Title

Tuning Electrostatic Potentials for Imaging the Quantum Properties of Massless Dirac Fermions in Graphene

Author

Wong, Dillon

Publication Date

2017

Tuning Electrostatic Potentials for Imaging the Quantum Properties of Massless Dirac Fermions in Graphene

by

Dillon Wong

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael F. Crommie, Chair
Professor Feng Wang
Professor Oscar D. Dubon

Spring 2017

Tuning Electrostatic Potentials for Imaging the Quantum Properties of Massless Dirac Fermions in Graphene

Abstract

Tuning Electrostatic Potentials for Imaging the Quantum Properties of Massless Dirac Fermions in Graphene

by

Dillon Wong

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Michael F. Crommie, Chair

Graphene, a two-dimensional (2D) honeycomb lattice of sp^2-bonded carbon atoms, is renowned for its many extraordinary properties. Not only does it have an extremely high carrier mobility, exceptional mechanical strength, and fascinating optical behavior, graphene additionally has an interesting energy-momentum relationship that is emergent from its space group symmetry. Graphene's low-energy electronic excitations consist of quasiparticles whose energies disperse linearly with wavevector and obey a 2D massless Dirac equation with a modified speed of light. This fortuitous circumstance allows for the exploration of ultra-relativistic phenomena using conventional tabletop techniques common to solid state physics and material science. Here I discuss experiments that probe these ultra-relativistic effects via application of scanning tunneling microscopy (STM) and spectroscopy (STS) to graphene field-effect transistors (FETs) in proximity with charged impurities.

The first part of this dissertation focuses on the ultra-relativistic Coulomb problem. Depending on the strength of the potential, the Coulomb problem for massless Dirac particles is divided into two regimes: the subcritical and the supercritical. The subcritical regime is characterized by an electron-hole asymmetry in the local density of states (LDOS) and, unlike in nonrelativistic quantum mechanics, does not support bound states. In contrast, the supercritical regime hosts quasi-bound states that are analogous to "atomic collapse" orbits predicted to occur in atoms with nuclear charge $Z > 170$. By using an STM tip to directly position calcium (Ca) impurities on a graphene surface, we assembled "artificial nuclei" and observed a transition between the subcritical and supercritical regimes with increasing nuclear charge. We also investigated the screening of these charged impurities by massless Dirac fermions while varying the graphene carrier concentration with an electrostatic gate.

The second part of this dissertation focuses on the ultra-relativistic harmonic oscillator. We developed a method for manipulating charged defects inside the boron nitride (BN) substrate underneath graphene to construct circular graphene p-n junctions. These p-n junctions were effectively quantum dots that electrostatically trapped graphene's relativistic charge carriers, and we imaged the interference patterns corresponding to this quantum

confinement. The observed energy-level spectra in our p-n junctions closely matched a theoretical spectrum obtained by solving the 2D massless Dirac equation with a quadratic potential, allowing us to identify each observed state with principal and angular momentum quantum numbers.

The results discussed here provide insight into fundamental aspects of relativistic quantum mechanics and into graphene properties pertinent to technological applications. In particular, graphene's response to electrostatic potentials determines the scope in which its charge carriers can be directed and harnessed for useful purposes. Furthermore, many of the results contained in this dissertation are expected to generalize to other Dirac materials.

To science.

Contents

List of Figures

List of Abbreviations

1D	One-Dimensional
2D	Two-Dimensional
2DEG	Two-Dimensional Electron Gas
3D	Three-Dimensional
a.c.	Alternating Current
BN	Hexagonal Boron Nitride
CVD	Chemical Vapor Deposition
DFT	Density Functional Theory
dI/dV	Differential Tunneling Conductance
DOS	Density of States
E_D	Dirac Point Energy
E_F	Fermi Energy
FET	Field-Effect Transistor
LDOS	Local Density of States
LT	Low Temperature
MMA	Methyl Methacrylate
rms	Root Mean Square
RPA	Random Phase Approximation
STM	Scanning Tunneling Microscopy *or* Scanning Tunneling Microscope
STS	Scanning Tunneling Spectroscopy
UHV	Ultra-High Vacuum
V_g	Backgate Voltage
V_s	Sample Bias

Acknowledgments

There are many people who I want to thank for their contributions to my PhD research. First, I thank my advisor, Prof. Mike Crommie, for all the guidance that he has given me in the last seven years. Mike taught me how to ask the right questions, how to communicate effectively, and how to do good science. His enthusiasm for physics was inspiring, and he gave me an incredible amount of freedom to explore whatever I found interesting. None of this work would have been possible without his support.

I also want to thank Profs. Steven Louie and Oscar Dubon, who were on my qualifying exam committee and made the exam a surprisingly pleasant experience. Special thanks goes to Prof. Feng Wang, who chaired the committee and is a person who I immensely admire. Feng gave me a lot of useful advice throughout my time as a graduate student.

Yang Wang taught me how to perform STM experiments. He is one of the smartest people I know, and I am grateful for the opportunity to learn from him. We had a very fruitful collaboration, and even our failures (of which we had many) made me a better scientist. Yang was incredibly hardworking and patient with my blunders. I couldn't have asked for a better mentor and coworker.

My overlap with Victor Brar was brief, but I also learned a lot about STM from him. Despite being so far away (in Pasadena), he was always available to give advice when I needed it. Victor is a better experimentalist than I will ever be, and I'm glad to have known him. Plus, he's super fun to be around.

Jairo Velasco Jr. introduced me to the world of microfabrication, and I introduced Jairo to STM. When Jairo first arrived in Berkeley as a brand new postdoc, I didn't pay much attention to what he had to offer me scientifically – we were friends but not collaborators. It wasn't until a year later that we started working together. I regret that we hadn't started working together sooner. Jairo is an academic of the highest caliber, and I hope that we will have opportunities to collaborate in the future.

I enjoyed working with Juwon Lee, who I mentored and taught STM. Juwon has a better attention to detail than anyone else I know. We talked a lot about our political ideologies and worldviews, which I found enlightening even when we didn't see eye to eye.

I'm grateful to have worked with Long Ju, who I think is the most talented experimental physicist of our generation. I'm also glad to have worked with Zahra Pedramrazi, who is a good friend and a natural in lab. I think we must expect great things from her. Hsin-Zon Tsai is a logistical genius, and I don't know how he does it. In a different era, Hsin-Zon would have made a great military strategist. Salman Kahn is an undergrad, but he might as well be a graduate student – Salman deserves a PhD for his work.

I also want to give thanks to Ryan Yamachika, Aaron Bradley, and Kacey Meaker, who were of great assistance whenever our STM broke. Ryan doesn't say much, but behind his reticence is a fountain of knowledge. Ryan simply knows everything. Aaron, for better or worse, became an expert in instrumentation while in graduate school, and so he was very helpful when I needed to fix things. Kacey is unnaturally clumsy except (strangely) when near valuable scientific equipment.

I appreciate all of the people who kept me company down in Birge b207. Otherwise, the solitary nature of my work might have driven me insane. Chad Germany was (and still is) a dear friend who I wasted too much time chatting with in lab. Although I was happy for him, I was sad to see him leave Berkeley for graduate school at UIUC. I had a good time with Maider Ormaza, that spiky pineapple! And I'm sorry that I never went parasailing with Sajjad Tollabi. It was also nice talking philosophy, politics, linguistics, or whatever with Patrick Forrester (who convinced me to donate blood) and Joe Costello. I wish Patrick and Joe luck in graduate school. And thanks to Brandon Giles and Sergio Pezzini as well for wonderful and stimulating discussions. Daniel Rizzo didn't work in b207, but he stopped by enough that I feel compelled to acknowledge that **** here. Thank you, Dan.

I want to thank a few graduate students and postdocs from our group who were good friends to me. Those people are Trinity Joshi, Yen-Chia Chen, Miguel Moreno Ugeda, Alex Riss, Chen Chen, Yi Chen, and Chenggang Tao. Thanks to Trinity for inviting me to her baby shower, although her dog Jasper was terrifying. Thanks to Yen-Chia for hanging out with me every now and then. Thanks to Miguel just for being awesome. Thanks to Alex for being so funny. Thanks to Chen, Yi, and Chenggang for interesting conversations. I'm also glad to have met Danny Haberer, Christopher Bronner, Jiong Lu, Arash Omrani, Giang Nguyen, Xiaowei Zhang, Ivan Pechenezhskiy, and Sebastian Wickenburg, who I wish all the best. Unfortunately, I didn't get much of a chance to get to know Franklin Liou, Dimas de Oteyza, Régis Decker, Sarah Burke, Chuck Sleasman, Keenan Pepper, and Peigen Cao, but I have to say that all of my interactions with them have been more than just pleasant.

Many undergrads have come and gone from our lab during my seven years in Mike's group. A few of them stood out to me. In no particular order, they are Ramin Khajeh, Andrew Aikawa, Aditya Arun, Youngkyou Kim, Zaw Htet, Dhruv Desai, Yi-Shiou Duh, Griffin Rodgers, Edgar Olivera, Nick Kau, Steven Munn, Viska Wei, Abigail Iturra, and Bobby Ge. I don't really know Hoil Kim very well, but I remember him because Mike said, "Hoil will toil for you." I would like to mention Han Sae Jung and Won-Woo Choi separately from the others because they were involuntarily conscripted into the Korean military, which I believe is a waste of human talent (of which Han Sae and Won-Woo had plenty of) and is an example of government overreach into the lives of private individuals (#Freedom). Mandy Huo was an undergrad who I didn't know very well while she was in our group (although she was with me when I accidentally kicked that baby at the APS March Meeting in Denver...), but I keep in regular contact with her now. Melissa Panlasigui, Drew Edelberg, and Changmin Lee were undergrads in Mike's group when I was an undergrad. Drew showed me how to make PtIr tips, which I will discuss in this dissertation. Geun Ho Ahn was an undergrad from Prof. Ali Javey's group who made us some samples. He was very diligent, but I'm sorry to say that the project didn't work.

There were also many visitors from far away lands who came to Mike's group to learn STM. They are Prof. Caihong Jia, Wei Ruan, Prof. Pintu Das, Dian Shi, Yun Zhou, Qiong Wu, Mark Edmonds, Erik Piatti, Sulthan Rashid, Irene Rodriguez, Jakob Jørgensen, Hannes Böckmann, Ana Recio, and Juanjuan Feng. Wei Ruan is a genius, and that's all I need to say about him. Qiong helped make the CVD graphene samples that were crucial for our early experiments. Erik has an amazing opera voice, but he should refrain from speaking near STMs. Sulthan cooked me some beef, which was cool. They were all very helpful during my PhD studies, and I wish them the best of luck with whatever they're doing now. The visitor who was most important to me, however, was Prof. Roland Kawakami. He is not only a mentor to me, but I also consider him a friend.

I would like to thank members of the Zettl group who have helped me with various tasks. In alphabetical order, they are Aidin Fathalizadeh, Aiming Yan, Allen Sussman, Anna Goldstein, Anna Zaniewski, Ashley Gibb, – wow, you're actually reading this excessively long acknowledgments section – Claudia Ojeda-Aristizabal, Haider Rasool, Matt Gilbert, Seita Onishi, and Thang Pham. In Feng Wang's group, I would like to thank Bo Zeng, Chaw Keong Yong, Chenhao Jin, Jason Horng, Jonghwan Kim, Sufei Shi, Yaqing Bie, and Zhiwen Shi. I thank Nityan Nair from the Analytis group and Yingjie Zhang from the Salmeron group. I thank Shujie Tang from Prof. Zhi-Xun Shen's group at Stanford for providing us with some interesting materials to look at.

My research was greatly enhanced by the work of wonderful theorists. Profs. Leonid Levitov and Andrey Shytov developed the theory for atomic collapse in graphene. Profs. Shaffique Adam and Jeil Jung worked on twisted bilayer graphene. Arash Mostofi, Fabiano Corsetti, and Johannes Lischner calculated the LDOS of doped graphene near a calcium atom. Joaquin Rodriguez-Nieva and Prof. Levitov worked on the confined modes in p-n junctions. I also had fruitful discussions with Justin Song and Mervin Roy. All of these theorists provided substantial physical insight for guiding our experimental efforts and explaining our observations.

During my ten years (a whole decade...) at Berkeley, there were a few professors who were especially helpful to me. They are Profs. Hartmut Häffner, Dmitry Budker (always pleasant to talk to), Kam-Biu Luk, Ron Shen (my faculty mentor), Joe Orenstein (funny, except when he makes fun of me), Robert Lin (who got me a fantastic programming gig with James McFadden), James Analytis, Marjorie Shapiro, Alex Zettl (who has always been positive about my work), Alessandra Lanzara, Peter Yu (gave valuable information about defects in materials), Yury Kolomensky, and William Holzapfel. I also want to thank the physics department staff, specifically Don Orlando, Elaine Quiter, Donna Sakima, Claudia Trujillo, Eleanor Crump, Anthony Vitan, Joseph Kant, Dave Nguyen, and Warner Carlisle. Special thanks to Anne Takizawa and Lisa Partida, who were amazingly helpful and fun to talk to. How does Anne remember everything?

I know I've forgotten to include some people who were vital to my work. I apologize if you're amongst the forgotten few (or forgotten many). There are also people whose names I don't know, but I know their faces. I am acknowledging them in my thoughts now. Most importantly, I'd like to thank my friends (in academia and outside academia, neither of which I will list because there's too many names in here already... you know who you are...) and my family for supporting me on my academic journey. Amongst my family, I especially want to thank my father Kalvin, my mother Shu Ying, and my sister Daisy. Thank you.

Chapter 1

Introduction

1.1 The Coulomb Potential and the Harmonic Oscillator

Two of the most basic and important problems in physics are the Coulomb potential and the harmonic oscillator potential. These two potentials are ubiquitous and, because they are amongst the few that we know how to solve exactly, are often used as a starting point to model and understand the complicated world around us. The Coulomb problem has a long history, stretching back to when Issac Newton first demonstrated that Kepler's laws of planetary motion could be derived from a $1/r^2$ force. Much later, Niels Bohr and Erwin Schrödinger incorporated the Coulomb potential into the framework of a quantum theory to obtain the energy levels of a hydrogen atom. The quantum theory of the Coulomb potential is by no means limited in applicability, as it easily explains many aspects of semiconductor impurity states, excitons, and other phenomena. The usefulness of the Coulomb field is only surpassed by the harmonic oscillator, which is truly everywhere in all areas of science and engineering. Almost everything – including the Coulomb potential – is a harmonic oscillator or a system of coupled oscillators.

In nonrelativistic quantum mechanics, particle motion in a time-independent potential is governed by the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V(\mathbf{r})\Psi = E\Psi, \tag{1.1}$$

for which $V(\mathbf{r}) = Ze^2/r$ for the Coulomb potential, and $V(\mathbf{r}) = \frac{1}{2}m\omega^2 r^2$ for the harmonic oscillator. The Schrödinger equation can be solved for both potentials using the same method. Separation of variables yields a radial Schrödinger equation, which is then transformed into a Sturm-Liouville equation with a complete set of orthogonal eigenfunctions. These eigenfunctions are obtained by plugging in a power series into the differential equation and deriving a recursive formula that determines the coefficients of the power series. The power series is then given a name: the associated Laguerre polynomials for the three-dimensional (3D) Coulomb problem and the Hermite polynomials for the one-dimensional (1D) harmonic oscillator. The corresponding discrete energy eigenvalues are

$$E = -\frac{Z^2 e^4 m}{2\hbar^2 n^2}, \tag{1.2}$$

with $n = 1, 2, 3, ...$ for the 3D Coulomb potential, and

$$E = \hbar\omega(n + \frac{1}{2}), \tag{1.3}$$

with $n = 0, 1, 2, 3, ...$ for the 1D harmonic oscillator.

All of this was understood as early as the mid-1920s, and this is now taught in standard undergraduate courses in quantum mechanics. Less well known, however, is the behavior of a relativistic particle in a Coulomb or harmonic oscillator potential. As we shall see in the subsequent section, graphene provides a pathway for experimentally exploring how relativistic electrons react to these elementary potentials.

1.2 Graphene as a Platform for Relativistic Physics

Graphene, an atomically thin sheet of graphite, was first isolated on insulating substrates in 2004 by Andre Geim and Konstantin Novoselov [1], and its two-dimensional (2D) nature and π Berry phase were immediately verified via quantum Hall measurements [2, 3]. Graphene's electronic structure, however, was already theoretically understood as far back as 1947, when Philip Wallace first calculated the band structure using the tight-binding method [4]. Here, I present the tight-binding model for graphene and show that its electrons obey a 2D massless Dirac equation.

1.2.1 Tight-Binding Theory of Graphene

The carbon atoms in graphene are arranged in a honeycomb lattice that is composed of two interpenetrating triangular sublattices, hereby denoted as the A and B sublattices (see Fig. 1.1a). Each triangular sublattice can be constructed by the primitive lattice vectors

$$\mathbf{a}_1 = a\left(-\frac{\sqrt{3}}{2}, \frac{3}{2}\right),\ \mathbf{a}_2 = a\left(\frac{\sqrt{3}}{2}, \frac{3}{2}\right), \tag{1.4}$$

where $a = 1.42$ Å is the distance between an atom on the A sublattice and its nearest neighbors on the B sublattice. The vectors to the nearest neighbors of each B atom are

$$\boldsymbol{\delta}_1 = a\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right),\ \boldsymbol{\delta}_2 = a\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right),\ \boldsymbol{\delta}_3 = a\left(0, -1\right). \tag{1.5}$$

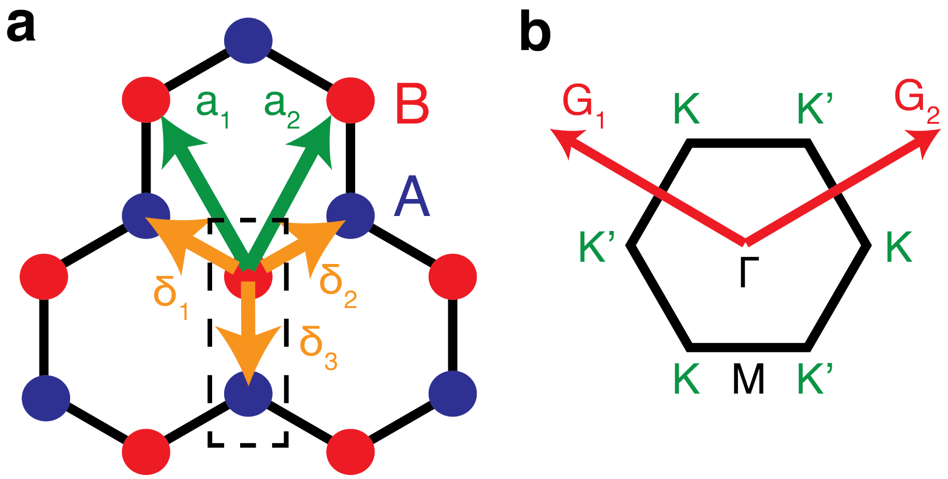

Figure 1.1: Graphene unit cell. (a) A honeycomb lattice consists of two triangular sublattices A and B. Each unit cell consists of a basis with an A atom and its nearest neighbor on B at $-\boldsymbol{\delta}_3$ away from the A atom. (b) The Brillouin zone is a hexagon. The corners of the Brillouin zone are denoted K and K'.

The reciprocal lattice vectors are

$$\mathbf{G}_1 = \frac{2\pi}{a}\left(-\frac{1}{\sqrt{3}}, \frac{1}{3}\right), \; \mathbf{G}_2 = \frac{2\pi}{a}\left(\frac{1}{\sqrt{3}}, \frac{1}{3}\right). \tag{1.6}$$

As depicted in Fig. 1.1b, the Brillouin zone for a triangular lattice is a hexagon with high symmetry points K and K' on its corners:

$$\mathbf{K} = \frac{2\pi}{a}\left(\frac{2}{3\sqrt{3}}, 0\right), \; \mathbf{K}' = \frac{2\pi}{a}\left(-\frac{2}{3\sqrt{3}}, 0\right). \tag{1.7}$$

The atoms are σ-bonded by sp^2 orbitals, while p_z orbitals stick out of the graphene plane, forming π and π^* bands [5]. These bands, which form the conduction and valence bands of graphene, can be modeled by the following Hamiltonian with hopping amplitude t:

$$H = -t\sum_{\langle i,j\rangle}\left(a_i^\dagger b_j + \text{h.c.}\right) + U\sum_i\left(a_i^\dagger a_i - b_i^\dagger b_i\right). \tag{1.8}$$

I have suppressed the spin index and have neglected hopping further than nearest neighbor. U is half the energy difference between an electron sitting on sublattice B as opposed to A, which is obviously zero. Here, $\langle i,j\rangle$ denotes nearest neighbors, and the operators $a_i^\dagger$ and a_i are the creation and annihilation operators for site i on sublattice A (and similarly, $b_j^\dagger$ and b_j are for site j on B).

Since graphene has discrete translational symmetry, we can construct Bloch states indexed by crystal momentum $\mathbf{k}$:

$$a_i = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_i} a_{\mathbf{k}}, \tag{1.9}$$

$$b_i = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}_i} b_{\mathbf{k}}. \tag{1.10}$$

$\mathbf{R}_i$ is the vector to the ith unit cell, and N is the number of unit cells. Plugging these two expressions into the Hamiltonian in Eq. 1.8 gives

$$\begin{aligned} H &= -\frac{t}{N} \sum_{\langle i,j \rangle} \sum_{\mathbf{k},\mathbf{k}'} \left(e^{i\mathbf{k}'\cdot\mathbf{R}_j - i\mathbf{k}\cdot\mathbf{R}_i} a^\dagger_{\mathbf{k}} b_{\mathbf{k}'} + \text{h.c.} \right) + \frac{U}{N} \sum_{i,\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'-\mathbf{k}\right)\cdot\mathbf{R}_i} \left(a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} - b^\dagger_{\mathbf{k}} b_{\mathbf{k}'} \right) \\ &= -t \sum_{\mathbf{k}} \left[\left(1 + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2} \right) a^\dagger_{\mathbf{k}} b_{\mathbf{k}} + \text{h.c.} \right] + U \sum_{\mathbf{k}} \left(a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} - b^\dagger_{\mathbf{k}} b_{\mathbf{k}} \right). \end{aligned} \tag{1.11}$$

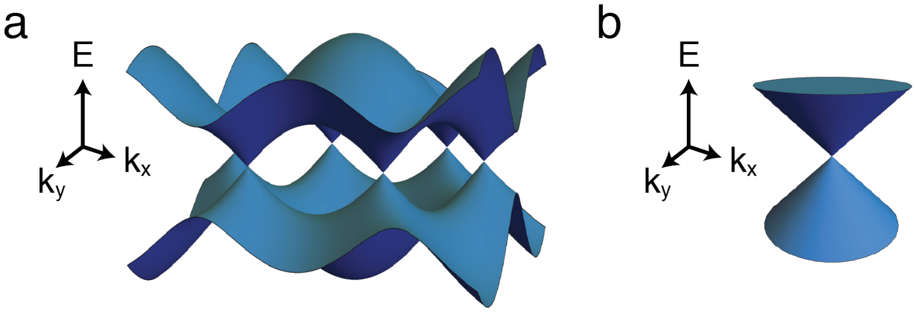

Figure 1.2: Graphene band structure. (a) Tight-binding energy-momentum dispersion of graphene. The upper (π^*) and lower (π) bands touch at the K and K' points. (b) The electronic dispersion relation near the K point is an isotropic Dirac cone. The energy disperses linearly with momentum away from the singularity, which is known as the Dirac point. The Fermi level coincides with the Dirac point for neutral (undoped) graphene.

If we organize the amplitudes for being on each sublattice into a two-component spinor, the Hamiltonian can be written as a matrix for each $\mathbf{k}$:

$$H(\mathbf{k}) = -t \begin{pmatrix} U & 1 + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2} \\ 1 + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2} & -U \end{pmatrix}. \tag{1.12}$$

For $U = 0$, the energy eigenvalues for this Hamiltonian matrix are $E = \pm|1 + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2}|$, which is plotted in Fig. 1.2a. For undoped graphene, the upper band is an empty conduction band, while the lower band is a filled valence band.

1.2.2 Dirac Cones and the Dirac Point

Eq. 1.12 can be expanded for small $\mathbf{q} = (q_x, q_y)$ around the K point (i.e. $\mathbf{k} = K + \mathbf{q}$) to obtain

$$H_K(\mathbf{q}) = \hbar v_F \left(q_x\sigma_x + q_y\sigma_y \right) + U\sigma_z, \tag{1.13}$$

where σ_x, σ_y, and σ_z are the Pauli matrices, and $v_F = \frac{3at}{2\hbar} \approx 10^6$ m/s is the graphene Fermi velocity. These Pauli matrices do not act on spin but rather on the sublattice degree of freedom. If we then define ϕ such that $\mathbf{q} = (q_x, q_y) = q(\cos(\phi), \sin(\phi))$, the Hamiltonian in Eq. 1.13 (for $U = 0$) can be rewritten as

$$H_K(\mathbf{q}) = \hbar v_F q \begin{pmatrix} 0 & e^{-i\phi} \\ e^{i\phi} & 0 \end{pmatrix}, \tag{1.14}$$

which has eigenvectors

$$\Psi_{\mathbf{K},\pi^*}(\mathbf{q}) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi/2} \\ e^{i\phi/2} \end{pmatrix}, \ \Psi_{\mathbf{K},\pi}(\mathbf{q}) = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\phi/2} \\ -e^{i\phi/2} \end{pmatrix}. \tag{1.15}$$

This is similar to a spin pointing in the ϕ-direction. For this reason, the sublattice degree of freedom is referred to as “pseudospin.” However, unlike the real spin, pseudospin is locked parallel or antiparallel to $\mathbf{q}$. It is said that graphene electrons have “helicity” (also inappropriately named “chirality”) equal to +1 in the conduction band (because pseudospin and $\mathbf{q}$ are parallel) and −1 in the valence band (because pseudospin and $\mathbf{q}$ are antiparallel).

If the A and B sublattices are energetically inequivalent (such as in hexagonal boron nitride (BN)), the σ_z term in Eq. 1.13 (often called the “mass term”) produces a band gap of width $2U$. Since the A and B sublattices in graphene are energetically equivalent (and ignoring spin-orbit coupling, which is small for carbon), graphene is gapless. This leads to a linear dispersion $E = \pm\hbar v_F q = \pm p v_F$ quite like that of a massless particle (except the speed of light has been replaced by v_F). The positive eigenvalue $E = \hbar v_F q$ is for $\Psi_{\mathbf{K},\pi^*}$, and the negative eigenvalue $E = -\hbar v_F q$ is for $\Psi_{\mathbf{K},\pi}$. Fig. 1.2b depicts this dispersion relation, which is composed of two symmetric conical bands that touch at an $E = 0$ singularity called the Dirac point E_D. As a consequence, graphene also has a linear electronic density of states (DOS) per unit area given by

$$\mathrm{DOS}(E) = \frac{2|E|}{\pi(\hbar v_F)^2}. \tag{1.16}$$

This expression includes both the spin degeneracy and “valley degeneracy” described in the next paragraph.

The Hamiltonian in Eq. 1.12 can also be expanded around the K' point to obtain

$$H_{K'}(\mathbf{q}) = \hbar v_F \left(-q_x\sigma_x + q_y\sigma_y\right) + U\sigma_z. \tag{1.17}$$

The Hamiltonian at K' is constrained to take this form because (for $U = 0$) time reversal and inversion symmetries require that $H_{K'}(-\mathbf{q}) = H_K^*(\mathbf{q})$ and $H_{K'}(-\mathbf{q}) = \sigma_x H_K(\mathbf{q})\sigma_x$. $H_K(\mathbf{q})$ and $H_{K'}(\mathbf{q})$ can be combined into a single Hamiltonian

$$H(\mathbf{q}) = \hbar v_F \left(q_x\sigma_x \otimes \tau_z + q_y\sigma_y\right) + U\sigma_z. \tag{1.18}$$

Here, τ_z is the third Pauli matrix acting on the K and K' degree of freedom, sometimes referred to as valley pseudospin. In this dissertation, we will not be concerned about the K' point except that it gives an additional degeneracy factor of 2 in Eq. 1.16.

1.2.3 Graphene Continuum Model

I henceforth drop the subscript K on H_K and abbreviate the Pauli matrices as $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$. We obtain a continuum model Hamiltonian for graphene

$$H = -i\hbar v_F \boldsymbol{\sigma} \cdot \nabla \tag{1.19}$$

by replacing the momentum $\hbar\mathbf{q}$ by $-i\hbar\nabla$ in Eq. 1.13. This is the 2D massless Dirac Hamiltonian and leads to ultra-relativistic phenomena such as the Klein paradox [6–11], Zitterbewegung [12], and atomic collapse [13–18].

In the presence of an external potential $V(\mathbf{r})$, the continuum model equation for graphene becomes

$$-i\hbar v_F \boldsymbol{\sigma} \cdot \nabla\Psi + V(\mathbf{r})\Psi = E\Psi. \tag{1.20}$$

We note here an important property of the above single-particle equation. If we replace Ψ with $\sigma_z\Psi'$, then by the anticommuting property of the Pauli matrices $\{\sigma_i, \sigma_j\} = \delta_{ij}$,

$$-i\hbar v_F \boldsymbol{\sigma} \cdot \nabla\Psi' - V(\mathbf{r})\Psi' = -E\Psi'. \tag{1.21}$$

This implies that if Ψ is a solution to the Dirac equation with potential $V(\mathbf{r})$ and energy E (Eq. 1.20), then $\Psi' = \sigma_z\Psi$ is a solution to the Dirac equation for $-V(\mathbf{r})$ and $-E$. So long as the Dirac equation remains a good description for charge carrier motion in graphene, the physics that occur in potentials $V(\mathbf{r})$ and $-V(\mathbf{r})$ are exactly the same (except the energy E is inverted).

1.3 The Klein Paradox

The fact that charge carrier motion in graphene can be described via a massless Dirac equation has many interesting consequences. The most famous is that of the Klein paradox, where an electron normally incident on a potential barrier has a 100% transmission probability [6]. This is highly counterintuitive because this Klein tunneling process occurs even when the kinetic energy is smaller than the potential barrier height. Suppose we have a Heaviside step potential

$$V(x) = 2V_0\Theta(x) - V_0 = \begin{cases} -V_0, & x < 0 \\ V_0, & x > 0. \end{cases} \tag{1.22}$$

Ignoring intervalley scattering, we can model graphene electrons using the Hamiltonian

$$H = -i\hbar v_F \boldsymbol{\sigma} \cdot \nabla + V(x). \tag{1.23}$$

There is translational invariance in the y-direction, so crystal momentum q_y is a good quantum number. Since we are considering an electron that is moving perpendicular to the barrier, we shall take $q_y = 0$. The effective Hamiltonian simplifies to

$$H = -i\hbar v_F \sigma_x \frac{\partial}{\partial x} + V(x). \tag{1.24}$$

This Hamiltonian commutes with σ_x, so pseudospin in the x-direction is conserved. As per the discussion in the previous section, pseudospin is parallel to group velocity. We thus conclude that if the pseudospin cannot be flipped by $V(x)$, the group velocity also cannot be reversed by the step potential. The electron travels through the barrier with 100% transmission probability and 0% reflection probability. Note that this is untrue if $q_y \neq 0$, i.e. for an electron not normally incident on the barrier.

The Klein paradox also occurs for massive relativistic fermions, in which the transmission coefficient approaches 1 as the potential barrier height approaches infinity.

1.4 The Relativistic Coulomb Potential

1.4.1 The Massive Case

According to Kepler's first law of planetary motion, planets move around the Sun in elliptical orbits. Newton explained this in terms of his $1/r^2$ force law, for which solutions to the equations of motion yield conic sections. Much later, with the advent of special relativity, Charles Darwin (not the famous one, but his grandson) solved the relativistic Kepler problem with energy

$$E = \sqrt{(pc)^2 + (mc^2)^2} + \frac{Ze^2}{r}. \tag{1.25}$$

Darwin found two regimes for different angular momentum L:

$$L > \frac{Ze^2}{c} \quad \text{and} \quad L < \frac{Ze^2}{c}. \tag{1.26}$$

The first regime is similar to that of Kepler's first law. There are elliptical orbits, although the ellipse may precess and form rosettes. The second regime is quite different and unfamiliar. These orbits spiral towards the center of the $1/r$ potential.

Heuristically, if we impose the quantization condition $L = \hbar$, then these two regimes become

$$Z < \frac{\hbar c}{e^2} = \frac{1}{\alpha} \quad \text{and} \quad Z > \frac{1}{\alpha}. \tag{1.27}$$

Here, $\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137}$ is the fine-structure constant in Gaussian units. We will dub the $Z\alpha < 1$ regime "subcritical" and the $Z\alpha > 1$ regime "supercritical," although the definitions of the subcritical and supercritical regimes will change slightly depending on the physical situation (for example, there is an additional factor of 1/2 in two dimensions).

The quantum version of the relativistic Kepler problem has been analytically solved. Suppose we have a hydrogen-like nucleus with atomic number Z. If m is the mass of an electron, then the energy eigenvalues of this system are given by [19]

$$E = \frac{mc^2}{\sqrt{1 + \frac{(Z\alpha)^2}{\left(n - \left(j + \frac{1}{2}\right) + \sqrt{\left(j + \frac{1}{2}\right)^2 - (Z\alpha)^2}\right)^2}}}, \tag{1.28}$$

where n and j are quantum numbers. This monstrous equation yields a complex (as opposed to purely real) eigenvalue when $Z\alpha > 1$. The non-Hermitian nature of this problem is due to the singularity at $r = 0$ and can be repaired by taking into account the finite size of the nucleus. With the $r = 0$ singularity regularized, real energy eigenvalues are obtained, but the condition for the supercritical regime is pushed upwards to $Z > 170$ [20].

When Z exceeds 170 in the supercritical regime, the binding energy exceeds $2mc^2$. The vacuum becomes unstable, and it is favorable to produce an electron-positron pair *ex nihilo*. The electron spirals into the nucleus and is confined in an "atomic collapse" state, while the positron flies away. There have been attempts to observe positron emission produced though this phenomenon by colliding two subcritical nuclei ($Z_1, Z_2 < 170$, such as uranium) that have combined atomic number $Z_1 + Z_2$ greater than 170 [21], but the experimental results have been inconclusive.

1.4.2 The Massless Case

Suppose we had a 2D massless charged particle with momentum p in a Ze^2/r Coulomb potential. The energy is

$$E = pc - \frac{Ze^2}{r}. \tag{1.29}$$

If we impose the Heisenberg uncertainty condition $pr > \hbar/2$, this expression becomes

$$E > \frac{\hbar c/2 - Ze^2}{r}, \tag{1.30}$$

which for $Z < \frac{\hbar c}{2e^2} = \frac{1}{2\alpha}$ gives $E > 0$. We conclude that, in the subcritical regime $Z\alpha < 1/2$, there are no bound states of the Coulomb potential. This argument is, of course, not rigorous (and I deliberately chose $pr > \hbar/2$ for the quantization condition to give the extra 1/2 that appears in the critical charge $Z_C = \frac{1}{2\alpha}$ in two dimensions). However, theoretical calculations by Antonio Castro Neto, Leonid Levitov, Michael Fogler, and their respective colleagues have shown that this is indeed true [13–17]. The reason that there are no bound states for massless electrons in a Coulomb potential is that they are repelled from the center by the centrifugal force. I will now give a handwavy argument to support this.

The 2D massless Dirac equation with a Coulomb potential is

$$-i\hbar c\boldsymbol{\sigma} \cdot \nabla\Psi - \frac{Ze^2}{r}\Psi = E\Psi. \tag{1.31}$$

Similarly to how the Schrödinger equation for the hydrogen atom is solved, separation of variables can be applied to separate the radial and angular parts of this partial differential equation. To do this, we employ the ansatz

$$\Psi_{nm} = \frac{e^{im\phi}}{\sqrt{r}} \begin{pmatrix} u_n^A(r)e^{-i\phi/2} \\ iu_n^B(r)e^{i\phi/2} \end{pmatrix}, \tag{1.32}$$

where n and m are principal and angular momentum quantum numbers, and $u_n^A(r)$ and $u_n^B(r)$ are functions only of the radial variable r. Plugging this ansatz wavefunction into Eq. 1.31 yields the radial Dirac equation

$$\frac{E}{\hbar c}\begin{pmatrix} u_n^A(r) \\ u_n^B(r) \end{pmatrix} = \left[i\sigma_y \frac{\partial}{\partial r} + \sigma_x \frac{m}{r} - \frac{Z\alpha}{r} \right] \begin{pmatrix} u_n^A(r) \\ u_n^B(r) \end{pmatrix}. \tag{1.33}$$

The second term in the brackets is a repulsive centrifugal term that is proportional to the half-integer angular momentum m. Unlike the centrifugal term in the radial Schrödinger equation for a central potential, which goes as $1/r^2$ because the kinetic energy goes as ∇^2, the repulsive centrifugal barrier in the Dirac equation goes as $1/r$. This is the same power law as the Coulomb term, and if $Z\alpha$ is less than the minimum unit of angular momentum (i.e. the subcritical regime), there are no bound states of the system because the centrifugal force dominates over the Coulomb attraction everywhere.

What happens in the supercritical regime $Z\alpha > 1/2$? The Hamiltonian in Eq. 1.31 is scale invariant, so there are no characteristic energy or length scales in the problem. Consequently, a bound state energy cannot be defined. This is similar to the case for the massive Dirac equation, where the $r = 0$ singularity caused problems. To patch this, we need to introduce a lower length scale cutoff r_0 to regularize the problem. Physically, r_0 could represent the size of the nucleus or the lattice dimensions.

In the supercritical regime (with cutoff r_0), the Coulomb attraction dominates over the centrifugal force, and bound states exist. The binding energy can be calculated via the Bohr-Sommerfeld quantization condition

$$\oint \mathbf{p} \cdot d\mathbf{r} = 2\pi n\hbar. \tag{1.34}$$

The resulting energies (labeled by integer n and angular momentum L) are [14]

$$E_n = -\frac{Ze^2}{r_0} e^{-\pi\hbar n/\gamma}, \tag{1.35}$$

where $\gamma = \sqrt{\left(\frac{Ze^2}{c}\right)^2 - L^2}$.

To summarize, for a massive fermion in a Coulomb potential, there are "normal" bound states for $Z < 170$ and special "atomic collapse" states for $Z > 170$. The massless case differs in that there are no bound states in the subcritical regime ($Z\alpha < 1/2$ in two dimensions), but the atomic collapse states exist for the supercritical regime. The interpretation of the atomic collapse states in both the massive and massless cases is the same: the bound states are resonant with a lower continuum of negative-energy states, allowing the production of electron-positron pairs.

