Spring 1-1-2013

Scanning Tunneling Microscopy Studies of Layered Correlated Electron Systems: Transition Metal Dichalcogenides and Iridates

Jixia Dai
University of Colorado at Boulder, daij@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/phys_gradetds

Recommended Citation
https://scholar.colorado.edu/phys_gradetds/84

This Dissertation is brought to you for free and open access by Physics at CU Scholar. It has been accepted for inclusion in Physics Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Scanning Tunneling Microscopy Studies of Layered Correlated Electron Systems: Transition Metal Dichalcogenides and Iridates

by

Jixia Dai

B.S., University of Science and Technology of China, 2006
M.S., University of Colorado at Boulder, 2011

A thesis submitted to the Faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Physics

2013
This thesis entitled:
Scanning Tunneling Microscopy Studies of Layered Correlated Electron Systems:
Transition Metal Dichalcogenides and Iridates
written by Jixia Dai
has been approved for the Department of Physics

________________________
Asst. Prof. Kyle McElroy

________________________
Prof. Daniel Dessau

Date ____________________

The final copy of this thesis has been examined by the signatories, and we find that both
the content and the form meet acceptable presentation standards of scholarly work in the
above mentioned discipline.
Dai, Jixia (Ph.D., Physics)

Scanning Tunneling Microscopy Studies of Layered Correlated Electron Systems:

Transition Metal Dichalcogenides and Iridates

Thesis directed by Asst. Prof. Kyle McElroy

Strongly correlated electron problems, including high temperature superconductivity, colossal magneto-resistivity, charge-density wave, heavy fermion phenomena, etc., have been major research focuses in condensed matter physics for the past several decades. Involvements of and competitions between different degrees of freedom and energy scales present considerable difficulties for understanding these problems.

In this thesis, I have used home-built variable-temperature high resolution (spatial and energy) spectroscopic imaging scanning tunneling microscopes (SI-STM) to study the topography and the local single-particle tunneling spectroscopy in a family of two-dimensional charge-density wave (2D-CDW) materials (TaSe$_2$, TaS$_2$, and NbSe$_2$) and in a newly discovered 5d Mott insulator Sr$_2$IrO$_4$ with strong spin-orbit coupling. In the 2D-CDW materials, we have shown the strong lattice distortion and weak electron modulation accompanying CDW transition, directly opposite to Peierls mechanism. We have also discovered that the 2D-CDW transitions in TaSe$_2$ and TaS$_2$ show involvement of the orbital degree of freedom, shedding light on a long standing puzzle with their electrical resistivity anomalies. In the novel 5d Mott insulator Sr$_2$IrO$_4$, we have found a large insulating gap of 620 meV, indicating a stronger-than-expected correlation effect for a 5d electron system.
Dedication

To my wife Ting Yan, our son Yian, my parents and my sister.
Acknowledgements

I firstly would like acknowledge my graduate advisor Prof. Kyle McElroy for his generous patience when teaching me about all kinds of experimental and theoretical details and for treating me as an independent researcher from the beginning. His rigorousness in experiments and analyses will affect and benefit me for life.

Thanks to Prof. Daniel Dessau for his great help with my experiments, for all of the samples he shared with us, and for his encouragements during my hard times. Thanks to Prof. Minhyea Lee for encouraging me to write my papers. Thanks to Prof. Michael Hermele and Prof. Dmitry Reznik for sharing their thoughts on my papers. Thanks to Prof. Scott Bunch for being nice and spending his precious time serving in my thesis committee. Thanks to Prof. Chuck Rogers for allowing me to use many of his equipments.

I would also like to acknowledge our sample growers for providing us their high quality single crystal samples, including Dr. Xiangde Zhu and Dr. Yuping Sun at High Magnetic Field Laboratory (Hefei, China), Dr. Thomas Wolf at Karlsruher Institut für Technologie (Germany), Dr. Helmuth Berger at Ecole Polytechnique Federal de Lausanne (Switzerland), and Prof. Gang Cao at University of Kentucky.

Thanks to my labmates Eduardo Calleja and Jacob Alldrege for helping me with my experiments and my papers, and for discussing with me about many scientific and non-scientific topics through out these many years. Thanks to Yue Cao for working closely with me and for sharing his thoughts with me about research. Thanks to Xiaqing Zhou for reading my poorly written drafts and making suggestions on them. Thanks to Zhe Sun,
Qiang Wang, Jason Gray, Long He, Ted Reber, Nick Plumb, Justin Griffith and many many others for their great encouragements and for helping out with my research.

Thanks to Tracy Buxkemper and Charlie Bowen in the Precision Instrument Shop. My experiments are not possible without their precise and beautiful work. Thanks to Sid Gustafson and Craig Joy in the Trades Teaching Lab for teaching me all of the details in the machine shop.

Lastly and most importantly, I would like to acknowledge my parents for making all of this possible. Thanks to my wife Ting Yan for understanding my pursuit in academia and for supporting me in the past five years. As a husband, I owe her too much to be described anywhere. Thanks to my son Yian, who always ‘reminds’ me of my responsibility and ‘urges’ me to work hard.
Contents

Chapter

1 Introduction to Scanning Tunneling Microscopy and Spectroscopy 1
   1.1 Tunneling phenomena ................................................. 1
   1.2 Theory of Scanning Tunneling Microscope ........................ 4
      1.2.1 Constant current topographic imaging ....................... 4
      1.2.2 Spectroscopy measurement and imaging .................... 6
   1.3 Design of the UHV-STM .............................................. 7
      1.3.1 Walker and scanner ........................................... 7
      1.3.2 Tip and sample preparation .................................. 12
      1.3.3 Electronics: pre-amplifiers and STM controller .......... 14
      1.3.4 Ultra-high Vacuum: chamber, load-lock and pumps .......... 15
      1.3.5 Cryostat ..................................................... 16
      1.3.6 Vibration isolation ........................................... 17
      1.3.7 Other experimental details .................................. 20

2 Charge-density wave in transition metal dichalcogenides 22
   2.1 Introduction of charge-density wave ............................. 22
      2.1.1 Peierls mechanism and Fermi surface nesting ................ 22
      2.1.2 Mean-field approach ........................................ 24
      2.1.3 LDOS with charge modulation ............................... 25
2.2 Charge-density waves in quasi-2D materials ........................................... 28
   2.2.1 Introduction to 2H-MX$_2$ ....................................................... 28
   2.2.2 Recent progresses and remaining issues ........................................ 32
   2.2.3 CDW gap opening ................................................................. 36
   2.2.4 Doping and superconductivity .................................................. 37
2.3 Summary ................................................................................. 38

3 Strong periodic lattice distortion in 2D charge-density wave systems .......... 40
   3.1 Introduction ................................................................. 40
   3.2 Previous STM study of topography in CDW systems ..................... 42
   3.3 Strong lattice distortion in 2H-MX$_2$ .................................... 43
   3.4 Topographic images on Cu$_x$TaS$_2$ ............................................. 52
   3.5 Summary ................................................................. 54

4 Orbital-cooperative CDW transition in 2H-MX$_2$ .................................... 56
   4.1 Introduction ................................................................. 56
   4.2 Averaged LDOS on TaSe$_2$ and TaS$_2$ .................................... 58
   4.3 Spatially resolved LDOS study on TaSe$_2$ and TaS$_2$ ................. 61
   4.4 Summary ................................................................. 65

5 Local density of states study of 5d Mott insulator Sr$_2$IrO$_4$ ..................... 66
   5.1 Introduction to Mott physics .................................................. 66
       5.1.1 Mott insulator .......................................................... 66
       5.1.2 Hubbard model .......................................................... 68
   5.2 The novel $J_{\text{eff}} = \frac{1}{2}$ Mott insulator Sr$_2$IrO$_4$ ..................... 68
       5.2.1 Spatial extensions of d orbitals ........................................ 68
       5.2.2 Spin-orbit coupling ...................................................... 69
       5.2.3 Spin-orbit Mott insulators .............................................. 70
5.3 Study of insulating gap in Sr$_2$IrO$_4$ with STS ................................. 72
5.4 Large insulating gap measured in Sr$_2$Ir$_{0.96}$Rh$_{0.04}$O$_4$ .................. 78
5.5 Study of impurity effects on insulating gap ........................................ 78
5.6 Summary ......................................................................................... 82

Appendix

A Vibration test of Room E2B37 ............................................................ 83

B Pressure vs. time in a UHV pumpdown ............................................ 85

C Feedback stability ............................................................................. 89

Bibliography ........................................................................................ 92
Tables

Table

2.1 CDW and superconductivity transition temperatures of 2H-MX$_2$ 35
Figures

Figure

1.1 Quantum tunneling of electrons. ............................................. 1
1.2 Tunneling current as integration of density of states .................. 2
1.3 Schematic diagram of a Scanning Tunneling Microscope ............... 5
1.4 An overview of the UHV-STM system .................................. 8
1.5 Solidworks drawing of the STM scan head .............................. 9
1.6 Walker piezo stack .......................................................... 10
1.7 Tubular piezo scanner ...................................................... 11
1.8 Images of tip and the etching process .................................. 12
1.9 Images of sample holder .................................................. 13
1.10 Modeling the noise of the pre-amplifier ............................... 14
1.11 Single-stage vibration isolation modeling ............................. 18
1.12 Transfer function of a single-stage isolation system ................. 19
1.13 Diagram for thermal sink with electrical isolation ................... 20

2.1 Lindhard response function for free electron systems ............... 23
2.2 Ideal Peierls transition for a 1D metallic system .................... 24
2.3 LDOS evolution with the periodicity of CDW from mean-field theory calculation. 26
2.4 Representative LDOS spectra in an electron-driven CDW. .......... 27
2.5 Topography and crystal structure of the 2H-NbSe$_2$ .................... 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6</td>
<td>Topography of the 2H-TaS$_2$ (a) and TaSe$_2$ (b)</td>
<td>30</td>
</tr>
<tr>
<td>2.7</td>
<td>Two CDW transitions in TaSe$_2$</td>
<td>31</td>
</tr>
<tr>
<td>2.8</td>
<td>Possible fermi surface nesting in 2H-TaSe$_2$</td>
<td>32</td>
</tr>
<tr>
<td>2.9</td>
<td>Phonon anomaly in 2H-NbSe$_2$</td>
<td>33</td>
</tr>
<tr>
<td>2.10</td>
<td>Conductivity enhancement with CDW transitions in 2H-MX$_2$</td>
<td>34</td>
</tr>
<tr>
<td>2.11</td>
<td>Local density of states measurement in NbSe$_2.$</td>
<td>36</td>
</tr>
<tr>
<td>2.12</td>
<td>Phase diagrams of Cu intercalation into TiSe$_2$ and TaS$_2$ (Cu$_x$TiSe$_2$ and Cu$_x$TaS$_2$)</td>
<td>37</td>
</tr>
<tr>
<td>3.1</td>
<td>Previous STM images of Ge(111) surface covered by 1/3 monolayer of Sn</td>
<td>42</td>
</tr>
<tr>
<td>3.2</td>
<td>Previous STM images of NbSe$_3$</td>
<td>43</td>
</tr>
<tr>
<td>3.3</td>
<td>Topographic image of 2H-TaS$_2$ including a CDW mismatch.</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>Topographic images showing symmetrization and antisymmetrization on 2H-TaS$_2$ at 52 K</td>
<td>47</td>
</tr>
<tr>
<td>3.5</td>
<td>Topographic images of NbSe$_2$ and TaSe$_2$ at 5 K.</td>
<td>49</td>
</tr>
<tr>
<td>3.6</td>
<td>Topographic images on 2H-NbSe$_2$ at $T = 6$ K.</td>
<td>51</td>
</tr>
<tr>
<td>3.7</td>
<td>Topographic images with various bias voltages on Cu$_x$TaS$_2$ ($x = 0.04$).</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>Band structure of TaSe$_2$ from first principle calculation.</td>
<td>57</td>
</tr>
<tr>
<td>4.2</td>
<td>LDOS measurements on TaSe$_2$ and TaS$_2$.</td>
<td>59</td>
</tr>
<tr>
<td>4.3</td>
<td>High voltage LDOS on TaSe$_2$.</td>
<td>60</td>
</tr>
<tr>
<td>4.4</td>
<td>Ratio map of TaSe$_2$ and intra-unit-cell variation</td>
<td>62</td>
</tr>
<tr>
<td>4.5</td>
<td>Local distortion and related shifts in orbital energies.</td>
<td>63</td>
</tr>
<tr>
<td>4.6</td>
<td>Ratio map and intra-unit-cell variation of LDOS</td>
<td>64</td>
</tr>
<tr>
<td>4.7</td>
<td>Ratio map at 100 mV showing the asymmetry is intrinsic.</td>
<td>65</td>
</tr>
<tr>
<td>5.1</td>
<td>Density of states change in Mott transition.</td>
<td>67</td>
</tr>
<tr>
<td>5.2</td>
<td>Spatial extensions of d orbitals</td>
<td>69</td>
</tr>
<tr>
<td>5.3</td>
<td>Splitting of 5d levels in Sr$_2$IrO$_4$</td>
<td>71</td>
</tr>
</tbody>
</table>
5.4 Topography and atomic structure of Sr$_2$IrO$_4$ .......................... 73
5.5 The 620 meV intrinsic energy gap in Sr$_2$IrO$_4$. .............................. 75
5.6 Bias dependent topography in Sr$_2$IrO$_4$. .................................. 77
5.7 Large insulating gap measured on Sr$_2$Ir$_{0.96}$Rh$_{0.04}$O$_4$. ................. 79
5.8 Tip induced band bending in Sr$_2$IrO$_4$. ...................................... 81

A.1 Vibrational test of Room E2B37 .................................................. 83

B.1 Plot of pressure versus time of the UHV-STM system. ...................... 86
B.2 Plot of pressure versus time of the UHV-STM system during normal operation. 87

C.1 STM feedback loop ................................................................. 89
C.2 Transient response of the tip .................................................... 90
1.1 Tunneling phenomena

Quantum tunneling, a classically forbidden process of electrons (particles) penetrating through energy barriers, was experimentally verified by Leo Esaki in semiconductors [1] and by Ivar Giaever in superconductors [2] respectively. Their work of single particle tunneling led towards the prediction of Cooper pair tunneling (Josephson effect) by Brian D. Josephson [3]. This same process is also the essence of nuclear alpha decay, as initially pointed out by George Gamow [4], where the microscopic particle that tunnels is the alpha particle.

![Figure 1.1: Quantum tunneling of electrons between two metals separated by a vacuum gap.](image)

Tunneling is made possible by the fact that in quantum mechanics particles are described by wave functions that are extended in space. A simplified picture of tunneling in one dimension is shown in Fig. 1.1. When the two metals are brought close to each other (3 – 10 Å for STM junctions and a few nm or more for planar junctions), the exponential
tails of wave functions overlap before point contact takes place. In the low energy limit (kinetic energy much smaller than barrier height), the electron wave functions between the two metals are linked via an exponential decaying term,

\[ |\psi_2|^2 = e^{-2\kappa z} |\psi_1|^2, \quad (1.1) \]

where \( \psi_1 \) and \( \psi_2 \) denote the electron wave function on the two sides. The characteristic decay length of the wave function is \( \kappa = \sqrt{\frac{2m\phi}{\hbar}} \simeq 0.513 \sqrt{\phi(eV)} \text{Å}^{-1} \), \( m \) is the mass of electron and \( \phi \) is the height of the vacuum barrier, which is approximately the (averaged) work function of the metals. Since work functions for most materials are in the range of \( 4 - 6 \text{ eV} \), much larger than the electron kinetic energy, it is a good approximation to treat the two sides of the junction with the same work function (\( \sim 5 \text{ eV} \)). This approximation of work function values results in a roughly unchange decay constant, \( \kappa \simeq 1 \text{ Å}^{-1} \). In terms of electron density (magnitude square of wave function), the exponential decay is around \( e^2 \simeq 7.4 \) times per Å: the tunneling current will decrease 7.4 times when the distance between two metals is increased by 1 Å.

Figure 1.2: Tunneling current as integration of density of states. The tunneling process is only possible between occupied states on one side to the unoccupied states of the other side. The difference of the Fermi levels is set by the bias voltage.
When the two sides of the junction are connected with no voltage (equal Fermi levels $E_F$), the electrons can hop back and forth due to thermal activation around $E_F$, but there is no net current. However, when a voltage $V$ is applied between the two sides, a net current can be induced. The current can be calculated using the Fermi’s golden rule [5],

$$I = \frac{4\pi e}{h} \int_{-\infty}^{+\infty} [f(E_F - eV + \epsilon) - f(E_F + \epsilon)] \rho_s(E_F - eV + \epsilon) \rho_t(E_F + \epsilon) |M|^2 d\epsilon,$$

where the $f(E)$ is the Fermi-Dirac distribution function, and $\rho_s$ and $\rho_t$ are the density of states (DOS) of the two sides. This equation is visualized in Fig. 1.2.

