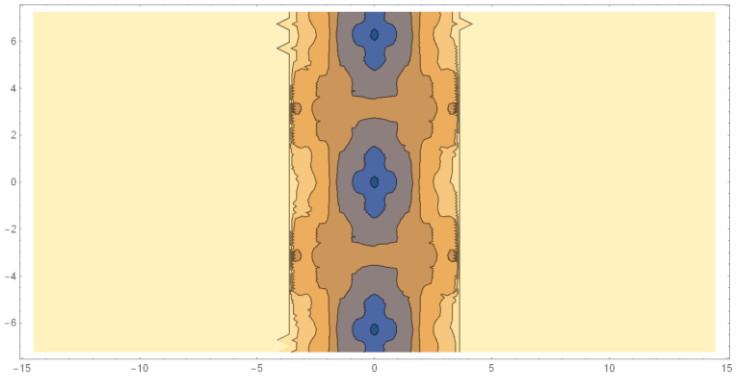

Figure A2.17: Distortion Analysis of Generated Hexagonal Soliton

This results in a somewhat different pattern than in the case of the square lattice. We see that the addition of dependent vectors lends phase structure to the inside of the soliton. This looks much more like a moire pattern confined to the soliton itself, corresponding with variations from the other principle vectors also share the x-axis as a component.

A2.7 Discrete Matching a Soliton –Y Junction

A soliton junction entails a much more interesting system however these have no simple analytical form. Our objective therefore is to transform our continous equations into a discrete numeric form (with additional use for parsing data). We start by simply producing the soliton lattice from vectorized positions of atomic centers.

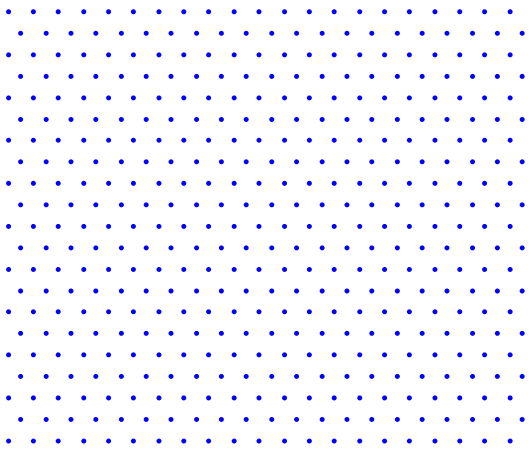

Figure A2.18: Generated Discrete Triangular Lattice

Start a Perfect Lattice (for proof of principle):

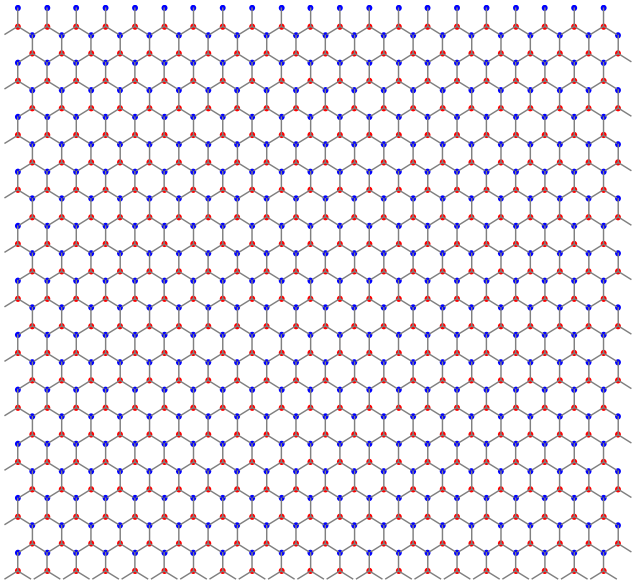

Figure A2.19: Generated Full $MoSe_2$ Lattice

A single soliton junction, where we show the first two layers, top (blue) with soliton and bottom (red) without soliton. This is as we would expect on the crystal itself:

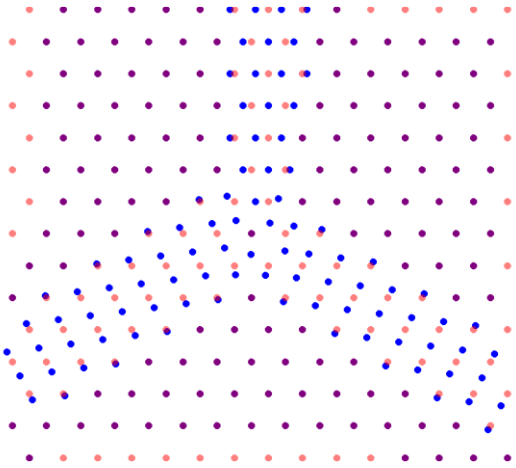

Figure A2.20: Soliton Junction Discrete Model

Now let's turn this into something that we can sample for a pixel density, as opposed to a discrete list of points. We'll try putting a gaussian at the center of each atomic position:

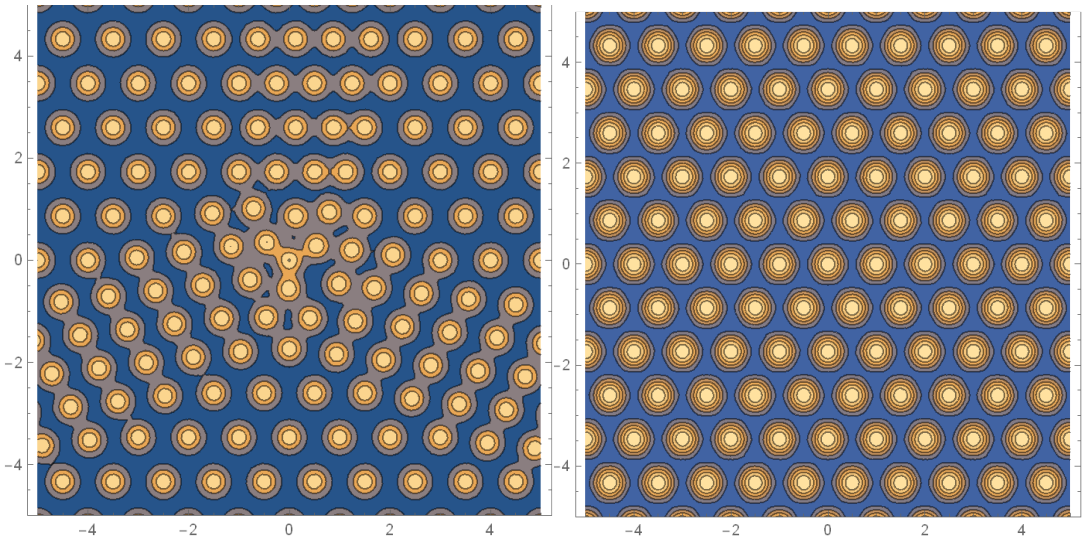

Figure A2.21: Comparison of Soliton Model and Perfect Lattice

For the perfect lattice we have no need to compromise so we can use a waveform to describe it over all space. Now we discretize the integration area.

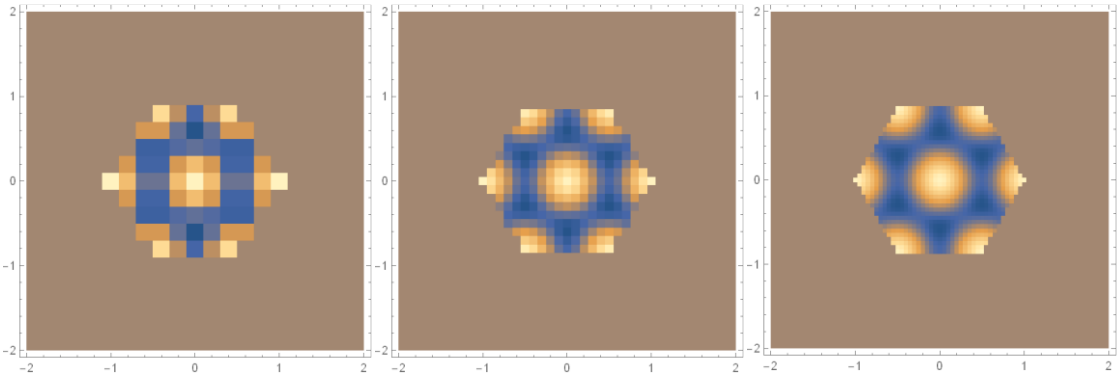

Figure A2.22: Discretized Integration Filter for Lattice Distortion Measurements

With even as low as 10 pixels from atom to atom the accuracy of the phase match metric M is ± 0.025. This is acceptable error for the phase we wish to detect (difference between 0 and 1). Now we need to sweep the integration area over all of the space containing the soliton junction. Presto, we have a phase matched junction:

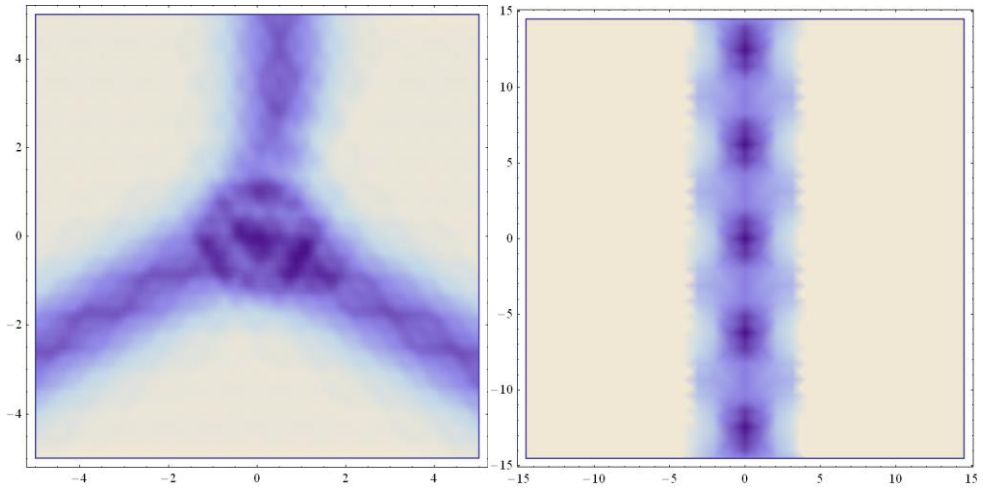

Figure A2.23: Distortion Map of a Junciton and Linear Soliton