At the present time that I am writing this dissertation, the periodic table ends at $Z = 118$. We are far from synthesizing a $Z = 137$ nucleus (dubbed Feynmanium, after Richard

Feynman who recognized that $Z > 1/\alpha$ was problematic), let alone $Z > 170$ elements. Therefore, we turn to graphene to realize the supercritical atomic collapse phenomenon. The continuum model for graphene is the 2D massless Dirac equation, except the speed of light c has been replaced by the Fermi velocity v_F. Since v_F is approximately 300 times smaller than the speed of light, the effective fine-structure constant in graphene $\alpha_{\text{eff}} = \frac{e^2}{\hbar v_F}$ is 300 times larger than $\alpha = \frac{e^2}{\hbar c}$. Note that α_{eff} is not small (even if we introduce a dielectric constant), so it makes for a poor parameter for perturbative expansions. Also, α_{eff} is only relevant for electronic motion in graphene; the α in the famous $\pi\alpha$ optical absorption of graphene is the real fine-structure constant.

Since α_{eff} is of order unity, the supercritical regime is much more attainable, and graphene indeed serves as an excellent platform for simulating relativistic phenomena. By placing Coulomb impurities on graphene, we should be able to explore atomic collapse, which manifests as a bound state slightly below the Dirac point. The atomic collapse state will be resonant with the graphene valence band (which contains hole-like quasiparticles) instead of a positron continuum. I will discuss our experimental results on the subcritical regime in Ch. 5 and the supercritical regime in Ch. 6.

1.5 The Relativistic Harmonic Oscillator

The behavior of a 2D massless fermion in a harmonic oscillator potential is described by the Hamiltonian

$$H = -i\hbar v_F \boldsymbol{\sigma} \cdot \nabla + \kappa r^2, \tag{1.36}$$

where $r^2 = x^2 + y^2$. Like light and sound, a massless fermion has a linearly dispersive energy-momentum relationship. Therefore, many phenomena that occur for light and sound waves should also occur for graphene electrons. One such phenomenon is the so-called whispering-gallery mode, in which a wave travels along the perimeter of a circular cavity. For acoustic waves, one can whisper at one end of the circular dome at St. Paul's Cathedral (or at the Temple of Heaven or at that circular bench near the Berkeley Campanile), and a listener would be able to hear the whisper at the other end. For relativistic electronic waves, the Hamiltonian in Eq. 1.36 describes such whispering-gallery modes in a circular resonant cavity [22]. In Ch. 8, I will discuss the spatial imaging of electrostatically confined quasi-bound states for the Eq. 1.36 Hamiltonian.

Chapter 2

Principles of Scanning Tunneling Microscopy

Scanning tunneling microscopy (STM) is the main characterization technique used in the experiments described in this dissertation. Therefore, I must include an obligatory chapter discussing its basic principles. When two conductors are brought together, electrons can tunnel from one conductor to the other with exponential sensitivity to the distance between the conductors. If one conductor is a sharp metal tip, then the tip can be raster scanned across the other conductor's surface to produce an image of that surface. By using a carefully prepared tip, STM can provide detailed topographic and electronic information about conducting surfaces with atomic-scale resolution.

2.1 Theory of Quantum Tunneling

An electron can tunnel through a classically forbidden barrier because it is a wave. The simplest way to understand tunneling is to solve the Schrödinger equation for a 1D rectangular potential barrier. Imposition of the boundary conditions requires the electron to have some probability of being on the "wrong side" of the barrier. Although it is useful to keep this picture in mind, it is also important to have a theory that can make quantitative predictions for tunneling in STM. Here, I give a simplified treatment of the Bardeen [23] and Tersoff-Hamann [24, 25] theories of tunneling.

2.1.1 The Bardeen Theory

Suppose we have a system described by the Hamiltonian $H = H_0 + H'$. According to Fermi's golden rule [26], the transition rate from an H_0 eigenstate $|i\rangle$ (the initial state with energy E_i) to $|f\rangle$ (a continuum of final states around energy E_f) is given by

$$W = \sum_f \frac{2\pi}{\hbar} |\langle f|H'|i\rangle|^2 \delta(E_f - E_i). \tag{2.1}$$

We are interested in a system that comprises an STM tip and a conducting sample. If electrons are tunneling from sample states $|s\rangle$ (with energies E_s) to STM tip states $|t\rangle$ (with energies E_t), then the transition rate is

$$W = \sum_s \sum_t \frac{2\pi}{\hbar} |M_{t,s}|^2 \delta(E_t - E_s)(1 - f(E_t)) f(E_s), \tag{2.2}$$

where $f(E)$ is the Fermi-Dirac distribution, and $M_{t,s} = \langle t|H'|s\rangle$ is the tunneling matrix element.

Since the tip states form a continuum, we can replace the sum over t with an integral over $D_t(E_t)dE_t$ (for a tip density of states $D_t(E_t)$):

$$W \propto \sum_s \int_{-\infty}^{\infty} |M_{t,s}|^2 D_t(E_t) \delta(E_t - E_s)(1 - f(E_t)) f(E_s) dE_t. \tag{2.3}$$

If we apply a sample bias V (or equivalently, if we apply a voltage $-V$ on the tip), then we must change this expression for W to reflect the fact that the sample electrochemical potential shifts by $-eV$. After doing the integral over dE_t,

$$W \propto \sum_s |M_{t,s}|^2 D_t(E_s - eV)(1 - f(E_s - eV)) f(E_s). \tag{2.4}$$

The total tunneling current $I(V)$ is $-e$ multiplied by the difference of the transition rates from sample to tip and from tip to sample. Thus,

$$I(V) \propto \sum_s |M_{t,s}|^2 D_t(E_s - eV)(f(E_s - eV) - f(E_s)). \tag{2.5}$$

We assume the tip density of states is constant in the energy range of interest. Also, since our STM measurements are at low temperature, we approximate the Fermi-Dirac distribution as a step function, i.e. $\frac{\partial}{\partial V} f(E_s - eV) \approx e\delta(E_F - (E_s - eV))$, where E_F is the Fermi energy. Then, the differential conductance is

$$\frac{dI}{dV} \propto \sum_s |M_{t,s}|^2 \delta(eV - (E_s - E_F)). \tag{2.6}$$

Now, let's calculate the tunneling matrix element $M_{t,s} = \langle t|H'|s\rangle$. Define U_s to be the potential due to the sample and U_t to be due to the tip. Also assume that U_s and U_t do not overlap in space, i.e. $U_s = 0$ everywhere $U_t \neq 0$ and vice versa. Then, by simple application of the Schrödinger equation (with $E_t = E_s$) [27],

$$\begin{aligned} M_{t,s} = \langle t|H'|s\rangle &= \int_\tau \psi_t^* U_t \psi_s d^3\mathbf{r} = \int_\tau \left(\frac{\hbar^2}{2m} \nabla^2 \psi_t^* + E_t \psi_t^* \right) \psi_s d^3\mathbf{r} \\ &= \frac{\hbar^2}{2m} \int_\tau \left(\psi_s \nabla^2 \psi_t^* - \psi_t^* \nabla^2 \psi_s \right) d^3\mathbf{r} \\ &= \frac{\hbar^2}{2m} \int_\sigma \left(\psi_s \nabla \psi_t^* - \psi_t^* \nabla \psi_s \right) \cdot d^2\mathbf{r}, \end{aligned} \tag{2.7}$$

where ψ_t and ψ_s represent tip and sample wavefunctions, respectively. The above $d^3\mathbf{r}$ integrals are taken over the entire volume τ such that $U_t \neq 0$. I used Green's identity to convert the volume integral into a flux integral over a surface σ between the tip and the sample.

Putting everything together, we have

$$\frac{dI}{dV} \propto \sum_s \left| \int_\sigma (\psi_s \nabla \psi_t^* - \psi_t^* \nabla \psi_s) \cdot d^2\mathbf{r} \right|^2 \delta(eV - (E_s - E_F)). \tag{2.8}$$

2.1.2 The Tersoff-Hamann Theory

In the volume τ where the sample potential $U_s = 0$, $(\nabla^2 - \kappa^2)\psi_s = 0$, where $\hbar^2\kappa^2 = 2m(\Phi - E_s)$, and Φ is the sample work function. If we assume the tip wavefunction obeys $(\nabla^2 - \kappa^2)\psi_t = -\delta(\mathbf{r} - \mathbf{r}_0)$ and is spherically symmetric, then [28, 29]

$$\psi_t(\mathbf{r}) = \frac{e^{-\kappa|\mathbf{r}-\mathbf{r}_0|}}{4\pi|\mathbf{r}-\mathbf{r}_0|}. \tag{2.9}$$

In other words, we take the tip wavefunction to be an s-wave centered around the tip apex nucleus at $\mathbf{r}_0$. With this assumption, we find that $M_{t,s} = \langle t|H'|s\rangle \propto \psi_s(\mathbf{r}_0)$. Thus, the tunneling matrix element is proportional to the sample wavefunction evaluated at the tip apex.

Plugging in $M_{t,s} \propto \psi_s(\mathbf{r}_0)$ into Eq. 2.6,

$$\frac{dI}{dV} \propto \sum_s |\psi_s(\mathbf{r}_0)|^2 \delta(E_F + eV - E_s) = \mathrm{LDOS}(\mathbf{r}_0, E_F + eV). \tag{2.10}$$

Therefore, the experimentally measurable quantity dI/dV can be interpreted as the sample's total wavefunction probability density at position $\mathbf{r}_0$ and energy $E_F + eV$. This is called the local density of states (LDOS) and is a theoretically calculable quantity.

As depicted in Fig. 2.1, the total tunneling current is the integral of Eq. 2.10:

$$I(V) \propto \int_{E_F}^{E_F+eV} \sum_s |\psi_s(\mathbf{r}_0)|^2 \delta(E - E_s) dE. \tag{2.11}$$

2.1.3 Tunneling into Bloch States

According to the simplest 1D model for tunneling through a rectangular potential barrier, a sample state decays into vacuum as $\psi_s(z) = \psi_s(0)e^{-\kappa z}$, where $\kappa = (2m(\Phi - E_s)/\hbar^2)^{1/2}$, and Φ is the sample work function. However, a correction must be made for Bloch states in a crystal. For simplicity, suppose a Bloch state decaying into the vacuum has the form

$$\psi_s(x, y, z) = e^{i(k_x x + k_y y)} e^{-\gamma z}. \tag{2.12}$$

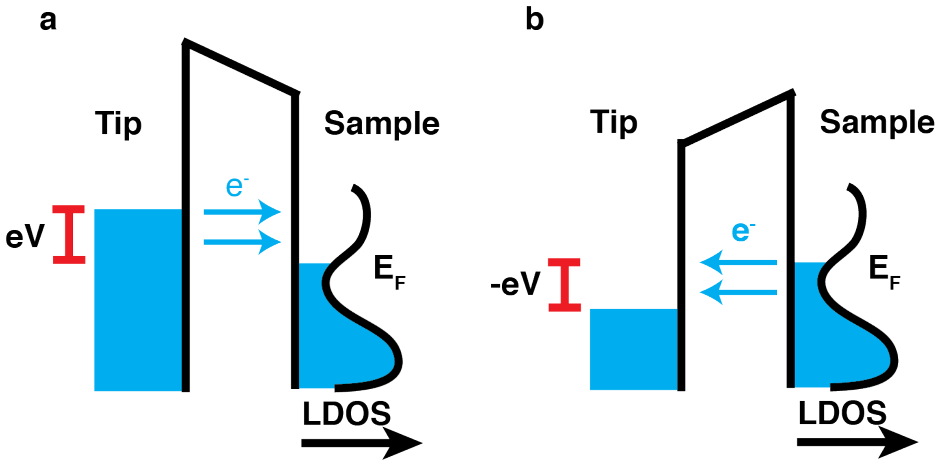

Figure 2.1: Tunneling between tip and sample. (a) The total tunneling current between an STM tip and a conducting sample is proportional to the integrated LDOS between E_F and $E_F + eV$. If $V > 0$, electrons tunnel from the tip to unoccupied states in the sample. (b) If $V < 0$, electrons tunnel from occupied states in the sample to the tip.

Then, plugging into the Schrödinger equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi = \kappa^2\psi \tag{2.13}$$

gives us

$$\gamma = \sqrt{\kappa^2 + k_{||}^2}, \tag{2.14}$$

where $k_{||} = \sqrt{k_x^2 + k_y^2}$ is the component of the crystal momentum parallel to the surface. Thus, states with a larger parallel component of momentum decay faster into the vacuum than states with a smaller parallel component of momentum. This will be important for interpreting dI/dV spectroscopy for graphene.

2.1.4 The Spectral Function and the Local Density of States

There is an interesting interpretation of the LDOS that is useful to be aware of. Define the "one-particle Green's function" as

$$G(\mathbf{r}', t', \mathbf{r}, t) = -i\langle 0|T\{\Psi(\mathbf{r}, t)\Psi^\dagger(\mathbf{r}', t')\}|0\rangle. \tag{2.15}$$

Here, $|0\rangle$ is the many-body ground state with N electrons, $\Psi^\dagger(\mathbf{r}', t')$ is a field operator that creates an electron at position $\mathbf{r}'$ at time t', and $\Psi(\mathbf{r}, t)$ destroys an electron at position $\mathbf{r}$ and time t . T is the time-ordering operator, which I can ignore if I take $t' = 0$ and $t > 0$. In the Heisenberg picture, the operators evolve as

$$\Psi(\mathbf{r}, t) = e^{iHt/\hbar}\Psi(\mathbf{r}, 0)e^{-iHt/\hbar}, \tag{2.16}$$

$$\Psi^\dagger(\mathbf{r},t) = e^{iHt/\hbar}\Psi^\dagger(\mathbf{r},0)e^{-iHt/\hbar}. \tag{2.17}$$

Putting this all together gives

$$G(\mathbf{r}',\mathbf{r},t) = -i\langle 0|\Psi(\mathbf{r},0)e^{-iHt/\hbar}\Psi^\dagger(\mathbf{r}',0)|0\rangle = -i\langle \mathbf{r}|e^{-iHt/\hbar}|\mathbf{r}'\rangle, \tag{2.18}$$

where the ground state energy is zero, and $|\mathbf{r}'\rangle = \Psi^\dagger(\mathbf{r}',0)|0\rangle$ is the ground state with an additional electron localized at $\mathbf{r}'$. Eq. 2.18 has a physical interpretation. It is the probability amplitude that an electron starting at position $\mathbf{r}'$ and subject to time evolution for time t will end up at position $\mathbf{r}$.

Suppose we have a complete, orthonormal set of $(N+1)$-electron energy eigenstates labeled by λ, i.e. states $\{c_\lambda^\dagger|0\rangle\}$ with excitation energies $\{E_\lambda\}$. Inserting the resolution of the identity twice yields

$$G(\mathbf{r}',\mathbf{r},t) = -i\sum_{\lambda\lambda'}\langle \mathbf{r}|\lambda\rangle\langle\lambda|e^{-iHt/\hbar}|\lambda'\rangle\langle\lambda'|\mathbf{r}'\rangle = -i\sum_{\lambda} e^{-iE_\lambda t/\hbar}\psi_\lambda^*(\mathbf{r}')\psi_\lambda(\mathbf{r}), \tag{2.19}$$

where we let $|\lambda\rangle = c_\lambda^\dagger|0\rangle$ and $\psi_\lambda(\mathbf{r}) = \langle\mathbf{r}|\lambda\rangle$.

Let $\mathbf{r}' = \mathbf{r}$ because the location where the STM injects electrons into the sample is also the measurement location. Also, STM measurements are "slow," so we are more interested in energy-resolved spectroscopy rather than time-resolved dynamics. Thus, we should Fourier transform the Green's function with respect to the time variable:

$$G(\mathbf{r},\omega) = -i\sum_\lambda\int_0^\infty e^{i(\omega-E_\lambda/\hbar)t}e^{-\eta t}|\psi_\lambda(\mathbf{r})|^2dt. \tag{2.20}$$

I put in $e^{-\eta t}$ to help the integral converge. Physically, $1/\eta$ represents the lifetime of a quasiparticle excitation. For non-interacting fermions, we can take the limit $\eta\to 0$ at the end of the calculation. Evaluating Eq. 2.20 gives us

$$G(\mathbf{r},\omega) = -i\sum_\lambda\left[\frac{e^{i(\omega-E_\lambda/\hbar)t}e^{-\eta t}}{i(\omega-E_\lambda/\hbar)-\eta}\right]_0^\infty|\psi_\lambda(\mathbf{r})|^2 = \sum_\lambda\left[\frac{1}{\omega-E_\lambda/\hbar+i\eta}\right]|\psi_\lambda(\mathbf{r})|^2. \tag{2.21}$$

We can define the "spectral function" $A(\mathbf{r},\omega)$ for interpreting STM data:

$$A(\mathbf{r},\omega) = -\frac{1}{\pi}\mathrm{Im}[G(\mathbf{r},\omega)] = \frac{1}{\pi}\sum_\lambda\frac{\eta}{(\omega-E_\lambda/\hbar)^2+\eta^2}|\psi_\lambda(\mathbf{r})|^2. \tag{2.22}$$

As a side note, there is an equivalent momentum-space spectral function for interpreting angle-resolved photoemission spectroscopy (ARPES) data:

$$A(k,\omega) = -\frac{1}{\pi}\mathrm{Im}[G(k,\omega)] = \frac{1}{\pi}\frac{\mathrm{Im}\Sigma}{(\omega-E_k/\hbar-\mathrm{Re}\Sigma)^2+(\mathrm{Im}\Sigma)^2}, \tag{2.23}$$

where I have introduced the so-called "self-energy" Σ.

For Eq. 2.22, take the $\eta \to 0$ limit:

$$\frac{1}{\pi} \frac{\eta}{(\omega - E_\lambda/\hbar)^2 + \eta^2} \to \delta(\omega - E_\lambda/\hbar). \tag{2.24}$$

Finally, we get an expression for the LDOS out of this formalism:

$$A(\mathbf{r}, \omega) \propto \sum_\lambda |\psi_\lambda(\mathbf{r})|^2 \delta(\hbar\omega - E_\lambda) = \text{LDOS}(\mathbf{r}, \hbar\omega). \tag{2.25}$$

The spectral function obeys certain sum rules. For example, its integral over $d\omega$ must give 1. Consequently, the LDOS obeys sum rules as well. The integral of the LDOS over all space must be the density of states (DOS):

$$\int \text{LDOS}(\mathbf{r}, E) d^3\mathbf{r} = \text{DOS}(E). \tag{2.26}$$

And of course, the integral of the DOS over all energies should give the total number of states.

2.2 Scanning Tunneling Microscopy Operation and Instrumentation

2.2.1 Scanning Tunneling Microscopy Basics

The experiments in this dissertation were performed using an ultra-high vacuum (UHV) Omicron low-temperature (LT) STM. The pressures were typically 10^{-10} torr or lower, and the temperatures were approximately 4 to 5 K (liquid-helium temperature). A photograph of our STM can be seen in Fig. 2.2.

A schematic diagram of how an STM works is depicted in Fig. 2.3. Using a piezoelectric motor, a sharp metal wire (the tip, usually made of PtIr or W) is brought less than 1 nm away from a surface. A bias $-V_s$ (conventionally, V_s is the voltage on the sample relative to the tip) is applied to the tip, causing a tunneling current I. The tunneling current is converted to a voltage via a preamplifier and is recorded by a computer. While the tip is scanned across the surface, a feedback loop controls the voltage on a piezo such that the tip-sample distance is adjusted to maintain a constant I. Changes in the tip-sample distance are recorded while the tip moves across the surface to produce a topographic image of the surface. This topographic image contains information about the positions of the atoms on the surface as well as information about the LDOS (since according to Eq. 2.11, the tunneling current depends on the LDOS). Since the tunneling current is exponentially sensitive to the tip height z, STM topographic images can have extremely fine vertical and lateral resolution.

The piezo motor in our Omicron LT-STM operates on stick-slip motion. Fig. 2.4b shows a photograph of our z coarse walker. Three little sapphire balls glued to shear piezos move

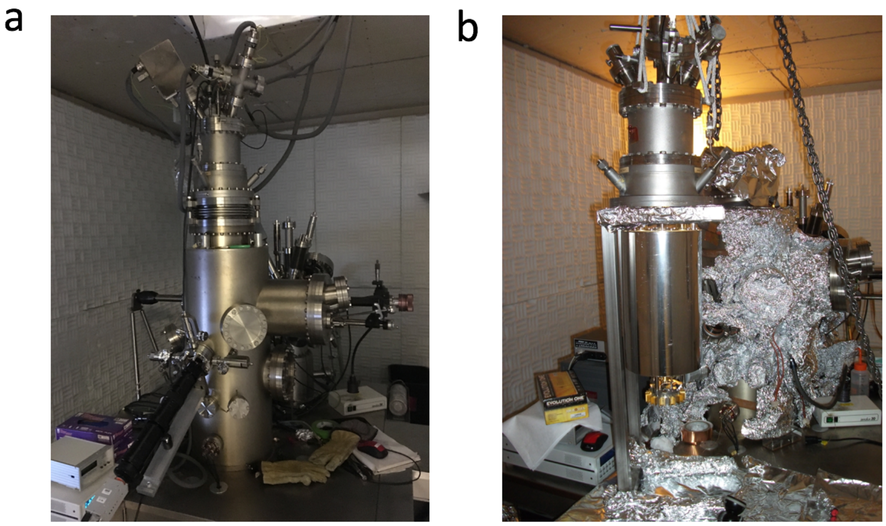

Figure 2.2: UHV Omicron LT-STM. (a) Photograph of an Omicron LT-STM system, which consists of an STM chamber, preparation chamber, and loadlock chamber. The tip-sample junction is visually accessible and can be seen via a long-distance optical microscope. (b) The STM is removed from the vacuum chamber. The STM rests inside a magnetic damping system hanging below liquid helium and liquid nitrogen cryostats.

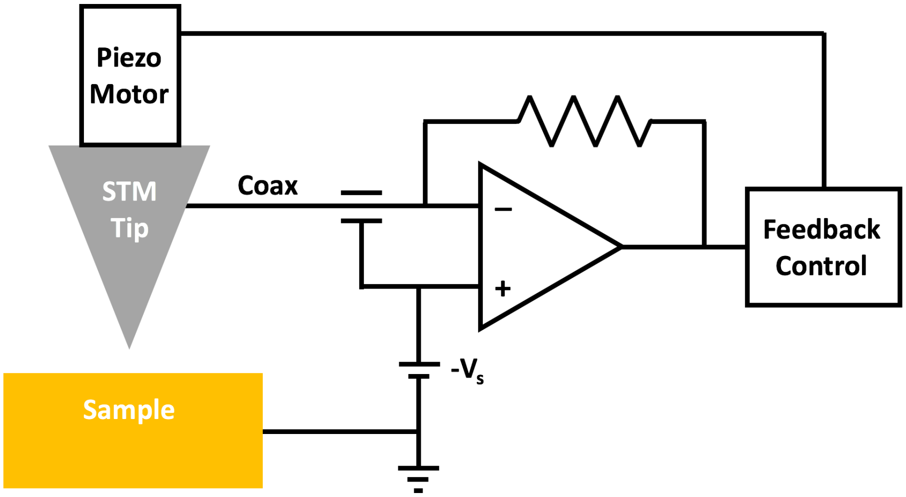

Figure 2.3: STM schematic diagram. An STM tip is placed ∼5 Å away from a sample surface. A bias voltage $-V_s$ is applied on the tip to induce tunneling. The tunneling current is measured, and a feedback loop is employed to adjust the tip-sample distance such that the tunneling current I is constant as the tip raster scans across the surface.

along three sapphire rails for coarse vertical alignment of the tip and the sample. An attached cylindrical “fine piezo” provides precision motion along the x-, y-, and z-directions for scanning the tip across the sample surface. The tip is magnetically attached to the very end of the fine piezo.

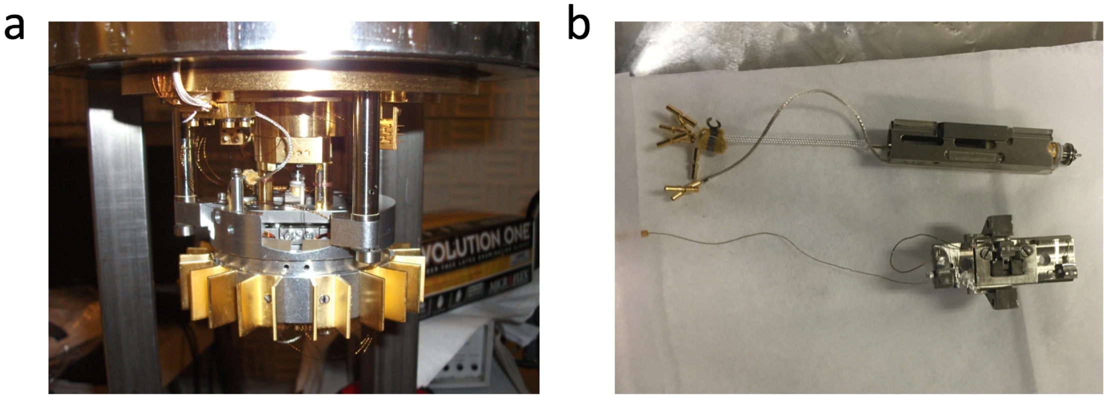

Figure 2.4: STM tip and piezoelectric actuator. (a) Photograph of piezoelectric motor sitting inside gold-plated fins used for magnetic damping. (b) A dismantled piezo motor. Sapphire balls are glued to a “coarse piezo.” The sapphire balls can freely slide against sapphire rails for motion in the z-direction. The STM tip is mounted atop a cylindrical “fine piezo” that provides picometer-resolved motion in the x-, y-, and z-directions.

In addition to obtaining topographic images of surfaces, STM is also a powerful tool for directly imaging electronic wavefunctions. In what is known as scanning tunneling spectroscopy (STS), we can add a small a.c. “wiggle” voltage (typically with amplitude 1 to 8 mV root mean square (rms) and frequency 400 to 700 Hz) to the tip bias and monitor the tunneling current response. This is achieved by feeding the current I into a lock-in amplifier set to filter out everything except the first harmonic, giving the derivative dI/dV.

To obtain a dI/dV point spectrum, the STM tip is fixed at a certain location (x, y, z), and the sample bias V_s is swept through a range of interest while the STM feedback loop is open (i.e. off). Since, as mentioned in Sec. 2.1.2, dI/dV is proportional to the LDOS, a dI/dV spectrum provides information about the distribution of sample eigenstates at (x, y, z) as a function of energy $E = E_F + eV_s$.

Alternatively, one can obtain spatial information about the LDOS at a fixed energy. This can be done in two ways: dI/dV grids and dI/dV maps. A dI/dV grid is simply a set of energy-resolved dI/dV spectra obtained at an array of spatial locations and is often very time-consuming to obtain. A dI/dV map, however, is obtained by acquiring dI/dV simultaneously while the tip raster scans across the surface with the feedback loop closed (i.e. on). Both methods, but especially the dI/dV map, do not give dI/dV on an isosurface of constant z. Rather, z is a function of (x, y) because the tip moves up and down to keep the tunneling current constant.

Suppose that we are interested in $\mathrm{LDOS}(x, y, z_0, E)$, which is a function of (x, y) for fixed z_0 and energy E. Since wavefunctions decay exponentially into a tunneling barrier, we assume that

$$\mathrm{LDOS}(x, y, z, E) = \mathrm{LDOS}(x, y, z_0, E)e^{-(z-z_0)/\lambda}, \tag{2.27}$$

where λ is the length scale of the exponential decay, which we assume to be constant (although it really is not). Here, z is the STM tip height, which is a function of (x, y) as the tip scans across the surface and the feedback loop adjusts z to keep the tunneling current at a constant I_0. The current is given by

$$\begin{aligned} I_0 &\propto \left| \int_0^{V_s} \mathrm{LDOS}(x, y, z, E_F + eV)dV \right| \\ &\propto \left| \int_0^{V_s} \mathrm{LDOS}(x, y, z_0, E_F + eV)dV \right| e^{-(z-z_0)/\lambda}. \end{aligned} \tag{2.28}$$

This implies that the measured dI/dV signal is

$$\frac{dI}{dV}(x, y, z)\bigg|_{V=V_s} \propto \mathrm{LDOS}(x, y, z, E_F + eV_s) \propto \frac{\mathrm{LDOS}(x, y, z_0, E)}{\left| \int_0^{V_s} \mathrm{LDOS}(x, y, z_0, E_F + eV')dV' \right|}. \tag{2.29}$$

Therefore, dI/dV maps measure the LDOS with a spatially varying proportionality factor. Caution is advised when interpreting dI/dV and comparing to theoretical LDOS calculations [30, 31].

2.2.2 Tip Preparation

A crucial aspect of STM is preparing the tip. Without a good tip, an STM measurement can be worthless. Indeed, since dI/dV is proportional to a product of the tip DOS and the sample LDOS (see Eq. 2.5), one can produce wildly different dI/dV signals by simply changing the tip DOS. A lot of care is required to ensure a constant tip DOS such that dI/dV can be interpreted as the sample LDOS.

Our STM tips are made of either W or an 80% platinum/20% iridium alloy. In this section, I discuss how we prepare PtIr tips, since this comprises most of the tips that we use. W tips are produced by etching W wires in NaOH.

To make a PtIr tip, first dissolve anhydrous $\mathrm{CaCl_2}$ powder in ultrapure water (distilled water is probably sufficient) until saturation. The $\mathrm{CaCl_2}$ solution will get hot as you mix the $\mathrm{CaCl_2}$ and water. Then, dilute the saturated $\mathrm{CaCl_2}$ solution with an equal volume of water (a 2:1 ratio is probably better, but I can never remember if it is 2:1 $\mathrm{CaCl_2}$ to water or the reverse). Bend a W wire into a ring shape, and immerse it just under the surface of the solution. Dip the very end of a PtIr wire into the solution at the center of the W ring. Use a transformer to step down the wall power to 35 to 50 V rms, and apply this voltage between the PtIr wire and W ring. The apex of the PtIr wire should emit a sound (and possibly also light). When the sound stops, the end of the PtIr wire will have been etched by the $CaCl_2$ solution. Check the PtIr wire under an optical microscope to see if the PtIr is sufficiently sharp to be used as an STM tip (this varies depending on one's requirement for the experiment). If the PtIr is not sufficiently sharp, dip it back into the solution, and try again. If the PtIr is sufficiently sharp, then squirt ethanol or isopropanol onto the PtIr to clean it. Be careful not to damage the new tip apex.

For some experiments, we have also electron-beam heated the tip in UHV. This is called "flashing" and is typically not necessary. The PtIr tip is placed into the STM and needs to be calibrated for dI/dV spectroscopy. This is achieved by "poking" and "pulsing" the tip on an Au(111) crystal (prepared by sputtering with 0.5 keV Ar^+ ions and annealing to 375°C).

To poke a tip is to controllably crash the tip $\sim$1 nm into the Au(111) surface. The selected depth of the poke varies depending on how "bad" the tip is. This is assessed by examining the surface of Au(111) underneath the poke location. A poke typically leaves behind a protrusion in Au(111). An ideal poke leaves a single round protrusion with an aspect ratio of 4 nm in diameter per 1 nm in height. Unfortunately, the poke is often not ideal and leaves behind a strangely shaped protrusion – or worse, multiple protrusions. By a principle of reciprocity, an STM topograph is simultaneously an image of the tip and the sample. A misshapen poke is indicative of an oddly shaped tip that will produce anomalous distortions of objects in topographic images. Multiple protrusions from a poke is indicative of a "double tip" or a "multi-tip," which can lead to multiple copies of the same object in a single image. These "bad tips" are undesirable and must be poked away. An example of a protrusion from a "good poke" can be seen at bottom of Fig. 2.5.

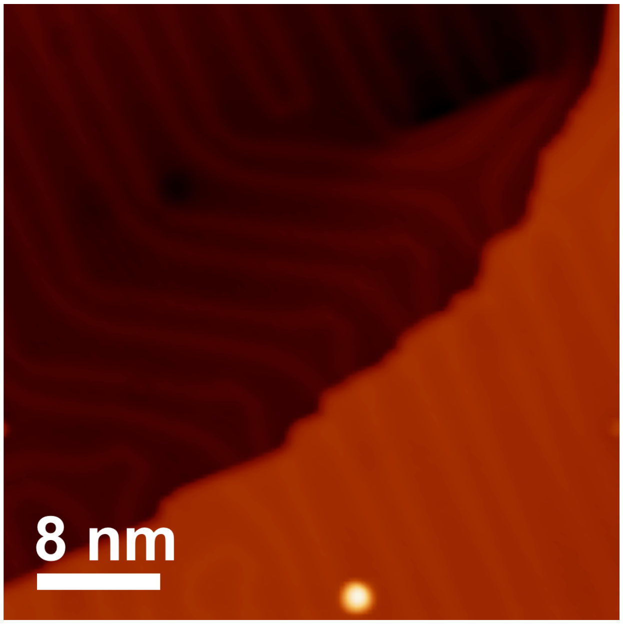

Figure 2.5: Au(111) topographic image. The twinned stripes are a surface "herringbone" reconstruction. A protrusion resulting from a poke is seen at the bottom.

If a tip is “really bad,” i.e. poking does not appear to improve the tip, one can “pulse” the tip. This entails applying a short 10 V pulse on the tip (with duration usually less than 1 second) while the STM is in tunneling regime. If the tip is “really, really bad,” i.e. the pulses do not appear to improve the tip, one can scratch the tip into tantalum foil.

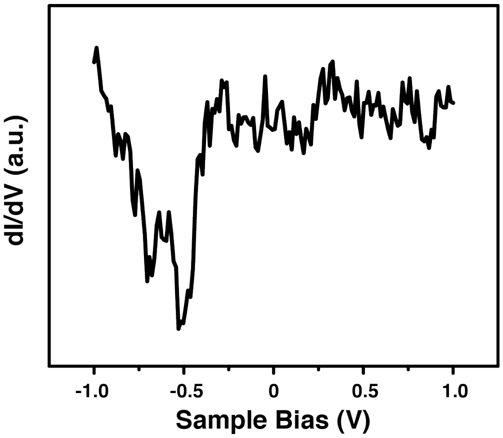

Figure 2.6: Au(111) dI/dV spectrum. The onset of the Au(111) surface state is around $V_s = -0.5$ V. Since dI/dV is proportional to the LDOS and we are often uninterested in the proportionality constant, the dI/dV signal is plotted in arbitrary units (a.u.). The Fermi energy is $V_s = 0$ V.

Usually, poking and pulsing suffices for fixing an STM tip. Once a desirable tip shape is obtained (by examining the protrusion that results from a poke), one acquires a dI/dV spectrum on Au(111). The spectrum should look like that in Fig. 2.6, with a sharp increase at $V_s = -0.5$ V (the onset of the Au(111) surface state) but relatively featureless otherwise [32]. If the obtained dI/dV spectrum does not look like Fig. 2.6, poke until it does. Afterwards, I usually check the dI/dV spectrum again after ∼15 minutes to make sure the tip is stable.

2.3 Scanning Tunneling Microscopy of Graphene

Fig. 2.7 shows an atomically resolved STM topographic image of graphene/BN. The interference of the graphene and BN atomic lattices creates a superlattice structure known as a moiré pattern [33, 34].

As mentioned in Sec. 1.2.2, the DOS of graphene is proportional to $|E|$. Therefore, as depicted in Fig. 2.8b, a dI/dV spectrum acquired on graphene should be V-shaped. However, that is not the case for a tip that is properly calibrated against an Au(111) surface state. In Sec. 2.1.3, we showed that states with larger crystal momentum decay faster into the vacuum. This implies that STM is much more sensitive to tunneling into states at the Γ point of the Brillouin zone compared to the low-energy states at K and K'.

The low-energy states at K and K', however, are mixed with higher-energy Γ-point states in the σ and σ^* bands by electron-phonon coupling. We can interpret this as a

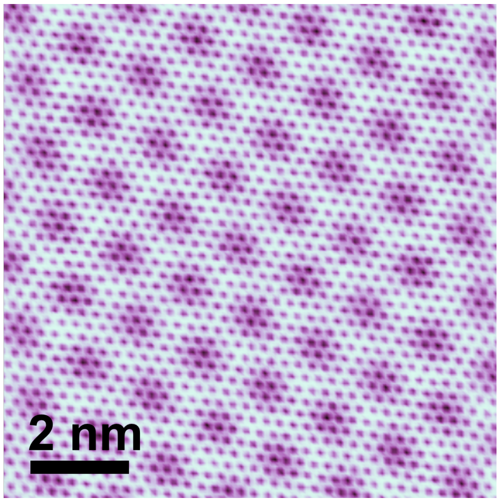

Figure 2.7: Graphene/BN topographic image with atomic lattice and moiré pattern.

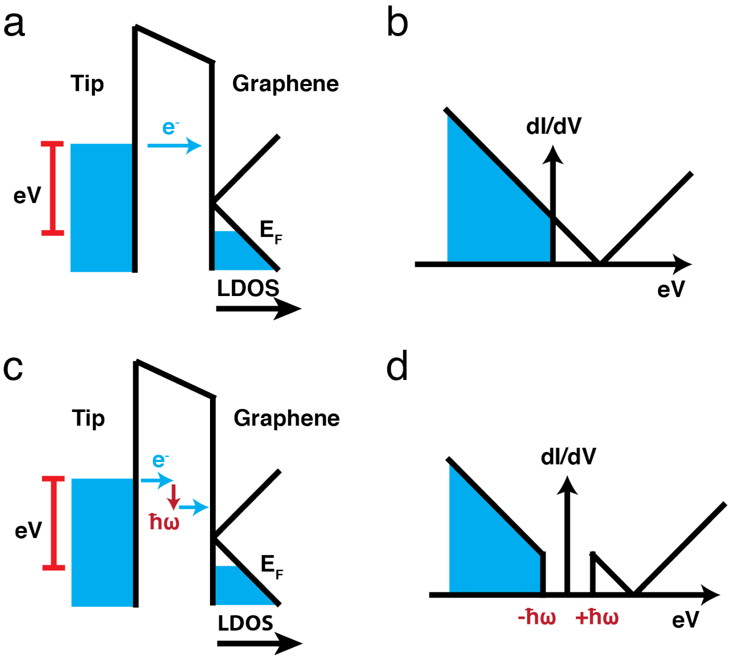

Figure 2.8: Elastic and inelastic tunneling into graphene. (a) An electron elastically tunneling from the tip to an unoccupied state in graphene. (b) Since graphene has a linear DOS, the expected dI/dV spectrum due to the process depicted in (a) will be V-shaped with a minimum at the graphene Dirac point. (c) An electron can also inelastically tunnel from tip to graphene (or vice versa) with the emission (or absorption) of a phonon with energy $\hbar\omega$. (d) The contribution to dI/dV due to the inelastic process in (c) has a "phonon gap" of energy $2\hbar\omega$.

second-order inelastic tunneling process (depicted in Fig. 2.8c) in which an electron makes a virtual transition between a state in the σ or σ^* band at Γ and a state in the π or π^* band at K or K'. To conserve momentum, this must be accompanied with the emission or absorption of a K- or K'-point excitation (which will be an $\hbar\omega = 60 \sim 70$ meV phonon). If this phonon-assisted inelastic tunneling process dominates over the elastic channel, dI/dV will look like Fig. 2.8d, which has a $2\hbar\omega = \sim$130 meV gap-like feature centered at the Fermi energy ($V_s = 0$ V) because it costs a minimum of $\hbar\omega$ energy to excite the phonon. It is important to note that such a gap is the result of inelastic tunneling and is not an actual electronic structure feature intrinsic to graphene. More information about phonon-assisted inelastic tunneling can be found in Refs. [35] and [36].