In Fig. 1.2, we have drawn a relatively featureless DOS on the right side of the junction, which is a normal scenario for the study in this thesis, since we usually prepare our probe tips such that there is no structure in its DOS. Furthermore, we can simplify Eqn. (1.2) by assuming that temperature is low comparing to energy scale of interest and the tunneling matrix element $M$ is energy independent (following Bardeen [6]), $M \sim e^{-\kappa z}$. With these approximations, we can reduce Eqn. (1.2) as,

$$I = \frac{4\pi e}{h} |M|^2 \rho_t(E_F) \int_{0}^{eV} \rho_s(E_F - eV + \epsilon) d\epsilon,$$

from which we can see that,

$$\frac{dI}{dV} \propto \rho_s(E_F - eV).$$

The differential conductance measurement is simply proportional to the density of states. In this thesis, I will refer to local density of states (LDOS) as $dI/dV$ measurement, but one should always be aware that they are only equal to each other with the approximations mentioned above, at best.

The density of states in superconducting lead, as measured by Giaever [2], provided strong evidence for the BCS theory of superconductivity and triggered a lot of interests in studying electronic structures with tunneling experiments. However, since planar junctions are only able to measure DOS on average, it became highly desirable to take advantage of the same effect and measure local density of states. In 1982, it was finally possible because of the invention of Scanning Tunneling Microscope by Binnig and Rohrer [7].
1.2 Theory of Scanning Tunneling Microscope

A Scanning Tunneling Microscope (STM) takes advantage of the electron tunneling effect between an atomically sharp tip and the sample under investigation. It usually consists of four basic units: coarse motion actuator, fine motion actuator, current amplifier, and electronic control unit. As shown in Fig. 1.3, during the operation of a STM, a bias voltage is firstly applied to the sample and then the metallic tip is brought close to the surface of the sample by the coarse motion actuator, and then feedback is turned on immediately when a tunneling current is detected by the control unit. The tiny current (∼ pA–nA) is detected by a transimpedance amplifier with gain at 0.1 – 10 GΩ, which converts this small tunneling current signal to a voltage.

1.2.1 Constant current topographic imaging

Due to the high sensitivity of tunneling dependence on the distance between tip and sample, it is usually beneficial to operate the STM in constant current mode, in which the tunneling current is maintained constant by a feedback mechanism. In the constant current mode, the STM controller takes the voltage signal from the current amplifier, compares it with the pre-set value and adds the difference to the output z-piezo voltage with a gain. This forms a feedback loop involving the z-piezo, tunnel junction, current amplifier, and the control unit, and assures that the tunneling current is constant. More details about the feedback system can be found in Appendix C.

While the feedback loop is active, scanning voltages can be applied on x and y-piezo to move the tip line by line across the sample surface. Since the z-piezo is controlled by the feedback loop so it closely follows the height profile of the surface, namely the topography. Recording the height profile with the x and y information, one would end up with a 2D image of the surface height under the constant tunneling current condition, namely constant current imaging.
Figure 1.3: Schematic diagram of a Scanning Tunneling Microscope. The STM controller outputs a bias voltage $V_{bias}$, five high voltage signals for controlling the tube scanner ($z$, $+x$, $-x$, $+y$, $-y$), and other voltages if necessary. A sharp metallic tip is firmly attached to the tube scanner. When the tip is a few Å away from the surface and the sample voltage $V_{bias}$ is nonzero, a tunneling current will be induced between the virtually grounded tip and the sample. This current is converted to voltage signal by a transimpedance amplifier, and sent over to the STM controller. The controller then compares the tunneling current with the set point value (set by user), and adds the error to the $z$-piezo voltage output with a gain (set by user). This feedback loop is the essence of the constant current mode. While the feedback loop is active, voltages can be sent to x/y scanner to scan the tip across the sample and the tip height information during a scan is essentially a constant current topographic image. A lock-in amplifier is used to perform spectroscopic measurements.
Since the tunneling process is only sensitive to electrons/holes with energy between $E_F$ and $E_F + eV_{\text{bias}}$ and feedback loop further constrains the tunneling current to be constant,

$$I(x, y) = I_0 e^{-z/z_0} \int_{0}^{eV} LDOS(x, y, \epsilon)d\epsilon = \text{const.},$$  \hspace{1cm} (1.5)

constant current imaging essentially reflects the integrated local density of states and morphology of the surface through $z$,

$$z(x, y) = -z_0 \ln \left( \frac{I}{I_0} \right) + z_0 \ln \left( \int_{0}^{eV} LDOS(x, y, \epsilon)d\epsilon \right),$$  \hspace{1cm} (1.6)

where $z_0$ is converted from the $\kappa$ in Eqn. (1.1), $z_0 = \frac{1}{2\kappa}$.

### 1.2.2 Spectroscopy measurement and imaging

Besides the constant current imaging mode, STM is also often operated in the spectroscopy mode. In this mode, the $xy$ scanning and the feedback loop is temporarily disabled (constant height), and the bias voltage is swept from one value to another one, while the tunneling current ($I - V$ curve) is recorded. More often, $dI/dV$ is measured at the same time with a lock-in amplifier by adding an AC voltage with small amplitude ($\sim 1 - 10 \text{ mV}_{\text{rms}}$) to the bias voltage and measuring the signal in tunneling current with the same frequency as the added signal. From the Taylor expansion of the tunneling current,

$$I(V_0 + V_m \cos(\omega t)) = I(V_0) + \frac{dI}{dV} \bigg|_{V_0} V_m \cos(\omega t) + ..., \hspace{1cm} (1.7)$$

we can see that the lock-in amplifier actually measures this quantity,

$$V_{\text{lock-in}} = \frac{dI}{dV} \cdot V_m \cdot R \cdot \frac{10V}{\text{sens}}, \hspace{1cm} (1.8)$$

where $V_m$ is the rms of the ac modulation, $R$ is the gain of the preamplifier, and $\text{sens}$ is the sensitivity of the lock-in amplifier, usually set to $50 - 200 \text{ mV}$. If the lock-in amplifier is set to output ‘Display’ instead of ‘X’ or ‘Y’, the last term should be dropped, but the signal noise ratio coming out of the lock-in will be affected due to the signal reduction from $V$ to $mV$ scale.
Furthermore, for deliberately designed systems with long continuous operating time, like the ones in our lab, it is possible to run with the spectroscopic imaging mode. In this mode, $dI/dV$ measurement is performed on every point of a pixel grid. The feedback loop is activated when scanning the tip in between points and deactivated when taking the $dI/dV$ measurement. This process is usually called as ‘spectroscopic imaging’, ‘$dI/dV$ mapping’ or simply ‘mapping’. With our UHV-STM, a $128 \times 128$ points $dI/dV$ map with 256 energy points takes about 16 hours.

1.3 Design of the UHV-STM

Our UHV-STM system includes a cryostat purchased from CreaTec, a STM chamber customized with A&N corp, a load-lock with a turbo-molecular pump, an Ion/TSP pump from Varian Inc. and a home-made STM scan-head.

1.3.1 Walker and scanner

A coarse motion actuator is needed for bringing the sharp tip to a small distance (usually $\sim 100$ nm) away from the sample. There are more than a few types of coarse approach mechanisms, including the walker [8], the beetle [9], etc. In our UHV-STM, a Pan-style walker design was adopted. Each leg of the walker is consisted of four sheets of piezo stacked together with alternating polarization directions, as shown in Fig. 1.6. The size of each piezo sheet is $0.20'' \times 0.20'' \times 0.020''$.

The shear-mode piezo operates in the manner shown in Fig. 1.6a. When an electric field $E$ perpendicular to the polarization $P$ is applied to the piezo pad, the corresponding motion of the piezo pad is to shear between the top and bottom surfaces. The equation that describes this motion is as follows,

$$\Delta x = d_{15}E_y h = d_{15}V_y,$$  \hspace{1cm} (1.9)

where $\Delta x$ is the amount of shear motion. This $\Delta x$ is often not directly proportional to
Figure 1.4: An overview of the UHV-STM system. (a) Image of the entire system. From top to bottom are: cryogen and gas ports, electric feedthroughs, body of the dewar, STM chamber and the table attached on top of 4 air springs, the ion pump and titanium sublimation pump. On the backside, we can see part of the loadlock and turbo pump system. On top of the table is a Model 1212 current preamplifier (DL instruments), which is usually replaced by our home-made pre-amplifier. (b) Image taken through the viewport. From inner out, we see the body of the STM, the inner, the outer cryo-shield and the UHV chamber with a viewport. (c) Image of the STM and cryo-shields when it is detached from its chamber. (d) Image of the STM suspended by the springs. The bottom copper piece is for Eddy current damping.
Figure 1.5: A solidworks drawing of the STM scan head. The top and bottom ‘gray’ pieces are made out of titanium, while the center ‘blue’ body is macor. The numbers indicate the important small parts, including (1) sample holder, (2) piezo scan tube and tip assembly, (3) sapphire prism, (4) one of the six walker piezo stacks, and (5) the scan tube holder, to which (2) is glued.
the single-step size of the walker, since the walker is also sensitive to other conditions like surface cleanness and the tension exerted by the BeCu spring (the front cover in Fig. 1.5). The walking step size of the UHV-STM is \( \sim 200 \text{ - } 300 \text{ nm} \) at room temperature and \( \sim 40 \text{ - } 60 \text{ nm} \) at helium temperature with the driving voltage set to 320 V.

![Diagram](image)

Figure 1.6: Operation of shear mode piezo (top). Walker piezo stack with side view (left) and 3D view (right). The piezo sheets are glued together with copper foils by conducting epoxy (H21D) and the copper tabs sticking out of the stack are electrical contacts.

The mostly used fine motion actuator is the tubular scanner invented by Binnig and Smith [10]. Shown in Fig. 1.7 is a schematic diagram of such a tubular scanner and how the electrodes are connected. The dimension of the scanner in our UHV-STM is as follows, \( L = 0.25" \), \( OD = 0.125" \) and \( ID = 0.080" \). When voltage is applied to electrode \( z \), the scanner will extend or shrink along its length direction. The total \( z \)-range of the scanner at 77 K is about \( \pm 180 \text{ nm} \) with \( \pm 420 \text{ V} \) on the \( z \) voltage. When voltage is applied to the \( xy \) electrodes, the scanner will tilt in the corresponding directions. For instance, when voltage is applied to \( \pm x \), the scanner tilt in the \( x \) direction. The detailed equations describing the motion of the scanner are as follow,

\[
\Delta x = \frac{2\sqrt{2}d_{31}VL^2}{\pi Dh}, \quad \Delta z = d_{31}V\frac{L}{h},
\]

(1.10)

where the voltages applied on \( \pm \) quadrants (of \( x \) or \( y \)) are assumed to be with the same magnitude but opposite signs.
Figure 1.7: A diagram for the tubular piezo scanner. The piezo is polarized in the radial direction and coated with layers of metal (usually Au or Ni) on both inner and outer surfaces of the tube. The outer conductor is further split into four quadrants, $\pm x$ and $\pm y$, along the length direction. Applying pos. (neg.) voltage on $z$ will extend (shorten) the scan tube in the length direction. Applying voltage on $x/y$ quadrants (opposite sign on $\pm$ quadrants) will bend the scan tube in $x/y$ directions.
All of the piezos, including shear-mode pads and tubular scanner, are of type PZT8, purchased from companies including EBL Products, Inc. and Boston Piezo Optics, Inc. PZT8 is chosen for its high temperature stability and low mechanical hysteresis (creeping). The important parameters of PZT8 for our applications are the piezoelectric constants $d_{15} = 330 \text{ pm/V}$ and $d_{31} = -95 \text{ pm/V}$. The piezo constants of PZT8 changes slowly with temperature. The relative ratio of the piezo constants at 4 K, 77 K and room temperature is roughly, $2/3:1:2$. This ratio comes from calibration using the XY piezo calibrations, so it depends on the tip length and has an error bar as high as 10% or more.

1.3.2 Tip and sample preparation

Figure 1.8: Images of tip and the etching process. (a) Image of the tip etching process. The center wire is a tungsten wire with 0.25 mm diameter and the outer ring is made out of stainless steel. A positive dc voltage ($2 \sim 6 \text{ V}$) is applied on the tungsten wire. (b) Image of an etched tungsten tip taken under a microscope with $100 \times$ magnification. (c) Image of the tip while in the tip holder and ready for performing experiment. The tungsten tip is bent to a zigzag shape in order to offset the tunneling location from the rotational center by $\sim 1.5 \text{ mm}$.

We use electrochemically etched tungsten tips in our tunneling experiments. During an etching process, a tungsten wire of diameter 0.25 mm is pulled through a film of 2M NaOH-water solution, and then a voltage is applied between the tungsten wire and the stainless steel ring (Fig. 1.8). The film is formed by dipping the stainless steel ring into the NaOH solution. The applied voltage is usually set to 2 to 6 Volts. Different etching voltages usually result in different etching rates. The etching rate can be estimated via the rate of bubble
formation around the stainless steel ring. Occasionally, the film bursts during the etching process, which requires reforming the film and restarting the etching process. This can result in different shapes of the tip apex, observable under optical microscopes.

Freshly etched tip needs to be cleaned by DI water and alcohol, and then checked under a microscope with magnification $100 \times$ or above. This process can assure the apex of the tip is less than a few micron or under the diffraction limit, which is sufficient for our experiments since field emission usually results in blunt tip apex with radius close to $\mu$m. Atomically sharp tips usually require further preparation after field emission. A commonly used method is to mechanically collide the tip with a gold surface and retract before re-engaging to tunneling. Further scanning on clean gold surface with high tunneling current (e.g. 1 nA) and high bias voltage (e.g. 1 V) can stabilize the tip. However, mostly importantly, the density of states of the tip should be relatively featureless before it can be used to investigate a sample. This is possible since the density of states on amorphous gold sample is flat for reference.

![Figure 1.9: Images of a sample and the sample holder. (a) An image of the sample holder used in the UHV-STM, glued with a piece of gold and a Bi-2212 sample. (b) Image of the cleaved surface of Bi-2212.](image)

Cleavable samples are usually attached to the sample holder with conducting epoxy (H21D). After the glue cured, a short rod is then attached directly on top of the sample. This short rod is knocked off to cleave the sample and yield a fresh surface, after the sample is introduced into vacuum. In our UHV-STM, this is realized by positioning the sample transfer manipulator so that the sample cleave rod sits at the position of the LN$_2$ shield door and
then the shield door is dropped to hit the cleave rod. The sample with the fresh surface is then inserted into our STM and the entire scan-head reaches temperature equilibrium a few hours.

Together with the sample, we can also glue a piece of press-flattened gold, as the target for field-emission tip preparation. With the two targets symmetrically attached around the rotational center and the tip bent carefully, we can easily switch between tunneling with either gold or another sample. With this method, we can reduce the sample exchange time from 4–5 hours down to about 1 hour. It also greatly improves the experimental success rate on rare samples, as long as the cleave is successful.

1.3.3 Electronics: pre-amplifiers and STM controller

In order to measure the tunneling current, a type of pre-amplifier called transimpedance amplifier is needed. This type of amplifier converts the tunneling current signal into voltage, so they are also called IVCs (I-V converters). We usually use two transimpedance pre-amplifiers: the first one is the Model 1212 from DL electronics and the second one is home-built with OPA627 or OPA140.

The amplification of a transimpedance amplifier is controlled by a feedback resistor which in our application is on the order of GΩ. The value of this resistor is also the gain of this amplifier, \( V = I \times R \) where \( I \) is the input and \( V \) is the output. In order to get a clean

![Figure 1.10: Modeling the noise of the pre-amplifier.](image-url)
tunneling signal, electronic noise of the amplifier needs to be considered. Shown in Fig. 1.10 is a minimum model for analyzing this electronic noise. The current source \( i \) is the tunneling current in the STM junction, \( i_n \) and \( v_n \) are the current and voltage noises of the op-amp (OPA627 or OPA140) itself, and the voltage noise \( v_{en} \) is the thermal noise of the feedback resistor, \( v_{en} = \sqrt{4k_B T R_f} \). Another very important parameter that limits the frequency at which the spectroscopy work can be performed is the shunt capacitance between the tip wire and the ground, \( C_S \). Combining all of these noises, we get

\[
i_{\text{noise}}^2 = \left[ \frac{1}{R_f^2} + \left( 2\pi f C_S \right)^2 \right] v_n^2 + i_n^2 + \frac{4k_B T}{R_f}.
\]

(1.11)

Here, the units of currents and voltages are \( A/\sqrt{Hz} \) and \( V/\sqrt{Hz} \). This equation can be used as an ultimate limit for electronic noise. However, most of the time larger noise comes from external sources like crosstalk and grounding loops, for which more detailed diagnoses will be needed.

In order to move the STM tip in the \( xyz \) directions, high voltages need to be applied on the inner and outer electrodes of the scanner (Fig. 1.7). These voltages are provided by the STM controller (MFP-3D from Asylum Research), which is able to provide 3 pairs of high voltages (six channels in total) up to \( \pm 420 \) V with resolution of 24 bits. This bit resolution, when converted into real space \( xyz \) resolution of the scanner, is roughly \( \delta z = 0.022 \) pm and \( \delta x = 0.381 \) pm at \( T = 77 \) K.

### 1.3.4 Ultra-high Vacuum: chamber, load-lock and pumps

STM experiment generally requires ultra-high vacuum \(< 10^{-9} \) torr\) environment due to the high sensitivity of the vacuum junction to external objects like gas molecules, and the strong electric field within \( \sim 10^9 \) V/m\) makes the situation worse. The existence of ‘free’ gas molecules can cause noise spikes in the tunnel current and make the junction unstable. To achieve such a high vacuum environment, a great deal of effort is required, especially for STM experiment at temperature higher than \( \sim 10 \) K, where cryogenic vacuum is not
available.

In our UHV-STM, a 70 l/s turbo pump (with a Xtr-dry 250 backing pump) is used to evacuate the chamber from ambient pressure down to the $10^{-7}$ torr range, before an ion pump and the heater for bakeout are turned on.

Due to fact that most materials will absorb water or air molecules on their surfaces and they will outgas when pressure is reduced, particular care is needed for choice of materials for UHV experiments. Most metals, including stainless steel, iron, nickel, copper, aluminum, titanium, even tin and lead, are fine for vacuum, but materials with high vapor pressure (room temperature or during bake-out), like zinc and cadmium, should definitely be avoided. This excludes a very commonly used material, brass. Stainless steel, aluminum, and copper are widely used in our experiments. Aluminum bronze (Al-Cu alloy) is used for our sample holders, due to its high strength and machinability. Plastics are usually bad for vacuum since they trap gas molecules deeply into the bulk and have low ‘melting’ points. In our UHV systems, there are only three types of plastics that have been used, including PEEK, kapton (polyimide) and teflon. Following this same reasoning, most of the glues are not suitable for UHV and the two types we used are Torr-Seal high vacuum epoxy and H21D silver conducting epoxy. A relatively complete reference for UHV compatible material choice is Ref. [11].

To achieve UHV condition, we need to firstly bake-out the chamber which basically expedites the out-gassing from the chamber walls by heating up the entire system to 120 to 140 °C. A bake-out is very effective for removing water molecules from the UHV system, which is the biggest limiting factor of vacuum in the range of $10^{-4}$ to $10^{-9}$ torr. A normal pumpdown of the UHV-STM system can be found in Appendix B.

1.3.5 Cryostat

The UHV-STM cryostat, purchased from Createc, is consisted of two cryogen baths, including a 14 liters of LN$_2$ bath and a 4 liters of LHe/LN$_2$ bath. With the heat load from
wires in the current configuration, the outer bath can last for \( \sim 72 \text{h} \) and the inner bath can last for 36 hours with LHe or more than 20 days with LN\(_2\).

The inner dewar can also be pumped to low pressure (e.g. 1 torr) to lower the cryogen boiling point. The pumping is carried out by simply connecting a rotary pump to the outlet of the inner dewar with a long hose (> 30 feet). In order to reduce the vibration from the pump, we place the pump out of the STM room and press the hose firmly by lead bricks in 2 - 3 positions. With LN\(_2\) in the inner dewar, we can pump it so that the base temperature is lowered to 50 K instead of 77 K. At such a low temperature nitrogen is solid and the cooling power thereafter comes from the sublimation of solid nitrogen.

1.3.6 Vibration isolation

Mechanical vibration from the environment is ubiquitous, and a lot of experimental equipments like vacuum pumps, cryogenic dewars, electronic boxes have their own vibration that cannot be eliminated. For a scanning tunneling microscope, the crucial vibration is the relative motion between the tip and the sample. When the feedback loop is turned off, relative vibration as small as 1 pm could be seen in the tunneling current. In order to reduce this relative motion, there are several steps in general that need to be done. The first one is to build a compact and rigid scan head so that the acoustic path between the tip and sample is short and stiff. The walker [8] is a good example of rigid design with a high resonance frequency. The second important thing to do is to remove as much of the vibrational sources (vacuum pumps, electronic boxes, human beings, etc.) from the vicinity of the experiment. The third to do is to isolate the environmental vibration from the STM (air springs, acoustic isolation room, etc.).

In our UHV-STM system, we have placed the experiment in an acoustically isolated room (sound room), with most of the electronic boxes outside. We also turn off the turbo pump on the load-lock during tunneling experiments. Furthermore, we have used an air spring system, and two spring-damping stages (Fig. 1.4). The air springs separate the UHV
chamber from the floor, while the two spring-dampening systems isolate the scan-head from the UHV chamber at two different levels. In our first spring-damping system, the spring is actually a soft tube to which the inner cryogen bath is attached, and the damping is provided by a copper disk moving within a magnetic field (Fig. 1.4), known as Eddy current damping. The second spring-damping system is composed of four stainless steel springs with which the scan head is suspended and a copper cross sitting inside magnetic field provided by SmCo$_5$ permanent magnets on the bottom of the inner cryo-shield.

Here we describe the mechanism of a spring-damping isolation system with a simple model. Suppose there is a spring-mass system attached to a wall (see Fig. 1.11). The mass could be viewed as our experiment and the wall is the environment, which is moving by itself $y = y(t)$. Our objective is to relate the motion of the mass $x = x(t)$ to $y(t)$. The equation of motion is simply,

$$m\ddot{x} + \gamma(\dot{x} - \dot{y}) + k(x - y) = 0,$$

where $k$ is the spring constant and $\gamma$ is the dampening term. Solving this equation with trial solution $x(t) = x_0 e^{-\omega t}$ and $y(t) = y_0 e^{-\omega t}$,

$$\frac{x_0}{y_0} = \frac{i\gamma\omega + k}{-m\omega^2 + i\gamma\omega + k},$$

(1.13)
The magnitude of this ratio is defined as the transfer function,

\[ T_0(\omega) = \sqrt{\frac{k^2 + \gamma^2 \omega^2}{(k - m\omega^2)^2 + \gamma^2 \omega^2}} = \sqrt{\frac{1 + \beta^2 \omega^2/\omega_0^2}{(1 - \omega^2/\omega_0^2)^2 + \beta^2 \omega^2/\omega_0^2}}, \tag{1.14} \]

Here \( \beta \) is the reduced damping factor \( \beta = \gamma \omega_0/k \) and \( \omega_0 \) is the resonance frequency of the spring-mass system \( \omega_0 = \sqrt{k/m} \).

Two extreme cases for the Eqn. 1.14 and the resonance amplitude could be easily extracted and could also be seen in Fig. 1.12. At zero frequency, the transfer function is simply one. At high frequency, \( T_0(\omega) = \beta \omega_0/\omega \) which means the lower \( \beta \) is the better the isolation at high frequency will be. At the resonance frequency \( \omega = \omega_0 \), \( T_0 = \sqrt{1 + \beta^2/\beta^2} \) which indicates that small \( \beta \) will cause sharp resonance and therefore should be avoided. To balance the resonance at \( \omega_0 \) and isolation at high frequency an ideal condition could be roughly \( \beta = 1 \). In our UHV-STM adjusting \( \beta \) is realized the via changing the Eddy current damping. For example the thickness of the copper piece and the strength of the permanent magnets both could be varied to increase to decrease the damping factor \( \beta \). However, since the conductivity of copper changes with temperature, it is quite hard to meet the ideal requirement at all temperatures.

![Figure 1.12: Transfer function of a single-stage vibration isolation system.](image)

After all, the above discussions are all based on ideal situations, while in reality vibrational shorts (e.g. wires that connect the cryogen bath and the STM scan-head) are also
very crucial for vibrational noise performance. Checking for these shorts will be absolutely necessary to minimize the vibrational noise for tunneling experiments.

1.3.7 Other experimental details

For a STM experiment at high temperature (e.g. 77 K) the differential thermal expansion between different parts of the STM scan-head will cause significant amount of thermal drift between the tip and sample due to both experiment-related operations and temperature fluctuation caused by weather or ambient pressure changes. Thermal stabilization on the scan-head thus becomes necessary. In our UHV-STM we have used an home-made resistive heater and a Lakeshore temperature controller to stabilize the STM temperature within a few milli-Kelvin of the set temperature. At the meantime, we have thermally sunk the STM with the cryogen bath with a copper braid. The size of the copper braid can be adjusted to balance in between the base temperature, the cooling rate, and the boil-off rate.

![Diagram for making thermal sink with electrical isolation for the top-plate of STM scan-head.](image)

Figure 1.13: Diagram for making thermal sink with electrical isolation for the top-plate of STM scan-head. The sample bias voltage could be as high as 300 V during field emission. For good thermal sinking the size of the sapphire could be as large as 0.5 by 0.5 inch and the thickness of it could be as small as 0.02 inch. If sapphire is not available, quartz could also be used instead. Since PEEK is not a strong material, tightening of these screws should be very gentle, usually less than finger strength, but the performance of this thermal sink at low temperature is still quite good.

While thermally connecting the entire STM scan-head is necessary for the low-drift requirement, it is also necessary to electrically isolate some parts. For example, the top-plate of the scan-head where sample bias voltage is attached needs to be separate from all
of the other components including the cryogen bath, the aluminum shields, the heater coil and the temperature sensor. These isolations were done with PEEK screws/rings, sapphire pieces, and Kapton wire insulations. It is also necessary to connect the bottom-plate of the scan-head with the top-plate in order to reduce the thermal equilibrium time within the STM scan-head. This is done by separating the aluminum piece that connects the top and bottom plates from the top-plate by a piece of sapphire and clamped together with PEEK screws. The electrical isolation between the top-plate and the cryogen bath was done in a similar fashion at the interface between the copper braid and the cryogen bath. A simplified diagram of this type of thermal sinking and electrical isolation could be seen in Fig. 1.13.

In order to reduce the heat load on the STM and the cryogen baths we have used only manganin and stainless steel wires (mostly AWG 36 from MWS Wire Industries) connecting the top feedthroughs (Fig. 1.4) and the STM scan-head. Additional teflon heat-shrink tube were used to protect the manganin/SS wire braids against sharp parts of the vacuum chamber. The tip-line carrying the tunneling current of order pA to nA is a semi-rigid coaxial cable (SC-119 from Coax Co., LTD.). This coax cable uses stainless steel for both inner and outer conductors and polyimide for isolation. To avoid shorting the vibration isolation for the STM, semi-rigid cable was not used at the last section when connecting to the STM scan-head. This part is replaced with a coax cable (Type CC-SC-500 from Lakeshore).
Chapter 2

Charge-density wave in transition metal dichalcogenides

2.1 Introduction of charge-density wave

Charge-density wave breaks the lattice translational symmetry of a crystal by redistributing the electron density in a spatially periodic manner. It occurs in a wide range of materials, including both quasi-1D and quasi-2D systems. For quasi-1D systems [12], it has been found in NbSe$_3$ (and many siblings of it), (Ta$_2$Se$_4$I, blue/purple bronzes, many 1D organic conductors, 1D metallic chains on Si surface [13], etc. For quasi-2D CDW systems, the research has been mainly focused on the layered transition metal dichalcogenides (TMD or MX$_2$), Pb/Sn covered Ge(111) surface, etc. The rare-earth tritellurides, RTe$_3$ ($R$ = rare-earth element), are a class of materials that are quasi-two-dimensional in structure but the CDW in them is unidirectional [14, 15]. Review papers could be found in Ref. [12, 16, 17] for 1D systems and in Ref. [18] for 2D TMDs.

2.1.1 Peierls mechanism and Fermi surface nesting

We will start by introducing the Peierls mechanism in a simple 1D metal. For an atomic chain with partially filled band (e.g. one electron per site), the response of the system to external electric field is described by the Lindhard function,

$$\chi(q) = \int \frac{d^d k}{(2\pi)^d} \frac{f_k - f_{k+q}}{\epsilon_k - \epsilon_{k+q}},$$  \hspace{1cm} (2.1)
where \( \epsilon_k \) is the energy band (energy-momentum dispersion relation of electrons) and \( f_k = f(\epsilon_k) \) is the Fermi-Dirac distribution function, \( f(\epsilon) = \frac{1}{e^{\epsilon-\mu}/k_B T + 1} \). Results of calculation for free-electron systems with different dimensionalities could be found in Fig. 2.1. One would immediately notice that for the 1D system, the Lindhard response function is diverging at \( 2k_F \), which is not surprising since for a 1D system all parts of the Fermi surface (basically two points) are connected by the same wave vector to other parts of the Fermi surface. Similar divergence is possible for 2D systems too, since anisotropic 2D systems can have large sections of Fermi surface that are parallel to each other. This is the basic idea of ‘Fermi surface nesting’. Mathematically checking the nesting properties of a particular Fermi surface is possible by simply doing the auto-correlation of the spectral intensity map at the Fermi level (also called as the Fermi surface). If nesting condition is satisfied, it would show up as peak(s) in the auto-correlation map.

![Lindhard response functions for free-electron systems in different dimensions. These are qualitative sketches following Ref. [12].](image)

As a result of the divergence in \( \chi(q) \), the system undergoes a structural phase transition; i.e., the system enters into another lower energy state, the CDW state. In this new state, the electron density is redistributed according to the wave vector \( 2k_F \) and the ions are displaced due to electron phonon interaction (periodic lattice distortion). However, since lattice distortion increases the elastic energy, it is only a second order effect.
The concept of CDW, in its more general form of "charge ordering", is widely used in other correlated electron systems like cuprate and iron-based high-$T_c$ superconductors [19, 20]. It also has been suggested that the checkerboard in BSCCO is intimately tied with a CDW [21].

2.1.2 Mean-field approach

This part is a summary of the important results in Ch.3 of the book by Grüner [12].

To describe the charge density wave formation, we start by introducing the Fröhlich Hamiltonian including both electronic and phononic contributions,

$$H = \sum_k \epsilon_k a_k^\dagger a_k + \sum_q \hbar \omega_q b_q^\dagger b_q + \sum_{k,q} g_q a_{k+q}^\dagger a_k (b_q + b_{-q}^\dagger),$$

where $a_k$ ($b_q$) and $a_k^\dagger$ ($b_q^\dagger$) are electron (phonon) annihilation and creation operators.

After the CDW order sets in, the phonon mode at $2k_F$ becomes macroscopically occupied. The average occupation on this mode is the order parameter of the CDW transition,

$$|\Delta| e^{i\phi} = g \left( \langle b_{2k_F} \rangle + \langle b_{-2k_F}^\dagger \rangle \right),$$

The renormalized phonon frequency (the phonon frequency that can be measured experimentally),

$$\omega_{ren,q}^2 = \omega_q^2 + \frac{2g^2 \omega_q}{\hbar} \chi(q,T),$$
from which we can see the softening of phonon (since $\chi<0$). In the mean field approximation, the Fröhlich Hamiltonian is simplified as,

$$H = \sum_k \epsilon_k a_k^\dagger a_k + 2g \sum_k \left[ a_{k+2k_F}^\dagger a_k \langle b_-^{2k_F} \rangle + a_{k-2k_F}^\dagger a_k \langle b_-^{2k_F} \rangle \right] + 2\hbar \omega_{2k_F} \langle b_2k_F \rangle^2, \quad (2.5)$$

This quadratic Hamiltonian can be diagonalized using a Bogoliubov transformation,

$$\gamma_{1,k} = U_k e^{-i\phi/2} a_{1,k} - V_k e^{i\phi/2} a_{2,k}, \quad \gamma_{2,k} = V_k e^{-i\phi/2} a_{1,k} + U_k e^{i\phi/2} a_{2,k}, \quad (2.6)$$

with the diagonalizing condition being,

$$2\epsilon_k U_k V_k + |\Delta| (U_k^2 - V_k^2) = 0, \quad U_k^2 + V_k^2 = 1, \quad (2.7)$$

which yield the following solutions,

$$V_k^2 = \frac{1}{2} \left( 1 + \frac{\epsilon_k}{E_k} \right), \quad U_k^2 = \frac{1}{2} \left( 1 - \frac{\epsilon_k}{E_k} \right). \quad (2.8)$$

As a result, the band dispersion is BCS-like,

$$E_k = \text{sign}(k - k_F) \left( \epsilon_k^2 + \Delta^2 \right), \quad \epsilon_k = \hbar v_F (k - k_F), \quad (2.9)$$

and the ground state wave function is given as,

$$|\phi_0\rangle = \left( \prod_{|k|<k_F} \gamma_{1,k}^\dagger \gamma_{2,k}^\dagger \right) |0\rangle, \quad (2.10)$$

### 2.1.3 LDOS with charge modulation

In a CDW state, the electronic density is modulated with the wave length $2\pi/2k_F$. Since it is also expected that the local summation of occupied and unoccupied states results in a constant, it would be interesting to study how the local density of states is modified with the CDW transition.

Electronic density can be calculated from the wave function Eqn. (2.10), as the summation over all of the occupied states,

$$\rho(x) = \sum_{|k|<k_F} \left[ 1 - 2U_k V_k \cos(2k_F x + \phi) \right], \quad (2.11)$$
Figure 2.3: LDOS evolution with the periodicity of CDW from mean-field theory calculation, with the horizontal axis being energy and vertical axis being real-space position. Three clear features are: 1. There is a gap around $E_F$; 2. States around $E_F$ is affected the most; 3. The LDOS is enhanced on the filled and empty sides alternatively.
In a STS experiment, one measures density of states not only in real space but also with energy resolution, so it is more appropriate to rewrite the density of states as follows,

$$
\rho(x, E) = \left[ 1 - \frac{\Delta}{E} \cos(2k_F x + \phi) \right] \frac{|E|}{\sqrt{|E|^2 - |\Delta|^2}},
$$

which is very similar to the quasi-particle density of states of a superconductor (BCS theory). This equation is only valid for $|E| > |\Delta|$, and for the energy window $|E| < |\Delta|$, there is no density of states $\rho = 0$. This corresponds to the gap opening shown in Fig. 2.2.