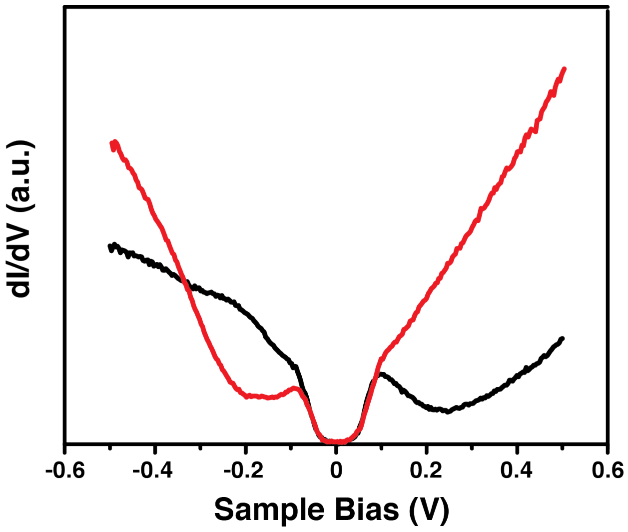

Figure 2.9: Graphene dI/dV spectrum. The black curve is for p-doped graphene, while the red curve is for n-doped graphene. A ~130 meV gap-like feature can be seen straddling the Fermi level.

In our experiments, the contribution of inelastic tunneling to dI/dV appears to be much larger than that of elastic tunneling. The dI/dV curves in Fig. 2.9 each show a symmetric ~130 meV gap-like feature around E_F ($V_s = 0$ V). There is an additional local minimum in each dI/dV curve. In the black curve, which is for p-doped graphene, a local minimum appears to the right of E_F. In the red curve, which is for n-doped graphene, a local minimum appears left of E_F. These are interpreted as minima in the LDOS due to the Dirac point and are shifted by $\hbar\omega$ relative to the true energy locations of the Dirac point.

Chapter 3

The Dirac Equation

I include here a chapter on the Dirac equation itself as it is understood in relativistic quantum mechanics. This chapter strays far from condensed matter physics, but I feel that it is relevant for understanding the context of the phenomena discussed in this dissertation.

A good treatment on the Dirac equation can be found in J. J. Sakurai's "Modern Quantum Mechanics" [19], although I personally used an amazing set of lecture notes by Prof. Robert Littlejohn as a reference for writing this chapter.

3.1 The Klein-Gordon Equation

At the turn of the 20th century, Albert Einstein and Henri Poincaré showed that spacetime transformations in inertial frames respected the following "Minkowski" invariant:

$$(ct)^2 - x^2 - y^2 - z^2 = (ct')^2 - x'^2 - y'^2 - z'^2. \tag{3.1}$$

A spacetime "4-vector" that has coordinates (ct, x, y, z) in one frame of reference will have coordinates (ct', x', y', z') in another frame of reference. Such transformations (if they keep the origin fixed) are called the Lorentz transformations and include rotations and "boosts" (transformations between frames moving at constant velocity with respect to each other).

Eq. 3.1 is often written in a shorthand

$$x^\mu x_\mu = x'^\nu x'_\nu, \tag{3.2}$$

where $x^0 = x_0 = ct$, $x^1 = -x_1 = x$, $x^2 = -x_2 = y$, $x^3 = -x_3 = z$ (and similar for the primed variables). Summation over the Greek indices is implied.

Also components of a 4-vector, the energy E and momentum $\mathbf{p}$ must obey the same relation. If $p^0 = p_0 = E/c$, $p^1 = -p_1 = p_x$, $p^2 = -p_2 = p_y$, and $p^3 = -p_3 = p_z$, then

$$p^\mu p_\mu = p'^\nu p'_\nu. \tag{3.3}$$

This invariant is the mass squared: $(mc)^2 = p^\mu p_\mu$. For $p = |\mathbf{p}|$,

$$E^2 = (pc)^2 + (mc^2)^2. \tag{3.4}$$

This is the heralded energy-momentum relationship for a relativistic particle.

To obtain a quantum theory, we can replace E by $i\hbar\frac{\partial}{\partial t}$ and $\mathbf{p}$ by $-i\hbar\nabla$. This gives us the Klein-Gordon equation

$$\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} - \nabla^2\Psi = -\left(\frac{mc}{\hbar}\right)^2\Psi. \tag{3.5}$$

If we write $\partial_\mu = \frac{\partial}{\partial x^\mu}$ and $\partial^\mu = \frac{\partial}{\partial x_\mu}$, then

$$\partial^\mu\partial_\mu\Psi = -\left(\frac{mc}{\hbar}\right)^2\Psi. \tag{3.6}$$

Notice that, unlike the Schrödinger equation, the Klein-Gordon equation treats time and space on equal footing. The equation has second derivatives in both time and space.

3.2 The Three-Dimensional Dirac Equation

3.2.1 Derivation from the Klein-Gordon Equation

The fact that the Klein-Gordon equation has a second derivative in time leads to an unfortunate property: it allows solutions with negative probability density. To remedy this, Paul Dirac invented an equation that only has first derivatives in time and space:

$$i\hbar\frac{\partial\Psi}{\partial t} = -i\hbar c\sum_{j=1}^{3}\alpha_j\frac{\partial\Psi}{\partial x_j} + mc^2\beta\Psi = 0. \tag{3.7}$$

Here, α_j and β are to be determined by imposing the energy-momentum relation Eq. 3.4. The wavefunction Ψ must be a solution to the Klein-Gordon equation (Eq. 3.5), so we apply the Dirac Hamiltonian $H = c\boldsymbol{\alpha}\cdot\mathbf{p} + mc^2\beta$ twice to get an equation with second derivatives:

$$-\hbar^2\frac{\partial^2\Psi}{\partial t^2} = -\hbar^2c^2\sum_{jk}\alpha_j\alpha_k\frac{\partial^2\Psi}{\partial x_j\partial x_k} - i\hbar mc^3\sum_j\left(\alpha_j\beta + \beta\alpha_j\right)\frac{\partial\Psi}{\partial x_j} + m^2c^4\beta^2\Psi = 0. \tag{3.8}$$

Comparing this to Eq. 3.5, we arrive at the anticommutation relations

$$\begin{aligned}\{\alpha_j, \alpha_k\} &= 2\delta_{jk},\\ \{\alpha_j, \beta\} &= 0,\\ \{\beta, \beta\} &= 2.\end{aligned} \tag{3.9}$$

We can deduce from these relations that α_j and β are $N \times N$ traceless, Hermitian matrices that square to 1. If they square to 1, then their eigenvalues are ± 1, so N must be even (or else they would have nonzero trace).

The obvious guess is $N = 2$, for which $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$ forms a basis for the space of 2×2 Hermitian matrices. This set is not big enough to construct α_j (three matrices) and β (one matrix). It turns out $N = 4$ is the smallest matrix size allowed for the 3D Dirac equation.

There are many possible equivalent choices for α_j and β. The most popular are Dirac-Pauli, Weyl, and Majorana. The Dirac-Pauli representation is

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \; \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \tag{3.10}$$

Here, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

The Dirac equation is often written as

$$i\hbar\gamma^\mu \partial_\mu \Psi = mc\Psi, \tag{3.11}$$

which is just Eq. 3.7 with β multiplied on both sides. The γ^μ are defined as

$$\boldsymbol{\gamma} = \beta\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}, \; \gamma^0 = \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \tag{3.12}$$

This form of the Dirac equation is often preferred because it is manifestly Lorentz covariant. Immediately following from the anticommutation relations in Eq. 3.9, the γ^μ obey

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{3.13}$$

This is called the Clifford algebra. $g^{\mu\nu}$ is the metric tensor such that $x^\mu = g^{\mu\nu} x_\nu$:

$$g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.14}$$

We can multiply Eq. 3.11 on both sides by $\Psi^\dagger\gamma^0$ to get $i\hbar\Psi^\dagger\gamma^0\gamma^\mu\partial_\mu\Psi = mc\Psi^\dagger\gamma^0\Psi$. Then taking the conjugate transpose of Eq. 3.11 (and using an easily verified identity $\gamma^0 (\gamma^\mu)^\dagger = \gamma^\mu\gamma^0$), we get $-i\hbar\partial_\mu\Psi^\dagger\gamma^0\gamma^\mu\Psi = mc\Psi^\dagger\gamma^0\Psi$. Subtracting these expressions gives

$$\Psi^\dagger\gamma^0\gamma^\mu\partial_\mu\Psi + \partial_\mu\Psi^\dagger\gamma^0\gamma^\mu\Psi = 0 \tag{3.15}$$

or equivalently,

$$\partial_\mu \left(\Psi^\dagger\gamma^0\gamma^\mu\Psi\right) = \frac{1}{c}\frac{\partial}{\partial t}\left(\Psi^\dagger\Psi\right) + \nabla \cdot \left(\Psi^\dagger\boldsymbol{\alpha}\Psi\right) = 0. \tag{3.16}$$

This looks just like a continuity equation for

$$\rho = \Psi^\dagger\Psi, \tag{3.17}$$

$$\mathbf{J} = c\Psi^\dagger\boldsymbol{\alpha}\Psi. \tag{3.18}$$

We can interpret $\Psi^\dagger\Psi$ as the wavefunction probability density (just like in nonrelativistic quantum mechanics) and $c\boldsymbol{\alpha}$ as a velocity operator. We conclude that the Dirac equation is a bona fide wave equation that conserves probability.

3.2.2 Positive and Negative Energy Eigenvalues

It is straightforward to check that the Dirac Hamiltonian has positive and negative energy eigenvalues (see also Fig. 3.1):

$$E = \pm\sqrt{(pc)^2 + (mc^2)^2}. \tag{3.19}$$

The solutions of the Dirac equation with positive energy eigenvalues are easily understood as particle states with momentum $\mathbf{p} = (p_x, p_y, p_z)$ and mass m. The solutions with negative energy eigenvalues, however, are not as easy to interpret. If these negative-energy states exist, it could be possible for Dirac particles to infinitely cascade to lower and lower energies by emitting electromagnetic radiation. This is clearly absurd, so Dirac hypothesized that all of the negative-energy states were already occupied. The Pauli exclusion principle would then forbid fermions from falling into lower-energy states, and the negative-energy states would no longer be a concern.

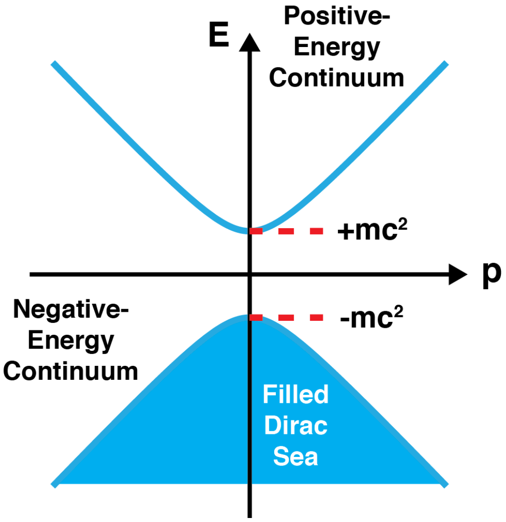

Figure 3.1: Positive and negative energy eigenvalues of the Dirac Hamiltonian. The Dirac equation has both positive-energy and negative-energy solutions, with $E = \sqrt{(pc)^2 + (mc^2)^2}$ and $E = -\sqrt{(pc)^2 + (mc^2)^2}$, respectively. The positive energies comprise a continuum of electron states, while the negative-energy continuum is a filled "Dirac sea" of positron states. There is a $2mc^2$ gap in the spectrum.

This interpretation of the negative-energy states has an interesting consequence. If energy at least $2mc^2$ is supplied, a particle in a negative-energy state can be promoted into a positive-energy state. This leaves behind a hole, an unoccupied state in this "Dirac sea" of negative-energy particles. This hole would behave like a particle with the same mass m but opposite charge.

A more modern interpretation does not invoke an infinite number of particles occupying an infinite number of negative-energy states. The Dirac equation is not actually an equation for a single particle but rather for a many-body quantum field. The quantum field gives rise to particles (e.g. electrons) and "antiparticles" (e.g. positrons). However, the Dirac sea picture is perfectly valid in solids, where there are valence bands that actually are completely filled with electrons.

3.2.3 Lorentz Symmetry of the Dirac Equation

Since the Dirac equation is a relativistic wave equation, it must be invariant under the Lorentz transformations. Let Λ be a rotation, a boost, or some combination of rotations and boosts that transforms a spacetime 4-vector x into x':

$$x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}. \tag{3.20}$$

Also define $D(\Lambda)$ such that

$$\Psi'(x) = D(\Lambda)\Psi(\Lambda^{-1}x) \tag{3.21}$$

is the Lorentz-transformed wavefunction. Since Ψ' must also be a solution to the Dirac equation, we have

$$i\hbar\gamma^{\mu}\frac{\partial}{\partial x^{\mu}}\left[D(\Lambda)\Psi(\Lambda^{-1}x)\right] = mc\left[D(\Lambda)\Psi(\Lambda^{-1}x)\right]. \tag{3.22}$$

Multiply both sides of this equation by $D^{-1}(\Lambda)$, and let $y^{\mu} = (\Lambda^{-1})^{\mu}_{\nu}x^{\nu}$. We obtain

$$i\hbar D^{-1}(\Lambda)\gamma^{\mu}D(\Lambda)\frac{\partial}{\partial x^{\mu}}\Psi(y) = mc\Psi(y). \tag{3.23}$$

By the chain rule, $\frac{\partial}{\partial x^{\mu}} = \frac{\partial}{\partial y^{\nu}}\frac{\partial y^{\nu}}{\partial x^{\mu}} = \frac{\partial}{\partial y^{\nu}}(\Lambda^{-1})^{\nu}_{\mu}$,

$$i\hbar D^{-1}(\Lambda)\gamma^{\mu}D(\Lambda)(\Lambda^{-1})^{\nu}_{\mu}\frac{\partial}{\partial y^{\nu}}\Psi(y) = mc\Psi(y). \tag{3.24}$$

Comparing Eq. 3.24 to the Dirac equation, we obtain the following identity:

$$D^{-1}(\Lambda)\gamma^{\mu}D(\Lambda) = \Lambda^{\mu}_{\nu}\gamma^{\nu}. \tag{3.25}$$

As long as $D(\Lambda)$ obeys this relation, the Dirac equation will have the symmetries of the Lorentz group. The heuristic (but not quite true) explanation of Eq. 3.25 is that the γ^{μ} matrices rotate like a 4-vector.

3.3 The Two-Dimensional Dirac Equation

Much of the formalism developed in the previous section on the 3D Dirac equation also applies to two dimensions. The 2D Dirac equation is the same as Eq. 3.7 except without the α_3 term:

$$i\hbar\frac{\partial\Psi}{\partial t} = -i\hbar c\sum_{j=1}^{2}\alpha_j\frac{\partial\Psi}{\partial x_j} + mc^2\beta\Psi = 0. \tag{3.26}$$

The anticommutation relations in Eq. 3.9 continue to apply, meaning that α_1, α_2, and β must be of even dimension. However, unlike the 3D Dirac equation, we can use simpler 2×2 matrices because the Pauli matrices themselves are a set of three matrices that satisfy the anticommutation relations.

$$\alpha_1 = \sigma_x,\ \alpha_2 = \sigma_y,\ \beta = \sigma_z. \tag{3.27}$$

The Dirac Hamiltonian is thus

$$H = -i\hbar c\boldsymbol{\sigma}\cdot\nabla + mc^2\sigma_z. \tag{3.28}$$

Here, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$.

The Dirac-Pauli γ^μ matrices become

$$\gamma^0 = \sigma_z,\ \gamma^1 = i\sigma_y,\ \gamma^2 = -i\sigma_x, \tag{3.29}$$

and the continuity equation (Eq. 3.16) continues to hold in two dimensions by a similar construction.

Notice that the 2D and 3D Dirac equations have $\boldsymbol{\sigma}$ in them. Unlike the Schrödinger equation, where the spin degree of freedom needs to be kludged in, spin naturally emerges from the Dirac equation. Hence, the Dirac equation is an equation of motion for spin $\frac{1}{2}$ particles called "Dirac fermions." Many spin-related phenomena such as spin-orbit coupling, hydrogen fine structure, and the electron g-factor also originate from the Dirac equation.

3.3.1 The Nonrelativistic Limit

We can couple the Dirac equation to electric and magnetic fields:

$$i\hbar\frac{\partial\Psi}{\partial t} = c\boldsymbol{\sigma}\cdot\left[-i\hbar\nabla - \frac{q}{c}\mathbf{A}\right]\Psi + mc^2\sigma_z\Psi + q\Phi\Psi, \tag{3.30}$$

where q is the charge of the Dirac fermion ($-e$ for an electron), $\mathbf{A}$ is the magnetic vector potential, and Φ is the electric scalar potential. Using the Heisenberg equation of motion, one could derive the Lorentz force law. Furthermore, one could also obtain the Schrödinger equation as the nonrelativistic limit of the Dirac equation. It is rather tedious to do it, so I only quote the result here. Let us write Ψ as

$$\Psi = e^{-imc^2t/\hbar}\begin{pmatrix}\psi\\ \zeta\end{pmatrix}. \tag{3.31}$$

Then, after a bunch of work and judiciously chosen approximations...

$$i\hbar\frac{\partial\psi}{\partial t} = \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}\mathbf{A}\right)^2\psi + q\Phi\psi - \frac{\hbar}{2mc}qB\psi, \tag{3.32}$$

where $B = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial_y}$. The first term on the right-hand side is obviously the kinetic energy, the second term is the electric potential energy, and the third term is the Zeeman term.

3.3.2 Rotations and Boosts

We can obtain analytic forms for the $D(\Lambda)$ in Sec. 3.2.3 by expanding them in terms of "infinitesimal generators." I will not do that here, but I will quote the results for the 2D Dirac equation. To rotate a wavefunction Ψ by an angle θ, multiply by

$$D(\theta) = e^{-i\theta\sigma_z/2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}. \tag{3.33}$$

As an example, it turns out that the spin of a massless particle is locked parallel to its momentum, so the wavefunction Ψ of a positive-energy massless Dirac fermion with momentum $p_x\hat{x}$ is

$$\Psi = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-ip^\mu x_\mu/\hbar}, \tag{3.34}$$

where $p^\mu = (p_x, p_x, 0, 0)$. We can then obtain Ψ' for a massless particle moving in any other direction by multiplying by $D(\theta)$ and replacing x_ν with $(\Lambda^{-1})^\mu_\nu x_\mu$:

$$\Psi' = \begin{pmatrix} e^{-i\phi/2} \\ e^{i\phi/2} \end{pmatrix} e^{-i\Lambda^\mu_\nu p^\nu x_\mu/\hbar}. \tag{3.35}$$

Likewise, the D-matrix for boosting by velocity v in the x-direction is

$$D(\hat{x},\lambda) = e^{\lambda\sigma_x/2} = \begin{pmatrix} \cosh\lambda/2 & \sinh\lambda/2 \\ \sinh\lambda/2 & \cosh\lambda/2 \end{pmatrix}, \tag{3.36}$$

where $\tanh\lambda = v/c$. Boosting by v in the y-direction is given by

$$D(\hat{y},\lambda) = e^{\lambda\sigma_y/2} = \begin{pmatrix} \cosh\lambda/2 & -i\sinh\lambda/2 \\ i\sinh\lambda/2 & \cosh\lambda/2 \end{pmatrix}. \tag{3.37}$$

3.3.3 Angular Momentum

The Dirac equation has rotational symmetry in the 2D plane, so angular momentum should be conserved. Define the total angular momentum operator as the sum of the orbital and spin angular momentum operators:

$$J_z = xp_y - yp_x + \frac{\hbar\sigma_z}{2}. \tag{3.38}$$

We can compute the commutator of J_z with the Hamiltonian $H = -i\hbar c(\sigma_x p_x + \sigma_y p_y) + mc^2\sigma_z$.

$$\begin{aligned}
[H, J_z] &= -i\hbar c\sigma_x p_y[p_x, x] + i\hbar c\sigma_y p_x[p_y, y] - \frac{i\hbar^2 c}{2} p_x[\sigma_x, \sigma_z] - \frac{i\hbar^2 c}{2} p_y[\sigma_y, \sigma_z] \\
&= -\hbar^2 c\sigma_x p_y + \hbar^2 c\sigma_y p_x - \hbar^2 c\sigma_y p_x + \hbar^2 c\sigma_x p_y \\
&= 0
\end{aligned} \tag{3.39}$$

We used the identity $[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l$. Since the commutator $[H, J_z] = 0$, angular momentum J_z is conserved, and we can find simultaneous eigenstates of H and J_z. This does not change if we add a central potential $V(r)$ to the Dirac Hamiltonian.

The orbital angular momentum $L_x = xp_y - yp_x$ and spin angular momentum $S_z = \frac{\hbar\sigma_z}{2}$ are not separately conserved. It is only the total angular momentum $J_z = L_z + S_z$ that is conserved.

3.4 The Massless Dirac Equation

Although the 2D and 3D Dirac equations are similar, there are significant differences between them. Whereas the 3D Dirac equation respects parity and time reversal, the mass term $mc^2\sigma_z$ in the 2D Dirac Hamiltonian (Eq. 3.28) does not. In fancy-pants language, parity maps a $(\frac{1}{2}, 0)$ left-handed Weyl spinor into a $(0, \frac{1}{2})$ right-handed Weyl spinor (and vice versa). A 3D Dirac fermion is a $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ bispinor representation of the Lorentz group, so the 3D Dirac equation is invariant under parity. On the other hand, the 2D Dirac equation is only symmetric under proper, orthochronous Lorentz transformations, which exclude parity $P : (x, y) \to (x, -y)$ and time reversal $T : t \to -t$. The 2D massless Dirac equation in Eq. 3.40 does not have this problem and is symmetric under both parity and time reversal.

$$-i\hbar c\left[\sigma_x \frac{\partial}{\partial x} + \sigma_y \frac{\partial}{\partial y}\right]\Psi = E\Psi \tag{3.40}$$

This Dirac equation describes the physics of graphene.

3.4.1 Parity Transformation

As mentioned above, the 2D parity operator P takes (x, y) into $(x, -y)$. Note that I am defining parity P as a mirror reflection and not as the full spatial inversion $(x, y) \to (-x, -y)$ because inversion through a point is the same as a rotation by 180°, which is continuously connected to the identity element of the Lorentz group. Let us write Ψ as

$$\Psi = e^{i(p_x x + p_y y)/\hbar}\chi, \tag{3.41}$$

where χ is a spinor. To be a solution of Eq. 3.40,

$$\chi = \begin{pmatrix} e^{-i\phi/2} \\ e^{i\phi/2} \end{pmatrix} \tag{3.42}$$

for positive-energy solutions, or

$$\chi = \begin{pmatrix} e^{-i\phi/2} \\ -e^{i\phi/2} \end{pmatrix} \tag{3.43}$$

for negative-energy solutions. Note that this is the same as Eq. 1.15. Here, ϕ is defined to be the angle between the x-axis and $\mathbf{p} = (p_x, p_y)$.

Parity takes ϕ to $-\phi$. Notice that the σ_x matrix does exactly this to χ. Therefore, we can implement the parity transformation on Ψ by acting σ_x on it, i.e. $P\Psi(x, y) = \sigma_x\Psi(x, -y)$.

If Ψ is a solution to the 2D massless Dirac equation (Eq. 3.40), then $\Psi'(x, y) = P\Psi(x, y)$ is also a solution. Parity is a symmetry of the 2D massless Dirac equation.

3.4.2 Time Reversal Symmetry

Time reversal is also a symmetry of the 2D massless Dirac equation. To verify this, we need to find an operator that maps

$$\begin{pmatrix} e^{-i\phi/2} \\ e^{i\phi/2} \end{pmatrix} \to \begin{pmatrix} -ie^{-i\phi/2} \\ ie^{i\phi/2} \end{pmatrix} \tag{3.44}$$

and

$$\begin{pmatrix} e^{-i\phi/2} \\ -e^{i\phi/2} \end{pmatrix} \to \begin{pmatrix} -ie^{-i\phi/2} \\ -ie^{i\phi/2} \end{pmatrix} \tag{3.45}$$

because time reversal takes ϕ to $\phi + \pi$. It is easy to check that $T = \sigma_y K$ does this, where K is defined to be an operator that implements complex conjugation.

Suppose Ψ' and $\Psi = T\Psi' = \sigma_y K\Psi'$ are related to each other by time reversal. If we plug in Ψ into the 2D massless Dirac equation (Eq. 3.40), we get

$$-i\hbar c \left[\sigma_x \frac{\partial}{\partial x} + \sigma_y \frac{\partial}{\partial y}\right] \sigma_y K\Psi' = E\sigma_y K\Psi'. \tag{3.46}$$

By the properties of the Pauli matrices,

$$-i\hbar c \left[\sigma_x \frac{\partial}{\partial x} + \sigma_y \frac{\partial}{\partial y}\right] \Psi' = E\Psi'. \tag{3.47}$$

Therefore, if Ψ is a solution to the 2D massless Dirac equation, then its time-reversed partner Ψ' is also a solution to the 2D massless Dirac equation.

Notice that $T^2 = -1$. This implies that $\Psi \neq T\Psi$ and that each energy level is at least doubly degenerate. This is Kramers theorem.

Chapter 4

Graphene Device Fabrication

In this chapter, I discuss how we fabricate our graphene/BN devices. The monolayer graphene samples used in the experiments described in Chs. 5 and 6 (the relativistic Coulomb potential) were produced via chemical vapor deposition (CVD), while those in Chs. 7 and 8 (the relativistic quantum harmonic oscillator) were made by mechanical exfoliation (the Scotch tape trick). Since the CVD process is described in great detail in a video in Ref. [37], I will focus on the method involving exfoliated graphene. The technique for transferring exfoliated graphene onto BN was innovated by Cory Dean in the groups of Philip Kim, Kenneth L. Shepard, and James Hone [38], but we use a variant of a different procedure discussed in Zomer *et al.* [39].

4.1 Dry Transfer Procedure

Graphene (and other 2D materials) can be placed on top of other layered materials to create heterostructures with atomically clean interfaces. This can achieved through a variety of transfer methods, one of which is schematically depicted in Fig. 4.1 and is described below:

1. Use a diamond scribe to dice a glass microscope slide into three pieces, and clean one of these pieces with acetone and then with ethanol or isopropanol.

2. Affix a piece of single-sided transparent tape to that glass slide piece, and clean again with acetone and isopropanol.

3. Drop nine droplets of 10% methyl methacrylate (MMA) in methyl isobutyl ketone (MIBK) solution onto the tape/glass slide, and spin it at 1300 to 1400 revolutions per minute (RPM) for 65 seconds.

4. Mechanically exfoliate graphite onto the MMA side of the MMA/tape/glass slide structure, and search for a desirable monolayer graphene flake (see Fig. 4.2 for an optical microscope image of graphite and graphene flakes on MMA).

5. Cut the MMA/tape into a ∼2 mm by 2 mm area.
6. Dice a heavily doped silicon wafer with a 285 nm SiO_2 layer into a ∼1 cm by 1 cm square chip.
7. Mechanically exfoliate BN onto the SiO_2/Si chip, and find a BN flake with a desirable thickness. The thickness of a BN flake an be ascertained through its color under an optical microscope. Our BN flakes are usually 60 to 100 nm thick.
8. Anneal BN/SiO_2/Si in a 350°C Ar/H_2 forming gas.
9. Using an optical microscope with a large working distance and suitable micromanipulators (such as one depicted in Fig. 4.3), align the graphene/MMA/tape/glass slide structure over the BN/SiO_2/Si chip such that the desired graphene flake is directly above the desired BN flake. Slowly lower the glass slide until the graphene makes contact with the BN.
10. Heat the silicon chip to ∼90°C for 5 to 10 minutes. Let it cool for another 5 to 10 minutes.
11. The silicon chip should now be stuck to the glass slide. Place the entire structure in 45°C CH_2Cl_2 for 20 to 30 minutes to dissolve the tape and MMA.
12. Look at the graphene/BN heterostructure under an optical microscope. The graphene should have adhered to the BN, creating the desired heterostructure. If the graphene is not stuck to BN, then toss the chip and start over. There may be a few bubbles (possibly made of hydrocarbons trapped between the graphene and BN interface), but hopefully not many.
13. Anneal the graphene/BN/SiO_2/Si structure in a 350°C Ar/H_2 forming gas.
14. Scan the graphene with an ambient atomic force microscope (AFM) to check the sample cleanliness. If the sample is too dirty, then throw it away and start over. Scratching the graphene surface with the AFM tip may remove unwanted contaminants and improve the sample quality.

4.2 Sample Mounting and Contacting

Following the procedure laid out in Sec. 4.1 produces a graphene/BN/SiO_2/Si heterostructure that needs to be electrically contacted and mounted onto a sample holder. To electrically contact the graphene, we use either standard electron-beam lithography (with polymethyl methacrylate (PMMA) 495 and 950 as the resist) or shadow mask techniques. Shadow masks are preferred over lithography because PMMA can be difficult to remove from the heterostructures. In either case, we deposit 10 nm of Ti and 100 nm of Au for the contact

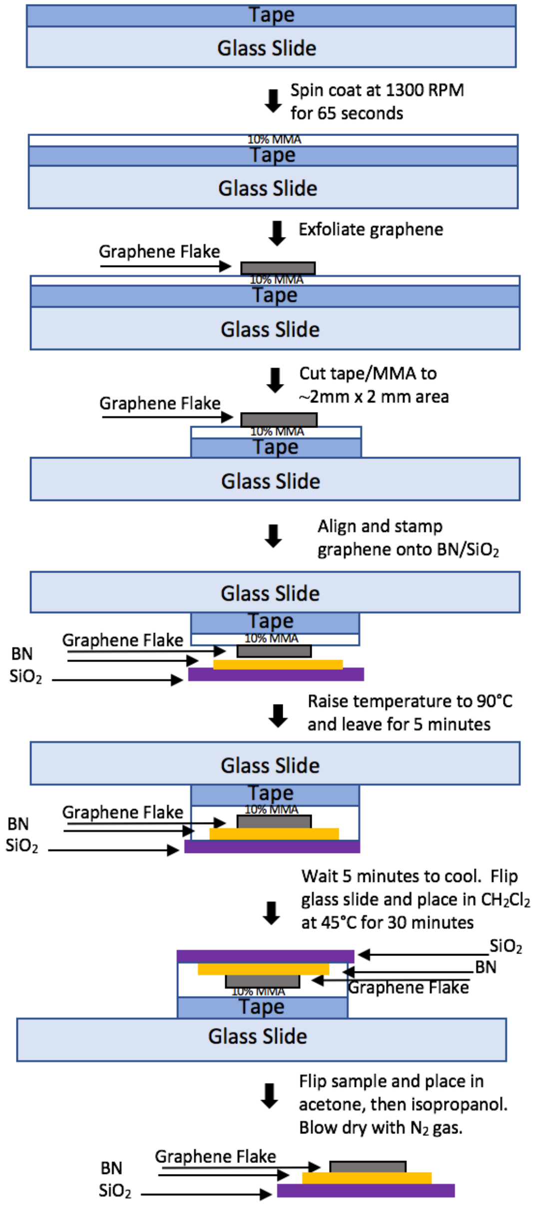

Figure 4.1: Graphene/BN dry transfer procedure. First, a transparent piece of tape is attached to a microscope glass slide cut into a square shape. MMA is then spin coated onto the tape/glass slide. Graphene is mechanically exfoliated onto the MMA. The MMA/tape is cut into a ∼2 mm by 2 mm area. BN is exfoliated onto a ∼1 cm by 1 cm SiO_2/Si chip, and the glass slide is aligned above the SiO_2/Si chip such that the desired graphene flake will be stamped on top of the desired BN flake. Upon stamping the graphene/MMA/tape/glass slide structure onto the BN/SiO_2/Si, heat the chip to ∼90°C for 5 to 10 minutes. Afterwards, dissolve the tape/MMA in dichloromethane at 45°C for 20 to 30 minutes. If successful, the graphene should adhere to the BN, forming a graphene/BN heterostructure.

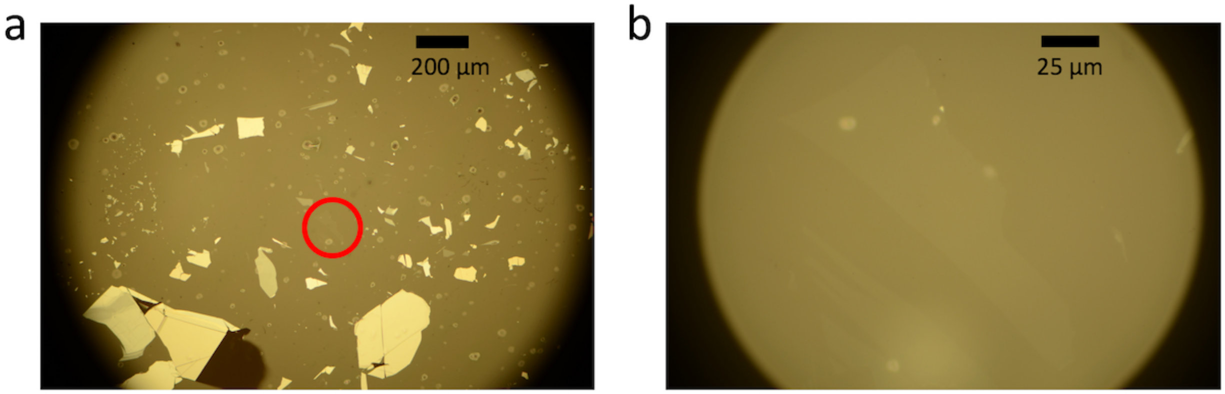

Figure 4.2: Optical micrograph of graphene/MMA. (a) Image of graphite flakes exfoliated onto MMA. The thinner graphite flakes have color that is more similar to the background than the thicker flakes, with monolayer graphene barely distinguishable from the background. (b) Zoomed-in image of graphene flake in the red circle in (a).

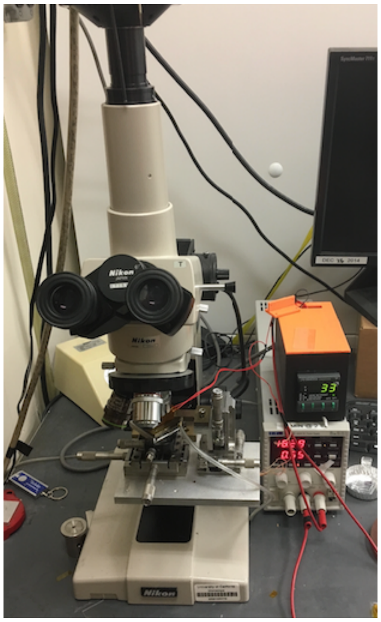

Figure 4.3: Optical microscope with micromanipulators used for aligning 2D materials for dry transfer and for aligning stencil masks to samples. This optical microscope also has a custom-built heater stage powered by the adjacent power supply.

electrode (or alternatively, Cr/Au if we are making a two-terminal device and are concerned about contact resistance).

To mount the heterostructure onto a sample holder (see Fig. 4.4), we cleave the Si chip into a narrow strip and clamp the strip down with tantalum foil. The tantalum foil is held in place by nuts and threaded rods. (Before 2015, we also placed a sapphire spacer between the Si chip and the sample plate. Sample mounting was often quite frustrating with this extra piece of sapphire, and so we got rid of it.)

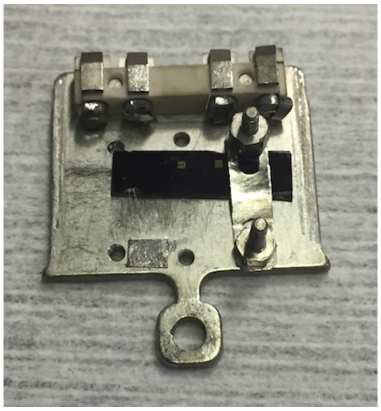

Figure 4.4: Sample plate. Silicon chips are rigidly held onto sample holders via tantalum clamping foil. Each sample holder has four contacts that are electrically isolated from the sample plate. A wire connects one of these contacts to a gold electrode on the chip, which grounds a graphene sample.

The sample holder has four electrical contacts in addition to the baseplate itself. We use wire bonding (with Si/Al wire) to connect the Ti/Au electrode to one of the four electrical contacts (again, see Fig. 4.4). In our experiments, we ground graphene by grounding this contact. If the sample of interest is a two-terminal device with source and drain contacts, caution is required to prevent electrostatic discharge from destroying the graphene.

The Si layer of the heterostructure is used as a backgate. Applying a gate voltage V_g to the Si allows us to tune the charge carrier density in the graphene because the graphene and Si layer form a parallel plate capacitor with capacitance per unit area $C = \epsilon/d$. Here, ϵ is the dielectric constant of the insulating layers between graphene and Si, and d is the distance in between. The dielectric constant is $\epsilon \approx 4$ (or $4\epsilon_0$ in the SI unit system) for both SiO_2 and BN. The distance d is the sum of the SiO_2 thickness (285 nm) and the BN thickness. The charge density n induced in graphene by a gate voltage V_g is

$$e(n - n_0) = C(V_g - \phi_Q) \approx \epsilon V_g/d, \tag{4.1}$$

where $e = |e|$ is the magnitude of the electron charge, ϕ_Q is the electric potential on graphene (assumed to be small compared to V_g because the graphene is grounded), and n_0 is the charge density when $V_g = 0$ V. The density n_0 is a residual charge density in graphene caused by doping from contaminants, substrates, tip-induced contact potentials, and other extrinsic sources. Note that n is related to the Fermi energy and Dirac point energy via

$$n = \frac{|E_F - E_D|\,(E_F - E_D)}{\pi(\hbar v_F)^2}. \tag{4.2}$$

Therefore, $|E_F - E_D| \propto \sqrt{|V_g - V_{CNP}|}$, where V_{CNP} is known as the charge neutrality point (i.e. the gate voltage such that graphene is undoped). In measurements of conductance (or resistance) against V_g, V_{CNP} appears as a minimum (maximum).

In order to actually use the Si backgate, we must connect it to a power supply (we use a Keithley 2400 SourceMeter or a similar model). To do this, we electrically connect the Si layer to the sample holder baseplate and apply V_g to the baseplate. Although the Si layer and the baseplate are already in physical contact, we ensure electrical contact by using a diamond scribe to scratch off some of the SiO_2. This exposes the Si layer underneath the scratched location, and we then use wire bonding to attach a wire between the scratch and the baseplate. Wire bonding to Si can be difficult, so the bond is made such that the wire adheres to SiO_2 while physically touching the exposed Si.

4.3 Approaching Graphene Devices

After mounting the sample and attaching wires to the Ti/Au electrode and Si gate, we move the sample and sample holder into the UHV chamber. Before we approach the sample with our STM tip, however, we anneal the device overnight at 200°C to 400°C. This removes contaminants and adsorbed water.