![Figure 2.4: Representative LDOS curves in an electron-driven CDW. Here we have assumed the phase $\phi$ of CDW is zero, $\phi = 0$.](image)

An important feature of Eqn. (2.12) is that the coefficient of the cosine term changes its sign in between occupied ($E < 0$) and unoccupied states ($E > 0$). This results in either enhancement or suppression of density of states on the occupied/unoccupied states, which varies with the spatial location (relative to the CDW). Figs. 2.3 and 2.4 shows the spatial evolution of the LDOS with the periodicity of CDW. As we mentioned in Ch. 1, the STM imaging is not only sensitive to the positions of the ionic cores but also the electronic density of states around the Fermi level (see Eqn. (1.6) for details). From this we can easily see that in the presence of this spatially switching enhancement/suppression in LDOS the STM topography will show a contrast inversion between the images of occupied and unoccupied
states (set by the polarity of tunneling bias voltage). This contrast inversion phenomenon has been well demonstrated in various CDW systems, including NbSe$_3$ [22], Sn(Pb) covered Ge(111) surface [23, 24], CaC$_6$ [25], etc.

2.2 Charge-density waves in quasi-2D materials

Quasi-2D materials that exhibit CDW states include the layered transition metal dichalcogenide (TMD) [26], $R$Te$_3$ [14], Sn (Pb) covered Ge(111) surfaces, etc. The TMD is a wide class of materials that mainly includes two structure phases (the 2H and the 1T phases) of many transition metal dichalcogenides. For example, 2H-NbSe$_2$, 2H-TaSe$_2$, 1T-TaS$_2$ and 1T-TiSe$_2$ are among the most studied TMDs. 2H-NbSe$_2$ shows both superconductivity at 7 K and CDW at 35 K. 2H-TaSe$_2$ has two CDW phases with transition temperatures at 122 K and 90 K, together with a superconducting transition at 0.1 K. 1T-TaS$_2$ exhibits a rich temperature-pressure phase diagram including numerous CDW phases and superconducting phase [27]. In addition, the checkerboard phenomena in superconducting cuprates is also argued as a strong candidate for CDW [28, 21], while cuprates are widely accepted as two-dimensional systems.

2.2.1 Introduction to 2H-MX$_2$

Transition metal dichalcogenides are among the earliest materials studied by scanning tunneling microscopes [29, 30, 31, 32, 33, 34], likely due to that they are easily cleaved and their surfaces are relatively inert even under ambient condition. In this thesis, I will mainly focus on the trigonal prismatic structure phase (2H) of TMDs, including 2H-TaSe$_2$, NbSe$_2$ and TaS$_2$. The structure of the 2H phase is fairly simple; it consists of chalcogen layers and transition metal layers with the smallest building block being two chalcogen layers sandwiching one transition metal layer. As shown in Fig. 2.5, the bonding between the two adjacent chalcogen layers is of van der Waals type and thus can be mechanically cleaved. The topmost atomic layer post cleavage can be imaged with STM, as shown in Fig. 2.5. Each
individual object is a chalcogen atom and the $3 \times 3$ superstructure is due to the formation of CDW. Similar atomic resolving images of TaSe$_2$ and TaS$_2$ are shown in Fig. 2.6.

Figure 2.5: Topography and crystal structure of the 2H-NbSe$_2$. (a) Topographic image of NbSe$_2$ surface at 5 K and 50 mV and 80 pA. (b) Crystal structure of 2H-MX$_2$ with the dashed line indicating the cleave plane. Black/red atoms are Nb/Se. (c) A line profile indicating a discontinuity of CDW order and visualizing the incommensurability of CDW in NbSe$_2$.

The CDW transition in 2H-TMD have been studied in numerous experiments, including transport and specific capacity measurements, scattering experiments (neutron, X-ray, electron diffractions, etc.), real-space imaging (STM), etc. The difference in momentum-space electronic structure can also be measured using angle-resolved photoemission spectroscopy (ARPES) experiments.

Specific heat measurement of the CDW transitions in TaSe$_2$ done by Craven et al. is shown in Fig. 2.7 [35]. As we can see that there is a well defined 2nd order phase transition at temperature 122 K and there is another transition at 90 K that is hysteretic and hence 1st order. The transition at 122 K is a transition from normal state to an ICDW (incommensurate CDW), while the transition at 90 K was thought to be a lock-in transition from incommensurate to commensurate state (see Fig. 2.7). However, recent study with high-resolution X-ray diffraction and better temperature control reveals a more complicated phase diagram between 90 and 122 K [36].
The CDW phases in NbSe$_2$ and TaS$_2$ on the other hand are both incommensurate with respect to the lattice periodicity down to the lowest temperatures studied thus far. The CDW wave vector in NbSe$_2$ is measured to be $q_{\text{CDW}} = 0.328a^*$ by neutron scattering, where $a^*$ is the reciprocal lattice vector [37]. In TaS$_2$ it is observed that $q_{\text{CDW}} = 0.338a^*$ by electron diffraction experiments [38].
Figure 2.7: Specific heat and neutron scattering measurements of the two CDW transitions in TaSe$_2$. (a) Specific heat measurement shows a second order transition at 122K and a hysteretic first order transition at 90K. (from ref. [35].) (b) Elastic neutron scattering measurement shows the appearance of two super-lattice peaks around $4/3a^*$, which at lower temperature merge into one peak sit exactly at $4/3a^*$. (from ref. [39].)
2.2.2 Recent progresses and remaining issues

Electronic originating CDW is a natural result when one considers the different CDW wave vectors $q_{CDW}$ in different materials that vary with temperature, since the most obvious difference between them is first of all the difference in Fermi surface geometry. The incommensurate behaviors in CDW further support the role electrons in the CDW transitions, due to the general expectation that $2k_F$ will not be a rational fraction of the lattice reciprocal vectors. When the Fermi surface of TaSe$_2$ is mapped out by ARPES experiment [40] and the auto-correlation map is generated, a peak corresponding to $q_{CDW}$ was indeed found there, as shown in Fig. 2.8.

![Figure 2.8: Fermi surface map obtained by fitting ARPES data and possible nesting mechanism. (a) Fermi surface obtained by fitting the ARPES data with a tight-binding model. (b) Auto-correlation map of (a). (c) Line profile along the ΓM direction. Data from: [40]](image)

The purely electron-driven Peierls mechanism is a beautiful theory since it only involves consideration about the geometry of the Fermi surface. However, the applicability of this theory has been questioned for long time. Chan and Heine [41], after considering the Coulomb and exchange interactions, concluded that a CDW is not possible when there is not periodic lattice distortion (PLD). If there were no PLD, a CDW is never possible also because that spin-density wave (SDW) instability will set in first. The criterion for a CDW/PLD to occur is,

$$\frac{4\eta_q^2}{\hbar\omega_q} \geq \frac{1}{\chi_q} + (2U_q - V_q),$$  \hspace{1cm} (2.13)

where $\omega_q$ is the undressed phonon frequency, $\chi_q$ is the Lindhard response function, $\eta_q$ is the
electron-phonon coupling strength, $U_q$ and $V_q$ are the Coulomb and exchange interaction. Furthermore they found that CDW/PLD can win over SDW only when $\frac{4\pi^2}{\hbar\omega_q} - 2U_q > 0$ and $V_q < \frac{1}{\chi_q}$.

More recently, Johannes et al., after studying the exact same systems including TaSe$_2$ and NbSe$_2$ using first principle calculations [42, 43], found that no CDW can be formed by the electron systems if the ions were held in place. Their study further extends to other materials like ErTe$_3$ and artificial Na atomic chains with the same conclusion that lattice distortion is a prerequisite for CDW formation.

Using inelastic X-ray scattering that directly probes the phonon dispersions, Weber et al. discovered that in NbSe$_2$ the phonon mode corresponding to $q_{\text{CDW}}$ exhibits unconventional behavior with the CDW transition [44]. Unlike the expected cusp-shape phonon softening near $q_{\text{CDW}}$ [12], the phonon energy was reduced to zero over an extended range of $q$ vector. This is more consistent with a wavevector-dependent electron-phonon coupling which is capable of driving a structure phase transition.

Figure 2.9: Phonon anomaly in 2H-NbSe$_2$. The extended region where the phonon energy is reduced to zero cannot agree with conventional fermi surface nesting related Khon anomaly. Data from [44]

Another difficulty with the applicability of Peierls mechanism in 2H-NbSe$_2$, TaSe$_2$ and TaS$_2$ lies in the results of transport measurements (see Fig. 2.10). All three of the materials show kinks in $\rho(T)$ at the CDW transition temperatures that indicate better electrical
conductivity in the CDW phases. This directly conflicts with the prediction of Peierls mechanism that an energy gap should open up at the Fermi level and the system becomes an insulator. This could be explained if the CDW state is carrying electrical current as the cooper pairs in a superconductor. However, the pairing in a CDW is between electrons and holes [12] which is charge neutral. Furthermore, the lattice distortion associated with the CDW also prohibits it from sliding freely, making it impossible to contribute to the increase of conductivity. The enhancement of electrical conductivity in these compounds at $T_{\text{CDW}}$ has remained mysterious thus far, to the knowledge of the author.

![Figure 2.10: Conductivity enhancement with CDW transitions in 2H-MX$_2$. Data from [35] and [45].](image)

It has also been reported that resistivity anisotropy increase with a larger slope with respect to temperature in the CDW phases [46, 47]. It appears like the CDW transitions alter the in-plane electronic scattering mainly. In TaSe$_2$, the electronic scattering rate from the Drude term follows the same trend as the resistivity [48]. From thermal transport measurement in NbSe$_2$, a large Nernst singal has been observed in the CDW state [49], which also points to a change in the electronic scattering along with the transition.

Shown in Table 2.1 is a summary of the transition temperatures of both CDW and superconductivity in four 2H-TMDs. The CDW transition temperatures are much higher with the Ta compounds, and superconductivity transition temperatures are much higher.
Table 2.1: CDW and superconductivity transition temperatures of 2H-MX$_2$

<table>
<thead>
<tr>
<th></th>
<th>TaSe$_2$</th>
<th>TaS$_2$</th>
<th>NbSe$_2$</th>
<th>NbS$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{CDW}$ (K)</td>
<td>122</td>
<td>78</td>
<td>35</td>
<td>no CDW</td>
</tr>
<tr>
<td>$T_C$ (K)</td>
<td>0.1</td>
<td>0.8</td>
<td>7.2</td>
<td>6.1</td>
</tr>
</tbody>
</table>

with the Nb ones. Among the differences between niobium and tantalum are that Ta has 5d electrons and Nb is the 4d counterpart and that Nb has slightly higher electronegativity. It is also known that in NbSe$_2$ the metal ion is not fully tetravalent and the Fermi surface comprise a small pocket of Se p band [50], while it is not the case for TaSe$_2$ [40]. Do 5d electrons favors CDW while 4d electrons favor superconductivity? Are electron-phonon coupling dominating these transitions? Our STM results (Figs. 2.5 and 2.6) show that the CDW in Ta compounds is more rigid compared to NbSe$_2$, in which long range CDW order is often interrupted. This could not be due to different impurity levels in the sample growing processes, since the STM images show no direct relation between the atomic defects and the defects in CDW superstructures. However, within the system of NbSe$_2$, Iwaya et al. indeed reported that Ta dopants disturb the CDW order [45].
2.2.3 CDW gap opening

One of the most prominent signature of a Peierls type CDW is the opening of an energy gap around the Fermi level. In 2H-TMDs the search for such kind of a gap has been very frustrating for physicists. Three of the most commonly used methods are single-particle tunneling experiments, optical conductivity measurement and angle-resolved photoemission spectroscopy. These experiments give results that are not in agreement.

An early study by Hess et al. in NbSe$_2$ revealed the kinks at 35 meV in tunneling dI/dV spectrum, which were then attributed to the CDW formation Hess [34]. Firstly, the finite density of states around zero bias voltage means the gap opening is not around the entire Fermi surface. Secondly, the measured gap value of 35 meV seemed to be too large for mean-field theory to be applied; $2\Delta/k_B T_{CDW} = 23.3$ which is one order of magnitude larger than the predicted value from mean-field calculation, $2\Delta/k_B T_{CDW} \approx 3.52$ [12]. Recently, it has been reported that this 35 meV gap persists at temperature above $T_{CDW}$ [51], indicating that it is probably not the correct CDW energy gap.

![Figure 2.11: Local density of states measurement on NbSe$_2$, showing the two kinks at ±35 mV measured at 4 K (red arrows). The lower part shows the superconducting gap measured at 50 mK. Data from [34].](image)

ARPES experiments, capable of resolving band structure in the $k$–space, have reported
momentum-dependent gap structures across the Fermi surface for many 2H-TMDs [40, 50, 52]. For TaSe$_2$, there is a $k-$dependent pseudogap which reduces with increasing temperature but persist up to room temperature with the gap value varying from 10 to 40 meV. In NbSe$_2$ there is also a similar $k-$dependent gap opening with gap value ranging from $2 - 6$ meV. Despite their difference in gap values, there is one interesting common feature between TaSe$_2$ and NbSe$_2$; the gap values are larger on the Fermi surface segments closer to K-point of the Brillouin zone and are smaller or zero on points closer to M-point or Γ-point. Studying the complicated gap formation in $k-$space might lead towards understanding of not only the CDW problems themselves but also other correlated behaviors of electrons.

Another commonly used experiment that can determine the gap size is by measuring the optical conductivity or dynamics. From both the dynamics study by Demsar et al. [53] and optical conductivity study by Vescoli et al. [48], gaps of $2\Delta = 70$ meV have been measured in TaSe$_2$, although their gap-temperature dependence are not in agreement.

2.2.4 Doping and superconductivity

Figure 2.12: Phase diagrams of Cu intercalation into TiSe$_2$ (a) and TaS$_2$ (b) (Cu$_2$TiSe$_2$ and Cu$_x$TaS$_2$), from [54, 55].
Intercalating dopants into the space between the chalcogen layers bonded by van der Waals force usually introduces interesting changes to the physical properties of the layered structure materials like TMD [56]. For example, intercalation of copper into 1T-TiSe$_2$ and 2H-TaS$_2$ have both resulted in superconductivity with $T_c$ above 4 K (Cu$_x$TiSe$_2$ [54] and Cu$_x$TaS$_2$ [55]). Their doping phase diagrams (Fig. 2.12) are very similar to that of high-$T_c$ superconducting cuprates [57] and pnictides with CDWs replaced by pseudogap, antiferromagnetic or spin-density wave phases.

Not only metals can be intercalated into the van den Waals gaps of the layer-structured materials, (organic) molecules can do very similar things [56]. An interesting fact is that methylamine (CH$_3$NH$_2$) intercalated TaS$_2$ has the highest superconducting transition temperature ($T_c = 5.6$ K) among the intercalation compounds of TaS$_2$. The dimensionality change due to intercalation or the suppression CDW order caused by the disorder is probably responsible for the rising of superconductivity in these materials.

More interestingly nickel dopants in Ni$_x$TaS$_2$, which is also superconductor ($T_c = 4$ K), have been reported as replacing Ta instead of intercalation between the S-Ta-S sandwiches [58]. Atomic resolving STM imaging could serve as a powerful tool to distinguish between these two cases. Furthermore, studying the different intercalating/substituting dopants with respect to the reminiscent CDW order will be helpful for studying the microscopic origin of the CDW in the TMDs.

2.3 Summary

Charge-density wave is a long-standing topic in condensed matter physics and yet is still a very intriguing and puzzling one, which is closely related to other topics like high-$T_c$ superconductivity, colossal magneto-resistivity, etc. In the 2H-TMDs, the driving force of the CDW transition is still under debate, while many experimental facts (e.g. transport and phonon) have clearly shown their unconventional behaviors compared to 1D CDW systems. Using an atomically resolving STM, together with its high energy resolution of measuring
local density of states of both the occupied and unoccupied levels, seems to be a promising tool for understanding the microscopic physics of charge-density waves in the 2H-TMDs. Together with the recently reported data from ARPES and high-resolution scattering experiments, we will try to push the understanding of TMD-CDW systems to a better level.
Chapter 3

Strong periodic lattice distortion in 2D charge-density wave systems

3.1 Introduction

In condensed matter physics only a handful of models have both elegance in their construction and are thought to describe the behavior of numerous materials. The Peierls transition [59], for electrons crystallizing into charge-density waves (CDW), is one such model that has been successfully applied to many 1D systems [12]. However, in the quasi-2D electron systems of the layered transition metal dichalcogenides (TMD) there is still a controversy about the nature of the transitions to CDW phases, i.e. whether they are described by a Peierls-type mechanism or by a lattice driven model. By performing scanning tunneling microscopy (STM) experiments on the canonical TMD-CDW systems, we have successfully imaged the electronic modulation and the lattice distortion separately for the first time. Across three materials (2H-TaS$_2$, TaSe$_2$, and NbSe$_2$), we found dominant lattice contributions instead of the electronic modulation expected from Peierls transitions, in contrast to the CDW states that show the hallmark of contrast inversion, in 2D [23, 24, 25] and 1D [22] systems. Our results imply that the periodic lattice distortion (PLD) plays a vital role in the formation of CDW phases in the TMDs, which is in agreement with the theoretical studies [41, 42, 43], especially recent work by Johannes et al. [42, 43]. Our results illustrate the importance of taking into account the more complicated lattice degree of freedom, and they will prove useful in the studies of other correlated electron systems.

The CDW instability in low dimensional systems was initially explained by the Peierls
transition [59, 12] in which a divergence in electronic response function at a particular wave vector \( (2k_F \text{ for 1D}) \) results in a periodic charge redistribution in real space. That is, a low dimensional electronic system can lower its energy by opening an energy gap at the Fermi level, piling up electrons periodically and reducing its lattice transitional symmetry. As a result of the periodic electron modulation (eMod) and the electron-phonon interaction, the lattice itself undergoes a similar modulation, which causes an increase in the lattice elastic energy [12]. As a consequence, the lattice distortion will be minimized and acts as a secondary effect [43]. The eMod is embedded in the change of the local density of states (LDOS), which we have calculated for real-space from the Peierls model (Supplementary part I). Another way of visualizing the problem is by using the sum rule in quantum mechanics, where the summation of the numbers of filled and empty states on every lattice site is a constant. So that wherever there are excess electron states, there will be less hole states on the same site. The LDOS calculation we have performed also shows the conservation of total states.

The tunneling current in a STM junction is proportional to the sample’s integrated density of states and a matrix element with exponential dependence on the tip-sample distance, $I \sim e^{-2\kappa s} \int_0^{eV} LDOS(\epsilon) d\epsilon$ [5]. In the constant current imaging mode, the tunneling current is maintained at a fixed value with the tip-sample distance being adjusted and recorded. For a CDW system that only contains eMod and no lattice distortions (like the one in Fig. 2.3), the topographic image of the filled states will be complementary with the empty states due to the nature of LDOS variation and uniformity of atomic positions. Therefore, the constant current height images of filled and empty states, in such a CDW system, will accordingly be complementary. This phenomenon has been well demonstrated by STM experiments on many CDW systems (see next Section for details). In each of these cases the modulations, with the CDW periodicity, are of opposite phases for electron- and hole-like states in the topography. This is predicted for a Peierls transition.
3.2 Previous STM study of topography in CDW systems

Figure 3.1: Previous STM images of Ge(111) surface covered by 1/3 monolayer of Sn, showing the contrast inversion of topography between filled and empty states. Images were taken with $\pm 1\,\text{V}$ for filled/empty states at room temperature (upper) and low temperature (lower). The image on the right is an averaged image between the two images taken at low temperature, which fully recovers the atomic order with no super-lattice structure. Data from [24].

As we mentioned in Ch. 2 and in the previous section, a hallmark of the CDW phase is the contrast inversion between images of filled and empty states; the electron-like states are redistributed in a periodic manner, namely charge-density wave. At the same time, the hole-like states will also be redistributed but with an opposite phase; the summation of electron-like and hole-like states on a particular lattice site has to be conserved. This hallmark contrast inversion phenomena have been reported on many different CDW systems, including the quasi-1D NbSe$_3$ (Fig. 3.2 [22]), the 2D systems of Sn/Pb covered Ge(111) surfaces (Fig. 3.1 [23, 24]), and the quasi-2D CaC$_6$ with 1D modulation [25].

In the two-dimensional systems of Sn/Pb covered Ge(111) surface, the triangular lattice of Sn/Pb atoms show a uniform atomic lattice at room temperature. Upon cooling to low temperature, both Sn and Pb systems exhibit a reconstruction with $\sqrt{3}$ times larger lattice periodicity (Fig. 3.1). More importantly, STM imaging showed that the filled and
empty states were perfectly complementary, indicating the reconstruction is with valence electron redistribution instead of lattice distortion, which is predicted as a charge-density wave. However, detailed study about the electronic structure in the framework of Fermi surface nesting were not able to predict this charge instability [24], implying the limitation of nesting related CDW origins.

In NbSe$_3$, a one-dimensional chain compound that exhibits two CDW transitions at low temperatures, very similar contrast inversion has been observed by Brun et al., as shown in Fig. 3.2. In this material, there are three types of one-dimensional chains (type I, II and III in Fig. 3.2), lying along the $b$-axis. The features along the chain direction are the CDW minima and maxima. With the blue lines indicating the same locations in different images, we can clearly see the inversion of apparent height between the images of filled and empty states ($\mp 200$ mV).

3.3 Strong lattice distortion in 2H-MX$_2$

The TMDs (e.g. 2H-NbSe$_2$, TaSe$_2$, 1T-TaS$_2$, and TiSe$_2$) have drawn a lot of attention due to the coexistence and competition between the CDW and the superconductivity phases which are induced by the complicated interactions between the internal degrees of freedom.
freedom, including charge, lattice and orbital [54, 55, 27, 60]. However, whether or not an electron-driven mechanism can be the cause of the formation of CDWs in the TMDs with the trigonal prismatic structure (2H) is still under debate. The existence of incommensurate CDW phases does lend support for the role of electronic interactions as the cause, since $2k_F$ is generally not expected to be a rational fraction of the lattice reciprocal vectors. Furthermore, the electronic origin is also supported by photoemission experiments [40, 50]. By measuring the Fermi surface with ARPES and calculating the electronic susceptibility through the autocorrelation, peaks were found at wave vectors corresponding to those of the CDW [40, 50]. However, inelastic X-ray scattering experiment [44] reveals that the lattice dynamics exhibit unconventional behavior and hence may be dominating the transition. In a recent real space study, Soumyanarayanan et al. [51] have shown that there is a close relationship between CDW formation and the local strain related to the surface layer buckling in NbSe$_2$. Furthermore, existing studies have proposed a variety of different mechanisms, including the Fermi surface nesting [40], saddle band driven susceptibility divergence [61], f-wave gapping and marginal Fermi liquid [62], etc. Rossnagel et al. has, in addition, argued that the spin-orbit interaction is also important for 5d-electron materials like TaSe$_2$ and that the band structure calculation is able to reproduce $k$-dependent gap structure [63]. Moreover, recent LDA calculations have shown the difficulties with Fermi surface nesting and have suggested that PLD, instead of eMod, is the essential ingredient [42, 43]. Because of this debate, it is necessary to study the nature of the 2D CDWs in the TMDs with the simultaneous measurement of any intrinsic lattice distortion while spatially resolving the electronic component across this family of compounds.

We have examined three 2H-TMDs, TaS$_2$, TaSe$_2$, and NbSe$_2$. Figure 3.3a is a representative constant current topographic image of a 2H-TaS$_2$ surface showing both the atomic corrugation and the 3×3 CDW superlattice. TaS$_2$ is being studied here for the first time by a STM. Fig. 3.3b shows crystal structure with the neutral cleave plane (dashed line) between the two layers of chalcogen atoms. The atomic lattice shown in Fig. 3.3a is the resulting
Figure 3.3: (a) Atomically resolved topographic image of 2H-TaS$_2$ with CDW modulation (3×3 superstructure). This image is obtained at 60 K and with setup condition being $-100$ mV and 200 pA. Inset: zoom in to show the atomic structure of the 3×3 unit cell with red circles indicating the sulfur atoms. (b) Crystal structure of trigonal prismatic (2H) TMD. The red/black atoms are chalcogen/metal atoms, and the dashed line is where it cleaves. (c) Fourier transform of (a). Blue (red) circles indicate the primary peaks of CDW (atomic) modulation, and $b_1$ and $b_2$ are the atomic lattice reciprocal vectors.
triangular chalcogen lattice, while the evident $3 \times 3$ superlattice is due to the formation of the CDW. The Fourier transform (Fig. 3.3c) of Fig. 3.3a shows that the atomic and the CDW signals are the main features of this surface. After excluding the atomic corrugation by Fourier filtering, we find that the height of the top layer of atoms changes on the order of picometer in the $z$ direction due to the formation of CDW, for all three of the 2H-TMDs. This change is consistent with the distortion measured by both x-ray diffraction [64] and neutron scattering experiments [39]. This agreement strongly indicates that our surface sensitive measurements of the CDW show similar behavior as the bulk.

In order to probe the differences between the electron and hole (filled and empty) states of the CDW phases in these materials topographic images were taken at opposite sample biases. In Figs. 3.4 and 3.5, our topographic images obtained on 2H-TaS$_2$, NbSe$_2$ and TaSe$_2$ are shown. Figs. 3.4a and 3.4b are images of the filled and empty states (as indicated by negative and positive sample biases) in TaS$_2$ while Figs. 3.5a-b and 3.5e-f are images of the empty and filled states in NbSe$_2$ and TaSe$_2$. Prior to analysis, these images are processed with a drift-correction algorithm [65] so that images taken at different times can be aligned on a subatomic length scale. The surface defects are used to ensure accuracy of the drift correction algorithm. This process, combined with the high thermal stability of our experiment, allows for an extremely precise measurement of position (with less than 50 pm of drift per day) and enables us to uncover the true structure of the CDW state in these materials for the first time.

Comparing our results with those from known CDW systems [23, 24, 22], which include STM results on NbSe$_2$ [66], we find qualitative discrepancies. In our measurements, the filled/empty states topographic images of the TMD samples do not show the expected out-of-phase or spatial “inversion” that the Peierls model predicts. This contrasts with previous observations of 2D systems such as Pb (or Sn) coated Ge(111) [23, 24]. Instead, the topographic images taken at +50 and -50 mV sample bias voltages on TaS$_2$ are very similar to each other when one compares them with the naked eye. In order to rigorously check this...
Figure 3.4: Topographic images showing symmetrization and antisymmetrization on 2H-TaS$_2$ at 52 K. (a) and (b): Topographic images with −/+50 mV and 100 pA. These two images were individually corrected for drift and aligned with sub-atomic precision. (c) and (d): The symmetrized ($S = (a + b)/2$) and anti-symmetrized ($AS = (a - b)/2$) images of (a) and (b). The contrast in (d) is much smaller compared to (c). (e): The line profiles in (a-d) indicated by colors.
similarity, we have calculated the normalized cross correlation coefficients (NCC) between
two images \( A_{ij} \) and \( B_{ij} \),
\[
NCC = \frac{\langle (A_{ij} - \mu_A) (B_{ij} - \mu_B) \rangle}{\sigma_A \sigma_B}
\]
where \( \mu_A \) and \( \mu_B \) are the average values of the two images and \( \sigma_A \) and \( \sigma_B \) are the standard
deviations. NCC is expected to be +1 for two identical images, and −1 for two images
with perfect contrast inversion. For an eMod-dominating CDW, where spatial inversion
between filled and empty states is expected, such as Pb/Ge(111), one would expect the
NCC between the filled and empty states images to be close to −1. What we find when
we carry out this analysis on TaS\(_2\), NbSe\(_2\) and TaSe\(_2\) is that the NCCs are +0.62, +0.79
and +0.95 respectively, showing that the majority of the topographic images are in-phase
between the two biases. Considering that only the CDW corrugation will be out-of-phase
while the atomic corrugation is actually in-phase, we can remove the atomic signals by
Fourier filtering them and keep only the CDW signal. After removing the atomic signals,
the measured NCC are still +0.37, +0.67 and +0.97 for TaS\(_2\), NbSe\(_2\) and TaSe\(_2\) respectively,
in sharp contrast to the −1 expected for a traditional CDW. It is important to note that this
result differs from previous STM results on NbSe\(_2\) [66] where spatially out-of-phase electron
and hole components were seen. This discrepancy is due to our superior ability to control
and correct for drift at the subatomic level which allows us, for the first time, to uncover
the true nature of these materials. Interestingly our measured positive valued NCCs do not
agree with the conventional pure electronic CDW model [12] and instead show that there is
another non-electronic, lattice effect that is important to the formation of the 2D CDW in
TMDs.

PLD is a natural explanation for the positive NCCs we measure since STM measure-
ment is a combination of the atomic structure (via the tip-sample distance) and the integrated
density of states. If the ionic cores, around which the valence electron clouds are centered,
are displaced with the CDW transition, i.e. lattice distortion occurs, then the filled-state and
Figure 3.5: Topographic images of NbSe$_2$ and TaSe$_2$ at 5 K. (a/b): Topographic images of NbSe$_2$ with $-/+100$ mV and 100 pA. (e/f): topographic images of TaSe$_2$ with $-/+50$ mV and 40/30 pA. (c/g) and (d/h): Symmetrized and antisymmetrized images of (a/e) and (b/f). The contrast of (c/g) is much larger than (d/h), same as TaS$_2$. 
empty-state topographic images will be in-phase with each other. This is because they only 
enter into the tunneling current through the electron-hole symmetric matrix element. In 
other words, if there is only a periodic lattice distortion then the filled-state and empty-state 
topographic images will be the same, except with a possible difference in the amplitudes of 
the height profiles, which will not affect the NCC. This is exactly what we see in our STM 
images of the 2H-TMDs.

By calculating the symmetrized (Figs. 3.4c, 3.5c and 3.5g, \( S = (a + b)/2 \) where \( a \) 
and \( b \) are the images of filled and empty states) and anti-symmetrized (Figs. 3.4d, 3.5d 
and 3.5h, \( AS = (a - b)/2 \) images, we can directly access the PLD and eMod parts of the 
data respectively. Results of these calculations, shown in Figs. 3.4 and 3.5 using the same 
height scaling, clearly show that the symmetrized channel has a larger z-range than that of 
the anti-symmetrized one, agreeing with our measured NCCs above. The two channels (\( S \) 
and \( AS \)) reveal the decomposed, spatially resolved, PLD and eMod components of the \( 3 \times 3 \) 
superlattice for the first time. We find similar signals for the two other members of the 2H 
family that we have measured, NbSe\(_2\) and TaSe\(_2\). In the framework of Peierls instability, the 
lattice distortion acts as a secondary effect which has to be minimized because it increases 
the elastic energy of the ground state. Contrasting with the results of the previous STM 
study [66], we have demonstrated that the PLD is playing an important, if not the primary, 
role in the CDW formation in 2H-TMDs. This is calling out a strong entanglement of the 
PLD and the eMod, which is not to be expected in a purely electronic Peierls picture.

Our results agree with the calculations which show that without the lattice involvement 
a CDW in these systems cannot exist. It has long been argued that without lattice distortions 
electron correlations are too strong for a purely electronic CDW to form [41] and that there 
must be corresponding lattice distortions. Our observations clearly show that these lattice 
distortions are present and dominate the observed structure in the 2H-TMDs, which is in 
contrast to previous reports to the contrary [66]. In addition, recent theoretical calculations 
in this family of materials have shown that the observed CDW wave vectors do not correspond
Figure 3.6: Topographic images on 2H-NbSe$_2$ at $T = 6$ K. Scan size is about $20 \times 20$ nm$^2$. (a) and (b): Topographic images with -50 mV and +50 mV. (c) and (d): Symmetrized and anti-symmetrized image of (a) and (b).
to the peaks in the electronic susceptibility [42]. This calls for a more complicated model that includes nontrivial phonon involvement. Such complicated origins of the charge ordering transition in such a simple family of materials demonstrates the necessity of taking the lattice contribution into account in describing the phase diagrams of everything from the simple systems like TMDs, to the more complex correlated electron materials like cuprates and manganites.

The symmetrized topographic images do not contain the eMod signal because the eMod will cause asymmetric changes in LDOS (Fig. 2.4 in Ch. 2). Instead the symmetrized topographic images contain the PLD signal since it symmetrically affects the filled and empty states. It is possible, however, that the anti-symmetrized images do contain more information than just that of the eMod. In this sense, the relative intensity of PLD versus eMod we present is only a lower bound on the lattice contribution to the CDW. As we mentioned earlier, the topography is only sensitive to the energy window which is set by tunneling bias voltage, so high energy studies might reveal a more complete picture of CDW in 2H-MX$_2$. However, even with our low bias voltage images presented here, our study presents more than enough evidence to show the violation of the Peierls picture in the 2D CDWs in these materials. Furthermore, we have indeed checked these results by measuring over a range of bias voltages form 25 to 200 mV and have seen no inconsistencies.

3.4 Topographic images on Cu$_x$TaS$_2$

We have further extended our study to Cu intercalated TaS$_2$ sample, Cu$_x$TaS$_2$ ($x = 0.04$). In this sample, the copper atoms sit inside the van der Waals gap between the two S layers, and after cleavage they shows up as ad-atoms on the surface (Fig. 3.7a). Our first finding was that the copper dopants largely disturbs the 3 × 3 super structure (weakening the CDW order) which is consistent with the reduction of CDW transition temperature by copper intercalation [55].

These copper atoms are mobile under the strong electric field inside a STM tunneling
Figure 3.7: Topographic images with various bias voltages on Cu$_x$TaS$_2$ (nominal doping $x = 0.04$) at T = 6 K. (a) A large area (120 × 120 nm$^2$) scan of the Cu$_x$TaS$_2$ surface. (b-e) Images of the same area with various bias voltages, after the copper atoms were moved away from this area.
junction and therefore can be manipulated by the tip. To help understanding the role of copper atoms in the reduction of CDW order, we then removed ‘all’ of the copper from a small area. The removal of the surface copper does not recover the long range $3 \times 3$ order as those in pristine TaS$_2$, which points the local disorder to the Cu buried underneath the first sandwich layer of S-Ta-S.