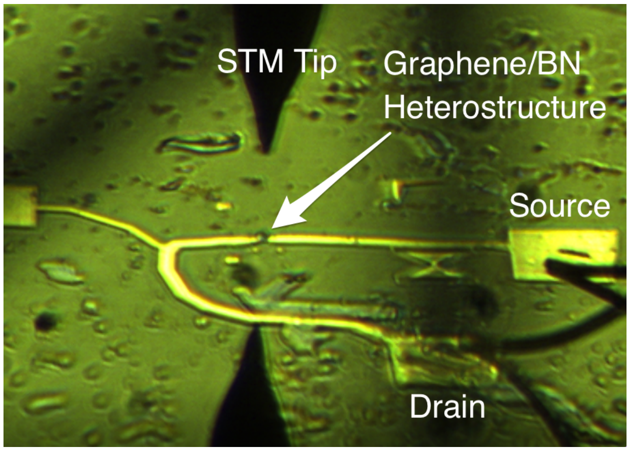

Figure 4.5: Optical image of graphene/BN heterostructure inside STM. The heterostructure is electrically connected to lithographically patterned source and drain electrodes, which in turn are connected to wire bonds. The STM tip is seen above the heterostructure, while a reflection of the tip is seen below.

We are then finally ready to approach the graphene/BN heterostructure with our STM tip. If the graphene/BN is larger than 30 µm by 30 µm, we can visually approach the heterostructure because it can be resolved by our Infinity K2/SC long-distance optical microscope with PixeLINK CCD camera. If the graphene is smaller, then we must use another technique to find the heterostructure. We have used two different techniques to navigate our tip to graphene: a capacitance-based method pioneered by Eva Y. Andrei [40] and a scanning gate technique invented by Brian J. LeRoy. We find the latter method to be more effective, so I will now discuss our version of it. This method, which only works for devices in which the conductance G can be measured, exploits the fact that the STM tip is a movable top gate that is most effective when it is very near graphene. The procedure is described below:

1. Bring the STM tip as close as possible to the graphene/BN heterostructure without crashing the tip. This can be done by approaching the source or drain electrode.
2. Set V_s to the largest possible value without field emission. We use 10 V or -10 V.
3. Measure the conductance G as a function V_g.
4. Set V_g at the value where $\left|dG/dV_g\right|$ is largest. Minimize any visible light that may be illuminating the graphene/BN heterostructure.
5. While monitoring G, walk the STM tip towards graphene. Move the tip to the location where G changes the most from its background value.
6. Repeat step 5, but walk the tip in a perpendicular direction.
7. Iterate steps 5 and 6 until the tip converges onto the location of the graphene/BN heterostructure.

When the tip has been successfully navigated to graphene, approach and scan. The remainder of this dissertation discusses the results of our experiments on these graphene/BN devices.

Chapter 5

Charge Carrier Screening of a Subcritical Impurity

This chapter describes our work on the electronic screening of subcritical charged impurities, particularly its dependence on the carrier density. The results presented here are based on an unpublished manuscript titled "Spatially resolving density-dependent screening around a single charged atom in graphene" by Dillon Wong, Fabiano Corsetti, Yang Wang, Victor W. Brar, Hsin-Zon Tsai, Qiong Wu, Roland K. Kawakami, Alex Zettl, Arash A. Mostofi, Johannes Lischner, and Michael F. Crommie. The STM measurements were performed by Yang Wang and myself; the samples were made from CVD graphene by Hsin-Zon Tsai and Qiong Wu; and the theoretical calculations were done by Fabiano Corsetti, Arash A. Mostofi, and Johannes Lischner. The manuscript has recently been accepted for publication in Physical Review B.

5.1 Introduction

Since electrons in graphene behave like massless Dirac fermions, graphene serves as an excellent platform for simulating ultra-relativistic physics. By placing a single charged ion on graphene, we can create what is effectively a relativistic hydrogen atom and image its eigenstates with STM. As discussed in Ch. 1, massless Dirac fermions have drastically different responses to Coulomb potentials depending on the strengths of the potentials as quantified by the effective charge Z. In the subcritical regime, Z is less than a critical charge Z_C, and there are no bound states of this relativistic Coulomb system. Strictly speaking, however, the potential of a charged ion on graphene is not Coulombic because charge carrier screening changes the shape of the potential. Nevertheless, a subcritical regime is still observed.

By studying the electronic response of graphene quasiparticles to a subcritical charged ion while varying the graphene carrier density, we can obtain important information about the screening of electrostatic potentials in graphene as well as deeper understanding of the electron-electron interactions therein. This is important because knowledge of screening processes and of the wavevector-dependent static dielectric function $\epsilon(q)$ is critical for understanding many material properties. The ultra-relativistic nature of its charge carriers cause graphene to have very unusual screening behavior: undoped graphene behaves like a dielectric while doped graphene behaves like a metal [41, 42].

The screening of charged impurities by massless Dirac fermions is of particular technological importance to the performance and construction of graphene-based field-effect transistors (FETs). Charged impurities limit carrier mobility [43–45], shift the chemical potential [46], induce phase transitions [47–49], and split Landau levels [50]. In this chapter, I present an STM/STS study of the local screening response of gated graphene to positively charged calcium (Ca) atoms. We find that charged defects are screened by graphene's ultra-relativistic carriers over an unusually long length scale on the order of ten nanometers. The electronic screening length is highly dependent on the charge carrier density and is thus tunable via gate voltage. Our experimental observations, which provide important insight into electron-electron interactions in graphene, are in good agreement with theoretical simulations of the electronic structure of doped graphene in the presence of a screened Coulomb potential.

5.2 Charged Calcium on Graphene

The experiments were performed in a UHV Omicron LT-STM at 4.8 K using PtIr STM tips calibrated against the surface state of Au(111) [32]. dI/dV was measured using standard lock-in detection of the a.c. tunneling current modulated by a 6 mV rms, 500 – 700 Hz signal added to V_s.

Monolayer graphene sheets were grown through CVD [51] and were subsequently transferred onto BN crystals exfoliated onto SiO_2/Si wafers. The BN crystals provide atomically smooth surfaces [33, 34] with reduced charge inhomogeneity compared to SiO_2 [52, 53]. The graphene samples were electrically contacted by Ti/Au electrodes and were then cleaned by annealing at 400°C in UHV for several hours. These experiments were repeated on multiple gate-tunable graphene devices for many different calcium atoms.

Calcium atoms were deposited onto graphene by thermally heating a calcium getter source (Alvatec and Trace Sciences International) calibrated by a mass spectrometer (SRS Residual Gas Analyzer). Before each deposition, the graphene was first checked using STM for surface cleanliness. The STM tip was then retracted out of line-of-sight of the getter source to avoid contamination of the tip. After degassing the getter source, calcium atoms were evaporated directly onto the low-temperature graphene surface [37]. Fig. 5.1a shows a schematic of the experimental setup, and Fig. 5.1b shows a typical STM topographic image of graphene following a Ca deposition. Ca atoms appear as identical round protrusions that are surrounded by dark depressions caused by the rearrangement of LDOS spectral weight. These dark depressions are clear signatures of the charged nature of the Ca atoms [54].

We can quantitatively assess the charge donated by each Ca atom by plotting the carrier density in graphene against the density of Ca adatoms, which is shown in Fig. 5.2. The carrier density n is given by Eq. 4.2, where $|E_D - E_F|$ was measured through STS at $V_g = 0$ V.

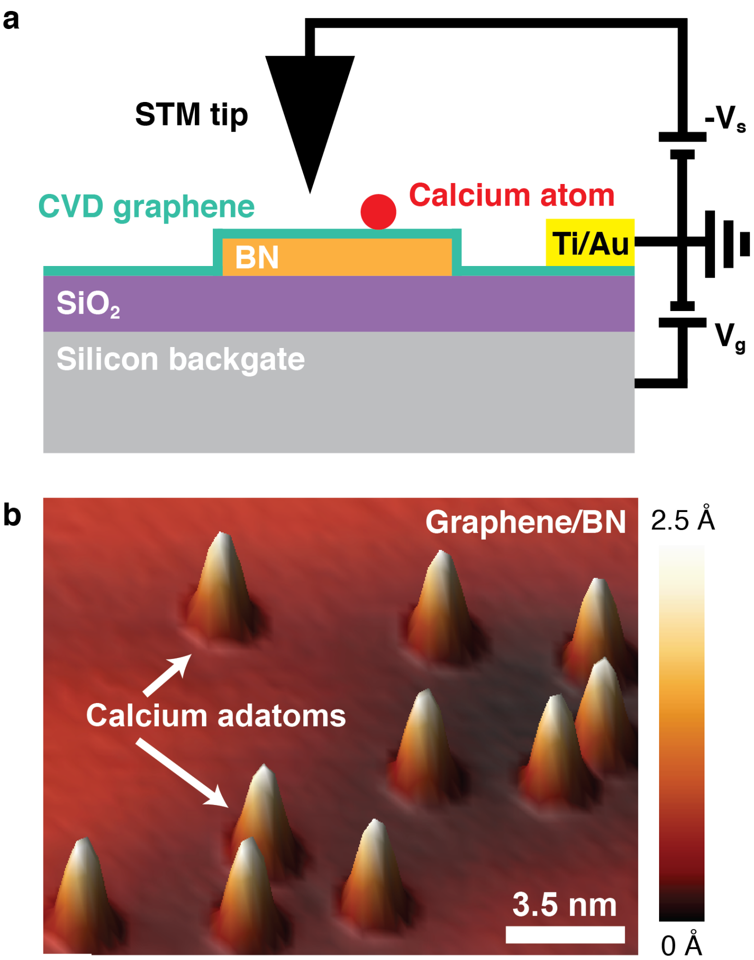

Figure 5.1: Calcium adatoms on graphene. (a) Schematic diagram of experimental setup. A calcium atom sits on a graphene/BN/SiO_2/Si heterostructure. A bias $-V_s$ is applied to the STM tip, while V_g is applied to the Si to gate the graphene. (b) STM topographic image of Ca adatoms on graphene/BN. Tunneling parameters: $V_s = -0.45$ V, $I = 2$ pA.

The density of Ca adatoms, which was estimated by counting the number of Ca adatoms in many different areas, was controlled via the time durations of the Ca depositions and by the number of depositions. By fitting a linear regression line (dashed red line in Fig. 5.2) to the data, we obtained a charge transfer of 0.7 ± 0.2 electrons per Ca adatom.

5.3 Electron-Hole Asymmetry

In order to determine the charge states of the Ca adatoms at different doping levels, we performed gate-dependent dI/dV spectroscopy on graphene at various distances away from an individual Ca adatom that was at least 20 nm away from every other Ca adatom. This

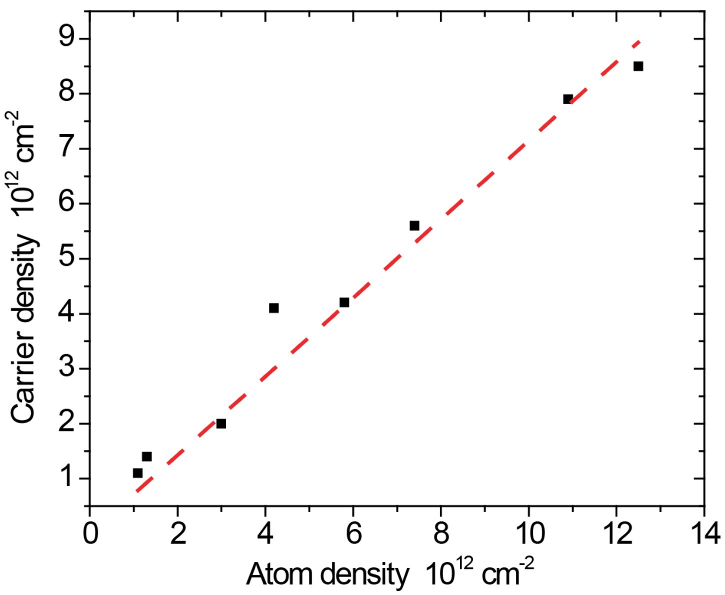

Figure 5.2: Carrier density versus calcium adatom density.

data is plotted in Figs. 5.3a-c for p-doped, nearly neutral, and n-doped graphene, respectively. Each dI/dV curve has been normalized by a different constant to account for the exponential dependence of the tunneling conductance on tip height, as was done in Ref. [54]. As explained in Sec. 2.3 and Refs. [35] and [36], the dI/dV curves all show a $\sim$130 meV gap-like feature at the Fermi energy caused by phonon-assisted inelastic tunneling. The p-doped (n-doped) spectra have local minima on the right (left) of the Fermi level that represent the graphene Dirac point. For the nearly neutral graphene spectra, the Dirac point is near the Fermi energy and is obscured by the $\sim$130 meV gap-like feature.

The dI/dV spectra in Figs. 5.3a-c all display a characteristic electron-hole asymmetry in which the dI/dV intensity at the energies above the Dirac point increases near the Ca atom (relative to far away from the impurity), whereas the dI/dV intensity at energies below the Dirac point decreases near the Ca atom. This is consistent with theoretical predictions that the LDOS of graphene near a positively charged Coulomb impurity should increase above the Dirac point and decrease below the Dirac point [13, 14]. Indeed, similar behavior has been seen in dI/dV spectroscopy near positively charged cobalt trimers [54]. Thus, we conclude that the Ca adatom is positively charged regardless of the graphene doping level achievable within our experimental conditions. The charge state of the Ca atom is stable and does not change during our measurements.

The observed electron-hole asymmetry can be simply understood as an attraction of the states above the Dirac point towards the positively charged atom and a repulsion of the states below the Dirac point away from the atom. In analogy to relativistic quantum mechanics in which the positive-energy continuum represents electrons and the negative-

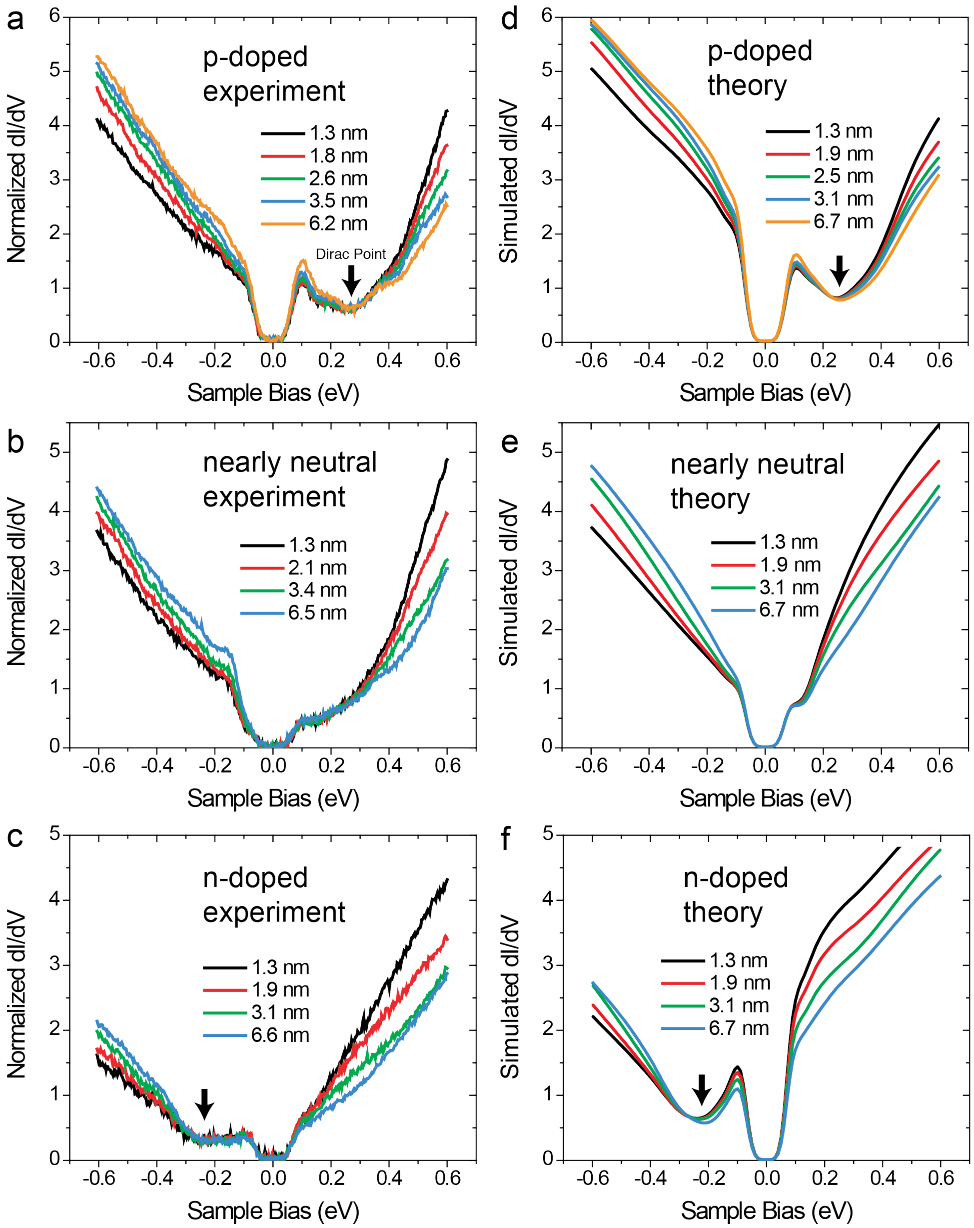

Figure 5.3: Experimental and simulated spectroscopy near isolated Ca adatom. (a) Normalized dI/dV spectra obtained at various distances away from a Ca atom on p-doped graphene. (b) Same as (a) for nearly neutrally doped graphene. (c) Same as (a) for n-doped graphene. These dI/dV curves show that Ca adatoms remain positively charged as graphene's carrier density is tuned via V_g. Initial tunneling parameters: $V_s = 0.6$ V, $I = 60$ pA. (d) Tight-binding simulation of dI/dV spectra on p-doped graphene away from a screened Coulomb potential. (e) Same as (d) for nearly neutral graphene. (f) Same as (d) for n-doped graphene. The Dirac points in (a), (c), (d), and (f) are indicated by black arrows.

energy continuum represents positrons, the conduction band states of graphene are electron-like, and the valence band states are hole-like. Thus, states above the Dirac point are negatively charged quasiparticles that are attracted to the positive Ca atom, increasing the LDOS and dI/dV in its vicinity. States below the Dirac point behave like positively charged particles that are repelled from the atom, lowering the LDOS and dI/dV.

An alternate way to interpret the electron-hole asymmetry is to realize that the LDOS of graphene in the presence of the atom is the LDOS of graphene without the atom but shifted towards lower energies by the local value of the potential produced by the atom. This explains the reduction of dI/dV below the Dirac point and the increase of dI/dV above the Dirac point. For the spectral weight at a given energy E, this is valid for distances much further from the atom than $\hbar v_F / |E - E_D|$. If the atom produces a Coulomb potential, the Dirac point itself is not shifted by the local value of the Coulomb potential because the massless Dirac Hamiltonian is scale invariant.

5.4 Dependence on Charge Carrier Density

Since the Ca adatoms are charge stable under our experimental conditions, we can study graphene's screening response to charged impurities over a wide range of carrier densities. We expect the spatial distribution of dI/dV to alter as V_g is changed because charge carrier screening, which controls the shape of the potential produced by a Ca adatom, should depend on the electron density. Figs. 5.4a-c and 5.5a-c show gate-dependent dI/dV maps near a Ca adatom for p-doped graphene and n-doped graphene, respectively. Fig. 5.4c (Fig. 5.5c) represents graphene that is more p-doped (n-doped) than in Fig. 5.4b (Fig. 5.5b), which is in turn more p-doped (n-doped) than in Fig. 5.4a (Fig. 5.5a). The sample bias V_s chosen for each dI/dV map varies with V_g such that we imaged the same states in each map. To optimize dI/dV contrast, the maps in Figs. 5.4a-c (p-doped graphene) imaged states 0.15 eV above the Dirac point, and the maps in Figs. 5.5a-c (n-doped graphene) imaged states 0.08 eV below the Dirac point. We avoided including the Ca atom in the dI/dV maps in order to minimize the risk of picking up the atom with the STM tip, so the location of the Ca adatom is marked by a red disk below each map. The color schemes of the dI/dV maps are adjusted to the maximum and minimum values of each map, with yellow representing high dI/dV intensity and blue representing low dI/dV intensity.

The overall appearance of the dI/dV maps is qualitatively consistent with a stable, positively charged Ca atom. In Figs. 5.4a-c, the dI/dV intensity is higher near the Ca atom because states above the Dirac point were probed. In contrast, the dI/dV intensity is lower near the Ca atom in Figs. 5.5a-c because states below the Dirac point were probed. I emphasize that whether dI/dV increases or decreases near the Ca atom does not depend on whether the graphene is n-doped or p-doped but rather on the position of V_s relative to the Dirac point. Additionally, comparison between Figs. 5.4a-c (and Figs. 5.5a-c) shows a systematic difference in the length scale of the dI/dV contrast. When the magnitude of V_g increases (for both n-doped and p-doped graphene), dI/dV returns to the background value

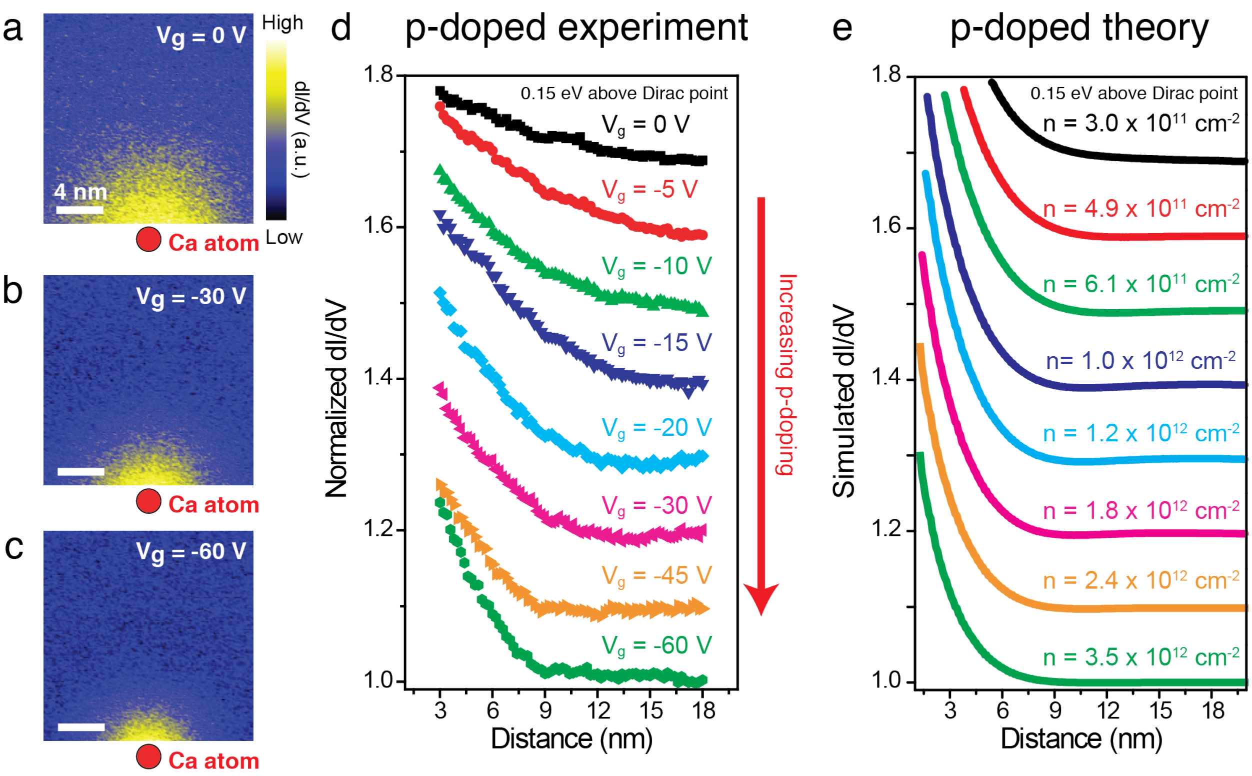

Figure 5.4: Gate-dependent dI/dV for p-doped graphene. (a-c) dI/dV maps (0.15 eV above the Dirac point) near a Ca atom (represented by red disks) on p-doped graphene. (d) Radially averaged dI/dV (0.15 eV above the Dirac point) as a function of distance away from a Ca atom on p-doped graphene. Curves are vertically offset for clarity, with the magnitude of the p-doping increasing from the top curve to the bottom curve. The values of dI/dV far away from the Ca atom are normalized to 1. (e) Simulated dI/dV (0.15 eV above the Dirac point) as a function of distance away from an RPA-screened Coulomb potential on p-doped graphene. The values of the charge carrier density n were chosen to correspond to the gate voltages in (d). Graphene was modeled via a nearest-neighbor-hopping tight-binding model. The values of the simulated dI/dV far away from the screened Coulomb potential are normalized to 1. Tunneling parameters: (a) $V_s = 0.28$ V, $I = 28$ pA; (b) $V_s = 0.38$ V, $I = 38$ pA; (c) $V_s = 0.45$ V, $I = 45$ pA.

more quickly with increasing distance from the Ca atom.

In order to better illustrate and accurately quantify the trend shown in Figs. 5.4a-c and 5.5a-c, we plot dI/dV as a function of distance away from a Ca atom for p-doped (Fig. 5.4d) and n-doped (Fig. 5.5d) graphene. The data in Figs. 5.4d and 5.5d were obtained by radially averaging dI/dV maps similar to those in Figs. 5.4a-c and 5.5a-c (with the Ca atom as the center) and then normalizing the data to account for tip-height variations caused by the STM feedback loop [54]. For each of Figs. 5.4d and 5.5d, the dI/dV curves for different V_g

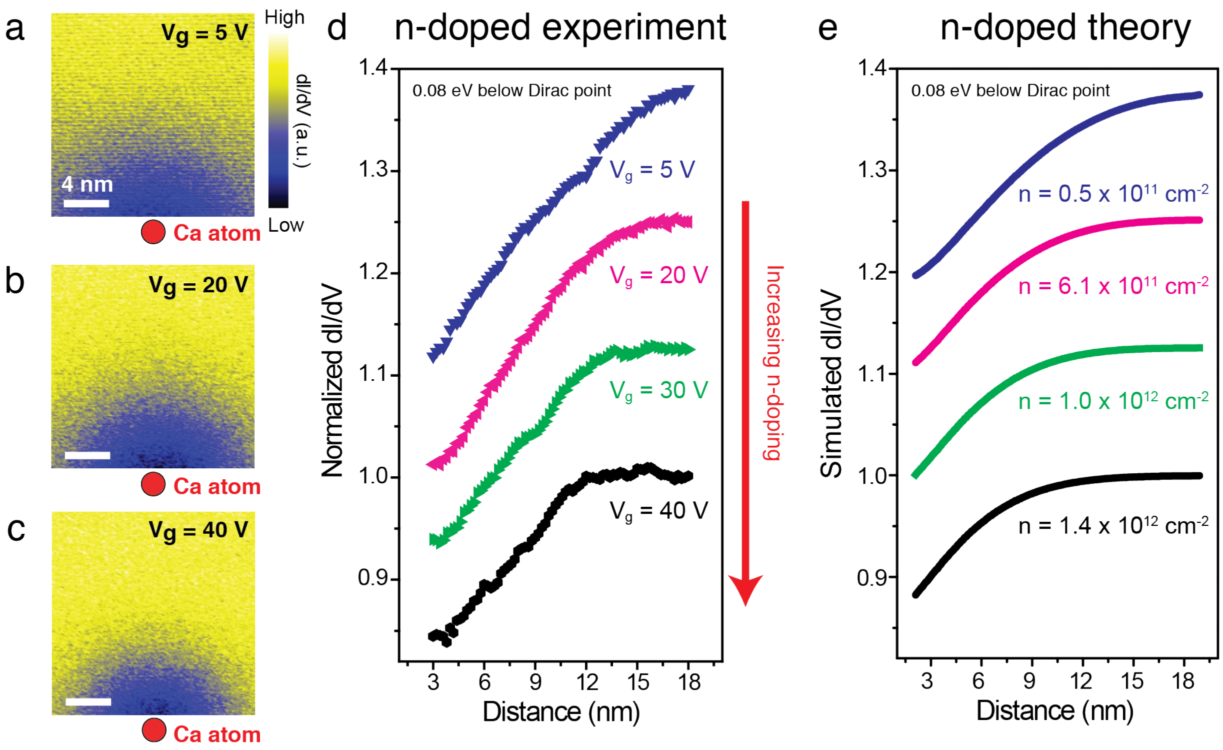

Figure 5.5: Gate-dependent dI/dV for n-doped graphene. (a) dI/dV maps (0.08 eV below the Dirac point) near a Ca atom (represented by red disks) on n-doped graphene. (d) Radially averaged dI/dV (0.08 eV below the Dirac point) as a function of distance away from a Ca atom on n-doped graphene. Curves are vertically offset for clarity, with the magnitude of the n-doping increasing from the top curve to the bottom curve. The values of dI/dV far away from the Ca atom are normalized to 1. (e) Simulated dI/dV (0.08 eV below the Dirac point) as a function of distance away from an RPA-screened Coulomb potential on n-doped graphene. The values of the charge carrier density n were chosen to correspond to the gate voltages in (d). Graphene was modeled via a nearest-neighbor-hopping tight-binding model. The values of the simulated dI/dV far away from the screened Coulomb potential are normalized to 1. Tunneling parameters: (a) $V_s = -0.16$ V, $I = 17$ pA; (b) $V_s = -0.22$ V, $I = 20$ pA; (c) $V_s = -0.28$ V, $I = 28$ pA.

are shifted vertically such that the magnitude of the graphene charge carrier concentration $|n|$ increases from the top curve to the bottom curve. As $|n|$ increases in magnitude, dI/dV more quickly returns to its unperturbed value (i.e. the value of dI/dV in the absence of the atom).

5.5 Models for the Dielectric Function

5.5.1 The Thomas-Fermi Model

Let's review the Thomas-Fermi theory of screening in a 3D electron system. Suppose we have a small, localized, and slowly varying external charge distribution $\rho_{\text{ext}}(\mathbf{r})$ that induces a change in the charge carrier density $n_{\text{ind}}(\mathbf{r}) = n(\mathbf{r}) - n_0$, where n_0 is the equilibrium density (i.e. without the external charge). We want to know the potential $V(\mathbf{r})$ produced by the external and induced charges, so we use Poisson's equation:

$$\nabla^2 V(\mathbf{r}) = -4\pi \left[\rho_{\text{ext}}(\mathbf{r}) - e n_{\text{ind}}(\mathbf{r})\right]. \tag{5.1}$$

I am working in Gaussian units here (not SI). Since the total electrochemical potential must be the same everywhere, $E_F(n_0) = E_F(n(\mathbf{r})) - eV(\mathbf{r})$, where E_F is a local chemical potential that is a function of $n(\mathbf{r})$. If we can assume $n_{\text{ind}}(\mathbf{r}) = n(\mathbf{r}) - n_0$ is small, then we can approximate $E_F(n(\mathbf{r})) - E_F(n_0) \approx \frac{dE_F}{dn}[n(\mathbf{r}) - n_0]$, so

$$n_{\text{ind}}(\mathbf{r}) = e\text{DOS}(E_F)V(\mathbf{r}), \tag{5.2}$$

where $\text{DOS}(E_F)$ is the density of states at the equilibrium Fermi energy. Plugging this into Eq. 5.1 gives

$$\nabla^2 V(\mathbf{r}) = -4\pi \left[\rho_{\text{ext}}(\mathbf{r}) - e^2\text{DOS}(E_F)V(\mathbf{r})\right]. \tag{5.3}$$

We can Fourier transform this equation to obtain

$$[-q^2 - 4\pi e^2\text{DOS}(E_F)]V(\mathbf{q}) = -4\pi\rho_{\text{ext}}(\mathbf{q}), \tag{5.4}$$

where q is the magnitude of the wavevector $\mathbf{q} = (q_x, q_y, q_z)$. $V(\mathbf{q})$ and $\rho_{\text{ext}}(\mathbf{q})$ are the Fourier transforms of $V(\mathbf{r})$ and $\rho_{\text{ext}}(\mathbf{r})$, respectively. Finally, we get the 3D Thomas-Fermi dielectric function

$$\epsilon(\mathbf{q}) = 1 + \frac{4\pi e^2\text{DOS}(E_F)}{q^2}. \tag{5.5}$$

This dielectric function is applicable for a 3D metal. Doped graphene is 2D, so we need to modify this formalism a little. We again start from Poisson's equation

$$\nabla^2 V(\mathbf{r}) = -4\pi e\delta(\mathbf{r}) + 4\pi e^2\text{DOS}(E_F)V(x, y, z = 0)\delta(z) \tag{5.6}$$

but put in a point charge at the origin as the explicit external charge. Note that the Thomas-Fermi approximation is not very good here because $\delta(\mathbf{r})$ is not slowly varying... but let's proceed anyway. I also put $\delta(z)$ into the induced-charge term because the induced charge is confined in the z-direction. Fourier transforming Eq. 5.6 yields

$$\left(q^2 + q_z^2\right) V(\mathbf{q}, q_z) = 4\pi e - 4\pi e^2\text{DOS}(E_F) \int \frac{dq_z}{2\pi} V(\mathbf{q}, q_z). \tag{5.7}$$

Now $\mathbf{q} = (q_x, q_y)$ and $q^2 = q_x^2 + q_y^2$. I intentionally separated q_z from $\mathbf{q}$ because

$$V(\mathbf{q}) = \int \frac{dq_z}{2\pi} V(\mathbf{q}, q_z) \tag{5.8}$$

is the Fourier transform of the potential in the 2D plane. The 2D Fourier transform of the Coulomb potential is $\widetilde{V}(\mathbf{q}) = 2\pi e/q$, and we want to compare the screened potential $V(\mathbf{q})$ to it.

It is straightforward to check that the solution of Eq. 5.7 is

$$V(\mathbf{q}) = \frac{2\pi e}{q} \frac{1}{\epsilon(q)}, \tag{5.9}$$

where the 2D Thomas-Fermi dielectric function is [5, 41, 55, 56]

$$\epsilon(\mathbf{q}) = 1 + \frac{2\pi e^2 \mathrm{DOS}(E_F)}{q}. \tag{5.10}$$

The main difference between Eqs. 5.5 and 5.10 is an extra factor of q. Screening in 2D materials is typically weaker than in 3D materials because electric field lines can leave the plane of a 2D material [57]. In Eq. 5.10, I have ignored the substrate dielectric constant. To include a BN substrate, replace the 1 with $(\epsilon_{\mathrm{BN}} + 1)/2$.

The characteristic length scale of Eq. 5.10 is the Thomas-Fermi screening length [42]

$$\lambda_{TF} = \frac{1}{2\pi e^2 \mathrm{DOS}(E_F)} = \frac{\hbar v_F}{4e^2 k_F}, \tag{5.11}$$

where $k_F = \sqrt{\pi|n|}$ is the magnitude of the Fermi wavevector relative to the K and K' points. Importantly, λ_{TF} depends sensitively on k_F and can therefore be tuned by application of a gate voltage. Additionally, λ_{TF} only depends on the absolute value of the carrier density n. Increasing the magnitude of the charge carrier density via the gate voltage V_g leads to a decrease of λ_{TF}, which explains the observed decrease of the decay length of dI/dV in Figs. 5.4d and 5.5d.

It is also noteworthy that a conventional 2D electron gas (2DEG) with parabolic bands has a $\mathrm{DOS}(E_F)$ and a λ_{TF} that are independent of n [58]. It is thus a consequence of the linear, relativistic-like bands of graphene that $\lambda_{TF} \propto 1/\sqrt{|n|}$.

5.5.2 The Random Phase Approximation

Thomas-Fermi screening theory qualitatively explains our experimental findings, but it does not include the effects of interband transitions between graphene's π and π^* bands, and it is only valid for slowly varying potentials. To obtain a quantitatively more accurate model for describing screening in graphene, we turn to the random phase approximation (RPA).

The static dielectric function is given by

$$\epsilon(\mathbf{q}) = 1 - \widetilde{V}(q)\Pi(\mathbf{q}). \tag{5.12}$$

$\widetilde{V}(q)$ is the Fourier transform of the Coulomb potential, and $\Pi(\mathbf{q})$ is the electron-hole polarization bubble diagram. Applying RPA for a 3D metal yields the famous Lindhard dielectric function. For a 2D graphene sheet with area L^2, $\Pi(\mathbf{q})$ is given by

$$\Pi(\mathbf{q}) = \frac{2}{L^2} \sum_{\mathbf{k},s,s'} \frac{f_{s,\mathbf{k+q}} - f_{s',\mathbf{k}}}{E_{s,\mathbf{k+q}} - E_{s',\mathbf{k}}} \left[1 + ss' \cos\left(\theta_{\mathbf{k+q}} - \theta_{\mathbf{k}}\right)\right], \tag{5.13}$$

where $s, s' = \pm 1$ are band indices, $f_{s,\mathbf{k}}$ is the occupation of band s at wavevector $\mathbf{k}$, $E_{s,\mathbf{k}}$ is the energy of band s at $\mathbf{k}$, and $\theta_{\mathbf{k+q}} - \theta_{\mathbf{k}}$ is the angle between $\mathbf{k} + \mathbf{q}$ and $\mathbf{k}$. Doing some difficult math for the case of undoped graphene (using the Dirac Hamiltonian and only considering interband transitions) gives a constant

$$\epsilon(\mathbf{q}) = 1 + \frac{\pi e^2}{2\hbar v_F}. \tag{5.14}$$

More math for doped graphene (now including intraband transitions) yields [41, 55, 59]

$$\epsilon(\mathbf{q}) = 1 + \frac{2\pi e^2 \mathrm{DOS}(E_F)}{q} F(q, k_F), \tag{5.15}$$

$$F(q, k_F) = 1 + \Theta(q - 2k_F) \left[-\frac{1}{2}\sqrt{1 - \left(\frac{2k_F}{q}\right)^2} + \frac{q}{4k_F} \cos^{-1}\left(\frac{2k_F}{q}\right) \right]. \tag{5.16}$$

Here, Θ is the Heaviside step function. Note that when $q < 2k_F$, $F(q, k_F) = 1$, so the RPA dielectric function is identical to the Thomas-Fermi dielectric function.

5.6 Tight-Binding Simulation of Calcium on Graphene

Now that we have the RPA dielectric function, we need to calculate the LDOS to quantitatively compare to experiment. My theory collaborators (Fabiano Corsetti, Arash A. Mostofi, and Johannes Lischner) used a nearest-neighbor tight-binding model for the graphene electrons and an on-site RPA-screened Coulomb potential to describe the Ca atom.