This further allows us to check for the lattice distortion in such a disordered CDW system. Figs. 3.7b-e show the topographic images of this area with different bias voltages. It is easy to observe that these four topographic images look very similar to each other except the change of overall contrast. The change of overall contrast is also consistent with the undoped samples. The similarity between these images prove two things: 1. the PLD is playing the major role again in this disordered CDW system; 2. the copper atoms are likely not substituting the Ta due to the atomic uniformity at all bias voltages. Moreover, we found that the local CDW order is not changed very much before and after the surface layer copper atoms were removed, indicating that the copper atoms are altering the CDW by creating the local strain instead of charge doping. The absence of charge doping is also confirmed by the weak bonding between the copper and TaS$_2$ (easiness of moving the copper around the surface).

### 3.5 Summary

In summary, we have demonstrated the importance of a high resolution, high stability, low drift STM to the understanding of the structure and origins of CDWs in the TMDs. We have shown that atomic length scale measurements are necessary in order to resolve and separate the structures of both the PLD and eMod in TaS$_2$, NbSe$_2$ and TaSe$_2$. This shows that the PLD dominates the CDW transition for the entire family of 2H-TMD and that the eMod, which has been previously been seen as the dominating mechanism, plays a secondary role instead. These conclusions were further confirmed by data from a copper doped TaS$_2$ sample. Our results indicate that pure electron driven CDW is unlikely to exist
in the 2H-TMDs and that the PLD plays an important role, which is in excellent qualitative agreement with the recent theoretical studies [43].
4.1 Introduction

The charge-density wave (CDW) transition breaks the lattice translational symmetry and piles up electric charge periodically in real space. In the simplest model for this, Peierls [12] pointed out that this transition, purely driven by electronic instability, will lower the energy of electrons by opening up a gap at the Fermi level and driving the system from metallic to insulating state. After being successful with many 1D charge/spin density wave systems [12], this model itself and similar ones which rely on Fermi surface nesting have been widely applied to correlated electron systems [19, 67, 21, 20]. However, a good understanding of the two dimensional charge-density wave (2D-CDW) in the layered transition metal dichalcogenides (TMD, MX$_2$) [26] is still lacking leaving many unanswered questions. For instance, enhancement of electric conductivity with the CDW transition has been reported for many MX$_2$ systems including 2H-TaSe$_2$, TaS$_2$ and NbSe$_2$ [35, 46, 60, 45], opposite to the prediction of Peierls mechanism. Partial gapping on Fermi surface [68] is only able to explain the metallic state below $T_{\text{CDW}}$, but not the enhancement of conductivity. The sliding CDW picture also lacks direct experimental support for these 2D systems. Moreover, measurement of the Nernst effect on NbSe$_2$ [49] also gave surprisingly large signal accompanying the CDW transition that has not been fully explained. These unconventional properties call for alternative explanations for the 2D-CDW in the TMDs. From theoretical perspectives, challenge to the Peierls mechanism has been brought up by Chan and Heine [41], who ar-
gued that a CDW cannot be stable without lattice distortion. Recent work by Johannes et al. [42, 43] further suggests that the low-energy approximations, like the BCS-type mean field calculation [12], are not enough to describe the TMD-CDW systems, and states away ($\gg 3.5k_B T_{\text{CDW}}$) from the Fermi level are also important for the transition. Furthermore, van Wezel [69] has pointed out that orbital degree of freedom could play an important role for TaS$_2$.

Two dimensional charge-density wave (CDW) materials like 2H-TaSe$_2$ and TaS$_2$ belong to the multi-orbital correlated systems, similar to manganites [71], Fe-based superconductors [72], ruthenates [73], iridates [74], etc. According to theoretical calculations [70, 75, 76], the conduction band is composed of three hybridized d-orbitals, with the $d_{z^2}$ at the center and $d_{xy}/d_{x^2-y^2}$ on the edges of Brillouin zone. These calculations have been confirmed with Fermi surface measurements by ARPES [40]. The orbital texture of the conduction band has been argued as nontrivial [69]. However, to date there has been no experimental report on the orbital structure of the energy bands, likely due to the complicated structure of the particular d-orbitals involved here. Fortunately, the dramatic difference expected in the spatial extension of $d_{z^2}$ and $d_{xy}/d_{x^2-y^2}$ orbitals offers an excellent opportunity for real-space resolving probes to look for their ordering. One such probe, spectroscopic imaging scanning

Figure 4.1: Band structure of TaSe$_2$ from first principle calculation (black and blue open circles). The color of the circles indicate the characters of $d_{z^2}$ (black) and $d_{xy}/d_{x^2-y^2}$ (blue) orbitals. From: ref. [70]
tunneling microscope (SI-STM), measures electronic density not averaging the entire sample but the component that extends into the vacuum by 5-10 Å [5], where tunneling process can occur but no significant perturbation to the intrinsic property is induced by the tip. With a simple atomic limit model (see supplementary information), we have shown that the exponential decay length of $d_z^2$ is much larger than for the $d_{xy}/d_{x^2-y^2}$ states leading to a much larger overlap at the location of the STM tip. Therefore LDOS measurement will sample $d_z^2$ more than $d_{xy}/d_{x^2-y^2}$ orbitals allowing a spatial mapping of the different orbital states. Moreover, STM is able to image the periodic lattice distortion and the electronic component of the CDW state separately (see Ch. 3), and its LDOS measurement covers both filled and empty states.

Among the existing studies of 2H-MX$_2$ the change in local anion coordination of metal atoms with CDW formation are rarely taken into account, due to the small amplitude and difficulty to be measured precisely [64]. In previous work with STM topographic images, we were able to measure the lattice distortion directly (see Ch. 3). Since the 5d orbitals are spatially extending and thus the crystal field effect is strong, we argue that the small amount of lattice distortion is important, by proving a positive correlation between the local anion distortion and energy level shifts measured from LDOS.

### 4.2 Averaged LDOS on TaSe$_2$ and TaS$_2$

To investigate these issues we have carried out STM and STS experiments on 2H-TaSe$_2$ and TaS$_2$ at various temperatures. Here, we present spatially averaged LDOS spectra in both materials and address on the kink feature that has been ascribed as the CDW gap in the past [34]. We then use the unique ability of SI-STM to study the atomically resolvable LDOS on TaSe$_2$ and show the intra-unit-cell LDOS variation and its relation with the local lattice distortion. Comparable results on TaS$_2$ will also be presented.

In order to verify whether there is a gap formation with CDW, we have measured the LDOS by differential conductance ($dI/dV$) measurement with the standard lock-in method.
LDOS spectra in TaS$_2$ and TaSe$_2$ (Fig. 4.2a-c) show very similar overall asymmetry, with the negative bias side being almost constant and positive side increasing quickly with voltage. Around zero bias, all of our spectra display some dip features with roughly symmetric kinks on both sides of the dips (indicated by the black vertical lines). For TaSe$_2$ at 5 K, the kinks in the LDOS occur at ±32 meV which is almost identical to the energy scale of a similar feature reported on NbSe$_2$ [34, 51]. The similarity of these features in these two different materials is surprising if they are related to the CDW transition, since their $T_{CDW}$ differs by a factor of 3 (122 K for TaSe$_2$ and 33 K for NbSe$_2$). According to mean field calculation [12], the size of CDW gap should be proportional to $T_{CDW}$ ($2\Delta = 3.5k_B T_{CDW}$). Furthermore, as the temperature approaches $T_{CDW}$, one expects the energy gap to decrease, in sharp contrast to our observation that their energy increases with increased temperature (Fig. 4.2b-c). Additionally, the low temperature LDOS in TaSe$_2$ includes two more kinks at ±10 meV within the dip, which have never been reported. All of the evidences stated above do not support the 32 mV feature being the CDW gap. Soumyanarayanan et al. [51] have also shown that these kinks do not vanish above $T_{CDW}$. Instead of being related to the CDW, an interesting possibility is that the kinks in LDOS can be explained by phonon assisted inelastic tunneling, which is indeed able to give rise to weak step-like increase in LDOS [77], and phonon density of states of NbSe$_2$ and TaSe$_2$ do share strikingly similar
energy scales around 10 and 25 meV [44, 76].

Figure 4.3: A high energy LDOS obtained by averaging 3600 spectra taken over an area of 4 × 4 nm² and junction setup at −200 mV and 40 pA. The fast rising at −400 mV originates from the Se p-bands, and the hump around +380 mV indicates a van Hove singularity in the Ta d-band with \( d_{z^2} \) orbital characteristics.

A major feature in the LDOS spectra is the large asymmetry between filled and empty states. Asymmetry is common in tunneling experiments but its origin differs in each case so material details are necessary [78]. To study the asymmetry in TaSe₂ we firstly measured LDOS in a larger energy range. In the high energy LDOS (Fig. 4.3), two large features can be observed and related to the energy bands [70]. The first feature is the fast rising below −400 mV originated from the Se p-bands, since Se atoms are closer to the tip and thus enhanced in the tunneling process. The second one is the big hump around +380 mV, the edge of which extends towards the Fermi level and results in the ‘asymmetry’ seen around zero bias. We identify this hump with the van Hove singularity around the Γ-point with \( d_{z^2} \) characteristic. As we previously discussed the \( d_{z^2} \) orbital will be enhanced in STS measurements in addition to the expected enhanced DOS from a van Hove singularity. Due to the fact that conduction band gradually shifts its weight between \( d_{z^2} \) and \( d_{xy}/d_{x^2−y^2} \) orbitals and also the \( k_z \) dispersion, the broad nature of the feature seen around +380 mV in STS rather than a narrow one is expected.
4.3 Spatially resolved LDOS study on TaSe$_2$ and TaS$_2$

To further study the LDOS asymmetry, we have also measured the LDOS on a 96 $\times$ 76 pixel grid over an area of 5 $\times$ 4 nm$^2$ with the tunneling junction set at $-200$ mV and 300 pA, and we found that the asymmetry in LDOS varies in a subtle fashion with the periodicity of the 3 $\times$ 3 CDW superlattice. To quantify this asymmetry, we have calculated the ratio map (Fig. 4.4a) of hole versus electron density of states,

$$R(r, V) = \frac{\int_{0}^{eV} LDOS(r, \epsilon)d\epsilon}{\int_{-eV}^{0} LDOS(r, \epsilon)d\epsilon},$$

where $r$ denotes the two spatial coordinates. The ratio map indeed reflects the same periodicity with the topographic image (Fig. 4.4b), and there is no further inhomogeneity associated with any larger length scales. It became reasonable to focus on one CDW unit cell (drawn as the red diamonds in Figs. 4.4a, 4.4b and inset of 4.4c) and plot out the LDOS of different high symmetry points (defined in the inset of Fig. 4.4c) within. The major difference between the spectra is the variation of the overall slope (Fig. 4.4c). The normalized spectra (Fig. 4.4d) show no evident energy feature within the energy window of $\pm 200$ mV, implying that higher energy states are responsible for the CDW formation [43]. The lack of any feature around $\pm 32$ mV in the normalized spectra further supports that the kinks (Fig. 4.2a-c) are not related to CDW, since the BCS-type mean field calculation predicts the dominant change in LDOS to be at $\pm \Delta$ (see Fig. 2.4 in Ch.2).

More interestingly, we found that the ratio map is anti-correlated with the height map (topography): the ratio is high when the tip height is low and vice versa. The normalized cross-correlation number is $-0.65$. Considering the multi-orbital involvement [70, 75, 76], we have established a strong relation between the change of LDOS asymmetry $R(r)$ and the local anion distortion. The lattice distortion occurs with $\Sigma_1$ symmetry [39] in which the metal ions only move within the basal plane, and chalcogen atoms move in a mirror symmetric manner about the metal layer. Topographic image at $-100$ mV with the atomic corrugation filtered out (Fig. 4.5b) can be adopted to represent the distortion of the chalcogen atoms along
Figure 4.4: Ratio map and sub-unit-cell LDOS variation in TaSe$_2$ at $T = 5$ K. The scan size is $4 \times 5$ nm$^2$. (a) R-map as defined in Eqn. (4.1), with $V = 200$ mV. (b) Concurrent topographic image. Cross correlation number with (A) and (B) is $-0.65$, indicating an anti-correlation between them. (c) Averaged LDOS spectra over the equivalent sites of all the CDW unit cells. Inset: definition of point of interest (POIs) using the contours with different colors. POI 1–8 correspond to the colors from red to violet in a rainbow color scheme. LDOS spectra in (c) follow the same false color scale. (d) Same LDOS spectra in (c) normalized by the average LDOS of the entire region.
z-axis. Shifts in the angle between the M-X bond and z-axis (α in Fig. 4.5a) should predict shifts in the energy difference between $d_{z^2}$ and $d_{xy}/d_{x^2-y^2}$. When the angle α increases (left prism in Fig. 4.5a), the energy of $d_{z^2}$ ($d_{xy}/d_{x^2-y^2}$) will decrease (increase) and vice versa, similar to the Jahn-Teller effect. Within this picture the correlation between the height of chalcogen atoms and the corresponding LDOS asymmetry becomes natural. As shown in Figs. 4.4a and 4.4b, the ratio is enhanced in the left half of the CDW unit cell indicated by the diamond, which corresponds to a lowered $d_{z^2}$ level (and hence lower shoulder near the Fermi level) with the increased angle α (Chalcogen atoms move towards the Ta layer). The opposite situation is seen on the right half of that unit cell. The horizontal shift of the edges at +150 mV gives an approximate change in the local energy of the $d_{z^2}$ band, which is about 12 meV from lowest to highest. Interestingly, this value is very close to $k_B T_{\text{CDW}}$ of TaSe$_2$. This altering of the energy levels of the different orbitals in such a spatially periodic manner will accordingly result in a population imbalance between them, namely a partial orbital ordering.

Figure 4.5: (a) Local distortion to the prism that will cause the energy levels to shift. (Left/Right side) When the prism is compressed/elongated along the c-axis and expanded/compressed in basal plane, energy of $d_{z^2}$ will be lowered/increased. (b) Constant current image with atomic corrugation filtered out serves as a guide for local distortion in chalcogen layer.

Similar to the Jahn-Teller effect in manganites [71], the subtle change in local coordination will cause energy shift in the d orbitals and result in a partial orbital ordering.
The main difference between a charge ordering and an orbital ordering is that in the former electrons shift from one lattice site to another, while in the latter they transfer from one orbital to another on the same atom. Moreover, if the M-M bond length is also reduced [39], the overlap of the $d_{xy}/d_{x^2-y^2}$ orbitals will be further enhanced [63]. Both the partial orbital ordering and the enhancement of hopping could reduce the energy of electron system by pulling down the filled $d_{xy}/d_{x^2-y^2}$ band that extends deeply below $E_F$ [70], in accordance with high energy electron being involved [43]. It is a strong possibility that this orbital ordering will explain the enhanced conductivity as we mentioned earlier, since it can enhance the in-plane hopping of electrons. In the topographic image (Fig. 4.5b) we can see that the areas with enhanced in-plane hopping are connected, which is consistent with an increase of conductivity. It will also likely change the electronic scattering rate (for part of the Fermi surface) and result in the large Nernst signal related to CDW formation in the multi-band system NbSe$_2$ [49].

Figure 4.6: $10 \times 10$ nm$^2$ LDOS ratio map of TaS$_2$ at $T = 52$ K. The upper part is at a region with CDW misalignment, and the lower part shows very similar ratio structure with TaSe$_2$. 
Furthermore, we have performed similar experiments in TaS$_2$ and the intra-unit-cell variation of LDOS is very similar indicating that the orbital degree of freedom has a generic effect on the formation of CDW in this class of materials.

It is indeed possible that the asymmetry is varying due to variations in the tip height, if the filled and empty states decay into the vacuum with drastically different characteristic lengths. However, this possibility has been ruled out by comparing the asymmetry with another data set taken at a different setup tip height (Fig. 4.7) which shows identical results.

![Figure 4.7: Averaged LDOS spectra with two different setup conditions: $-200$ mV $300$ pA and $-100$ mV $100$ pA. The average difference in tip height is about $15$ pm, while the asymmetry difference is only $1\%$.](image)

**4.4 Summary**

In conclusion, we have shown that the common kink feature around $32$ mV is not related to CDW and suggested they can be explained by phonon-assisted inelastic tunneling. By correlating the LDOS asymmetry with the lattice distortion, we have provided strong evidences for a partial orbital ordering in the 2H-MX$_2$. Such an orbital ordering, although not as evident as the one in manganites, still has important implication on the unexpected transport behaviors and is very likely associated with the driving mechanism of CDW in 2H-MX$_2$. The discovery of this orbital ordering benefits from the tunneling process heavily favors one of the orbitals, which initially seemed to be a drawback. This methodology will be useful for future study of orbital physics in other correlated electron systems.
5.1 Introduction to Mott physics

5.1.1 Mott insulator

Mott insulators are insulating materials due to strong electron-electron correlation effects, which according to the conventional band theory are ought to be conductors with partially-filled conduction band [79]. Here, the electron-electron correlation effect usually refers to the intra-atomic (on-site) Coulomb repulsion interaction between electrons, defined as follows.

\[ U = \int \int |\phi_1(r_1)|^2 \frac{e^2}{r_{12}} |\phi_2(r_2)|^2 dr_1 dr_2 \]  

(5.1)

where \( \phi_1 \) and \( \phi_2 \) are the wave functions of the two electrons under consideration and \( r_{12} \) is the distance between them. From Eqn. (5.1), we can see why \( U \) will be large if the two electrons occupy orbitals that largely overlap with each other or (for most of the times) the same orbital. Furthermore, \( U \) will also be increased if the orbital is constrained to little space, i.e. localized, due to the decreased average distance between electrons. For instance, 3d electrons usually have much larger \( U \) than 4d and 5d electrons.

Early studies were related to explain the insulating behavior in many 3d transition metals, like CoO, where the localized the 3d orbitals play an important role of raising the on-site \( U \). Relatively recent studies focused on ternary transition metal compounds like La\(_2\)CuO\(_4\) and LaMnO\(_3\), where interesting physics like unconventional superconductivity and
colossal magneto-resistivity arise when the compound is chemically doped cite [80].

A simple view of the Mott insulator could be established by considering an artificial sodium crystal with N atoms, where the lattice constant $a$ is tunable. In this model the energy cost of double occupancy ($U$) is constant, but the electron kinetic energy $t$ is increasing with decreasing $a$, and hence the energy bandwidth ($W$). The energy gain of forming a half-filled band can be understood via the the process of two sodium atoms forming one Na$_2$ molecule; the two electrons only occupy the bonding orbital and leave the anti-bonding orbital empty and gain energy $2t$. Following this reasoning, one can find that energy gain by forming a half-filled band in the sodium crystal is $1/4NW$. At the mean time, since electrons distribute randomly between the lattice sites, a certain number of lattice sites will be doubly occupied and will increase the system energy by $1/4NU$. It could be naively concluded that a transition occurs at,

$$U = W.$$ (5.2)

When $U < W$, the system favors forming the band so it is metallic; when $U > W$, the system favors singly occupied state so it is insulating.

Figure 5.1: Density of states change during a Mott transition. When $U < W$ the system is metallic and when $U > W$ there is an energy gap between the lower and upper Hubbard bands.
5.1.2 Hubbard model

A commonly used theoretical model for describing Mott insulators is the Hubbard model,

\[ H = -t \sum_{\langle i,j \rangle} \left( c_i^{\dagger} c_j + c_j^{\dagger} c_i \right) + U \sum_i n_{i\uparrow} n_{i\downarrow}, \]  

(5.3)

where the \( c_i^{\dagger} \) and \( c_j \) are the electron creation and annihilation operators on lattice sites \( i \) and \( j \), \( \langle i, j \rangle \) means nearest neighboring sites and \( n_{i\sigma} \) is the number operator on lattice site \( i \).

This model gives the minimum description for realistic correlated systems, but is not exactly solvable in cases like the 2d/3d square lattice. However, some qualitative picture can be obtained by analyzing the limits of \( U \gg t \) and \( U \ll t \). When \( U \) is small, the system is simply metallic with bandwidth \( W = 2zt \), where \( z \) is the coordination number. When \( U \) is large, the system is insulating and energy gain via \( t \) is on the order of \(-t^2/U\), the result of a second-order perturbation, which only happens when neighboring electron spins are anti-parallel. This is the reason why a lot of Mott insulators are anti-ferromagnetic, as a result of strong electron-electron correlation.

5.2 The novel \( J_{\text{eff}} = 1/2 \) Mott insulator \( \text{Sr}_2\text{IrO}_4 \)

5.2.1 Spatial extensions of d orbitals

The fairly localized 3d electrons have large el-el correlation, which is lowered in the delocalized 4d and 5d electrons. The spatial extensions of d-orbitals with different principal quantum numbers could be seen in Fig. 5.2, with their radial wave functions being,

\[ R_{3d} = \frac{1}{9\sqrt{30}} \rho^2 Z^{3/2} e^{-\rho/2}, \]

\[ R_{4d} = \frac{1}{96 \sqrt{5}} (6 - \rho) \rho^2 Z^{3/2} e^{-\rho/2}, \]

\[ R_{5d} = \frac{1}{150 \sqrt{70}} (42 - 14 \rho + \rho^2) \rho^2 Z^{3/2} e^{-\rho/2}, \]

(5.4)

where \( \rho \) is the reduced radius \( \rho = 2Zr/na_0 \), \( Z \) is the effective nuclear charge, and \( a_0 = 0.53\text{Å} \) is the Bohr radius.
In Fig. 5.2, we have used effective charge $Z = 5$ for all three of them. We clearly see how the d orbital extensions change with respect to the principal quantum number $n$. The 5d orbital is almost 3 times (in length) large than the 3d orbitals; with a lattice constant of $a = 4\text{Å}$ which is reasonable most of the transition metal oxides, the 3d electrons are well isolated from each other, while the 5d electrons will be ‘touching’ one another. It should be noticed that the size of the orbitals are scaled by the effective nuclear charge, so it could vary a lot between elements even with the same principal quantum number $n$, or even the same element but with different oxidization states or crystal environment.

Following Eqn. 5.1, we can make some ‘naive’ comparison of the on-site Coulomb repulsion energies: $U_{3d}$ equals to a few times of $U_{5d}$. However, the exact numbers for ions in real solids are not easy to figure out from first principle and they will vary from one element to another even with the same principal quantum number. In reality, Hubbard-U is treated as an adjustable parameter for calculations for the most of the time.

### 5.2.2 Spin-orbit coupling

Spin-orbit interaction can be understood as the result of spin alignment in the effective magnetic field felt by the electron, while the magnetic field is introduced by Lorentz transformation of the nuclear electrical field. It is one of the relativistic effects inside atoms and
can be describe by the following Hamiltonian,

\[ \mathcal{H} = \zeta_{nl} \boldsymbol{L} \cdot \boldsymbol{s}, \]  

(5.5)

where \( \boldsymbol{L} \) and \( \boldsymbol{s} \) are the angular momenta of orbital and spin, and,

\[ \zeta_{nl} = \frac{\hbar^2}{2m^2c^2} \left( \frac{1}{r} \frac{dV_{nl}}{dr} \right), \]  

(5.6)

is an integral involving the electric field strength and the radial wave function \( R_{nl}(r) \). The exact form of this integral can be found in text book [81]. As pointed out by Tinkham [81], the integral of \( \zeta_{nl} \) favors the interior of the atoms, and hence the screening effect by other valence electrons is much less than when one is trying to compute the out-most electrons’ wave functions. As a result, \( \zeta \) increase rapidly with respect to \( Z \), \( \zeta \sim Z^4 \). For iridium with \( Z = 77 \), \( \zeta \) is about 0.5 eV in the atomic limit [82].

However, it does not necessary mean that spin-orbit coupling will always be large in heavy elements, since more often than not the orbital angular momentum is quenched in a solid via crystal field splitting of the \( nl \) manifolds. In the case of \( \text{Sr}_2\text{IrO}_4 \), the \( t_{2g} \) manifold with \( d_{xy}, d_{yz} \) and \( d_{zx} \) orbitals serves with an effective angular momentum \( l_{\text{eff}} = 1 \). It is the coupling of this effective angular momentum with the spin that causes the interesting physics we are talking here in this chapter.

### 5.2.3 Spin-orbit Mott insulators

As we mentioned above, the correlation effect (\( U \)) decreases in the d orbitals when the principal quantum number \( n \) increases. It is for this reason that most 4d transition metal oxides are good metals (e.g. \( \text{Sr}_2\text{RuO}_4 \) and \( \text{Sr}_2\text{RhO}_4 \)), unlike their 3d cousins (e.g. \( \text{La}_2\text{CuO}_4 \), \( \text{LaMnO}_3 \)). Iridium oxides like \( \text{Sr}_2\text{IrO}_4 \) and \( \text{Na}_2\text{IrO}_3 \), with 5d electrons, were expected to be more metallic than the 4d transition metal oxides. However, many iridium oxides, including the two we mentioned, are insulators [83, 84, 85, 86]. B. J. Kim et al. proposed the SOC-induced Mott insulating idea after studying \( \text{Sr}_2\text{IrO}_4 \) with angle-resolved
photoemission spectroscopy (ARPES), optical conductivity and various other methods [74].

In this picture, the $t_{2g}$ manifold is further split by SOC into two multiplets, a $J_{\text{eff}} = 1/2$ doublet and a $J_{\text{eff}} = 1/2$ quartet, with the atomic limit separation between them being $\frac{3}{2} \zeta$. With the 5d$^5$ electron configuration, the quartet band with lower energy is fully occupied and the doublet band become half-filled with a much smaller bandwidth (W). In this limit, a small electron correlation will cause the system to undergo Mott transition band become an insulator, as illustrated in Fig. 5.3. It is then very similar to the square lattice Hubbard model with exactly half-filling of the band.

After the discovery of this Mott insulator, physicists quickly realized that Sr$_2$IrO$_4$ is a very close analogy to La$_2$CuO$_4$, the parent compound of high-Tc superconductor [87]. This leads to the proposals like achieving superconductivity in this material via chemical doping [88, 89]. More interestingly, this material, together with the 213 compounds, are possible testbeds for exotic spin models like the quantum compass and Kitaev models [90, 91, 92].

However, one of the similarities to La$_2$CuO$_4$, which relies on Sr$_2$IrO$_4$ being in the strong SOC limit so that only the $J_{\text{eff}} = 1/2$ band need be considered, becomes less certain when
one realizes that the electron hopping energy \( t \approx 0.3 \text{ eV} \) \([88, 89]\)) is close to the SOC energy \( \zeta \sim 0.5 \text{ eV} \) \([82]\). Furthermore, instead of the Mott-type insulator, the magnetically originated Slater insulator has also been proposed \([93]\), supported by time-resolved optical measurements \([94]\).

A key to identify the nature of this insulator is to answer the important questions of what is the size of the insulating gap, \( \Delta \), and how does it compare to other energy scales (e.g. \( t \) and \( \zeta \)). Surprisingly, measurements of gap size from different probes in \( \text{Sr}_2\text{IrO}_4 \) vary widely. Gaps of \( \sim 0.1 \text{ eV} \) have been reported by fittings of the resistivity data with the thermal activation model \([95, 96]\) and by calculations using local density approximation (LDA) with SOC and \( U \) \([74, 97]\). On the other hand, angle-resolved photoemission spectroscopy (ARPES) measurements do not agree with this value since the maximum of the valance band is already lower than \( -0.1 \text{ eV} \) \([74, 98]\). Optical conductivity and resonant inelastic x-ray scattering (RIXS) studies both yield \( \Delta \leq 0.4 \text{ eV} \) \([87, 99]\). Importantly, these measurements are (either) indirect, susceptible to signals from other excitations like exciton or magnon, and/or can only measure the filled states, making the interpretations of their gap values difficult. It is highly desirable to measure the gap with an experiment that directly probes the free charge carriers.

5.3 Study of insulating gap in \( \text{Sr}_2\text{IrO}_4 \) with STS

STS is a direct probe for the electronic density of states and hence the insulating gap (if any) of a material, since it utilizes the single particle tunneling effect and is free of excitonic effect or magnetic excitation. Measurements of insulating gaps have been demonstrated in various places \([100, 101]\). We have performed STM and STS experiments on \( \text{Sr}_2\text{IrO}_4 \) at the temperature of 80 K with our UHV-STM system.

\( \text{Sr}_2\text{IrO}_4 \) has the \( \text{K}_2\text{NiF}_4 \)-structure and can be cleaved between the two adjacent and identical SrO layers. Fig. 5.4a shows a \( 20 \times 20 \text{ nm}^2 \) constant-current image taken with \( -300 \text{ mV} \) and \( 5 \text{ pA} \), in which we can see the Sr square lattice in the top layer (Fig. 5.4b). The Fourier transform (Fig. 5.4c) of this topographic image shows four strong atomic peaks,
confirming the high quality of this surface. The oxygen atoms do not appear in topography likely due to the tunneling voltage being far away from the O 2p levels at \(-3.3\) eV [102]. Together with the atomic lattice, the sample also shows a small number of defects with the most evident type being the dark patches (see below).

Figure 5.4: Topography and atomic structure. A 20×20 nm² topographic image taken with \(-300\) mV and 5 pA at 80 K (a) and its Fourier transform (c). The black arrows indicate the \(a\) and \(b\) axes of the crystal lattice. Red and blue arrows indicate two Sr-related defects. (b) Schematic diagram of the first SrO and second IrO₂ layers of a cleaved surface. The square lattice in (a) correspond to Sr in top layer. The image has been processed to remove vibrational noise at a certain frequency.

Shown in Fig. 5.5a is a LDOS spectrum measured by averaging ~1300 spectra taken over a 2×2 nm² area that did not include any apparent defects. The negative sample-bias side of the LDOS, corresponding to the occupied states, shows no distinct feature that can be related to the lower Hubbard band (LHB) of \(J_{\text{eff}}=1/2\) band down to \(-1\) eV (Fig. 5.5b), contrasting to the double-peak structure observed in optical conductivity measurement [99]. However, optical measurements in insulators are sensitive to excitons [103] and Kim et al. [104] have pointed out that the double-peak seen in the optical conductivity cannot be due to the \(J_{\text{eff}}=3/2\) (quartet) and LHB of \(J_{\text{eff}}=1/2\) bands. In Na₂IrO₃ and Li₂IrO₃, the importance of excitonic effect has been shown by a recent RIXS study [105]. It is therefore
strongly possible that the LHB of \( J_{\text{eff}} = 1/2 \) band is overlapping with rather than separate from the quartet band, in good agreement with calculations done by LDA+DMFT [93] and variational cluster approximation [89]. According to Watanabe et al. [89], the LHB is entirely enclosed by the quartet band. This overlap between LHB and quartet bands implies that the strong SOC limit and single band models are inadequate for \( \text{Sr}_2\text{IrO}_4 \) and many-body effect should be taken into account.

The insulating gap from our tunneling spectra is about 620 meV, with the valence band top at \(-150\) meV and conduction band bottom at \(+470\) meV (Fig. 5.5a). Here the energy gap is defined as the size of the energy window with no coherent density of states. The energy of the valence band top \((-150\) meV) is in excellent agreement with the one measured by ARPES [74, 98]. The slow rise in the LDOS between 0.2 and 0.5 eV varies with tunneling location and hence we believe it is extrinsic either due to local impurity states or varying tip-induced band bending. LDOS measurements on a set of points which are all far away from any defects show that the slow rise only appears in some (Fig. 5.5c) but not the rest of the spectra (Fig. 5.5b). On the other hand, the gap edges at \(-150\) and \(+470\) meV are independent of location. Therefore, the large 620 meV gap is due to the underlying nonlocal, energy bands and the inhomogeneous in-gap states are caused by the local defects. Furthermore, LDOS measurements in Rh-doped sample \( \text{Sr}_2\text{Ir}_{1-x}\text{Rh}_x\text{O}_4 \) \((x = 0.04)\) [106] also confirm this by showing that far from the intentional impurities a similarly sized large gap is present.

However, this large gap is not in good agreement with the gap values reported so far. In the similar material \( \text{La}_2\text{CuO}_4 \), the debate about the gap size has been lasting for long time [107, 108, 109] because of previously similar subtleties. The difficulty in extracting the correct energy scale of the intrinsic gap rises when other processes are involved, as stated above. STS, on the other hand, can measure the Mott gap directly through the single-particle tunneling process. It does require a tunneling current going through the sample and may have difficulty for materials with poor electrical conductivity, but we have overcome
Figure 5.5: The large intrinsic energy gap. (a) LDOS measured by averaging $36 \times 36$ spectra over the $2 \times 2 \text{ nm}^2$ area indicated by the black square in the topographic image (inset). The dashed lines are drawn to indicate the band edges at $-150 \text{ mV}$ and $+470 \text{ mV}$. Data was taken with $-300 \text{ mV}$ and $10 \text{ pA}$ and ac modulation of $8 \text{ mV}_{\text{rms}}$ at $2.2 \text{ kHz}$. (b) and (c) LDOS taken at different ‘intrinsic’ locations away from the defects, (b) without and (c) with the slow rise. (d) Diagram showing energy bands with two important features: the $620 \text{ meV}$ insulating gap and the overlap between the LHB and the $J_{\text{eff}} = 3/2$ bands.
this problem by performing the experiment at liquid nitrogen temperatures and with very low tunneling currents, minimizing the tips’ electric fields.

This 620 meV gap strongly implies that Sr$_2$IrO$_4$ is a relatively weak Mott rather than Slater insulator, since the magnetic coupling energy is merely 60–100 meV [87, 110], too small to account for the large gap. Additionally, this is also consistent with the transport measurement showing no insulator-to-metal transition at least up to 600 K [111]. Although the size of this gap is considerably smaller than the 3d cuprates [101, 107], it is still large compared to what one would expect from the extended 5d orbitals and the fact that Sr$_2$RhO$_4$ is actually a paramagnetic fermi liquid [112, 113]. The band minimum of the upper Hubbard band (UHB) of the $J_{\text{eff}} = 1/2$ band, is the major discrepancy between our STS result and the existing studies. LDOS measurements show that the band minimum of UHB is at +470 meV, implying that LDA calculation needs a larger Hubbard-U to capture this feature since it has been shown that the insulating gap depends on U sensitively [93]. In a recent study on Na$_2$IrO$_3$, Comin et al. [114] have shown that using U as large as 3 eV in LDA calculation is necessary to reproduce the 340 meV gap in that material. With this work a stronger than expected correlation effect has now been observed in two iridates. The measured gap in Sr$_2$IrO$_4$ is nearly twice the measured gap seen in Na$_2$IrO$_3$, indicating that in the material studied here the 5d electrons are even more correlated (larger U/W).