Fig. 5.6 shows the results of our tight-binding calculations for p-doped, nearly neutral, and n-doped graphene. To compare the simulation to our experimental data, we need to include the effects of phonon-assisted inelastic tunneling and finite quasiparticle lifetime. We follow a procedure outlined in Ref. [36] to add in these effects, and the results are displayed

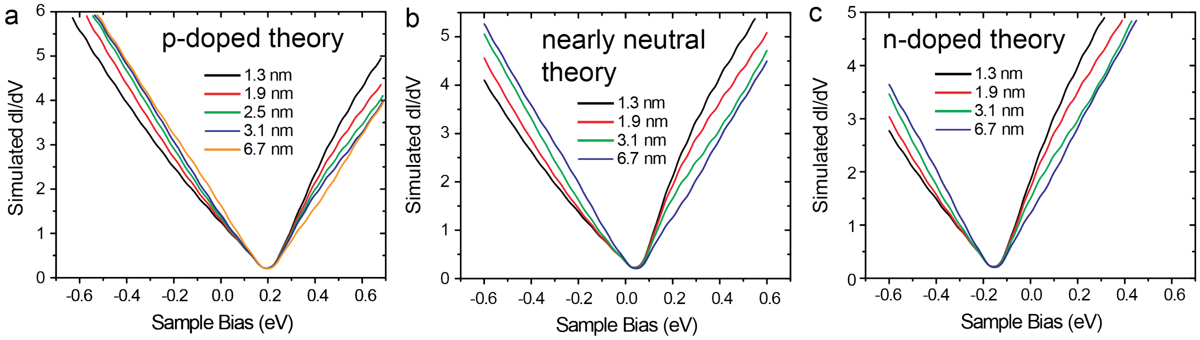

Figure 5.6: Calculated energy-dependent LDOS near screened Coulomb potential. (a) The LDOS of p-doped graphene ($|n| = 2.44 \times 10^{12}$ cm^{-2}) calculated via a tight-binding model for different distances away from a screened Coulomb potential. (b) Same as (a) for nearly neutral graphene ($|n| = 1.5 \times 10^{11}$ cm^{-2}). (c) Same as (a) for n-doped graphene ($|n| = 1.37 \times 10^{12}$ cm^{-2}).

in Figs. 5.3d-f. In agreement with the experimental data shown in Figs. 5.3a-c, the simulated spectra in Figs. 5.3d-f exhibit an electron-hole asymmetry in which dI/dV increases above the Dirac point and decreases below the Dirac point for locations closer to the center of the screened Coulomb potential.

To compare our calculations to Figs. 5.4d and 5.5d, we then analyzed the behavior of the theoretical tunneling spectrum as a function of distance away from the potential center for a fixed energy. Figs. 5.7a and 5.7b plot simulated dI/dV as a function of distance for p-doped and n-doped graphene, respectively. Again, we corrected for lifetime broadening and inelastic tunneling, and the results are plotted in Figs. 5.4e and 5.5e for p-doped and n-doped graphene, respectively. We chose the energies and carrier densities n for the calculations such that Fig. 5.4e (Fig. 5.5e) directly corresponds to Fig. 5.4d (Fig. 5.5d).

In agreement with the experimental observations seen in Figs. 5.4d and 5.5d, the theoretical tunneling spectra return to their unperturbed values more rapidly for higher doping concentrations for both p-doped and n-doped graphene, reflecting the decreased range of the potential caused by reductions in the screening length. Fabiano Corsetti, Arash A. Mostofi, and Johannes Lischner also carried out density functional theory (DFT) calculations for a Ca atom on doped graphene, and the DFT results corroborate the trends seen in the experimental data and in the tight-binding model.

It is apparent from comparing Figs. 5.4d and 5.5d to Figs. 5.4e and 5.5e that the theoretical tunneling spectra return to their unperturbed values more quickly than the experimental dI/dV. This may be evidence that RPA is not sufficient to fully explain screening in graphene. A theory of electron-electron interactions beyond RPA may be required to reconcile the differences between experiment and theory [60, 61].

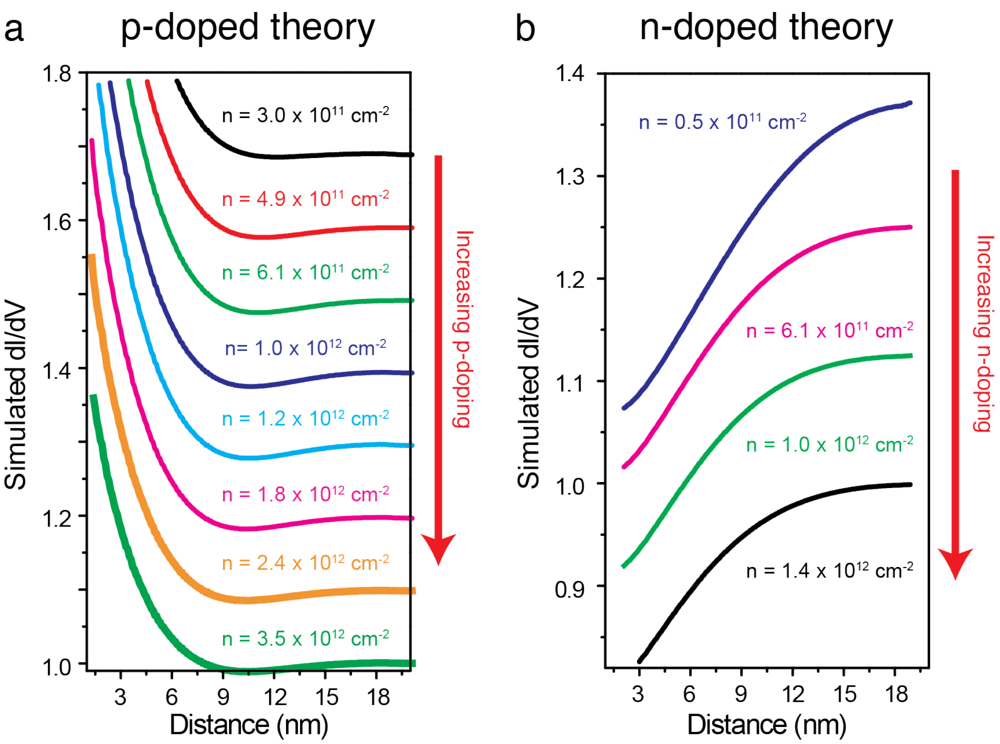

Figure 5.7: Calculated density-dependent LDOS near screened Coulomb potential. (a) Simulated dI/dV (calculated for energy 0.15 eV above the Dirac point) as a function of distance away from an RPA-screened Coulomb potential on p-doped graphene. (b) Same as (a) for n-doped graphene but at energy 0.08 eV below the Dirac point.

5.7 Conclusion

The results presented in this chapter confirm that the RPA model correctly describes electronic screening in graphene up to a point. RPA screening is already an essential ingredient in theoretical models for bipolar charge carrier transport in graphene, as it explains the gate dependence of the conductivity of graphene FETs [62–65]. Electronic screening also affects the energy levels of adsorbates on graphene [66, 67] as well as the coupling strengths of interlayer interactions [68–70]. Our data for the simplest possible system – an individual, isolated charged impurity on graphene – allows us to directly visualize this screening phenomenon on a microscopic scale.

To summarize, we have explored how relativistic charge carriers in graphene respond to a subcritical Coulomb potential. We also examined how these charge carriers screen a Coulomb potential at various carrier concentrations. STM/STS measurements of the local electronic structure of gate-tunable graphene in the presence of isolated Ca adatoms have allowed us to directly observe how the spatial distribution of the density of states around a charged defect changes with the charge carrier density. Unlike conventional parabolic-band 2DEGs, where the Thomas-Fermi screening length λ_{TF} is independent of the carrier density, graphene's Dirac-like band dispersion leads to $\lambda_{TF} \propto 1/\sqrt{|n|}$. Thus, increasing carrier density causes the LDOS (and dI/dV) near a charged impurity to return to its unperturbed value more quickly with increasing distance. This experimentally observed trend is confirmed by a tight-binding model of graphene incorporating an RPA-screened Coulomb potential. The fundamental behavior described here can be generalized to gain a deeper understanding into how density-dependent screening processes may affect the shape of other externally imposed electrostatic potentials.

Chapter 6

Atomic Collapse in Supercritical Artificial Nuclei

This chapter describes our work on the electronic response of massless Dirac fermions to supercritical clusters of charged impurities. The results presented here are based on Ref. [71]: Y. Wang *et al.*, "Observing atomic collapse resonances in artificial nuclei on graphene", Science **340** 734–737 (2013).

6.1 Introduction

An electron bound to a hydrogen-like atom with nuclear charge Z moves with speed $Z\alpha c$, where $\alpha \approx 1/137$ is the fine-structure constant. This clearly cannot be true for $Z > 1/\alpha$, and so a relativistic description of electron motion is required for superheavy atomic nuclei. According to relativistic quantum mechanics, when the charge Z of a nucleus exceeds a certain critical threshold ($Z_C \approx 170$ when the finite size of the nucleus is taken into account), the binding energy of an electron to the nucleus exceeds $2mc^2$ [20]. In this supercritical regime, the strong electric field of the Coulomb potential renders the vacuum unstable and leads to electron-positron pair production. The positron escapes, while the electron is bound to the nucleus. The bound state is known as an "atomic collapse" state because it represents the quantization of a semiclassical trajectory where the electron spirals inwards towards the nucleus [72].

Since the periodic table currently ends at atomic number 118, one might imagine that obtaining a $Z > 170$ supercritical nucleus is far outside our present technological reach. There have been many attempts to observe positron emission from transient supercritical nuclei created by colliding heavy ions [21, 73, 74]. The results of these attempts, however, have so far been inconclusive. There are still future experiments planned at the Facility for Antiproton and Ion Research (FAIR – a particle accelerator under construction in Darmstadt, Germany) that hope to observe atomic collapse behavior in heavy ion collisions [75, 76], but it remains to be seen whether these experiments will be successful.

Fortunately, graphene provides us with a convenient platform for exploring relativistic physics. Since quasiparticles in graphene behave like ultra-relativistic particles, there have been theoretical predictions that charged impurities on graphene could produce atomic collapse states [13–15], with holes in the valence band playing the role of positrons. Graphene's massless Dirac fermions move with speed v_F, so the effective fine-structure constant in graphene is enhanced by a factor $c/v_F \approx 300$. Consequently, the critical charge Z_C needed for supercritical behavior is drastically reduced, and observing atomic collapse in graphene does not require absurdly high values of the nuclear charge.

This chapter describes an experiment where we searched for atomic collapse states around "artificial nuclei" assembled on the surface of a gate-tunable graphene device. We created these artificial nuclei by using an STM tip to push ionized Ca dimers together. This STM atomic manipulation allowed us to create supercritical Coulomb potentials by combining subcritical charged impurities. Near our largest artificial nucleus, STS measurements revealed the emergence of a spatially extended electronic resonance with energy just below the Dirac point. We interpreted this resonance as the electron-like part of the atomic collapse wavefunction and compared our experimental observations with theoretical solutions of the massless Dirac equation with a Coulomb potential.

6.2 Construction of the Artificial Nuclei

The experiments were performed in a UHV Omicron LT-STM at temperature $T = 5$ K. Before all measurements, PtIr STM tips were calibrated against the surface state of an Au(111) crystal [32]. dI/dV was measured by standard lock-in detection of the a.c. tunneling current modulated by a 6 mV rms, 500 – 700 Hz signal added to the tunneling bias.

Our graphene samples were grown via CVD [51]. BN flakes were mechanically exfoliated onto heavily doped Si wafers coated with a 285 nm thermal oxide. Graphene monolayers were transferred on top of BN/SiO_2/Si, and electrical contact was made to graphene by depositing 10 nm Ti and 30 nm Au electrodes using a shadow mask. Completed heterostructures were annealed in UHV at $\sim$400°C overnight before STM/STS measurements.

We deposited Ca atoms onto a low-temperature ($T = 5$ K) graphene surface using a Ca getter source (Alvatec) that was calibrated by evaporating Ca atoms into a mass spectrometer (SRS Residual Gas Analyzer). Since Ca adatoms were difficult to manipulate with an STM tip, we used Ca dimers for our experiment. We obtained Ca dimers on graphene by warming the sample to $T = 16 \pm 3$ K for 1 to 2 minutes before returning the sample to $T = 5$ K. This caused Ca atoms to thermally diffuse and bind together to form dimers.

Fig. 6.1 shows an STM topographic image of Ca dimers and their manipulation. To move a Ca dimer with an STM tip, the tunneling parameters are first set to $V_s = 0.5$ V, $I = 2$ pA, and $V_g = -60$ V. Then, the tip is placed near the dimer on the side opposite to the desired direction of motion. Applying a $V_s = -1$ V pulse pushes the Ca dimer forward. Since Ca dimers are positively charged, they can be difficult to move near each other. The $V_s = -1$ V pulses sometimes push Ca dimers in wrong directions, especially if other Ca

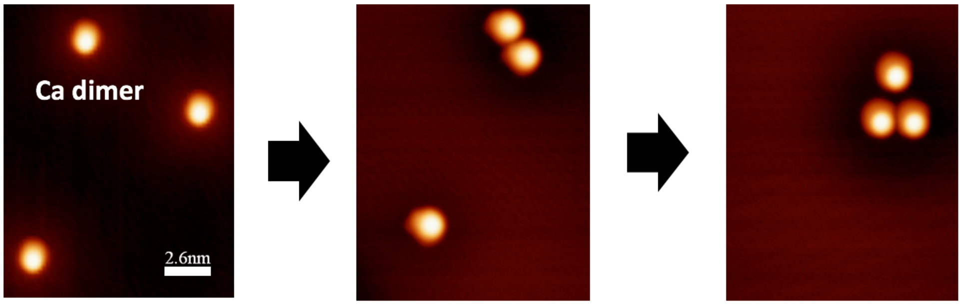

Figure 6.1: STM atomic manipulation of calcium dimers. Voltage pulses are applied to the STM tip to kick Ca dimers into clusters with larger charge.

dimers are nearby. Eventually, however, clusters of Ca dimers (as shown in Fig. 6.1) can be assembled with sufficient effort. We used atomic manipulation to assemble artificial nuclei containing one to five Ca dimers (see insets of Figs. 6.2a-e).

6.3 Spectroscopy Measurements Near the Artificial Nuclei

To characterize the effect that the artificial nuclei have on the motion of massless Dirac fermions in graphene, we measured dI/dV spectra as a function of lateral distance away from the center of each set of Ca dimers. These spectra (normalized for changes in tip height due to the STM feedback loop [54] and shown in Figs. 6.2a-e) all possess a ∼130 meV gap-like feature at the Fermi energy caused by phonon-assisted inelastic tunneling [35, 36]. The spectra obtained at locations far away from the artificial nuclei also have additional minima at around $V_s = 0.23$ V that are associated with the graphene Dirac point. Similar to the case for a single Ca monomer described in Ch. 5, there is also an electron-hole asymmetry in which the dI/dV intensities for states above the Dirac point increase near the clusters and decrease for states below the Dirac point. The electron-hole asymmetry indicates that the Ca-dimer clusters are positively charged because the electron-like states above the Dirac point are attracted to the cluster while hole-like states below the Dirac point are repelled from it.

There are noteworthy differences between each set of dI/dV spectra acquired for each artificial nucleus. Spectra taken near the two-dimer cluster (Fig. 6.2b) displayed greater electron-hole asymmetry than for the one-dimer system (Fig. 6.2a) as well as an extra oscillation at an energy above the Dirac point. Spectra obtained near the three-dimer cluster (Fig. 6.2c) showed even stronger electron-hole asymmetry, and the oscillation began to form a resonance-like structure near $V_s = 0.30$ V. For spectra taken near the four-dimer cluster (Fig. 6.2d), the resonance-like structure increased in intensity and coalesced into a sharp peak

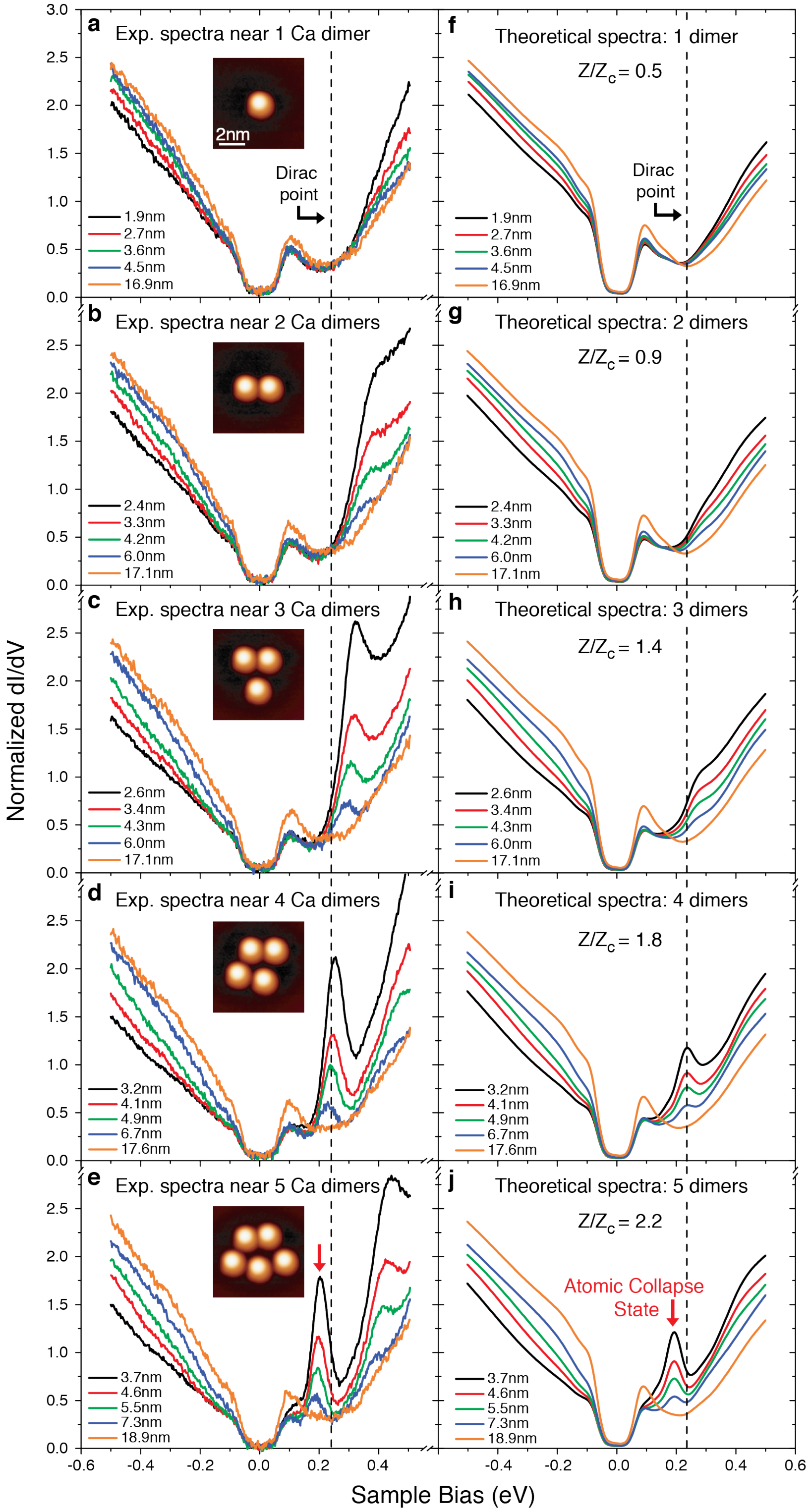

Figure 6.2: Evolution of artificial nuclei from subcritical to supercritical regime. (a-e) dI/dV spectra measured at different distances away from the centers of artificial nuclei composed of one to five Ca dimers. The center of an artificial nucleus is defined as the average of the coordinates of each dimer within a cluster. Initial tunneling parameters: $V_s = -0.5$ V, $I = 60$ pA, $V_g = -30$ V. The insets show STM topographic images of the artificial nuclei. (f-j) Theoretically simulated dI/dV spectra for graphene with a Coulomb potential at the same distances as for the artificial nuclei in (a-e) for (f) $Z/Z_C = 0.5$, (g) $Z/Z_C = 0.9$, (h) $Z/Z_C = 1.4$, (i) $Z/Z_C = 1.8$, and (j) $Z/Z_C = 2.2$. The black dashed lines indicate the location of the Dirac point, and red arrows point at the atomic collapse resonance.

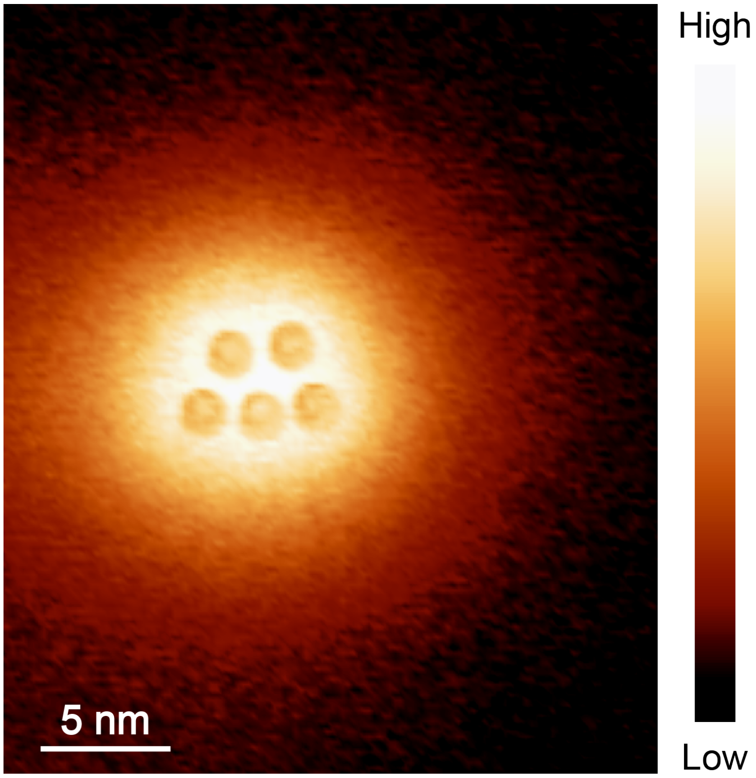

Figure 6.3: dI/dV map of atomic collapse state near a five-dimer artificial nucleus. The dimers appear as slightly darker disks near the center of the circular atomic collapse wavefunction. Tunneling parameters: $V_s = 0.20$ V, $I = 15$ pA, $V_g = -30$ V.

at a well-defined energy near the Dirac point. Finally, for the five-dimer cluster (Fig. 6.2e), the energy of the resonance shifted below the Dirac point. The formation of this resonance represents the transition from the subcritical regime to the supercritical regime and the emergence of a quasi-bound atomic collapse state.

To study the spatial dependence of the quasi-bound state that appears for a high-Z artificial nucleus, we performed dI/dV mapping at the resonance energy around the five-dimer cluster shown in the inset of Fig. 6.2e. This dI/dV map is shown in Fig. 6.3 and displays a radially symmetric distribution of the graphene spectral weight around the five-dimer cluster despite the asymmetric arrangement of the dimers. The state is spatially extended more than 10 nm from the nuclear center and is thus derived from the motion of quasiparticles in the pristine graphene away from the artificial nucleus.

6.4 Theoretical Calculations for Coulomb Potential on Graphene

The resonance that emerges as our cluster charge increases is the atomic collapse state of a supercritical nucleus. This explanation is supported by theoretical calculations of the LDOS of graphene. These LDOS calculations, performed by Andrey V. Shytov and Leonid S. Levitov, assume a 2D continuum Dirac model for undoped graphene in the presence of a radially symmetric potential $V(r)$. As a function of the distance r from the origin, $V(r)$ is Coulomb except for a cutoff at 1 nm, where the potential takes a constant value $V(r < 1\text{ nm}) = V(r = 1\text{ nm})$. The only essential fitting parameter in the calculations was Z/Z_C, where the "effective charge" Z is a measure of the strength of the Coulomb potential (including screening due to the substrate and graphene polarization), and $Z_C = \frac{\hbar v_F}{2e^2}$. More details about the calculations can be found in Refs. [14] and [15].

Fig. 6.4 shows the results of theoretical calculations for several values of Z/Z_C. To directly compare these simulations to our experimental data, we need to include the effects of phonon-assisted inelastic tunneling and quasiparticle lifetime. Figs. 6.2f-j show the results of the same calculation but taking into account inelastic tunneling and lifetime broadening by following a procedure outlined in Ref. [36].

Our experimental data suggests that the one-dimer system and the two-dimer cluster are in the subcritical regime, so we determine Z/Z_C for these two artificial nuclei by matching the strengths of the electron-hole asymmetry in experiment to theory in accordance to a method outlined in Ref. [54]. The three-dimer cluster appears to be transitioning into the supercritical regime, and the four-dimer and five-dimer clusters have fully entered the supercritical regime, as evidenced by the sharpness and energy location of the spatially localized resonance. For these three artificial nuclei, we determined Z/Z_C by matching the resonance energies in the experimental data to the theoretical calculations.

The values of Z/Z_C selected for the calculation results presented in Figs. 6.4 and 6.2f-j were obtained through the aforementioned fitting methods for the one-dimer ($Z/Z_C = 0.5 \pm 0.1$), two-dimer ($Z/Z_C = 0.9 \pm 0.1$), three-dimer ($Z/Z_C = 1.4 \pm 0.2$), four-dimer ($Z/Z_C = 1.8 \pm 0.2$), and five-dimer ($Z/Z_C = 2.2 \pm 0.2$) artificial nuclei. The main features observed in the experimental data are reproduced by their corresponding Dirac equation simulations, including the electron-hole asymmetry and the emergence of the resonance peak as the total charge increases. This confirms our interpretation of the resonance as the atomic collapse quasi-bound state of a supercritical charge.

As a sanity check, Sangkook Choi and Steven G. Louie used DFT to calculate the charge of a single Ca dimer. This calculation, which had no fitting parameters, yielded $Z/Z_C = 0.6 \pm 0.3$, consistent with $Z/Z_C = 0.5 \pm 0.1$ extracted from comparing our experimental data to the Dirac equation simulations. Additionally, we can rule out the possibility that our quasi-bound states are caused by multiple scattering of Dirac quasiparticles between the Ca dimers because the distance between dimers in an artificial nucleus is about 2 nm. The corresponding energy for multiple scattering is $E = \hbar v_F \pi/(2\text{ nm}) \approx 1$ eV, which is far too

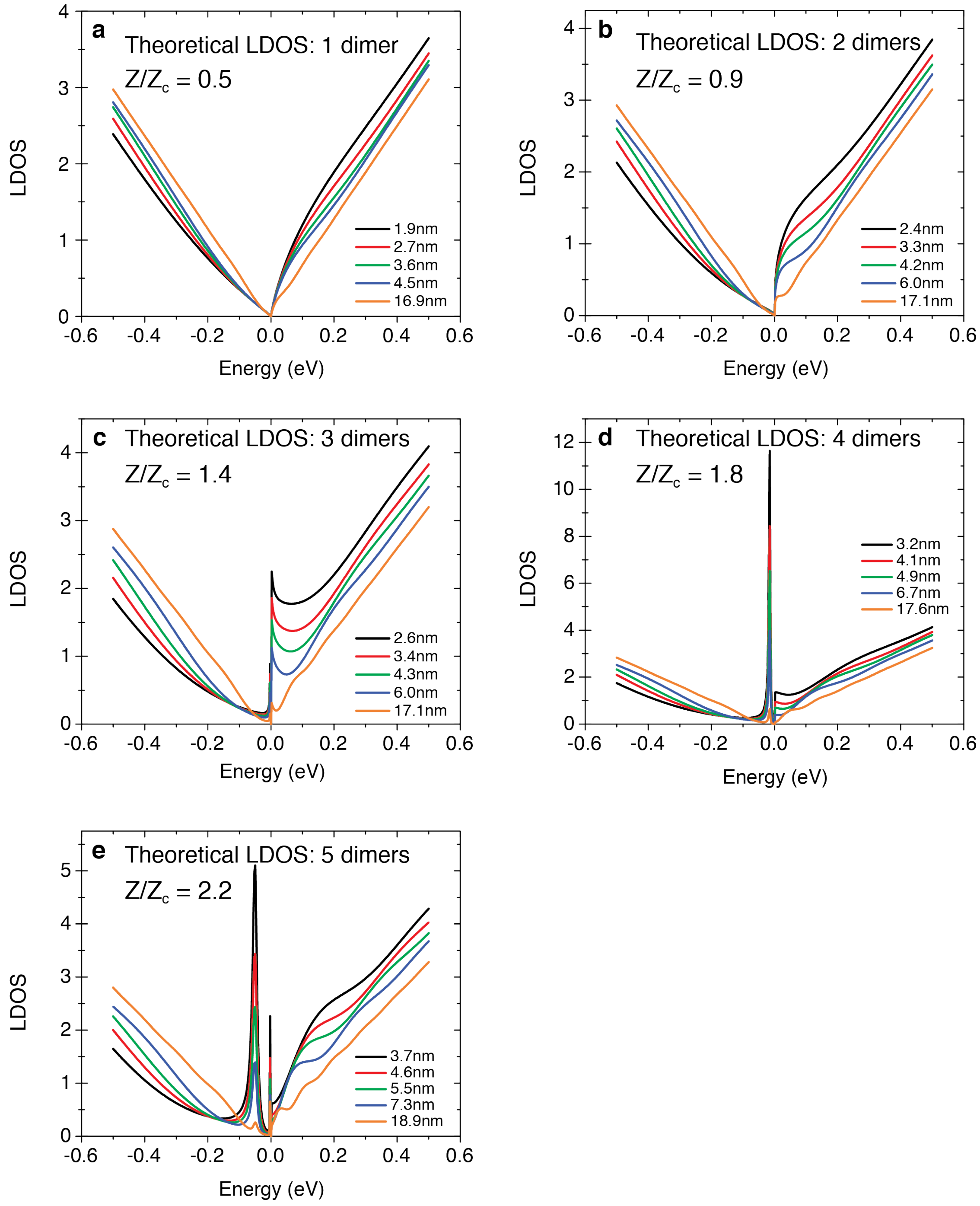

Figure 6.4: Theoretical LDOS of graphene with Coulomb potential calculated by solving the massless Dirac equation. These curves are for different distances from an impurity having charge (a) $Z/Z_C = 0.5$, (b) $Z/Z_C = 0.9$, (c) $Z/Z_C = 1.4$, (d) $Z/Z_C = 1.8$, and (e) $Z/Z_C = 2.2$.

large to explain our experimental observations. The large spatial extent of the quasi-bound state (see Fig. 6.3) also shows that the resonance arises from a spatially extended Coulomb potential and not from local impurity hybridization with the underlying carbon atoms in graphene.

The radially symmetric distribution of dI/dV seen in Fig. 6.3 is consistent with the predicted behavior for the atomic collapse state with lowest energy and lowest angular momentum (i.e. $j = 1/2$) [13–15]. In principle, just like the hydrogen atom, there should be an infinite number of quasi-bound states corresponding to different principal quantum numbers. However, because the energies of the atomic collapse states follow an exponential sequence and the spatial extent of each state is inversely proportional to its energy (as measured from the Dirac point), only the quasi-bound state with the lowest energy should be detectable because screening suppresses all states with sufficiently large spatial extent.

6.5 Carrier Density Dependence of Atomic Collapse Resonance

The observed behavior of the atomic collapse state depends on whether it is empty or occupied by electrons. This can be seen in the gate-dependent dI/dV spectra in Fig. 6.5, which were all measured 3.7 nm from the center of the five-dimer cluster in the inset of Fig. 6.2e. Each red arrow indicates the energy of the atomic collapse state for each gate voltage, and the black arrows indicate the energies of the Dirac point as extracted from measuring dI/dV on graphene at least 20 nm away from the cluster center. For p-doped graphene (-60 V $\leq V_g \leq -15$ V), the atomic collapse state is empty, and the resonance shifts to lower energy relative to the Dirac point as the magnitude of the carrier concentration decreases. For n-doped graphene (0 V $\leq V_g \leq$ 30 V), the resonance intensity is dramatically suppressed.

The strong doping dependence of the atomic collapse resonance can be partially explained as a screening effect. In the p-doped regime, electronic screening reduces the effective charge of the artificial nucleus, reducing the magnitude of the bound state energy. When the magnitude of the carrier density decreases, the screening of the impurity charge is reduced, so the resonance peak is shifted to lower energy.

The phenomenology in the n-doped regime is radically different and is difficult to explain. Here, the atomic collapse peak almost completely disappears. We hypothesize that this behavior is due to internal correlation among the localized charge carriers that inhabit the atomic collapse state. Normally, graphene states are fourfold degenerate due to the spin and valley degrees of freedom. However, the atomic collapse state cannot be inhabited by four electrons because each electron would screen the nucleus such that the effective charge Z would appear subcritical to the other electrons. Somehow, electron-electron interactions reduce the single-particle spectral function as measured by STS, but future studies are required to elucidate how this occurs.

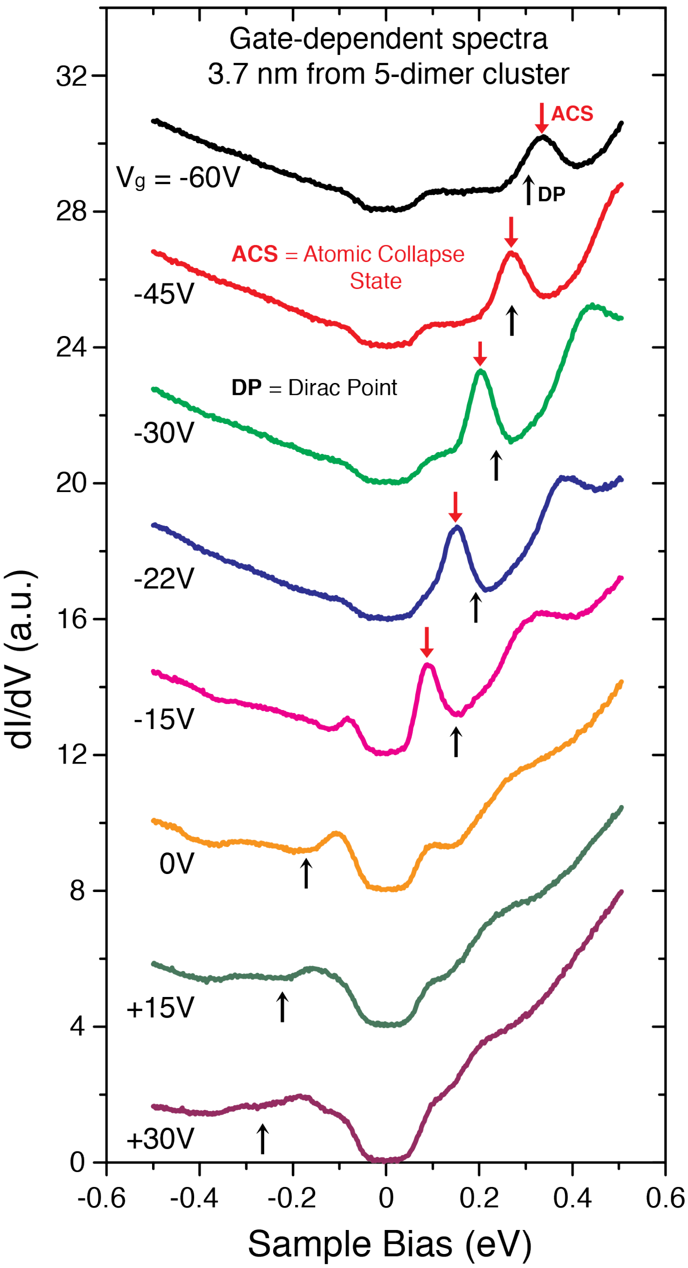

Figure 6.5: Doping dependence of atomic collapse state. Gate-dependent dI/dV spectra obtained at a lateral distance of 3.7 nm from the center of a five-dimer cluster. Curves are shifted vertically for clarity. Each red arrow indicates the energy of the atomic collapse state for each gate voltage, while black arrows indicate the Dirac point energies extracted by measuring dI/dV spectra at least 20 nm away from the cluster center. Initial tunneling parameters: $V_s = -0.5$ V, $I = 60$ pA.

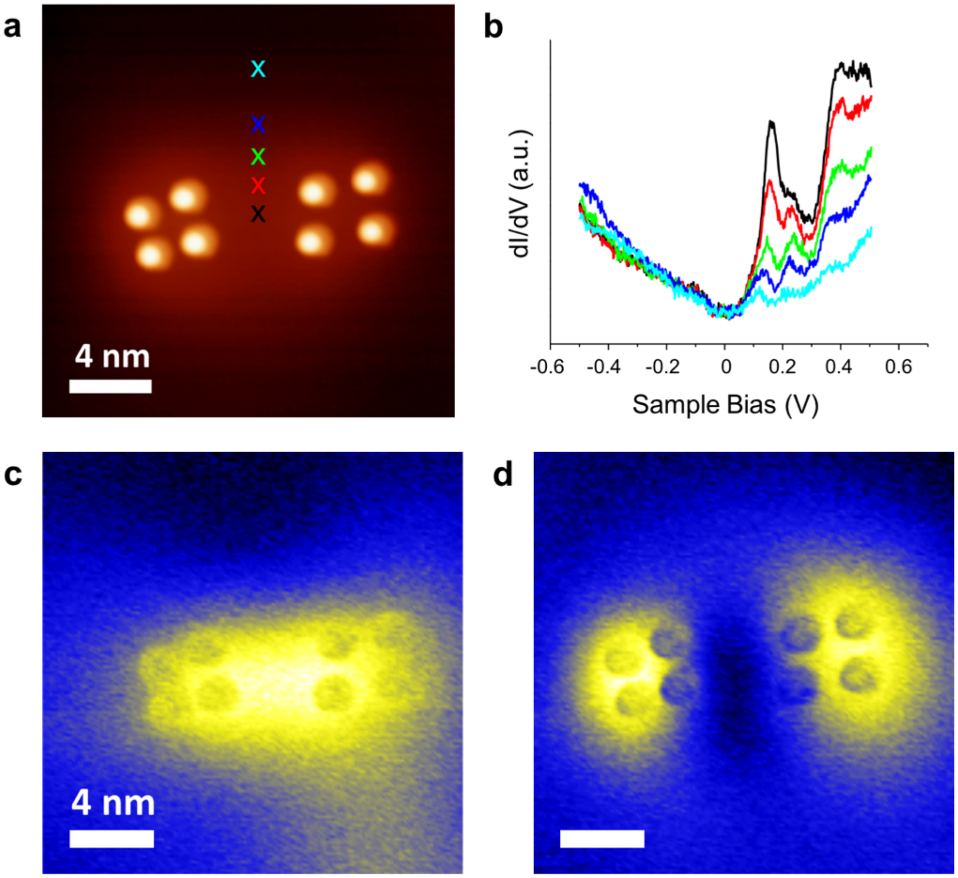

Figure 6.6: Atomic collapse molecule. (a) STM topographic image of two four-dimer clusters assembled near each other. (b) dI/dV spectra (not normalized) each acquired at the correspondingly color-coded “X” in (a). Initial tunneling parameters: $V_s = -0.5$ V, $I = 40$ pA, $V_g = -25$ V. (c) dI/dV map of the two four-dimer clusters. The dimers appear as slightly darker disks. Tunneling parameters: $V_s = 0.15$ V, $I = 10$ pA, $V_g = -25$ V. (d) Same as (c) but with different tunneling parameters: $V_s = 0.22$ V, $I = 20$ pA, $V_g = -25$ V.