The spatially extended nature of the 5d electrons can indeed largely reduce the on-site Coulomb repulsion, but at the meantime will increase the same type of interaction between the electrons sitting on neighboring Ir sites due to the increased overlap of their wave functions. The neighboring Coulomb interaction had been theoretically studied in the extended Hubbard model [115]. For Sr$_2$IrO$_4$, the Wannier functions calculated by Jin et al. [97] do show that a significant amount of electron weight is distributed on the four neighboring sites. It is thus strongly possible that the symmetric arrangement of the neighbors can result in a net repulsion that appears to be on-site. Interestingly, Ir$^{4+}$ in Sr$_2$IrO$_4$ has four nearest neighbors while it has only three in Na$_2$IrO$_3$, in agreement with their gap sizes, if the neigh-
Figure 5.6: (a-d) Images of defects with a wide range of junction resistances: $3.3 - 100 \, \text{G}\Omega$. The + signs indicate defects looking differently with different tunneling conditions. (e) LDOS measured on a dark spot (defined as the red cross in (a)). These images have been processed to remove vibrational noise at a certain frequency.
boring interaction is indeed causing a larger Hubbard-U. Moreover, Ba$_2$IrO$_4$, another iridate insulator [85] with a larger bandwidth W by not having the octahedra rotation, further confirms the strong entanglement of the energy scales in the 5d electrons.

5.4 Large insulating gap measured in Sr$_2$Ir$_{0.96}$Rh$_{0.04}$O$_4$

We have also measured a similarly sized gap in some area of a Rh-doped sample Sr$_2$Ir$_{0.96}$Rh$_{0.04}$O$_4$ (Rh4). In Fig. 5.7 we have shown a topographic image of the Rh4 sample and a gap map for the same area. The image in (a) was taken at -800 mV (20 pA), while the gap size in (b) was defined by energy locations where the measured dI/dV spectra pass a threshold value (0.05 in the this case). In Fig. 5.7c, we see a dI/dV curve measured by averaging the measurements on areas defined by green circles/ellipses in (a). This dI/dV clearly shows that the large gap survives at some areas of the Rh4 sample when they are far away from the dopants.

5.5 Study of impurity effects on insulating gap

To understand the effects caused by the defects, we have further studied topography at different junction resistances and LDOS on and off the defects. We see two types of Sr-related defects in the topography: Sr vacancy (the faint feature indicated by the red arrow in Fig. 5.4a) and Sr ad-atom (the bump in between four Sr atoms indicated by the blue arrow in Fig. 5.4a). These two types of defects could be caused by the cleaving process, but have little effects on the local electronic structure. In stark contrast the third type of defect we see, the dark patches, has a large effect on the local electronic structure. We identify these dark-patch defects in Sr$_2$IrO$_4$ as being excess/deficiency oxygen atoms, similar to those seen in manganite [116] and cuprate [117, 118]. This is in accord with the fact that Sr$_2$IrO$_4$ can have various levels of oxygen deficiency [119, 120], but more work is needed for a definitive answer.

Topographic images (Fig. 5.6) with junction resistance ranging from 3.3 to 100 GΩ
Figure 5.7: Large insulating gap measured on Sr$_2$Ir$_{0.96}$Rh$_{0.04}$O$_4$. (a) A $36 \times 36$ nm$^2$ topographic image of Rh-doped Sr$_2$IrO$_4$, (b) Map of gap size measured on the same area as (a) with the red squares indicating the locations of Rh dopants. The gap here is defined the points where dI/dV spectra crosses a threshold value. (c) A representative dI/dV curve that looks similar to that measured on the parent compound.
confirm that the patchiness originates from defects that are related to electronic inhomogeneity other than structure corrugation. The cross signs in Fig. 5.6a-d indicate the areas that look differently with different tip-sample separations. Such a high sensitivity in topography with respect to the tunneling condition reflects the fragile insulating nature of Sr$_2$IrO$_4$ in the presence of perturbation, and also the potential influence to the sample brought by O defects. Additionally, it is known that for some transition metal oxides, oxygen atoms in surface layers are volatile, but we didn’t observe any degradation of surface quality in several days at 80 K implying the stability of the surface. Therefore, these O related defects are likely intrinsic to the sample, in accord with O deficiency found in the system [119, 120]. This result naturally explains why transport measurements give a smaller gap [95, 96] and optical conductance shows a slow rise starting at 0.2 eV [99]. The presence of these defects leads to in-gap states, lowers the effective gap measured by transport and gives a lower energy scale for optical transitions. More importantly, the random distribution of the defects is in good agreement with the variable-range-hopping behavior in electrical conductivity [84]. The existence of intrinsic defects despite the good sample quality and our ability of measuring the intrinsic LDOS apart from the influence of defects, prove the necessity of using a local probe like STM to study this material.

Lastly, the LDOS study on top of a defect (Fig. 5.6e) shows how strongly it can change the density of states around $E_F$ by building up in-gap states. It should be admitted that, due to the dramatic difference in topography of this defect as we mentioned earlier, the STS measurement may not be reflecting the underlying LDOS accurately, but it should be sufficient to serve as evidence for the in-gap states. Fig. 5.8 shows the subtle variation in the measured spectra related to the tip-induced band bending, calling out the necessity of measuring with high junction resistances.
Figure 5.8: LDOS measured at a normal lattice site with different junction resistances: 10 GΩ (purple) and 1 GΩ (green). The difference could be accounted by tip-induced band bending. Inset: topographic image with the cross sign defining the spot of data taking.
5.6 Summary

As a summary, we have presented the first atomic-scale study in Sr$_2$IrO$_4$ with density of states measurement covering the entire Mott gap, which is 620 meV. The measurement of the occupied states has confirmed the overlap between LHB and quartet band. We concluded that Sr$_2$IrO$_4$ is a Mott but not Slater insulator by comparing the gap size with the magnetic coupling energy, implying the stronger-than-expected el-el correlation in this 5d system. Furthermore, we suggested that additional source of Coulomb repulsion need to be taken into account to quantitatively understand the 5d electron systems. Lastly, the junction dependence of the defects in the topography, together with the LDOS results about the defects, indicate their profound impacts on the local electronic structure of a 5d system. Our results strongly imply that in Sr$_2$IrO$_4$ the three energy scales (SOC, U and W) are strongly competing against each other and that studying the 5d electrons should put all of them on equal-footing.
Appendix A

Vibration test of Room E2B37

We have measured the background vibration of the room E2B37 with a geophone. Since the geophone is only sensitive to noise above 1 Hz, the measurements here were subjected to this low frequency cut-off. The voltage signals coming out of the geophone were then connected to a Fourier spectrum analyzer (SR760).

The tests have been done with several different conditions: normal is with people working inside the room; quiet is with all of the electronics turned off and only the test person staying inside; noisy outside is with people jumping in the hallway out of the room; and noisy inside is with people jumping inside the room. Furthermore, prior to the construction of our STM lab, there was a concrete pad on the floor that was isolated from the rest of the floor.

Figure A.1: Vibrational test of Room E2B37 measured by the geophone SN-1211 (Geo Space Corporation).
The pad was about 3 by 5 feet in area and 3 feet in depth into the ground. This pad was also tested for vibration. The test results of all of the above conditions were summarized in Fig. A.1. One major noise component in this room is at 30 Hz. As we can see the difference between ‘floor’ and ‘pad’ is mainly the overall shape of the vibrational noise: the pad tended to suppress high frequency vibration but enhance low frequency components (less than 60 Hz). Because STM experiment is very sensitive to noise from DC to a few hundred Hz, we decided that we won’t be able to benefit from the pad. It is also clear that noise on the hallway outside the room E2B37 is not increasing the vibration inside very much. Only when the noise is originated inside the room, there is a large enhancement of the vibration at high frequencies.
Appendix B

Pressure vs. time in a UHV pumpdown

In order to achieve ultra-high vacuum (UHV), we firstly need to pay attention to the materials that are used inside the vacuum chamber as we discussed in Ch. 1. An important criteria for choosing the materials is that they need to be sustainable with high temperature bake-out (at least up to 150 °C) without increased outgassing, chemical decomposition or losing mechanical strength. The bake-out process is crucial for achieving UHV environment; without a bake-out the UHV-STM system is only able to achieve $\sim 10^{-9}$ torr, but with a bake-out the base pressure is decreased by a factor of 10. With the further pressure reduction by cooling down the two dewars, the ultimate pressure of the system is about $5 \times 10^{-11}$ torr.

Shown in Fig. B.1 is a plot of pressure versus time for a normal bake-out process. The turbo pump is turned on at the beginning, with the ion pump and the bake-out heating turned on the next day. Here turning on the ion pump means opening the gate valve between the chamber and the ion pump, while the ion pump itself was powered on all the time and pressure sitting below $10^{-10}$ torr. For the first day the entire chamber is only pumped via the turbo pump connected to the load-lock with the pressure ended up with low $10^{-6}$ torr range. When the ion pump is turned on (opening of the gate valve between the ion pump and the STM chamber), the pressure dropped to $2 \times 10^{-7}$ torr immediately.

The next step is to turn on the heat for the bake-out, which results in the slow rising of chamber pressure back to $10^{-6}$ torr during the following few hours, with the ultimate pressure of the bake-out being $\sim 10^{-7}$ torr. During the bake-out the chamber pressure
Figure B.1: Plot of pressure versus time of the UHV-STM system during a bake-out process. The turbo pump was turned on at the beginning, while the ion pump and the bake-out heat were turned on at the second day. The pressure spikes were due to TSP flashes, which were made more frequent towards the end of the bake-out. The barely visible down-turn of the pressure at the very end of the curve is due to the termination of the bake-out heating, which resulted in a pressure decrease to low $10^{-9}$ torr within a day. Further pumping, together with TSP flashes, reduced the chamber pressure to $2 \times 10^{-10}$ torr prior to the cooling down with LN$_2$. 
decreased slowing to low $10^{-7}$ torr range in several days, which is mainly limited by the turbo pump speed and the low pumping conductance via the load-lock. If there were an additional turbo pump that is attached the chamber itself, this process will be largely shortened. In the process of the bake-out, the chamber temperature were raised to $\sim 120 - 140$ °C in a few hours but the STM scan-head temperature increased to $\sim 120$ °C in about 2 to 3 days. Further bake-out after this point won’t help reducing the pressure by much.

During the bake-out shown here in Fig. B.1, the pressure was further decreased using more frequent TSP flashes. During the normal operation when there is no heat on the chamber, one TSP flash is usually with 48 A current through the TSP filament for 1 to 2 minutes, with the frequency being once per day or even less. However, during a bake-out, the TSP cartridge inside the ion pump is already at temperature of 200 °C without any current, so current of 48 A will burn the TSP filament very quickly. Instead of 48 A, we have used current like 35 A which is sufficient for baking the filament itself. If we need to utilize the pumping power of the TSP, we usually use current of $\sim 43$ A.

![Figure B.2: Plot of pressure versus time of the UHV-STM system during normal operation.](image-url)

A pressure versus time chart is shown in Fig. B.2. We can see that the base pressure
during this period of time is about $7 \times 10^{-10}$ torr instead of $5 \times 10^{-10}$ torr, since only LN$_2$ were used and the STM temperature was set to 80 K. When the turbo pump is turned on for pumping the load-lock, the chamber pressure will increase slowly to $1 \times 10^{-10}$ torr. This slow rise of pressure is probably associated with the mechanical vibration of the turbo pump which can expedite the outgassing of gas molecules off the chamber. Every time when the gate valve between the load-lock and the chamber was opened for sample manipulation, there was a pressure spike ($2 - 5 \times 10^{-10}$ torr) even though the load-lock had been pumped for hours. The last part of that curve was showing the process of the outer dewar warming up due to outage of LN$_2$. This happens if the LN$_2$ is not refilled within 72 hours from the previous filling.
Appendix C

Feedback stability

We have done a transient response study of the STM feedback system, following Ref. [5]. A simplified diagram of the feedback system is shown in Fig. C.1. Signals in this feedback loop are all voltages except the tunneling current prior to the amplification.

Figure C.1: Feedback loop for the constant current imaging mode. The pre-amp converts tunneling current into voltage signal. The ‘log’, ‘comparison’ and ‘times gain’ operations are all done by the STM controller.

In this feedback loop, the tunneling current \( I = I_0 e^{- (z_s - z_t) / \kappa} \) is amplified by the pre-amp, \( V = I \times R_t \). The log of this signal is then compared with the set value, and the resultant error is then sent to the gain plant to produce the voltage signal that drives the z piezo (moves the tip back and forth). The gain of the loop is then defined by the following equation,

\[
G = 2\kappa \cdot K_V \cdot K_z \cdot K_I
\]
where $\kappa = 10 \text{ nm}^{-1}$ ($1 \text{ Å}^{-1}$), $K_V = \log_{10} e = 0.4343$, $K_z = 2.4 \text{ nm/V}$ and $K_I$ is the integral gain from user input.

The transfer functions of the system are as follows,

$$T_1(s) = 1 + t_1s + \frac{1}{A_1} \left(1 + \frac{s}{2\pi f_1}\right)(1 + t_3s)$$
$$T_2(s) = 1 + t_2s + \frac{1}{A_2} \left(1 + \frac{s}{2\pi f_2}\right)(1 + t_2s)$$ (C.2)

$$R(s) = \frac{G(1 + K_p s)}{G(1 + K_p s) + s \times T_1(s) \times T_2(s) \times \left(1 + \frac{s^2}{Q_w} + \frac{s^2}{w^2}\right)}$$

where $T_1(s)$ and $T_2(s)$ are the transfer function of the two stages in the current pre-amplifier and $R(s)$ is the total transfer function of this feedback loop. Using $R(s)$ we can do a transient response analysis by forward and inverse Laplace transformations for a set of given parameters.

![Graph](image.png)

**Figure C.2:** Transient response of the tip when scanning through a height step. The blue curve is the ideal response or the actual surface topography. The purple/red curves indicate situations with proper or unstable gain settings. This simulation was done by forward and inverse Laplace transformation in Mathematica.

In our case, both stages of the pre-amplifiers have some significant phase shift in the range of DC–10 kHz. In this simulation (Fig. C.2), we have used $t_1 = 50 \mu\text{sec}$, $t_2 = 43 \mu\text{sec}$ and $t_3 = 0.2 \text{ sec}$ (due to shunt capacitance $C_s = 200 \text{ pF}$ and feedback resistor $R_f = 1 \text{ G\Omega}$).
We have chosen $Q = 1000$ and $w = 30$ kHz for our PZT8 scantube ($L = 0.25$, OD = 0.125, and ID = 0.080 inch).

Another important fact is that lowering the value for $w$ can make the system unstable pretty quickly. The resonant frequency of the scantube decrease quickly as its length when other parameters are fixed. The lowest resonant frequency of a tubular structured object is as follows,

$$f = \frac{1}{2\pi} \sqrt{\frac{k}{m_t + 0.24m_c}},$$

(C.3)

where $k = \frac{3EI}{l^3}$, $I = \frac{\pi}{64}(d_o^4 - d_i^4)$, $m_t$ is the weight attached to the tube (In our case, the weight of the tip assembly.) and $m_c$ is the weight of the tube itself $m_c = \rho Al$. $l$ is the length of the tube, $E$ is the Young’s modulus of PZT8, $A$ is the cross sectional area, and $d_o/d_i$ is the outer/inner diameters of the scantube. If there were no attached weight, $m_t = 0$, the relation between $f$ and $l$ could be further simplified as follows,

$$f \sim \frac{1}{l^2}.$$  

(C.4)

Therefore, the length of the scantube is of critical importance in determining the stability of the STM feedback loop. Increasing the length will reduced the resonance frequency quickly as seen in Eqn. (C.4), although it is helpful in terms of increasing the full scan size of the STM (Eqn. (1.10)).
Bibliography


[53] J. Demsar, L. Forró, H. Berger, and D. Mihailovic. Femtosecond snapshots of gap-
forming charge-density-wave correlations in quasi-two-dimensional dichalcogenides 1T-

[54] E. Morosan, H. Zandbergen, B. Dennis, J. Bos, Y. Onose, T. Klimczuk, A. Ramirez,
July 2006.

[55] K. Wagner, E. Morosan, Y. Hor, J. Tao, Y. Zhu, T. Sanders, T. McQueen, H. Zand-
bergen, A. Williams, D. West, and R. Cava. Tuning the charge density wave and


[57] S. Hüfnner, M.A. Hossain, A. Damascelli, and G.A. Sawatzky. Two gaps make a high-

[58] L.J. Li, X.D. Zhu, Y.P. Sun, H.C. Lei, B.S. Wang, S.B. Zhang, X.B. Zhu, Z.R. Yang,


crystal growth and characterizations of Cu$_{0.03}$TaS$_2$ superconductors. J. Cryst. Growth,


[63] K. Rossnagel and N. Smith. Spin-orbit splitting, Fermi surface topology, and charge-

[64] C.D. Malliakas and M.G. Kanatzidis. Nb-Nb Interactions Define the Charge Density

of the high-T$_C$ copper-oxide pseudogap states. Nature (London), 466(7304):347–51,
July 2010.

variation of the local density of states in the charge density wave phase of 2H-NbSe$_2$.