6.6 Atomic Collapse Molecule

Now that we had some fun playing faux nuclear physics, we can also play pseudo-chemistry with the atomic collapse orbitals. Fig. 6.6a shows an STM topographic image of eight Ca dimers arranged in two groups of four. To probe the electronic structure of this coupled system of two artificial nuclei, we performed STS measurements at various locations near this Ca-dimer configuration. Fig. 6.6b shows a series of dI/dV spectra taken at points along a line between the two artificial nuclei. There appears to be two resonance peaks now.

Figs. 6.6c-d are two dI/dV maps taken at different energies. The map in Fig. 6.6d has a larger energy ($V_s = 0.22$ V) than the map in Fig. 6.6c ($V_s = 0.15$ V). The appearance of the dI/dV maps strongly suggests that Fig. 6.6c shows a bonding orbital and Fig. 6.6d shows an antibonding orbital. If this is indeed the case, then the double-four-dimer system is an "atomic collapse molecule." Then, by creating various arrangements of n-dimer clusters, we could engineer new and interesting quantum eigenstates for exploring $Z > 170$ ultra-relativistic chemistry.

6.7 Conclusion

In conclusion, we investigated the behavior of massless Dirac fermions in a Coulomb potential by using STM atomic manipulation to assemble clusters of Ca dimers on a gated graphene surface. We found that, as the number of dimers in a cluster increased, the system underwent a transition between a subcritical regime with no bound states to a supercritical regime with a quasi-bound atomic collapse state. This atomic collapse state observed in graphene is highly analogous to an electronic eigenstate of a $Z > 170$ atomic nucleus. Furthermore, the gate dependence of the atomic collapse resonance revealed very nontrivial behavior that we believe is caused by electron-electron interactions.

The atomic collapse state is the impurity state of a donor atom on graphene, and just like for donor and acceptor states in semiconductors, learning how to create and manipulate these impurity states may be useful for controlling the motion of charge carriers in graphene-based devices. Constructing donor and acceptor levels in graphene through atomic manipulation is quite difficult and time consuming, but there may be alternative methods for realizing these exotic quasi-bound states [77].

Chapter 7

Manipulation of Defects in Hexagonal Boron Nitride

In Ch. 8, I will discuss our work on the relativistic quantum harmonic oscillator. We created a parabolic potential in graphene by manipulating the charge states of defects in the underlying BN substrate. In this chapter, we examine some of the properties of such defects and describe a technique for manipulating them. The results presented here are based on Ref. [78]: D. Wong *et al.*, "Characterization and manipulation of individual defects in insulating hexagonal boron nitride using scanning tunnelling microscopy", Nature Nanotechnology **10**, 949–953 (2015). I also use content from Ref. [79]: J. Velasco Jr. *et al.* "Nanoscale control of rewriteable doping patterns in pristine graphene/boron nitride heterostructures", Nano Letters **16**, 1620–1625 (2016).

7.1 Introduction

Since defects have a major role in determining the properties and applications of many materials, characterizing them on a microscopic level is of crucial importance. Defects in BN are of particular interest because BN is a vital component in many new and technologically promising devices that incorporate 2D materials [38, 80]. Although BN defects have been investigated through cathodoluminescence [81], secondary ion mass spectrometry (SIMS) [82], electron paramagnetic resonance [83–85], and many other techniques [86–88], these studies have primarily been limited to spatially averaged defect behavior, and investigation of individual BN defects remains an outstanding challenge.

STM has long been used to image individual point defects in conductors [89, 90], semiconductors [91–94], and ultrathin films [95–99], but single-defect characterization remains an elusive goal for bulk insulators. In this chapter, I discuss the STM characterization and manipulation of defects inside insulating BN. It is normally difficult to use STM to obtain information about insulators because there are no conducting paths to drain the tunneling current, but this was not an issue in our experiment because BN was capped by monolayer

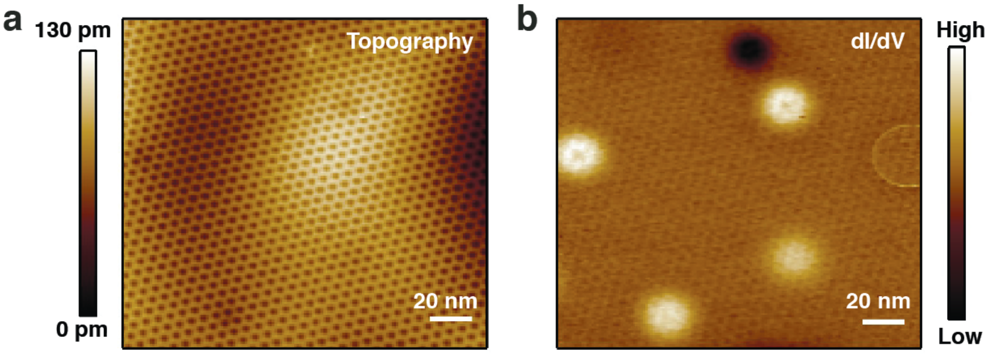

Figure 7.1: Detecting charged defects inside BN. (a) STM topographic image of a clean graphene/BN area with a ∼7 nm moiré pattern. (b) dI/dV map of the same location as (a). There are bright dots, a dark dot, and a ring structure that are presumably due to BN defects. Tunneling parameters: $V_s = -0.25$ V, $I = 0.4$ nA.

graphene. The atomically thin nature of graphene allowed us to visualize defects in the underlying BN layers and obtain charge and energy-level information about the defects.

We were also able to change the charge states of the defects through voltage pulses applied to our STM tip. By combining this tip-induced defect manipulation with an electric field from the Si backgate, we patterned nanoscale p-n junctions in graphene/BN heterostructures. The p-n junctions produced by this method were nonvolatile, erasable, and (most importantly for their use in Ch. 8) clean.

7.2 Detecting Subsurface Defects

Our samples were fabricated using the transfer technique described in Ch. 4. We exfoliated BN crystals (synthesized by Takashi Taniguchi and Kenji Watanabe) onto heavily doped Si chips each capped by a 285 nm SiO_2 dielectric layer. Graphene monolayers were exfoliated from graphite onto MMA and were each subsequently transferred onto a 60 to 100 nm thick BN flake. The samples were electrically contacted with standard electron-beam lithography and were then annealed in flowing Ar/H_2 forming gas at 350°C. We measured the electrical conductance of completed devices with a standard a.c. voltage bias lock-in technique with a 50 µV signal at 97.13 Hz. Samples that exhibited acceptable bipolar charge carrier transport were transferred into the preparation chamber of our UHV Omicron LT-STM, where they were annealed for several hours at ∼300°C before moving the devices into the STM chamber for measurements at temperature $T = 5$ K. Our PtIr STM tip was calibrated against the surface state of an Au(111) crystal [32]. All STM topographic images and dI/dV maps were obtained using constant-current feedback, and all STS measurements were acquired through lock-in detection of the a.c. tunneling current induced by a 6 mV, 613.7 Hz wiggle voltage added to V_s.

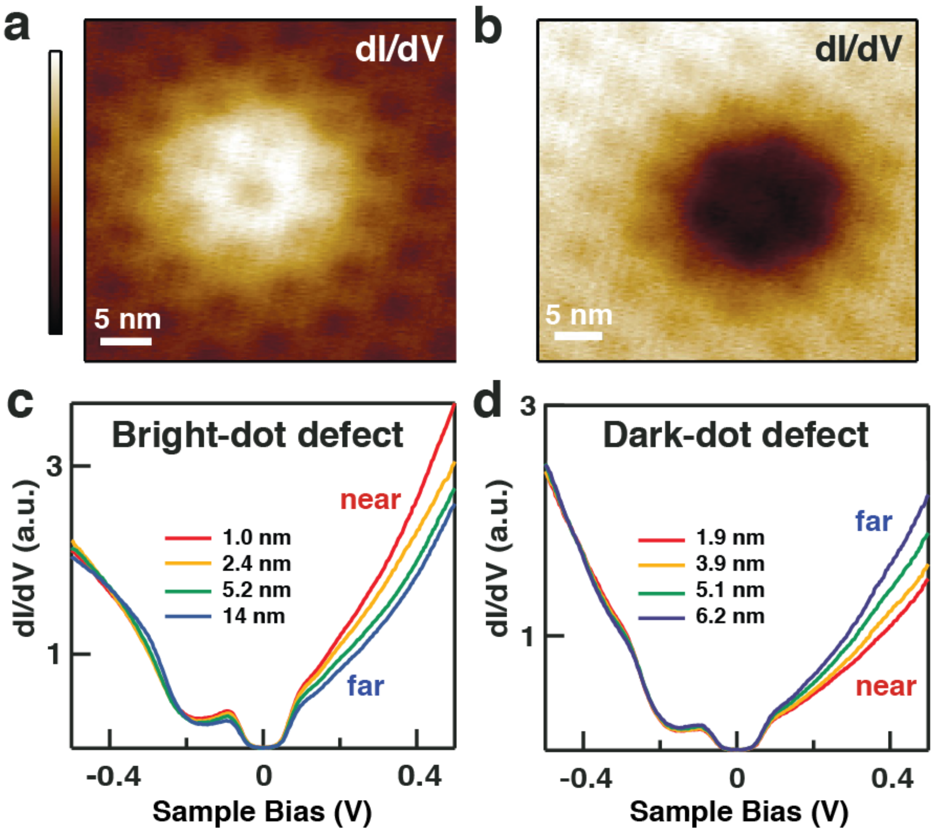

Figure 7.2: Identifying the charge states of BN defects. (a) dI/dV map of bright-dot defect. (b) dI/dV map of dark-dot defect. Tunneling parameters: $V_s = -0.3$ V, $I = 0.4$ nA, $V_g = 5$ V. (c) dI/dV spectra of graphene measured at various lateral distances from the center of the bright dot in (a). The spectra reveal that the bright dot is a positively charged defect. (d) Same as (c) but for the dark dot in (b). The spectra reveal that the dark dot is a negatively charged defect. Initial tunneling parameters: $V_s = -0.5$ V, $I = 0.4$ nA, $V_g = 20$ V.

Fig. 7.1a shows an STM topographic image of graphene/BN with a ∼7 nm moiré pattern caused by the interference of the graphene and BN atomic lattices [33, 34]. We can see localized shallow dips and a protrusion with height $\Delta z < 0.1$ Å. Fig. 7.1b presents a dI/dV map taken at the same location. The map reveals randomly distributed bright (high dI/dV) and dark (low dI/dV) circular dots that are ∼20 nm in diameter and have varying degrees of intensity. A sharp ring feature is also seen at the right edge of the map. None of the localized defect features seen in Fig. 7.1b appear to distort the moiré pattern at their corresponding locations in the topographic image in Fig. 7.1a.

Figs. 7.2a-b display higher-resolution dI/dV maps of representative bright and dark dots. The maps clearly show that the moiré pattern is not obscured by the defects. We cannot

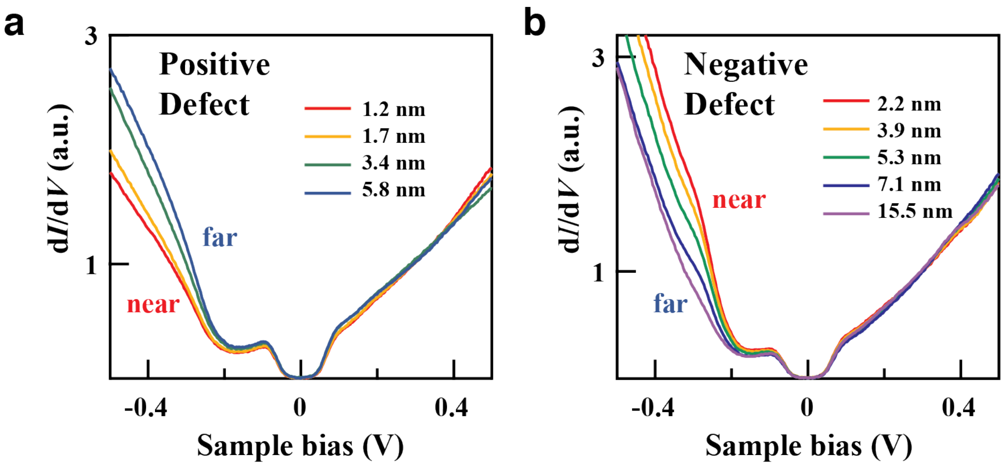

Figure 7.3: Electron-hole asymmetry in spectroscopy near BN defects. (a) dI/dV spectra of graphene measured at various lateral distances from the center of the bright-dot defect in Fig. 7.2. (b) Same as (a) but for the dark-dot defect. Initial tunneling parameters: $V_s = 0.5$ V, $I = 0.4$ nA, $V_g = 20$ V.

identify the charge states of these dots from the dI/dV maps alone because the appearances of the dot defects nontrivially depend on V_s and V_g. As mentioned in Sec. 2.2.1, as V_s and V_g are changed, the constant-current feedback loop adjusts the STM tip height such that the current I remains constant. This may cause bright-dot defects to appear dark (and vice versa) under different tunneling conditions.

In order to identify the charge states of these defects, we acquired dI/dV spectra at different distances away from the centers of the dots (plotted in Figs. 7.2c-d). The graphene Dirac point is seen at $V_s \approx -0.17$ V. For the bright-dot spectra (Fig. 7.2c), dI/dV above the Dirac point increases as the STM tip approaches the bright-dot center. Fig. 7.2d shows the opposite trend, in which dI/dV above the Dirac point decreases as the STM tip approaches the dark-dot center. As explained in Ch. 5, an increase of the LDOS above the Dirac point indicates that electron-like Dirac fermions are attracted to the center, and hence an increase in dI/dV is a signature of a positive charge. Likewise, a decrease of the LDOS above the Dirac point indicates that electron-like Dirac fermions are repelled from the center, and hence a decrease in dI/dV is a signature of a negative charge. Therefore, we can conclude that the bright dot is a positively charged defect while the dark dot is negatively charged.

The spectra in Figs. 7.2c-d are not normalized for changes in tip height due to the STM feedback loop [54]. Notice that the areas under the curves in Fig. 7.2c (as well as in Fig. 7.2d) integrated from $V_s = -0.5$ V to $V_s = 0$ V are equal. This is a consequence of the constant-current feedback condition. We can change the initial tunneling parameters such that the integrated areas between $V_s = 0$ V to $V_s = 0.5$ V are all equal. Fig. 7.3 shows precisely this. In these spectra, dI/dV below the Dirac point decreases near the positively charged defect (the bright dot) because hole-like Dirac fermions are repelled. In contrast, dI/dV below the Dirac point increases near the negatively charged defect (the dark dot) because hole-like Dirac fermions are attracted.

What are these dot defects? Three scenarios are possible: (i) adsorbates bound to the surface of graphene, (ii) adsorbates trapped at the interface between the graphene and BN, and (iii) intrinsic defects within the insulating BN substrate. Our observations imply that (iii) is the correct scenario. Scenario (i) can be ruled out because weakly bound adsorbates would have higher height profiles than the observed topographically small features ($\Delta z <$ 0.1 Å) and would also be swept away by the STM tip for the large tunneling currents involved in our measurements. Strongly bound adsorbates would also have taller height profiles and should disrupt the graphene honeycomb lattice, which we did not observe. Additionally, strongly bound adsorbates would lead to changes in the graphene spectroscopy due to the formation of localized bonding states, which are not seen. We can rule scenario (ii) out because an adsorbate trapped beneath graphene would cause a bump in graphene at least an order of magnitude larger than the $\Delta z < 0.1$ Å features observed here. A trapped adsorbate would also locally delaminate the graphene from the BN substrate, disrupting the moiré pattern, which is evidently not seen in Figs. 7.1 and 7.2a-b.

Scenario (iii) – charged defects inside the BN substrate – is therefore the most likely explanation for the defects observed here. Polycrystalline BN is known to host many types of charged defects, as inferred from thermoluminescence and electron paramagnetic resonance experiments [83–85]. Examples of such defects are nitrogen vacancies (which act as donors) and carbon impurities substituted at nitrogen sites (which act as acceptors). A SIMS investigation of high-purity, single-crystal BN synthesized at high pressure and temperature (which are used in the study described in this chapter) has also identified oxygen and carbon impurities [82]. A comparison between optoelectronic experiments performed on the high-purity, single-crystal BN [87, 88] and theoretical calculations [100, 101] demonstrates that the defects in the high-purity, single-crystal BN are consistent with nitrogen vacancies and carbon substitutions of nitrogen. Such defects, when ionized, could produce the bright dots and dark dots seen in Figs. 7.1 and 7.2. The defects appear as ∼20 nm clouds because they are indirectly observed through graphene's screening response to their electric potentials. The fact that these defects are embedded in BN layers underneath graphene explains why the dots have such small topographic deflection, why the moiré pattern is not disrupted, and also why no new states arise in graphene spectroscopy. The variations in the intensities of the bright and dark dots can be explained by BN defects lying at different depths relative to the graphene capping layer.

7.3 The Ring Defects

As seen at the right edge of Fig. 7.1, there are also ring defects. Similar rings observed in other systems have been attributed to the charging of impurities and quantum dots [92, 102–104]. We hypothesize that the ring in Fig. 7.1 is due to the charging of a defect in the topmost layer of BN. The defect hybridizes with graphene states, and the ring radius R

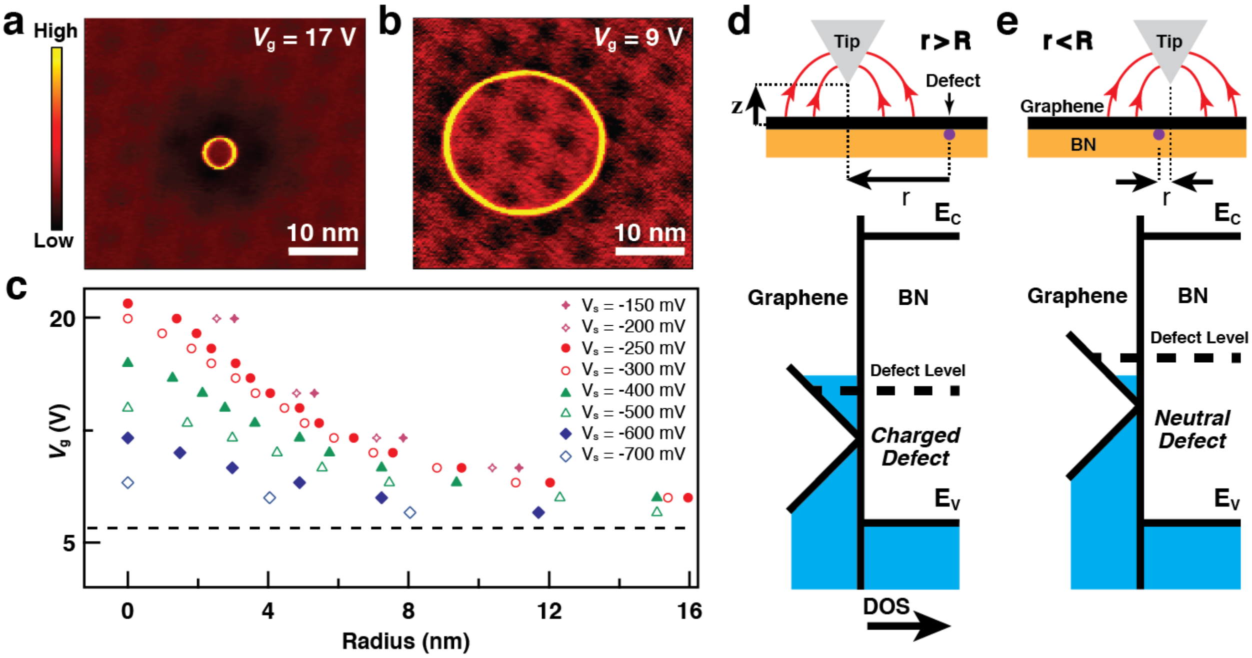

Figure 7.4: Schematic model for ring defect, along with data on the gate dependence of the ring radius. (a) dI/dV map of a ring defect at $V_g = 17$ V. Tunneling parameters: $V_s = -0.3$ V, $I = 0.4$ nA. (b) dI/dV map of the same ring defect as (a) but at $V_g = 9$ V. (c) Ring radius R for different values of V_s and V_g. The ring radius R was extracted from dI/dV maps of the same ring defect as in (a) and (b). (d-e) Schematic model (energies not to scale) for ring formation due to charge transfer between graphene and a defect in the top layer of BN. When the distance r between the STM tip and the defect is larger than the ring radius (i.e. $r > R$), the defect level is filled, and the defect is negatively charged. When the distance between the tip and the defect is smaller than the ring radius (i.e. $r < R$), local gating from the tip reduces the electron density around the defect such that the Fermi level is below the defect level, neutralizing the defect. If the tip potential is negative and the graphene is n-doped, lowering V_g shifts the unperturbed defect level closer to the Fermi energy, increasing R.

becomes dependent on the graphene carrier density. Figs. 7.4a-b show that the ring radius changes from 2 nm (Fig. 7.4a) to 11 nm (Fig. 7.4b) as the backgate voltage is changed from $V_g = 17$ V to $V_g = 9$ V at $V_s = -0.3$ V. Fig. 7.4c plots the ring radius R against V_g for various values of V_s. This data was acquired by measuring R from dI/dV maps taken at the same location as Figs. 7.4a-b but varying V_g and V_s. Although the value of R depends on the STM tip shape as well as the local graphene charge neutrality point, the qualitative trend seen in Fig. 7.4c is representative of most of the observed ring defects. For a constant V_s, R increases with decreasing V_g until a critical gate voltage ($V_C = 6 \pm 1$ V for the ring observed in Fig. 7.4), upon which the ring vanishes.

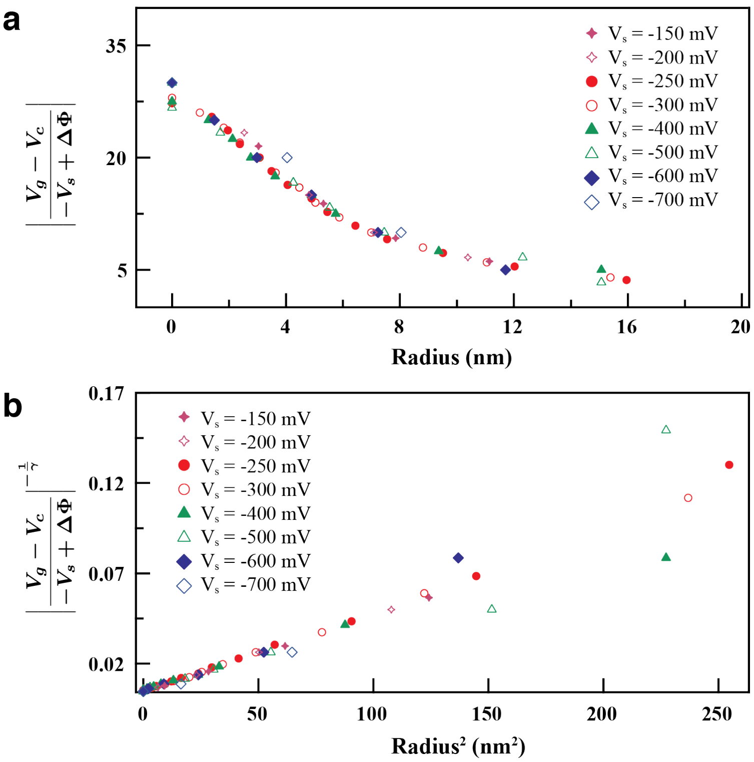

Figure 7.5: The gate dependence of the ring radius R for different values of V_s collapses onto a single curve when the gate voltage is divided by V_{tip}. In these plots, $V_C = 6$ V, $\Delta\Phi = -0.8$ V, and $\gamma = 0.63$.

I now present a model to explain the appearance and behavior of the ring. Suppose there is a defect in the topmost layer of BN such that the defect strongly interacts with graphene. Also suppose that the defect has an energy level that is below the Fermi energy when the tip is far away. The tip is a top gate that capacitively couples to the graphene region directly above the defect (this is called tip-induced band bending). If the electric potential at the tip is V_{tip}, then the change in the local carrier density δn induced by the tip is given by

$$e\delta n = C(r)V_{\text{tip}}. \tag{7.1}$$

Here, $C(r)$ is a capacitance per unit area that increases with decreasing tip-defect lateral distance r. The quantity V_{tip} is not the same as V_s because of the contact potential due to the difference between the PtIr and graphene work functions. The correct expression for V_{tip} is

$$V_{\text{tip}} = -V_s + \frac{\Phi_{\text{graphene}} - \Phi_{\text{PtIr}}}{e}, \tag{7.2}$$

where the work functions are $\Phi_{\text{graphene}} \approx 4.5$ eV [105] and $\Phi_{\text{PtIr}} \approx 5.2$ eV [106].

For the data acquired in Fig. 7.4, $V_{\text{tip}} < 0$ V, so electrostatic gating from the tip lowers the local electron density in the graphene. As the tip approaches the defect, $C(r)$ increases, and δn becomes more negative. The Fermi level eventually crosses the defect level when the tip is at some distance R^* away from the defect. The charge state of the defect and, consequently, the graphene LDOS abruptly change at $r = R^*$, causing a discontinuity in $I(V_s)$. This sudden change in the tunneling current manifests as a ring with radius $R = R^*$ in dI/dV maps. The ring radius R separates the region where the tip is too far from the defect to change its charge state ($r > R$) from the region where the tip is close enough to the defect to change its charge state ($r < R$). Figs. 7.4d and 7.4e sketch cartoons for the $r > R$ and $r < R$ cases, respectively. In these diagrams, the defect level is depicted as filled (and the defect is negatively charged) when $r > R$, while the defect level is depicted as empty (and the defect is neutral) when $r < R$. This picture is supported by the dI/dV maps in Figs. 7.4a-b, where the exterior of the ring is surrounded by a dark screening cloud that is absent in the interior.

If V_C^* is the backgate voltage where the Fermi level and the defect level coincide (without perturbation from the tip), then no ring should appear for $V_{\text{tip}} < 0$ V and $V_g < V_C^*$ because decreasing the electron density further increases the energy separation between the Fermi level and the defect level. Additionally, R should increase as V_g approaches V_C^* from above because the $|\delta n|$ required to neutralize the defect decreases. This is consistent with our observation that the ring radius diverges at $V_C = 6 \pm 1$ V. Hence, we deduce that $V_C^* = V_C$, which occurs when the Fermi level (and hence the defect level) is ~30 ± 10 meV above the graphene Dirac point. This puts the ring defect level at ~4 eV below the edge of the BN conduction band [107], which is similar to an energy level for carbon substitutions of nitrogen in bulk BN [84]. This may be a coincidence because the energy levels of defects at an interface are not the same as for defects in the bulk, and we expect hybridization between graphene and a carbon substitutional impurity to locally distort the graphene lattice.

We can test the validity of the model described above because $C(r) = e\delta n/V_{\text{tip}}$ is (to a very good approximation) only a function of r and is not a function of V_s. Hence, the data in Fig. 7.4c can be rescaled onto a single curve by plotting the quantity

$$\left| \frac{V_g - V_C}{-V_s + \left(\Phi_{\text{graphene}} - \Phi_{\text{PtIr}}\right)/e} \right| \tag{7.3}$$

against the ring radius R. This is evidently shown in Fig. 7.5a, where we take the work function difference to be $\Phi_{\text{graphene}} - \Phi_{\text{PtIr}} = -0.8$ eV. The data points do not fall onto a single curve as nicely for other values of $\Phi_{\text{graphene}} - \Phi_{\text{PtIr}}$.

Since the capacitance per unit area $C(r)$ must decrease with increasing r, let's assume the following functional form:

$$C(r) = \frac{\alpha}{(r^2 + \beta^2)^\gamma}, \tag{7.4}$$

where α, β, and γ are unknown geometric factors. Then, if we plot the quantity

$$\left| \frac{V_g - V_C}{-V_s + \left(\Phi_{\text{graphene}} - \Phi_{\text{PtIr}}\right)/e} \right|^{-\frac{1}{\gamma}} \tag{7.5}$$

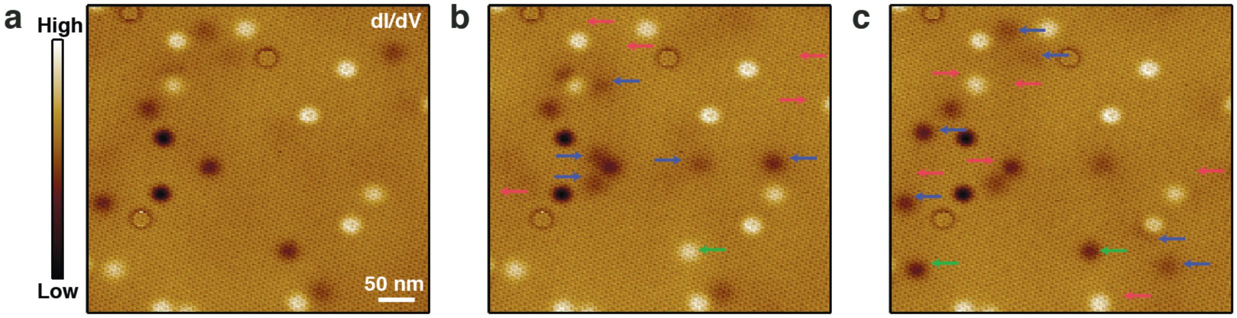

Figure 7.6: STM manipulation of BN defects. Voltage pulses are applied to the STM tip ($V_s = 5$ V for 10 seconds) to toggle the charge states of the dot defects. (a) dI/dV map of graphene/BN showing various dots and rings. Tunneling parameters: $V_s = -0.25$ V, $I = 0.4$ nA. (b) dI/dV map of the same location after the tip was moved to the center of the region in (a) and a voltage pulse was applied to the tip. (c) dI/dV map of the same location after another tip pulse. Red arrows mark the disappearance of dots relative to the previous image, blue arrows mark the appearance of dots, and each green arrow marks a defect that has changed the sign of its charge.

against R^2, the data points should fall onto a straight line for the correct value of γ. If the tip was an infinite parallel plate, $\gamma = 0$. If the tip was a perfect sphere hovering over a perfectly metallic surface, then $\gamma = 1$. Fig. 7.5b shows the data for $\gamma = 0.63$.

7.4 Defect Manipulation

7.4.1 Field Ionization of Defects

In addition to characterizing BN defects, we were also able to manipulate their charge states with the STM tip. Unlike the ring phenomenon discussed in Sec. 7.3, the defect manipulation presented in this section is nonvolatile, i.e. changes in the charge states persist when the tip is removed. STM control over defect charge has been demonstrated in other systems, such as for Si impurities in GaAs [108] and Au adatoms on NaCl films [95], but this is unprecedented in bulk insulators.

Fig. 7.6a shows a dI/dV map of graphene/BN with many charged defects. To change the charge states of defects in this map, we performed the following procedure:

1. Set $V_s = 0.5$ V, $I = 0.4$ nA, $V_g = 0$ V while the feedback loop is closed.
2. Open the STM feedback loop.
3. Withdraw the tip 1 nm.
4. Ramp V_s up to 5 V over a period of 30 seconds.

5. Wait 10 seconds.
6. Ramp V_s down to 0.5 V over a period of 30 seconds.
7. Close the STM feedback loop.

It is not necessary to follow this procedure exactly. The important step is to apply a relatively large bias on the tip to create a sufficiently large electric field in the tip-sample junction. The magnitude of the required bias depends on the tip-sample distance. Since graphene does not perfectly screen electric fields in the vertical direction [109, 110], defects in BN will experience a field that is strong enough to ionize them.

Fig. 7.6b presents a dI/dV map of the same region as Fig. 7.6a after performing this tip-pulse procedure, and Fig. 7.6c shows a dI/dV map after another tip voltage pulse. A close examination of Fig. 7.6 reveals that the defect configurations are significantly altered by application of these tip pulses. The defects are observed to reversibly switch between charged and neutral states as well as between states having opposite charge. We use colored arrows to highlight changes between each dI/dV map. A red arrow indicates the disappearance of a charged defect (when compared to the preceding image), a blue arrow marks the appearance of a charged defect, and a green arrow points out where a defect has changed the sign of its charge. Defects that disappear after a tip pulse always reappear in the same location after subsequent tip pulses. Furthermore, we note that dark dots (the negatively charged defects) disappear and reappear at a higher rate than bright dots (the positively charged defects). Ring defects, the brightest dots, and the darkest dots remain unchanged by tip pulses.

When a defect disappears after a tip pulse, it has changed from a charged state to a neutral state. Likewise, a defect that appears after a tip pulse has changed from a neutral state to a charged state. These tip-pulse-induced changes in the defects' charge states are caused by electric-field-induced emission of charge carriers from the BN defect states. By tilting the local potential landscape, the tip electric field causes charge carriers to tunnel through the ionization barrier between different defects, charging some while neutralizing others. Unfortunately, we do not know the precise mechanism whereby carriers are pulled from their defects, as there are multiple known processes for field-induced ionization of defects in insulators: direct tunneling [111], phonon-assisted tunneling [112], and thermionic emission (i.e. the Poole-Frenkel effect) [113, 114], just to name a few. Considering the low-temperature condition of our experiment, we suspect that our defects are ionized by direct tunneling, but we do not know for sure.

Field ionization of BN defects explains the disappearance and reappearance of dots at the same locations, which cannot be interpreted as defect migration through the BN lattice. The higher rate of switching for the dark dots compared to the bright dots suggests that it is more energetically favorable for neutral acceptors to emit holes than for neutral donors to emit electrons (and the same for the time-reversed processes). Hence, we hypothesize that the acceptor states may be closer to the BN valence band than the donor states are to the BN conduction band. Since the rings, the brightest dots, and darkest dots never change under tip pulses, we can infer that they are in the top layers of BN and are in direct

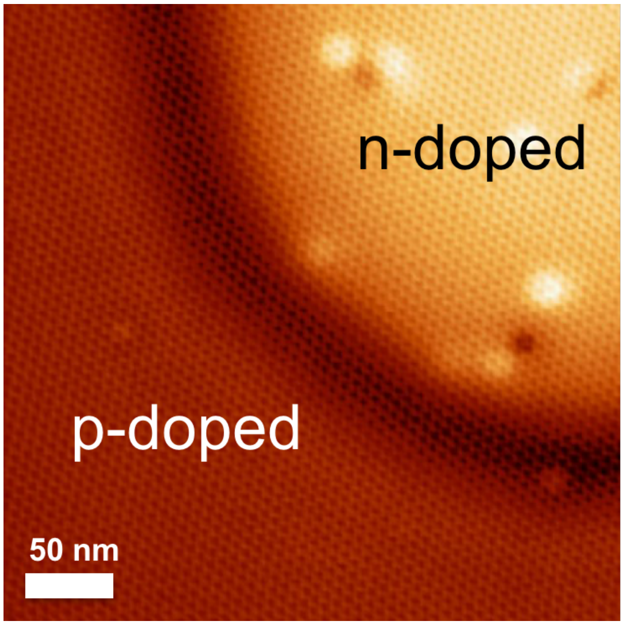

Figure 7.7: dI/dV map of a quadrant of a circular graphene p-n junction created by applying a $V_s = 5$ V pulse for 10 seconds at the top-right corner of the image. The tip was retracted 1.3 nm away from the graphene surface after opening the STM feedback loop, and the gate voltage applied during the pulse was $V_g = -20$ V. Tunneling parameters: $V_s = -0.25$ V, $I = 0.4$ nA, $V_g = -10$ V.

electrical contact with the graphene. The charge states of these unchanging defects depend only on graphene's local chemical potential and show no metastability or hysteresis with the electric field. The defects that switch charge states must be in lower BN layers, out of direct electrical contact with graphene.

7.4.2 Fabrication of p-n Junctions

The defect manipulation discussed in Sec. 7.4.1 does not lead to a net charge transfer between graphene and BN. However, if we use the Si backgate to impose a uniform electric field in the BN, a net charge could accumulate at BN defects to screen the gate electric field. The backgate electric field would also guide charge carriers towards graphene. Hence, we modify the procedure outlined in Sec. 7.4.1 to include a nonzero gate voltage during the tip pulse:

1. Set V_s between -1 V and 1 V while the feedback loop is closed. We typically use $V_s = -0.5$ V, $I = 0.5$ nA.

2. Set $V_g \neq 0$ V.

3. Open the STM feedback loop.

4. Withdraw the tip by a distance between 0.5 nm and 2.5 nm. You can choose to retract the tip by a distance that is less than 0.5 nm, but the probability of blasting a hole in the graphene sheet increases as the tip-sample distance decreases.
5. Increase V_s to 5 V.
6. Wait. We usually wait 10 – 60 seconds.
7. Decrease V_s back to the original value.
8. Close the STM feedback loop.

As I will discuss below, this procedure creates a nanoscale circular graphene p-n junction. More information about the tip-pulse procedure can be found in Refs. [79], [115], and [116]. The p-n junction can be erased by another tip pulse performed while holding $V_g = 0$ V. Alternatively, exposing the graphene device to visible or ultraviolet light also erases the p-n junction.

Fig. 7.7 is a dI/dV map that shows the result of applying this procedure with the tip located at the top-right corner. The gate voltage applied during the tip pulse was $V_g = -20$ V. The tip voltage pulse created a circular p-n junction in graphene with greater electron density in the interior than in the exterior. In this map (obtained with $V_g = -10$ V during data acquisition), the brighter region is n-doped graphene. It is surrounded by a dark circular boundary that is approximately charge neutral. The region outside of the dark circle is p-doped graphene.

Note that step 2 is an important part of this procedure. If $V_g < 0$ V during the tip voltage pulse, a circular graphene p-n junction with an n-doped interior is created. If $V_g > 0$ V during the tip pulse, we get a circular p-n junction with a p-doped interior. The distance in step 4 and the wait time in step 6 determine the magnitude of the doping. Longer wait times give rise to larger changes in the charge carrier density.

Suppose that the charge carrier concentration n in graphene is a function of V_g given by $n(V_g) = CV_g + n_0$, where C is proportional to the capacitance of the graphene device, and n_0 is a residual doping. After a sufficiently long wait time in step 6, the change in the carrier density at the center of the p-n junction will be $-CV_g^*$, where V_g^* is the gate voltage applied during the tip pulse. The carrier density at the center of the p-n junction will then be $n(V_g) = C(V_g - V_g^*) + n_0$, but the density far away from the p-n junction will remain $n(V_g) = CV_g + n_0$.

The creation of this tip-pulse-induced charge distribution can be explained by the field ionization of BN defects together with charge carrier movement within the BN insulator. The intense electric field of the tip pulse penetrates through the graphene into the BN below, causing a strong electric potential gradient around BN defects that results in defect field emission. If $V_g < 0$ V during a tip voltage pulse, the gate electric field causes electrons that have been released from defects to drift into the graphene, leaving behind a positive space charge layer in the BN. Likewise, if $V_g > 0$ V during a tip pulse, the gate electric field causes released holes to drift into the graphene, leaving behind a negative space charge layer in the BN. The space charge layer in the BN locally gates the graphene above it, changing the graphene charge carrier density and producing a p-n junction in the graphene. The net result is that both n-type and p-type doping profiles can be written into graphene with a spatial resolution determined by the electric potential gradient surrounding the STM tip.

7.5 Conclusion

In conclusion, we used STM/STS to image and manipulate point defects inside an insulating BN substrate underneath monolayer graphene. We identified the charge states of individual defects, and we extracted energy-level information for one type of negatively charged defect. Using voltage pulses applied to our STM tip, we were able to ionize, neutralize, and even switch the charge states of BN defects. By applying both a nonzero gate voltage and an STM tip voltage pulse, we created circular p-n junctions in graphene. In the next chapter (Ch. 8), I will discuss an experiment where we imaged the energy eigenstates of an ultra-relativistic quantum harmonic oscillator in these graphene p-n junctions.

As a side note, the technique developed in Sec. 7.4.2 for realizing artificial doping patterns is essentially a variant of charge-trap flash technology. In charge-trap flash devices, an electric pulse from a control gate causes electron accumulation in a charge-trapping layer (a dielectric, usually Si_3N_4), which in turn gates a metal-oxide-semiconductor field-effect transistor (MOSFET). We simply replaced the MOSFET with graphene, used the STM tip as the control gate, and deployed BN as the charge-trapping layer.

Chapter 8

The Relativistic Quantum Harmonic Oscillator

This chapter describes our work on the electronic response of massless Dirac fermions to a harmonic oscillator potential. The results presented here are based on Ref. [116]: J. Lee *et al.*, "Imaging electrostatically confined Dirac fermions in graphene quantum dots", Nature Physics **12**, 1032–1036 (2016).

Our work draws inspiration from Y. Zhao *et al.* [22] from Joseph A. Stroscio's group. They studied "whispering-gallery modes" (WGMs) caused by electron confinement in the electrostatic potential produced by their STM tip. However, they could not image the WGMs because their confinement potential moves with the tip. I had a pleasant chat with Dr. Stroscio at the American Physical Society March Meeting in San Antonio about these WGMs. At the time, Dr. Stroscio did not know that we could create circular graphene p-n junctions, nor did we detect signatures of WGMs in our p-n junctions. Dr. Stroscio told me that we were unlikely to see WGMs in our data because of broadening due to phonon-assisted inelastic tunneling, which immediately gave me the idea to look at the elastic channel within the phonon gap in graphene dI/dV spectroscopy. This is the reason that the data in this chapter focuses on the bias range $-0.1\text{ V} \leq V_s \leq 0.1\text{ V}$. I thank Dr. Stroscio for that very useful discussion, which ultimately enabled us to image the WGMs.

I also want to give a shout-out to C. Gutiérrez *et al.* [117] from Abhay N. Pasupathy's group, who investigated a closely related problem: Klein tunneling at sharp graphene p-n junctions. The primary difference between their study and ours is that their potentials are approximately step functions, while we have potentials that are roughly quadratic. Markus Morgenstern [118] and Lin He (at Beijing Normal University) have similar near-contemporaneous work as well.

Additionally, if you look at Figs. 2.4 and 2.5 of Yang Wang's PhD dissertation [18], you can see WGMs in the data. Yang and I, upon a suggestion from Mikhail E. Portnoi [119] back in 2011, briefly entertained the idea that the features in the data were quasi-bound states induced by the tip potential. We discarded that hypothesis because we were enamored with the idea that they were "plasmarons" [120–122]. Too bad for us... Sad!

8.1 Introduction

Quantum confinement provides a pathway to access and exploit the novel physical properties of many nanoscale materials. Since graphene holds great promise as a component in future electronic devices, it is important to devise methods for confining charge carriers in graphene. Graphene electrons have been localized in high magnetic fields [123], lithographically patterned structures [124–126], and chemically synthesized quantum dots [127–130]. An alternate route to trapping graphene's charge carriers is through electrostatic confinement, which is notoriously difficult because electric potentials are transparent to massless Dirac fermions at normal incidence [6–10, 131]. Although Klein tunneling allows Dirac fermions to penetrate potential barriers, several experimental and theoretical studies have demonstrated that circular graphene p-n junctions can still trap Dirac fermions [22, 132–136]. Spatial characterization of such confined relativistic states, however, has remained elusive.

In Sec. 7.4.2, we discussed a new technique for patterning circular p-n junctions in graphene/BN heterostructures. In an experiment described in this chapter, we spatially mapped the local electronic structure of these circular graphene p-n junctions. We found that these engineered p-n junctions form highly tunable cavities that strongly influence the motion of graphene's relativistic electrons. Within these p-n junctions, we observed energy levels corresponding to quasi-bound states, and we were able to directly visualize the quantum interference patterns of electrostatically confined electrons. Outside of the p-n junctions, Dirac fermions scatter off the cavity boundary, forming Friedel oscillations. We explained our findings by comparing our experimental data to a theoretical model for ultra-relativistic particles in a harmonic oscillator potential.

8.2 Electronic Structure of a Relativistic Harmonic Oscillator

8.2.1 Experimental Eigenstate Distribution

We fabricated graphene/BN devices using the transfer technique outlined in Ch. 4. Graphene monolayers were mechanically exfoliated from graphite and then subsequently transferred onto BN flakes (60 – 100 nm thickness) peeled onto SiO_2/Si chips. The graphene flakes were electrically grounded via Ti (10 nm)/Au (100 nm) electrodes deposited through electron-beam evaporation using a stencil mask. Samples were then annealed in Ar/H_2 forming gas at 350°C and in UHV at 200°C – 400°C.

STM measurements were carried out at temperature $T = 4.8$ K with PtIr tips calibrated against the surface state of Au(111) [32]. STS measurements were performed by lock-in detection of the a.c. tunnel current induced by a modulated voltage (1 – 8 mV at 613.7 Hz) added to V_s. We acquired $dI/dV(V_g, V_s)$ and $dI/dV(r, V_s)$ measurements by sweeping V_s (starting from a fixed set of initial tunneling parameters) and then incrementing the gate

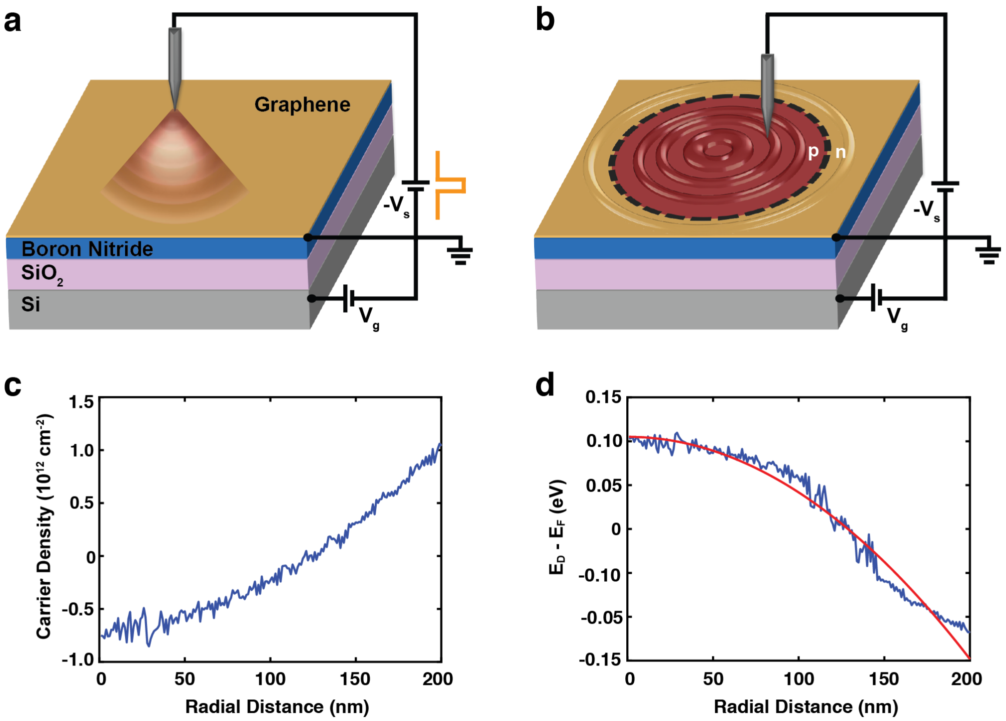

Figure 8.1: Creation and characterization of p-n junction. (a) Schematic diagram of the fabrication of a local embedded gate in a graphene/BN heterostructure. A voltage pulse $V_s = 5$ V is applied to the STM tip while V_g is held at a nonzero value. This creates a space charge in the BN layer, which gives rise to a circular p-n junction in graphene. (b) The STM probes the electronic eigenstates of the p-n junction. (c) Carrier density $n(r)$ as a function of the distance r from the center of a p-n junction fabricated using the technique sketched in (a-b). (d) The Dirac point energy E_D as a function of r. The red curve is a fit to the quadratic polynomial $E_D(r) = -\kappa r^2 + \text{constant}$, with the curvature $\kappa \approx 6 \times 10^{-3}$ meV/nm^2.

voltage V_g or the distance to the p-n junction center r. All figures containing d^2I/dV^2 plots were obtained by taking the simple difference between consecutive dI/dV values.

The procedure for fabricating circular graphene p-n junctions was explained in detail in Sec. 7.4.2. To summarize, we start with a graphene/BN heterostructure on an SiO$_2$/Si chip. The heavily doped Si layer is used as a global backgate while the BN layer acts as a tunable local embedded gate after being treated by a tip voltage pulse. To create a local embedded gate in the BN layer, the STM tip is first retracted by a distance Δz (approximately 2 nm) away from tunneling regime. A $V_s = 5$ V voltage pulse of duration Δt (usually about a minute) is then applied to the tip while simultaneously holding the backgate voltage at $V_g = V_g^*$ (for example, 40 V). The voltage pulse ionizes defects in the BN region directly underneath the tip, and the released charge is guided by the backgate electric field to the graphene. The charge migration leaves behind a space charge layer in the BN that screens the backgate electric field and behaves like a local embedded gate underneath graphene.

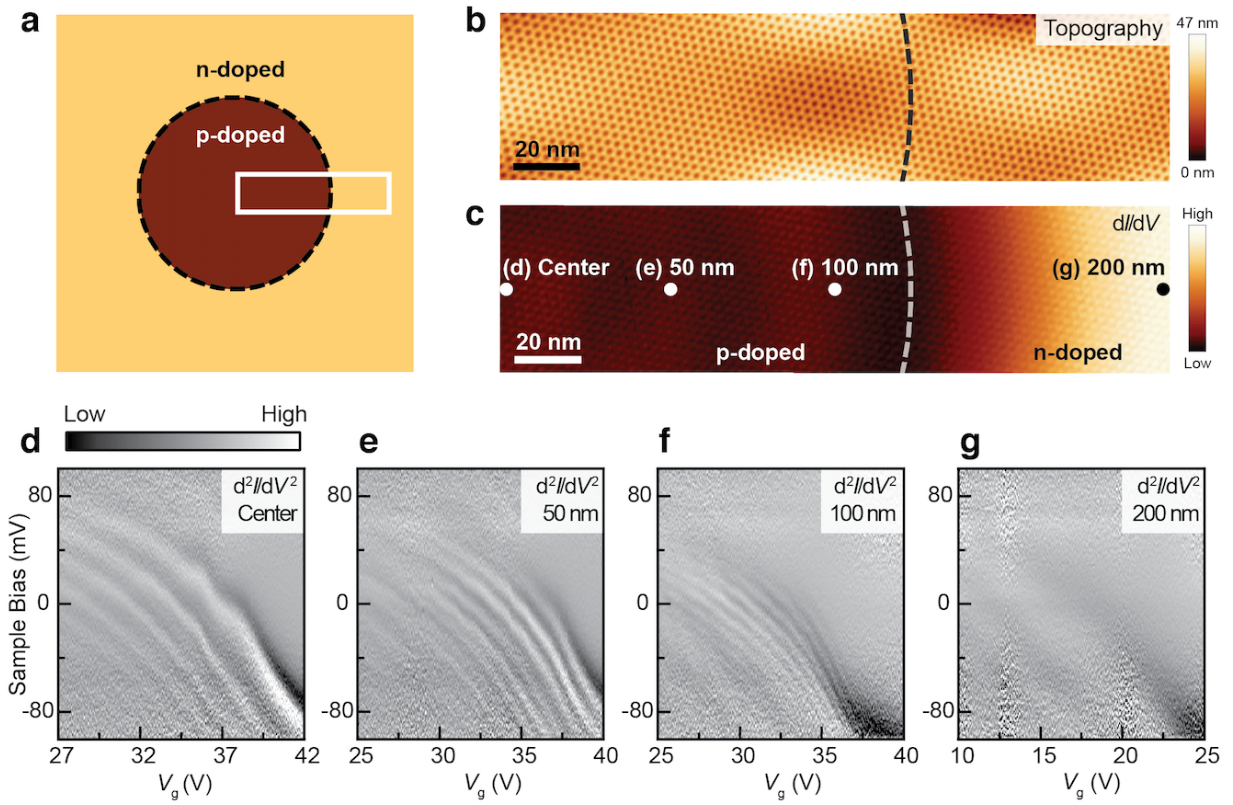

Figure 8.2: Gate-dependent electronic structure of circular graphene p-n junction. (a) Schematic diagram of a circular graphene p-n junction. The white rectangle indicates the measurement region. (b) STM topographic image of the region sketched in (a). (c) dI/dV map of the same region as shown in (b). Dashed lines are placed near the p-n junction boundary in (b-c). Tunneling parameters: $V_s = -0.25$ V, $I = 0.5$ nA, $V_g = 30$ V, 6 mV rms a.c. modulation added to V_s. (d-g) $d^2I/dV^2(V_g, V_s)$ measured at different distances away from the p-n junction center, as indicated in (c). Initial tunneling parameters: $V_s = -0.1$ V, $I = 1.5$ nA, 1 mV rms a.c. modulation. The observed resonances vary in energy E roughly according to the expected graphene dispersion $E \propto \sqrt{|V_g - V_{\mathrm{CNP}}|} + \text{constant}$. The energy spacing between resonances is seen to be larger at the center (d) than further out (e-f), and the resonances disappear outside the p-n junction boundary (g).

This process is depicted in the cartoon in Figs. 8.1a-b.

We applied the above procedure with $V_g^* = 40$ V, $\Delta z = 2$ nm, and $\Delta t = 1$ minute. Figs. 8.1c-d show plots of the resulting charge carrier density $n(r)$ and Dirac point energy $E_D(r)$ as a function of the distance r from the p-n junction center. The values of $n(r)$ and $E_D(r)$ were extracted from dI/dV curves. The Dirac point energy $E_D(r)$ (which represents the electric potential $U(r)$ in the graphene) appears roughly parabolic, so we fit $E_D(r)$ to a quadratic polynomial $E_D(r) = -\kappa r^2 + E_0$ (red curve in Fig. 8.1d). We determined the curvature $\kappa = 6.33 \times 10^{-3}$ meV/nm^2 and constant offset $E_0 = 1.05 \times 10^{-1}$ eV.

To study the local electronic structure of the p-n junction, we examined a rectangular sector of it as schematically represented in Fig. 8.2a. Fig. 8.2b shows an STM topographic image of the rectangular sector. The graphene surface is clean and has a 2.8 nm moiré pattern [33, 34]. A dI/dV map of the same region shown in Fig. 8.2c reflects changes in the LDOS caused by the spatially varying charge density distribution. To gain more direct information about the LDOS, we obtained $d^2I/dV^2(V_g, V_s)$ at four spatial locations at different distances away from the p-n junction center. This data is presented in Figs. 8.2d-g, and the four locations are marked in Fig. 8.2c. We plotted the derivative of $dI/dV(V_g, V_s)$ with respect to V_s in order to enhance the most important features: quasi-periodic resonances that disperse to lower energies with increasing V_g (the original dI/dV data can be found in the Supplementary Information of Ref. [116]). The energies E of the resonances roughly evolve as $E \propto \sqrt{|V_g - V_{\text{CNP}}|}$ + constant (where V_{CNP} is the local charge neutrality point), which is in accordance to the expected behavior of a 2D material with a linear band structure. The energy spacing between resonances ΔE decreases as the STM probes locations further from the p-n junction center; ΔE is 29 ± 2 mV at the center, 16 ± 2 mV at 50 nm from the center, and 13 ± 2 mV at 100 nm from the center for measurements performed at $V_g = 32$ V. The resonances disappear outside of the p-n junction in Fig. 8.2g.

Fig. 8.3a depicts a spatially resolved mapping of the energy-dependent eigenstate distribution for a fixed gate voltage $V_g = 32$ V. The data is plotted as $d^2I/dV^2(r, V_s)$ (see the Supplementary Information of Ref. [116] for the undifferentiated dI/dV data). The energy level structure and complex interior nodal patterns are clear and obvious. There are horizontal features enclosed in a parabolic dome, reminiscent of the solutions of the Schrödinger equation for the simple harmonic oscillator.

8.2.2 Theoretical Eigenstate Distribution

We can explain these observations by considering the response of massless Dirac fermions to a circular electrostatic potential. Klein tunneling causes normally incident Dirac quasiparticles to penetrate potential barriers with unity transmission. However, quasiparticles with large angles of incidence are reflected by potential barriers [6, 9, 10]. Therefore, inside a potential well with circular symmetry, charge carriers with large angular momenta (which are obliquely incident to the potential barrier) are reflected inwards, resulting in electron confinement and the formation of quasi-bound resonances [132–136]. As angular momentum is increased, quasiparticles are repelled from the center of the potential by the centrifugal force, leading to an increase in the number of resonances that should be observable in dI/dV spectroscopy measured away from the center [137]. This is consistent with Figs. 8.2d-f, in which the energy spacing ΔE at the center is approximately double the energy spacing at 100 nm away from the center.

We can test this qualitative explanation by comparing our data to theoretical calculations performed by Joaquin F. Rodriguez-Nieva and Leonid S. Levitov. Our theory collaborators

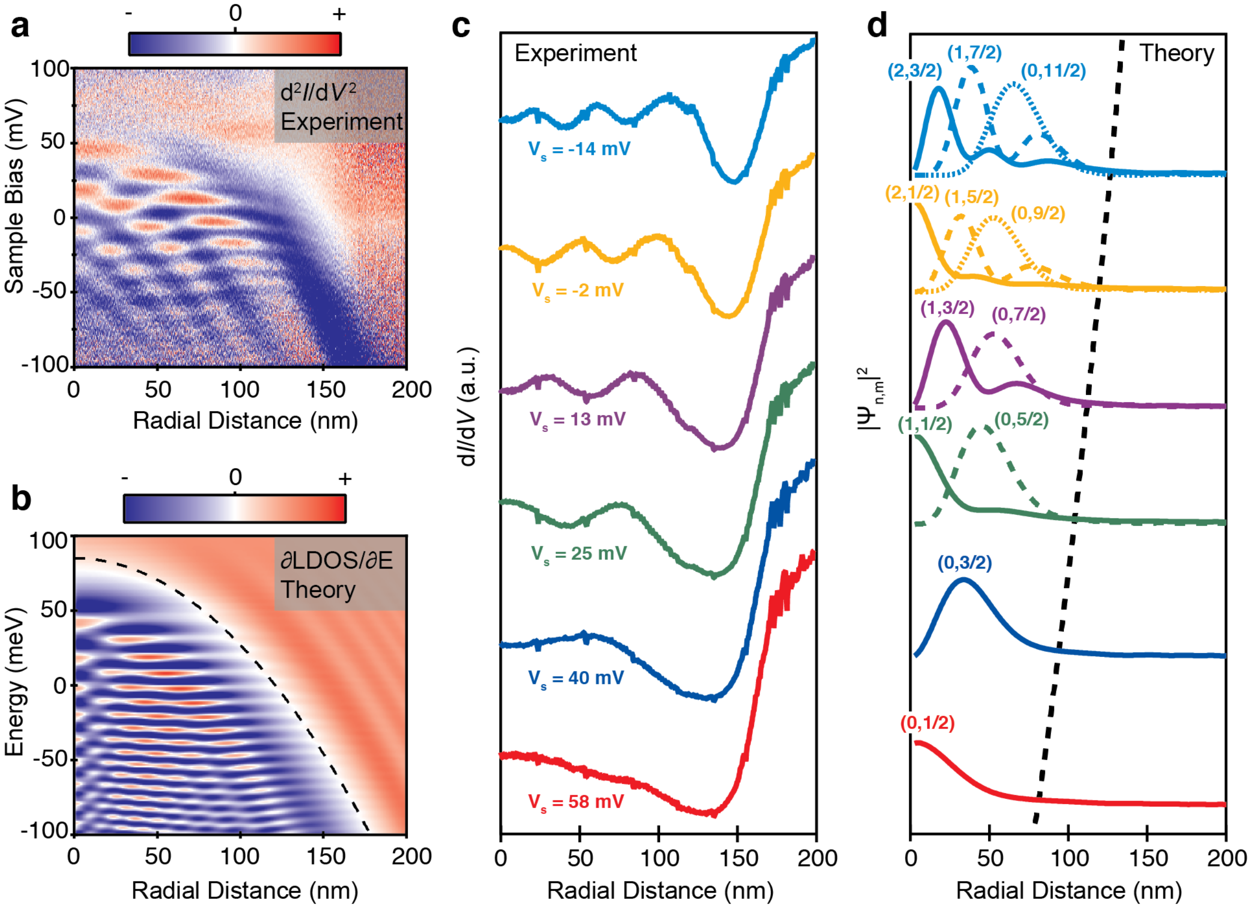

Figure 8.3: Spatially resolving harmonic oscillator energy levels. (a) $d^2I/dV^2(r, V_s)$ for a circular graphene p-n junction. The measurement was performed at a fixed gate voltage. Initial tunneling parameters: $V_s = -0.1$ V, $I = 1.5$ nA, $V_g = 32$ V, 1 mV rms a.c. modulation added to V_s. (b), Theoretical $\partial\text{LDOS}/\partial E$ as a function of energy E and radial distance r for a potential $U(r) = -\kappa r^2 + \text{constant}$ (shown as dashed line). (c) Experimental dI/dV radial line cuts at different V_s values for $V_g = 32$ V. (d) Theoretical $|\Psi_{n,m}|^2$ as a function of radial distance for quasi-bound states labeled by principal and angular momentum quantum numbers (n, m). Each theoretical curve has been vertically displaced by a quantity proportional to V_s for the correspondingly colored experimental curve in (c) to ensure that the black dashed line denotes the classical turning points of the potential $U(r)$.

solved the 2D massless Dirac equation with a quadratic potential:

$$-i\hbar v_F \boldsymbol{\sigma} \cdot \nabla\Psi - \kappa r^2 \Psi = E\Psi. \tag{8.1}$$

Note that, because of an argument presented in Sec. 1.2.3, solving the Dirac equation for $U(r) = -\kappa r^2$ is completely equivalent to solving the equation for $U(r) = \kappa r^2$. We used $\kappa = 6\times 10^{-3}$ meV/nm^2 for the curvature of the potential because this was the value extracted from the polynomial fit in Fig. 8.1d. The characteristic energy scale of this equation is $E^* = (\hbar^2 v_F^2 \kappa)^{1/3} \approx 15$ meV, and the characteristic length scale is $r^* = \hbar v_F/E^* \approx 50$ nm.

By assuming the ansatz in Eq. 1.32, we can obtain a radial Dirac equation that is only a function of the coordinate r:

$$\begin{pmatrix} \dfrac{-\kappa r^2 - E_{n,m}}{\hbar v_F} & \dfrac{\partial}{\partial r} + \dfrac{m}{r} \\ -\dfrac{\partial}{\partial r} + \dfrac{m}{r} & \dfrac{-\kappa r^2 - E_{n,m}}{\hbar v_F} \end{pmatrix} u_{n,m}(r) = 0, \tag{8.2}$$

where n is a principal quantum number, and m is an angular momentum quantum number. The LDOS is given by

$$\mathrm{LDOS}(E, r) = \sum_{n,m} \langle \left|u_{n,m}(r)\right|^2 \rangle_\lambda \delta(E - E_{n,m}), \tag{8.3}$$

$$\langle \left|u_{n,m}(r)\right|^2 \rangle_\lambda = \int_0^\infty dr' \left|u_{n,m}(r')\right|^2 e^{-(r-r')^2/2\lambda}. \tag{8.4}$$

Here, n represents quasi-bound and unbounded states. The LDOS was spatially broadened by $\lambda = 0.01r^*$.

The radial Dirac equation was solved through the finite difference method discretized on 1200 lattice sites in the interval $0 < r < L$. A large repulsive potential at $r = L = 12r^*$ was imposed as the boundary condition, and spurious states localized at $r = 0$ and $r = L$ were excluded. The sum in Eq. 8.3 was carried out for angular momenta in the range $-401/2 \le m \le 401/2$, and the delta function was approximated as a Lorentzian function of width $0.3E^*$ (this is sufficiently small such that the intrinsic quasiparticle lifetimes due to Klein tunneling are preserved).

The result of this calculation is shown in Fig. 8.3b, which plots $\partial\mathrm{LDOS}/\partial E$ (the energy derivative of the LDOS) with respect to the energy E and distance r. The theoretical eigenstate distribution in Fig. 8.3b closely resembles the experimental eigenstate distribution in Fig. 8.3a. Both theory and experiment have a characteristic parabolic envelope due to the confinement potential as well as a complex set of interior nodal patterns.

We can gain more insight about the quasi-bound states inside the p-n junction by comparing constant-energy experimental dI/dV line cuts (shown in Fig. 8.3c) to the theoretically simulated harmonic oscillator wavefunctions ($|\Psi_{n,m}|^2$ shown in Fig. 8.3d). The principal quantum numbers n are nonnegative integers, and the angular momentum quantum numbers m are half-integers because they represent a sum of the pseudospin and the orbital angular momentum (which are not separately conserved by the Dirac Hamiltonian – see Sec. 3.3.3).

We note two important properties of the energy eigenstates. First, although each wavefunction probability distribution $|\Psi_{n,m}|^2$ features $n+1$ maxima, most of its probability weight is concentrated in the first maximum. The position of this maximum is pushed further away from the center of the circular potential for larger magnitudes of the angular momentum $|m|$. Second, for massless Dirac fermions confined in a parabolic potential, we observe nearly perfect energy alignment of the energies $E_{n,m}$ and $E_{n-1,m+2}$ for small n and m, indicating an approximate accidental degeneracy. This degeneracy explains why different resonances originating from different $\Psi_{n,m}$ states form the horizontal rows seen in Figs. 8.3a-b. These two observations allow us to attribute each experimental dI/dV peak in Fig. 8.3c to a different $\Psi_{n,m}$ eigenstate because each state contributes most of its spectral weight to a single energy and radial location. This also explains why the energy spacing for the resonances close to the center is nearly twice as large as the energy spacing away from the center in Figs. 8.2d-g: only the lowest angular momentum states ($m = \pm 1/2$) have appreciable spectral weight at the center, while all other $\Psi_{n,m}$ states (for $m \neq \pm 1/2$) have their probability densities away from the center.

8.3 External Friedel Oscillations

If you compare the experimental data in Fig. 8.3a with the theory in Fig. 8.3b, you can see that there are oscillations outside of the parabolic dome in Fig. 8.3b that are absent in Fig. 8.3a. These are Friedel oscillations caused by the outward scattering of inbound Dirac fermions as well as by coherent electron-hole transmutation at the p-n junction (i.e. Klein tunneling). A circular p-n junction with a p-doped interior attracts hole-like carriers and repels electron-like carriers, while a p-n junction that is n-doped in the interior attracts electron-like carriers and repels hole-like carriers. We should therefore observe an external standing wave pattern due to the quantum interference of the repelled quasiparticles. Why do we not see these external Friedel oscillations?

The p-n junction studied in Figs. 8.1, 8.2, and 8.3 was p-doped in the interior and n-doped in the exterior. Fig. 8.4 presents $d^2I/dV^2(r, V_s)$ for various V_g for a circular graphene p-n junction of the opposite polarity, i.e. a p-n junction that was n-doped in the interior and p-doped in the exterior. The p-n junction was created by a $V_s = 5$ V tip voltage pulse performed with $V_g^* = -40$ V, $\Delta z = 1.8$ nm, and $\Delta t = 1$ minute. As is immediately obvious, the $d^2I/dV^2(r, V_s)$ in Fig. 8.4 are flipped upside down compared to that of Fig. 8.3a but are otherwise similar. Notably, however, is the appearance of the external Friedel oscillations. The parabola also has a different shape for different V_g measurements, presumably due to carrier-density-dependent electronic screening.

Figs. 8.5a-b show two dI/dV maps (acquired at two different energies) of a rectangular section of a circular graphene p-n junction. This p-n junction is different from that in Fig. 8.4 but is prepared in a similar manner ($V_s = 5$ V tip voltage pulse with $V_g^* = -40$ V,

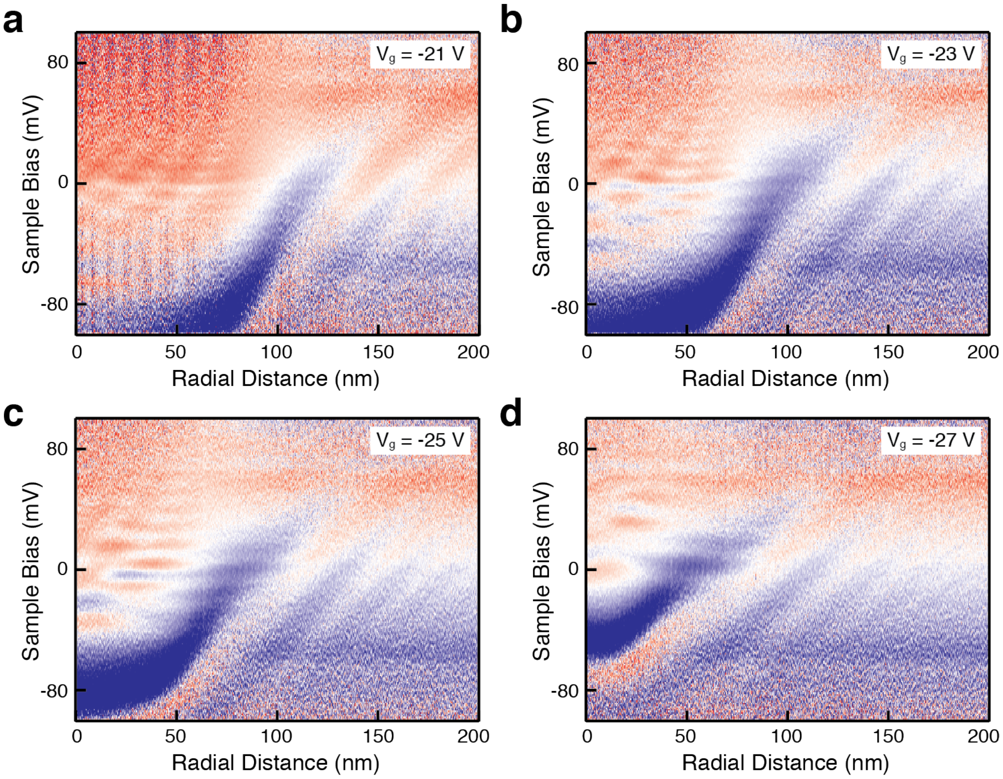

Figure 8.4: Electronic structure of graphene p-n junction with n-doped interior. (a) $d^2I/dV^2(r,V_s)$ for a circular graphene p-n junction that is n-doped at the center and p-doped outside. Initial tunneling parameters: $V_s = -0.1$ V, $I = 1.5$ nA, $V_g = -21$ V, 1 mV rms a.c. modulation added to V_s. (b-d) Same as (a) for $V_g = -23$ V, $V_g = -25$ V, and $V_g = -27$ V, respectively.

$\Delta z = 1.8$ nm, and $\Delta t = 30$ seconds at middle of the left edge of each map). The p-n junction is n-doped at the center and is p-doped outside. Inside the p-n junction, there are circular quantum interference patterns caused by the confinement of Dirac electrons. Outside of the p-n junction boundary, outwardly scattered hole-like quasiparticles form Friedel oscillations.

We do not understand why external Friedel oscillations are observed for p-n junctions that are n-doped in the interior and p-doped in the exterior but are not observed for p-n junctions that are p-doped in the interior and n-doped in the exterior. According to an argument presented in Sec. 1.2.3, for every solution $\Psi(r)$ (with energy eigenvalue E) of the Dirac equation with potential $U(r)$, there exists a solution $\Psi'(r)$ (with energy $-E$) of the Dirac equation with potential $-U(r)$ such that $|\Psi'(r)|^2 = |\Psi(r)|^2$. In principle, the electronic behavior of p-n junctions that are opposite in polarity should be the same. Then, what causes the lack of symmetry and the absence of external Friedel oscillations for p-n junctions with p-doped interiors? One possible factor to consider is that the STM tip is a movable top gate (with $V_{\text{tip}} < 0$ V under the conditions of our measurements – see Sec. 7.3)

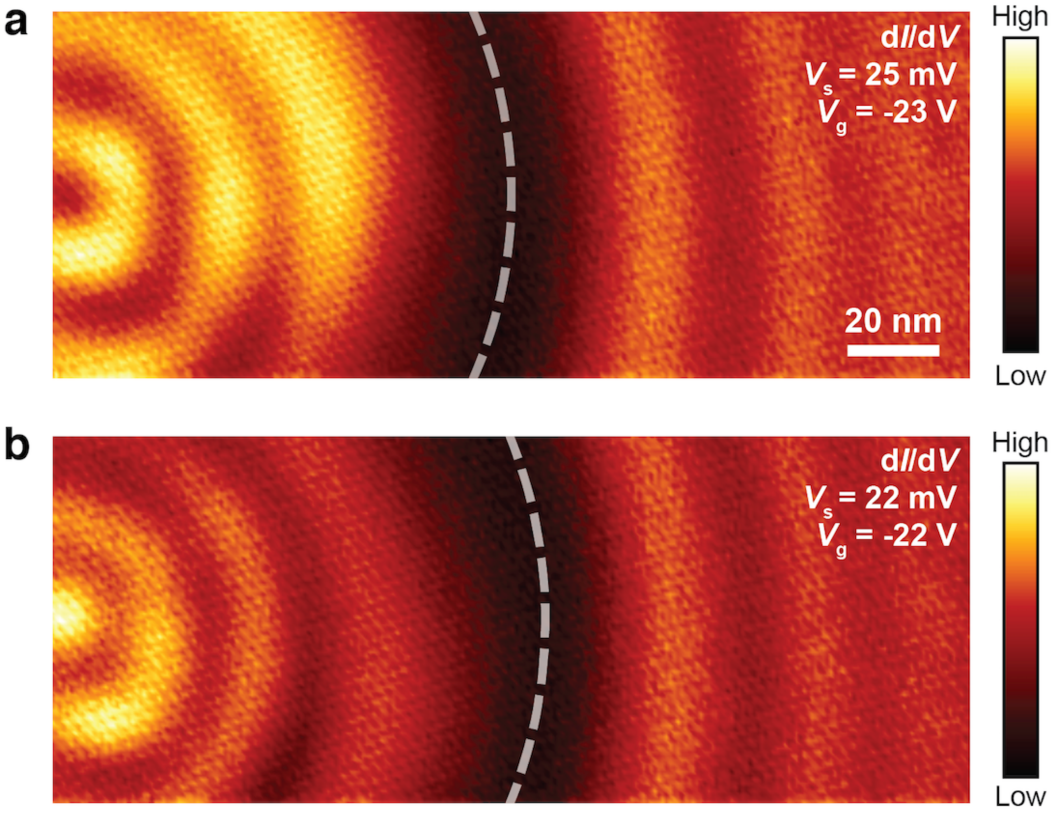

Figure 8.5: dI/dV map of quantum interference in graphene p-n junction. (a) dI/dV map measured for a p-n junction with an n-doped interior. Tunneling parameters: $V_s = 25$ mV, $I = 0.5$ nA, $V_g = -23$ V, 1 mV rms a.c. modulation added to V_s. (b) dI/dV map at the same location as (a) but for a different energy eigenstate showing a different spatial pattern. Tunneling parameters: $V_s = 22$ mV, $I = 0.4$ nA, $V_g = -22$ V, 1 mV rms a.c. modulation. The dark bands (low values of dI/dV) marked by additional dashed lines represent the classical turning points of the potential.

that produces a potential that must be combined with the stationary substrate-induced confinement potential. For a p-n junction with a p-doped interior, the tip potential and the stationary confinement potential are in the same direction (i.e. the tip reduces the carrier density, also p-doping graphene). For a p-n junction with an n-doped interior, the shapes of the tip potential and the stationary confinement potential are in opposite directions (i.e. the tip potential and the stationary confinement potential have curvatures with opposite signs). To better understand the influence of the tip potential on our dI/dV measurements, we would need simulations of graphene with a quadratic potential that include a movable tip.

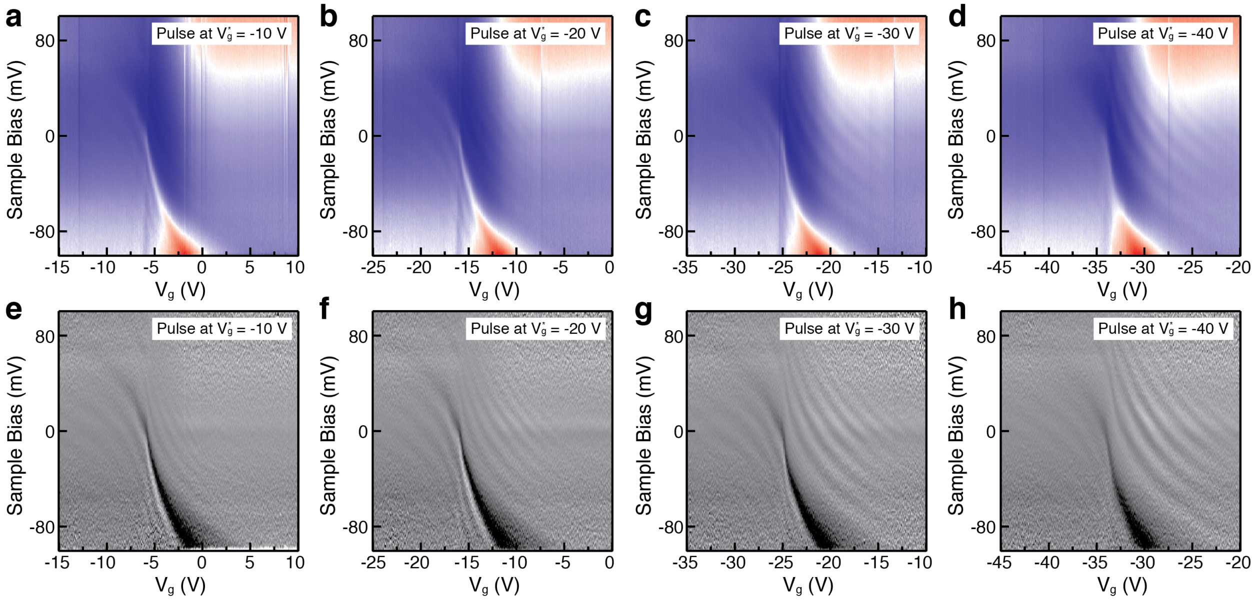

Figure 8.6: Gate-dependent behavior of graphene p-n junction with n-doped interior. (a) $dI/dV(V_g, V_s)$ at the center ($r = 0$ nm) of a circular graphene p-n junction that is n-doped in the interior and p-doped outside. The p-n junction was created by a $V_s = 5$ V tip voltage pulse applied while holding $V_g = V_g^* = -10$ V. Initial tunneling parameters: $V_s = -0.1$ V, $I = 1.5$ nA, 1 mV rms a.c modulation added to V_s. (b-d) Same as (a) except the tip pulse was applied for $V_g^* = -20$ V, $V_g^* = -30$ V, and $V_g^* = -40$ V, respectively. (e-h) $d^2I/dV^2(V_g, V_s)$ obtained by differentiating the plots in (a-d) with respect to V_s.

8.4 The Depth Dependence

The potential wells in our study are not truly quadratic. They have finite depths whose magnitudes can be controlled via the value of the gate voltage V_g^* used during a tip voltage pulse; increasing the magnitude of V_g^* increases the difference between the carrier density at the center and the carrier density far away from the p-n junction. Figs. 8.6a-d show $dI/dV(V_g, V_s)$ at the centers of circular graphene p-n junctions with sequentially larger well depths, and Figs. 8.6e-h show the corresponding $d^2I/dV^2(V_g, V_s)$. All of the p-n junctions are n-doped at their centers. The p-n junction for Figs. 8.6a and 8.6e was created via a $V_s = 5$ V tip pulse with $V_g^* = -10$ V, $\Delta z = 1.7$ nm, and $\Delta t = 10$ seconds. The p-n junction for Figs. 8.6b and 8.6f was created via a $V_s = 5$ V tip pulse with $V_g^* = -20$ V, $\Delta z = 1.68$ nm, and $\Delta t = 60$ seconds. The p-n junction for Figs. 8.6c and 8.6g was created via a $V_s = 5$ V tip pulse with $V_g^* = -30$ V, $\Delta z = 1.68$ nm, and $\Delta t = 60$ seconds. The p-n junction for Figs. 8.6d and 8.6h was created via a $V_s = 5$ V tip pulse with $V_g^* = -20$ V, $\Delta z = 1.68$ nm, and $\Delta t = 60$ seconds. By examining differences between these plots, one can obtain information about how the potential well depth influences the electronic structure of a p-n junction.

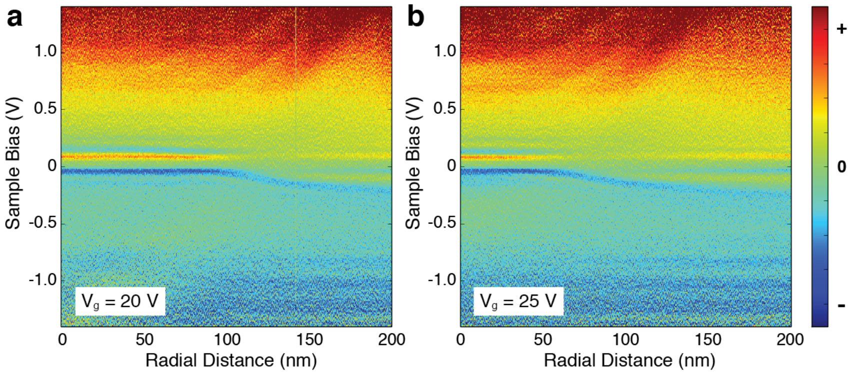

Figure 8.7: Spectroscopic features at higher sample biases. (a) $d^2I/dV^2(r, V_s)$ for a circular graphene p-n junction. The measurement was performed at a fixed gate voltage. The p-n junction was created via a $V_s = 5$ V tip voltage pulse with $V_g^* = 40$ V, $\Delta z = 1.9$ nm, and $\Delta t = 10$ seconds. Initial tunneling parameters: $V_s = -1.4$ V, $I = 1$ nA, $V_g = 20$ V, 8 mV rms a.c. modulation added to V_s. (b) Same as (a) for $V_g = 25$ V.

8.5 The Resonances at Higher Biases

According to the Tersoff-Hamann theory of tunneling (see Sec. 2.1.2), the tunneling current I is given by

$$I \propto \int_{E_F}^{E_F+eV_s} \mathrm{LDOS}(\mathbf{r}, E)dE. \tag{8.5}$$

Naively then, $dI/dV \propto \mathrm{LDOS}(\mathbf{r}, E_F + eV_s)$. This is not completely correct for a gate-tunable material because the chemical potential E_F can also depend on V_s through the effect of the tip as a top gate. Differentiating Eq. 8.5 with respect to V_s yields

$$\frac{dI}{dV} \propto \left(1 + \frac{1}{e}\frac{\partial E_F}{\partial V_s}\right) \mathrm{LDOS}(\mathbf{r}, E_F + eV_s) - \left(\frac{1}{e}\frac{\partial E_F}{\partial V_s}\right) \mathrm{LDOS}(\mathbf{r}, E_F). \tag{8.6}$$

The above equation implies that each localized state (with energy E) contributes to a dI/dV spectrum twice: once when (i) $E = E_F + eV_s$ and another time when (ii) $E = E_F$. Scenario (i) corresponds to the quasi-bound state resonances seen throughout the $-0.1 \text{ V} \leq V_s \leq 0.1$ V measurements in Secs. 8.2, 8.3, and 8.4. Fig. 8.7 shows features due to scenario (ii) in $d^2I/dV^2(r, V_s)$ plots obtained for the sample bias range $-1.4 \text{ V} \leq V_s \leq 1.4$ V. The features appear at the top of Figs. 8.7a-b and are extremely faint. If you are having trouble seeing the features, please hold this dissertation at least a meter from your eyes. Since the scenario (ii) features are caused by electrostatic gating from the tip, the signal intensity of these features strongly depends on the geometry and physical condition of the tip. On very rare occasions, the scenario (ii) features will be prominently visible in $d^2I/dV^2(V_g, V_s)$ at the same V_s values as the scenario (i) features, leading to apparent Coulomb diamonds.

8.6 Conclusion

In summary, we directly imaged the energy eigenstates of an ultra-relativistic quantum harmonic oscillator by fabricating circular graphene p-n junctions and performing STM/STS measurements of their local electronic structure. These circular p-n junctions electrostatically confined graphene's massless Dirac fermions, giving rise to quasi-bound states that we identified with a set of principal and angular momentum quantum numbers. We also observed Friedel oscillations caused by the outward scattering of inbound Dirac quasiparticles. The p-n junctions that we created were effectively highly tunable quantum dots for one type of graphene charge carrier and quantum antidots for the other type of carrier.

This concludes my dissertation on using STM to manipulate charged impurities in gate-tunable graphene/BN heterostructures for the purpose of creating electrostatic potentials that control the motion of charge carriers in graphene. We applied the methods developed and described herein to image the quantum wavefunctions of massless Dirac fermions in Coulomb and harmonic oscillator potentials.

Bibliography

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, "Electric field effect in atomically thin carbon films", Science **306**, 666–669 (2004).

[2] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, "Experimental observation of the quantum Hall effect and Berry's phase in graphene", Nature **438**, 201–204 (2005).

[3] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, "Two-dimensional gas of massless Dirac fermions in graphene", Nature **438**, 197–200 (2005).

[4] P. R. Wallace, "The band theory of graphite", Physical Review **71**, 622–634 (1947).

[5] A. H. Castro Neto, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene", Reviews of Modern Physics **81**, 109–162 (2009).

[6] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim, "Chiral tunnelling and the Klein paradox in graphene", Nature Physics **2**, 620–625 (2006).

[7] A. F. Young and P. Kim, "Quantum interference and Klein tunnelling in graphene heterojunctions", Nature Physics **5**, 222–226 (2009).

[8] N. Stander, B. Huard, and D. Goldhaber-Gordon, "Evidence for Klein tunneling in graphene p-n junctions", Physical Review Letters **102**, 026807 (2009).

[9] A. V. Shytov, M. S. Rudner, and L. S. Levitov, "Klein backscattering and Fabry-Pérot interference in graphene heterojunctions", Physical Review Letters **101**, 156804 (2008).

[10] V. V. Cheianov and V. I. Fal'ko, "Selective transmission of Dirac electrons and ballistic magnetoresistance of n-p junctions in graphene", Physical Review B **74**, 041403 (2006).

[11] L. M. Zhang and M. M. Fogler, "Nonlinear screening and ballistic transport in a graphene p-n junction", Physical Review Letters **100**, 116804 (2008).

[12] M. I. Katsnelson, "Zitterbewegung, chirality, and minimal conductivity in graphene", The European Physical Journal B - Condensed Matter and Complex Systems **51**, 157–160 (2006).

[13] V. M. Pereira, J. Nilsson, and A. H. Castro Neto, "Coulomb impurity problem in graphene", Physical Review Letters **99**, 166802 (2007).

[14] A. V. Shytov, M. I. Katsnelson, and L. S. Levitov, "Atomic collapse and quasi-Rydberg states in graphene", Physical Review Letters **99**, 246802 (2007).

[15] A. V. Shytov, M. I. Katsnelson, and L. S. Levitov, "Vacuum polarization and screening of supercritical impurities in graphene", Physical Review Letters **99**, 236801 (2007).

[16] M. M. Fogler, D. S. Novikov, and B. I. Shklovskii, "Screening of a hypercritical charge in graphene", Physical Review B **76**, 233402 (2007).

[17] A. Shytov, M. Rudner, N. Gu, M. Katsnelson, and L. Levitov, "Atomic collapse, Lorentz boosts, Klein scattering, and other quantum-relativistic phenomena in graphene", Solid State Communications **149**, 1087–1093 (2009).

[18] Y. Wang, "Scanning tunneling microscopy study of graphene electronic structures", PhD thesis (University of California, Berkeley, 2014).

[19] J. J. Sakurai, *Modern quantum mechanics* (Addison Wesley, Sept. 1993).

[20] J Reinhardt and W Greiner, "Quantum electrodynamics of strong fields", Reports on Progress in Physics **40**, 219 (1977).

[21] J. Silver, "Heavy-ion physics: giant nucleus at Darmstadt?", Nature **315**, 276–276 (1985).

[22] Y. Zhao, J. Wyrick, F. D. Natterer, J. F. Rodriguez-Nieva, C. Lewandowski, K. Watanabe, T. Taniguchi, L. S. Levitov, N. B. Zhitenev, and J. A. Stroscio, "Creating and probing electron whispering-gallery modes in graphene", Science **348**, 672–675 (2015).

[23] J. Bardeen, "Tunnelling from a many-particle point of view", Physical Review Letters **6**, 57–59 (1961).

[24] J. Tersoff and D. R. Hamann, "Theory and application for the scanning tunneling microscope", Physical Review Letters **50**, 1998–2001 (1983).

[25] J. Tersoff and D. R. Hamann, "Theory of the scanning tunneling microscope", Physical Review B **31**, 805–813 (1985).

[26] R. T. Yamachika, "Probing atomic-scale properties of organic and organometallic molecules by scanning tunneling spectroscopy", PhD thesis (University of California, Berkeley, 2009).

[27] C. J. Chen, *Introduction to scanning tunneling microscopy* (Oxford University Press, USA, 2007).

[28] V. W. Brar, "Scanning tunneling spectroscopy of graphene and magnetic nanostructures", PhD thesis (University of California, Berkeley, 2010).

[29] Y.-C. Chen, "Exploring graphene nanoribbons using scanning probe microscopy and spectroscopy", PhD thesis (University of California, Berkeley, 2014).

[30] J. Li, W.-D. Schneider, and R. Berndt, "Local density of states from spectroscopic scanning-tunneling-microscope images: Ag(111)", Physical Review B **56**, 7656–7659 (1997).

[31] C. Wittneven, R. Dombrowski, M. Morgenstern, and R. Wiesendanger, "Scattering states of ionized dopants probed by low temperature scanning tunneling spectroscopy", Physical Review Letters **81**, 5616–5619 (1998).

[32]W. Chen, V. Madhavan, T. Jamneala, and M. F. Crommie, “Scanning tunneling microscopy observation of an electronic superlattice at the surface of clean gold”, Physical Review Letters **80**, 1469–1472 (1998).

[33]R. Decker, Y. Wang, V. W. Brar, W. Regan, H.-Z. Tsai, Q. Wu, W. Gannett, A. Zettl, and M. F. Crommie, “Local electronic properties of graphene on a BN substrate via scanning tunneling microscopy”, Nano Letters **11**, 2291–2295 (2011).

[34]J. Xue, J. Sanchez-Yamagishi, D. Bulmash, P. Jacquod, A. Deshpande, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, and B. J. LeRoy, “Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride”, Nature Materials **10**, 282–285 (2011).

[35]Y. Zhang, V. W. Brar, F. Wang, C. Girit, Y. Yayon, M. Panlasigui, A. Zettl, and M. F. Crommie, “Giant phonon-induced conductance in scanning tunnelling spectroscopy of gate-tunable graphene”, Nature Physics **4**, 627–630 (2008).

[36]V. W. Brar, S. Wickenburg, M. Panlasigui, C.-H. Park, T. O. Wehling, Y. Zhang, R. Decker, Ç. Girit, A. V. Balatsky, S. G. Louie, A. Zettl, and M. F. Crommie, “Observation of carrier-density-dependent many-body effects in graphene via tunneling spectroscopy”, Physical Review Letters **104**, 036805 (2010).

[37]H. S. Jung, H.-Z. Tsai, D. Wong, C. Germany, S. Kahn, Y. Kim, A. S. Aikawa, D. K. Desai, G. F. Rodgers, A. J. Bradley, J. Velasco Jr., K. Watanabe, T. Taniguchi, F. Wang, A. Zettl, and M. F. Crommie, “Fabrication of gate-tunable graphene devices for scanning tunneling microscopy studies with Coulomb impurities”, Journal of Visualized Experiments, e52711 (2015).

[38]C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, and J. Hone, “Boron nitride substrates for high-quality graphene electronics”, Nature Nanotechnology **5**, 722–726 (2010).

[39]P. J. Zomer, S. P. Dash, N. Tombros, and B. J. van Wees, “A transfer technique for high mobility graphene devices on commercially available hexagonal boron nitride”, Applied Physics Letters **99**, 232104 (2011).

[40]G. Li, A. Luican, and E. Y. Andrei, “Self-navigation of a scanning tunneling microscope tip toward a micron-sized graphene sample”, Review of Scientific Instruments **82**, 073701 (2011).

[41]E. H. Hwang and S. Das Sarma, “Dielectric function, screening, and plasmons in two-dimensional graphene”, Physical Review B **75**, 205418 (2007).

[42]D. A. Siegel, W. Regan, A. V. Fedorov, A. Zettl, and A. Lanzara, “Charge-carrier screening in single-layer graphene”, Physical Review Letters **110**, 146802 (2013).

[43]J. H. Chen, C. Jang, S. Adam, M. S. Fuhrer, E. D. Williams, and M. Ishigami, “Charged-impurity scattering in graphene”, Nature Physics **4**, 377–381 (2008).

[44]K. Pi, K. M. McCreary, W. Bao, W. Han, Y. F. Chiang, Y. Li, S.-W. Tsai, C. N. Lau, and R. K. Kawakami, "Electronic doping and scattering by transition metals on graphene", Physical Review B **80**, 075406 (2009).

[45]K. M. McCreary, K. Pi, A. G. Swartz, W. Han, W. Bao, C. N. Lau, F. Guinea, M. I. Katsnelson, and R. K. Kawakami, "Effect of cluster formation on graphene mobility", Physical Review B **81**, 115453 (2010).

[46]L. Zhao, R. He, K. T. Rim, T. Schiros, K. S. Kim, H. Zhou, C. Gutiérrez, S. P. Chockalingam, C. J. Arguello, L. Pálová, D. Nordlund, M. S. Hybertsen, D. R. Reichman, T. F. Heinz, P. Kim, A. Pinczuk, G. W. Flynn, and A. N. Pasupathy, "Visualizing individual nitrogen dopants in monolayer graphene", Science **333**, 999–1003 (2011).

[47]K. C. Rahnejat, C. A. Howard, N. E. Shuttleworth, S. R. Schofield, K. Iwaya, C. F. Hirjibehedin, C. Renner, G. Aeppli, and M. Ellerby, "Charge density waves in the graphene sheets of the superconductor CaC_2", Nature Communications **2**, 558 (2011).

[48]S. Ichinokura, K. Sugawara, A. Takayama, T. Takahashi, and S. Hasegawa, "Superconducting calcium-intercalated bilayer graphene", ACS Nano **10**, 2761–2765 (2016).

[49]J. Chapman, Y. Su, C. A. Howard, D. Kundys, A. N. Grigorenko, F. Guinea, A. K. Geim, I. V. Grigorieva, and R. R. Nair, "Superconductivity in Ca-doped graphene laminates", Scientific Reports **6**, 23254 (2016).

[50]A. Luican-Mayer, M. Kharitonov, G. Li, C.-P. Lu, I. Skachko, A.-M. B. Gonçalves, K. Watanabe, T. Taniguchi, and E. Y. Andrei, "Screening charged impurities and lifting the orbital degeneracy in graphene by populating landau levels", Physical Review Letters **112**, 036804 (2014).

[51]X. Li, W. Cai, J. An, S. Kim, J. Nah, D. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S. K. Banerjee, L. Colombo, and R. S. Ruoff, "Large-area synthesis of high-quality and uniform graphene films on copper foils", Science **324**, 1312–1314 (2009).

[52]J. Martin, N. Akerman, G. Ulbricht, T. Lohmann, J. H. Smet, K. von Klitzing, and A. Yacoby, "Observation of electron-hole puddles in graphene using a scanning single-electron transistor", Nature Physics **4**, 144–148 (2008).

[53]Y. Zhang, V. W. Brar, C. Girit, A. Zettl, and M. F. Crommie, "Origin of spatial charge inhomogeneity in graphene", Nature Physics **5**, 722–726 (2009).

[54]Y. Wang, V. W. Brar, A. V. Shytov, Q. Wu, W. Regan, H.-Z. Tsai, A. Zettl, L. S. Levitov, and M. F. Crommie, "Mapping Dirac quasiparticles near a single Coulomb impurity on graphene", Nature Physics **8**, 653–657 (2012).

[55]T. Ando, "Screening effect and impurity scattering in monolayer graphene", Journal of the Physical Society of Japan **75**, 074716 (2006).

[56]T. Sohier, M. Calandra, and F. Mauri, "Density-functional calculation of static screening in two-dimensional materials: the long-wavelength dielectric function of graphene", Physical Review B **91**, 165428 (2015).

[57] P. Cudazzo, I. V. Tokatly, and A. Rubio, “Dielectric screening in two-dimensional insulators: implications for excitonic and impurity states in graphane”, Physical Review B **84**, 085406 (2011).

[58] S. Adam and S. Das Sarma, “Boltzmann transport and residual conductivity in bilayer graphene”, Physical Review B **77**, 115436 (2008).

[59] B. Wunsch, T. Stauber, F. Sols, and F. Guinea, “Dynamical polarization of graphene at finite doping”, New Journal of Physics **8**, 318 (2006).

[60] M. I. Katsnelson, “Nonlinear screening of charge impurities in graphene”, Physical Review B **74**, 201401 (2006).

[61] M. Ghaznavi, Z. L. Mišković, and F. O. Goodman, “Nonlinear screening of external charge by doped graphene”, Physical Review B **81**, 085416 (2010).

[62] E. H. Hwang, S. Adam, and S. D. Sarma, “Carrier transport in two-dimensional graphene layers”, Physical Review Letters **98**, 186806 (2007).

[63] K. Nomura and A. H. MacDonald, “Quantum transport of massless Dirac fermions”, Physical Review Letters **98**, 076602 (2007).

[64] S. Adam, E. H. Hwang, V. M. Galitski, and S. Das Sarma, “A self-consistent theory for graphene transport”, Proceedings of the National Academy of Sciences **104**, 18392–18397 (2007).

[65] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, “Electronic transport in two-dimensional graphene”, Reviews of Modern Physics **83**, 407–470 (2011).

[66] H.-Z. Tsai, A. A. Omrani, S. Coh, H. Oh, S. Wickenburg, Y.-W. Son, D. Wong, A. Riss, H. S. Jung, G. D. Nguyen, G. F. Rodgers, A. S. Aikawa, T. Taniguchi, K. Watanabe, A. Zettl, S. G. Louie, J. Lu, M. L. Cohen, and M. F. Crommie, “Molecular self-assembly in a poorly screened environment: F_4TCNQ on graphene/BN”, ACS Nano **9**, 12168–12173 (2015).

[67] S. Wickenburg, J. Lu, J. Lischner, H.-Z. Tsai, A. A. Omrani, A. Riss, C. Karrasch, A. Bradley, H. S. Jung, R. Khajeh, D. Wong, K. Watanabe, T. Taniguchi, A. Zettl, A. H. Castro Neto, S. G. Louie, and M. F. Crommie, “Tuning charge and correlation effects for a single molecule on a graphene device”, Nature Communications **7**, 13553 (2016).

[68] M. Yankowitz, J. Xue, D. Cormode, J. D. Sanchez-Yamagishi, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, P. Jacquod, and B. J. LeRoy, “Emergence of superlattice Dirac points in graphene on hexagonal boron nitride”, Nature Physics **8**, 382–386 (2012).

[69] G. Li, A. Luican, J. M. B. Lopes dos Santos, A. H. Castro Neto, A. Reina, J. Kong, and E. Y. Andrei, “Observation of Van Hove singularities in twisted graphene layers”, Nature Physics **6**, 109–113 (2010).

[70]D. Wong, Y. Wang, J. Jung, S. Pezzini, A. M. DaSilva, H.-Z. Tsai, H. S. Jung, R. Khajeh, Y. Kim, J. Lee, S. Kahn, S. Tollabimazraehno, H. Rasool, K. Watanabe, T. Taniguchi, A. Zettl, S. Adam, A. H. MacDonald, and M. F. Crommie, “Local spectroscopy of moiré-induced electronic structure in gate-tunable twisted bilayer graphene”, Physical Review B **92**, 155409 (2015).

[71]Y. Wang, D. Wong, A. V. Shytov, V. W. Brar, S. Choi, Q. Wu, H.-Z. Tsai, W. Regan, A. Zettl, R. K. Kawakami, S. G. Louie, L. S. Levitov, and M. F. Crommie, “Observing atomic collapse resonances in artificial nuclei on graphene”, Science **340**, 734–737 (2013).

[72]T. H. Boyer, “Unfamiliar trajectories for a relativistic particle in a Kepler or Coulomb potential”, American Journal of Physics **72**, 992–997 (2004).

[73]J. Schweppe, A. Gruppe, K. Bethge, H. Bokemeyer, T. Cowan, H. Folger, J. S. Greenberg, H. Grein, S. Ito, R. Schule, D. Schwalm, K. E. Stiebing, N. Trautmann, P. Vincent, and M. Waldschmidt, “Observation of a peak structure in positron spectra from U+Cm collisions”, Physical Review Letters **51**, 2261–2264 (1983).

[74]T. Cowan, H. Backe, M. Begemann, K. Bethge, H. Bokemeyer, H. Folger, J. S. Greenberg, H. Grein, A. Gruppe, Y. Kido, M. Klüver, D. Schwalm, J. Schweppe, K. E. Stiebing, N. Trautmann, and P. Vincent, “Anomalous positron peaks from supercritical collision systems”, Physical Review Letters **54**, 1761–1764 (1985).

[75]S. Hagmann, T. Stöhlker, Y. Litvinov, C. Kozhuharov, P. M. Hillenbrand, U. Spillmann, V. Shabaev, K. Stiebing, M. Lestinsky, A. Surzhykov, A. Voitkiv, B. Franzke, D. Fischer, D. Schneider, D. Jakubassa, A. Artiomov, E. DeFilippo, X. Ma, R. Dörner, and H. Rothard, “Few-body quantum dynamics of high-Z ions studied at the future relativistic high-energy storage ring”, Physica Scripta **2013**, 014086 (2013).

[76]P. M. Hillenbrand, S. Hagmann, T. Stöhlker, Y. Litvinov, C. Kozhuharov, U. Spillmann, V. Shabaev, K. Stiebing, M. Lestinsky, A. Surzhykov, A. Voitkiv, B. Franzke, D. Fischer, C. Brandau, S. Schippers, A. Mueller, D. Schneider, D. Jakubassa, A. Artiomov, E. DeFilippo, X. Ma, R. Dörner, and H. Rothard, “Future experiments using forward electron spectroscopy to study the quantum dynamics of high-Z ions at the ESR/CRYRING storage rings”, Physica Scripta **2013**, 014087 (2013).

[77]J. Mao, Y. Jiang, D. Moldovan, G. Li, K. Watanabe, T. Taniguchi, M. R. Masir, F. M. Peeters, and E. Y. Andrei, “Realization of a tunable artificial atom at a supercritically charged vacancy in graphene”, Nature Physics **12**, 545–549 (2016).

[78]D. Wong, J. Velasco Jr., L. Ju, J. Lee, S. Kahn, H.-Z. Tsai, C. Germany, T. Taniguchi, K. Watanabe, A. Zettl, F. Wang, and M. F. Crommie, “Characterization and manipulation of individual defects in insulating hexagonal boron nitride using scanning tunnelling microscopy”, Nature Nanotechnology **10**, 949–953 (2015).

[79] J. Velasco Jr., L. Ju, D. Wong, S. Kahn, J. Lee, H.-Z. Tsai, C. Germany, S. Wickenburg, J. Lu, T. Taniguchi, K. Watanabe, A. Zettl, F. Wang, and M. F. Crommie, “Nanoscale control of rewriteable doping patterns in pristine graphene/boron nitride heterostructures”, Nano Letters **16**, 1620–1625 (2016).

[80] A. K. Geim and I. V. Grigorieva, “Van der Waals heterostructures”, Nature **499**, 419–425 (2013).

[81] K. Watanabe, T. Taniguchi, and H. Kanda, “Direct-bandgap properties and evidence for ultraviolet lasing of hexagonal boron nitride single crystal”, Nature Materials **3**, 404–409 (2004).

[82] T. Taniguchi and K. Watanabe, “Synthesis of high-purity boron nitride single crystals under high pressure by using Ba-BN solvent”, Journal of Crystal Growth **303**, 525–529 (2007).

[83] M. Fanciulli and T. D. Moustakas, “Study of defects in wide band gap semiconductors by electron paramagnetic resonance”, Physica B: Condensed Matter **185**, 228–233 (1993).

[84] A. Katzir, J. T. Suss, A. Zunger, and A. Halperin, “Point defects in hexagonal boron nitride. I. EPR, thermoluminescence, and thermally-stimulated-current measurements”, Physical Review B **11**, 2370–2377 (1975).

[85] E. Y. Andrei, A. Katzir, and J. T. Suss, “Point defects in hexagonal boron nitride. III. EPR in electron-irradiated BN”, Physical Review B **13**, 2831–2834 (1976).

[86] U. Chandni, K. Watanabe, T. Taniguchi, and J. P. Eisenstein, “Evidence for defect-mediated tunneling in hexagonal boron nitride-based junctions”, Nano Letters **15**, 7329–7333 (2015).

[87] Z. Remes, M. Nesladek, K. Haenen, K. Watanabe, and T. Taniguchi, “The optical absorption and photoconductivity spectra of hexagonal boron nitride single crystals”, physica status solidi (a) **202**, 2229–2233 (2005).

[88] L. Ju, J. Velasco Jr, E. Huang, S. Kahn, C. Nosiglia, H.-Z. Tsai, W. Yang, T. Taniguchi, K. Watanabe, Y. Zhang, G. Zhang, M. Crommie, A. Zettl, and F. Wang, “Photoinduced doping in heterostructures of graphene and boron nitride”, Nature Nanotechnology **9**, 348–352 (2014).

[89] S. Heinze, R. Abt, S. Blügel, G. Gilarowski, and H. Niehus, “Scanning tunneling microscopy images of transition-metal structures buried below noble-metal surfaces”, Physical Review Letters **83**, 4808–4811 (1999).

[90] V. Madhavan, W. Chen, T. Jamneala, M. F. Crommie, and N. S. Wingreen, “Tunneling into a single magnetic atom: spectroscopic evidence of the Kondo resonance”, Science **280**, 567–569 (1998).

[91] R. M. Feenstra, J. M. Woodall, and G. D. Pettit, “Observation of bulk defects by scanning tunneling microscopy and spectroscopy: arsenic antisite defects in GaAs”, Physical Review Letters **71**, 1176–1179 (1993).

[92] K. Teichmann, M. Wenderoth, S. Loth, R. G. Ulbrich, J. K. Garleff, A. P. Wijnheijmer, and P. M. Koenraad, “Controlled charge switching on a single donor with a scanning tunneling microscope”, Physical Review Letters **101**, 076103 (2008).

[93] D. H. Lee and J. A. Gupta, “Tunable field control over the binding energy of single dopants by a charged vacancy in GaAs”, Science **330**, 1807–1810 (2010).

[94] D. Kitchen, A. Richardella, J.-M. Tang, M. E. Flatte, and A. Yazdani, “Atom-by-atom substitution of Mn in GaAs and visualization of their hole-mediated interactions”, Nature **442**, 436–439 (2006).

[95] J. Repp, G. Meyer, F. E. Olsson, and M. Persson, “Controlling the charge state of individual gold adatoms”, Science **305**, 493–495 (2004).

[96] N. A. Pradhan, N. Liu, and W. Ho, “Vibronic spectroscopy of single C_{60} molecules and monolayers with the STM”, The Journal of Physical Chemistry B **109**, 8513–8518 (2005).

[97] P. Avouris and R. Wolkow, “Scanning tunneling microscopy of insulators: CaF_2 epitaxy on Si (111)”, Applied Physics Letters **55**, 1074–1076 (1989).

[98] J. Repp, G. Meyer, S. Paavilainen, F. E. Olsson, and M. Persson, “Scanning tunneling spectroscopy of Cl vacancies in NaCl films: strong electron-phonon coupling in double-barrier tunneling junctions”, Physical Review Letters **95**, 225503 (2005).

[99] T. Choi, C. D. Ruggiero, and J. A. Gupta, “Incommensurability and atomic structure of $c(2 \times 2)$N/Cu(100): a scanning tunneling microscopy study”, Physical Review B **78**, 035430 (2008).

[100] C. Attaccalite, M. Bockstedte, A. Marini, A. Rubio, and L. Wirtz, “Coupling of excitons and defect states in boron-nitride nanostructures”, Physical Review B **83**, 144115 (2011).

[101] A. Zunger and A. Katzir, “Point defects in hexagonal boron nitride. II. theoretical studies”, Physical Review B **11**, 2378–2390 (1975).

[102] V. W. Brar, R. Decker, H.-M. Solowan, Y. Wang, L. Maserati, K. T. Chan, H. Lee, C. O. Girit, A. Zettl, S. G. Louie, M. L. Cohen, and M. F. Crommie, “Gate-controlled ionization and screening of cobalt adatoms on a graphene surface”, Nature Physics **7**, 43–47 (2011).

[103] N. A. Pradhan, N. Liu, C. Silien, and W. Ho, “Atomic scale conductance induced by single impurity charging”, Physical Review Letters **94**, 076801 (2005).

[104] M. T. Woodside and P. L. McEuen, “Scanned probe imaging of single-electron charge states in nanotube quantum dots”, Science **296**, 1098–1101 (2002).

[105] Y.-J. Yu, Y. Zhao, S. Ryu, L. E. Brus, K. S. Kim, and P. Kim, “Tuning the graphene work function by electric field effect”, Nano Letters **9**, 3430–3434 (2009).

[106] S. Zhou, Y. Liu, Y. Xu, W. Hu, D. Zhu, X. Qiu, C. Wang, and C. Bai, “Rectifying behaviors of Langmuir-Blodgett films of an asymmetrically substituted phthalocyanine”, Chemical Physics Letters **297**, 77–82 (1998).

[107] N. Kharche and S. K. Nayak, “Quasiparticle band gap engineering of graphene and graphone on hexagonal boron nitride substrate”, Nano Letters **11**, 5274–5278 (2011).

[108] J. K. Garleff, A. P. Wijnheijmer, C. N. v. d. Enden, and P. M. Koenraad, “Bistable behavior of silicon atoms in the (110) surface of gallium arsenide”, Physical Review B **84**, 075459 (2011).

[109] E. J. G. Santos and E. Kaxiras, “Electric-field dependence of the effective dielectric constant in graphene”, Nano Letters **13**, 898–902 (2013).

[110] L. Britnell, R. V. Gorbachev, R. Jalil, B. D. Belle, F. Schedin, A. Mishchenko, T. Georgiou, M. I. Katsnelson, L. Eaves, S. V. Morozov, N. M. R. Peres, J. Leist, A. K. Geim, K. S. Novoselov, and L. A. Ponomarenko, “Field-effect tunneling transistor based on vertical graphene heterostructures”, Science **335**, 947–950 (2012).

[111] J. G. Simmons, “Generalized formula for the electric tunnel effect between similar electrodes separated by a thin insulating film”, Journal of Applied Physics **34**, 1793–1803 (1963).

[112] S. Makram-Ebeid and M. Lannoo, “Quantum model for phonon-assisted tunnel ionization of deep levels in a semiconductor”, Physical Review B **25**, 6406–6424 (1982).

[113] J. Frenkel, “On pre-breakdown phenomena in insulators and electronic semi-conductors”, Physical Review **54**, 647–648 (1938).

[114] S. D. Ganichev, E. Ziemann, W. Prettl, I. N. Yassievich, A. A. Istratov, and E. R. Weber, “Distinction between the Poole-Frenkel and tunneling models of electric-field-stimulated carrier emission from deep levels in semiconductors”, Physical Review B **61**, 10361–10365 (2000).

[115] D. Wong, J. Velasco Jr., L. Ju, S. Kahn, J. Lee, C. Germany, A. Zettl, F. Wang, and M. Crommie, *Local doping of two-dimensional materials*, United States Patent 9,449,851, issued September 20, 2016.

[116] J. Lee, D. Wong, J. Velasco Jr., J. F. Rodriguez-Nieva, S. Kahn, H.-Z. Tsai, T. Taniguchi, K. Watanabe, A. Zettl, F. Wang, L. S. Levitov, and M. F. Crommie, “Imaging electrostatically confined Dirac fermions in graphene quantum dots”, Nature Physics **12**, 1032–1036 (2016).

[117] C. Gutiérrez, L. Brown, C.-J. Kim, J. Park, and A. N. Pasupathy, “Klein tunnelling and electron trapping in nanometre-scale graphene quantum dots”, Nature Physics **12**, 1069–1075 (2016).

[118] N. M. Freitag, L. A. Chizhova, P. Nemes-Incze, C. R. Woods, R. V. Gorbachev, Y. Cao, A. K. Geim, K. S. Novoselov, J. Burgdörfer, F. Libisch, and M. Morgenstern, “Electrostatically confined monolayer graphene quantum dots with orbital and valley splittings”, Nano Letters **16**, 5798–5805 (2016).

[119] C. A. Downing, D. A. Stone, and M. E. Portnoi, “Zero-energy states in graphene quantum dots and rings”, Physical Review B **84**, 155437 (2011).

[120] A. Bostwick, F. Speck, T. Seyller, K. Horn, M. Polini, R. Asgari, A. H. MacDonald, and E. Rotenberg, “Observation of plasmarons in quasi-freestanding doped graphene”, Science **328**, 999–1002 (2010).

[121] A. Principi, M. Polini, R. Asgari, and A. H. MacDonald, “The tunneling density-of-states of interacting massless Dirac fermions”, Solid State Communications **152**, 1456–1459 (2012).

[122] J. Lischner, D. Vigil-Fowler, and S. G. Louie, “Physical origin of satellites in photoemission of doped graphene: an *ab initio GW* plus cumulant study”, Physical Review Letters **110**, 146801 (2013).

[123] S. Jung, G. M. Rutter, N. N. Klimov, D. B. Newell, I. Calizo, A. R. Hight-Walker, N. B. Zhitenev, and J. A. Stroscio, “Evolution of microscopic localization in graphene in a magnetic field from scattering resonances to quantum dots”, Nature Physics **7**, 245–251 (2011).

[124] L. A. Ponomarenko, F. Schedin, M. I. Katsnelson, R. Yang, E. W. Hill, K. S. Novoselov, and A. K. Geim, “Chaotic Dirac billiard in graphene quantum dots”, Science **320**, 356–358 (2008).

[125] S. Schnez, J. Güttinger, M. Huefner, C. Stampfer, K. Ensslin, and T. Ihn, “Imaging localized states in graphene nanostructures”, Physical Review B **82**, 165445 (2010).

[126] K. Todd, H.-T. Chou, S. Amasha, and D. Goldhaber-Gordon, “Quantum dot behavior in graphene nanoconstrictions”, Nano Letters **9**, 416–421 (2009).

[127] D. Subramaniam, F. Libisch, Y. Li, C. Pauly, V. Geringer, R. Reiter, T. Mashoff, M. Liebmann, J. Burgdörfer, C. Busse, T. Michely, R. Mazzarello, M. Pratzer, and M. Morgenstern, “Wave-function mapping of graphene quantum dots with soft confinement”, Physical Review Letters **108**, 046801 (2012).

[128] S. K. Hämäläinen, Z. Sun, M. P. Boneschanscher, A. Uppstu, M. Ijäs, A. Harju, D. Vanmaekelbergh, and P. Liljeroth, “Quantum-confined electronic states in atomically well-defined graphene nanostructures”, Physical Review Letters **107**, 236803 (2011).

[129] S.-h. Phark, J. Borme, A. L. Vanegas, M. Corbetta, D. Sander, and J. Kirschner, “Direct observation of electron confinement in epitaxial graphene nanoislands”, ACS Nano **5**, 8162–8166 (2011).

[130] J. Lu, P. S. E. Yeo, C. K. Gan, P. Wu, and K. P. Loh, “Transforming C_{60} molecules into graphene quantum dots”, Nature Nanotechnology **6**, 247–252 (2011).

[131] A. De Martino, L. Dell’Anna, and R. Egger, “Magnetic confinement of massless Dirac fermions in graphene”, Physical Review Letters **98**, 066802 (2007).

[132] C. Schulz, R. L. Heinisch, and H. Fehske, “Scattering of two-dimensional Dirac fermions on gate-defined oscillating quantum dots”, Physical Review B **91**, 045130 (2015).

[133] J.-S. Wu and M. M. Fogler, “Scattering of two-dimensional massless Dirac electrons by a circular potential barrier”, Physical Review B **90**, 235402 (2014).

[134] J. H. Bardarson, M. Titov, and P. W. Brouwer, "Electrostatic confinement of electrons in an integrable graphene quantum dot", Physical Review Letters **102**, 226803 (2009).

[135] A. Matulis and F. M. Peeters, "Quasibound states of quantum dots in single and bilayer graphene", Physical Review B **77**, 115423 (2008).

[136] H.-Y. Chen, V. Apalkov, and T. Chakraborty, "Fock-Darwin states of Dirac electrons in graphene-based artificial atoms", Physical Review Letters **98**, 186803 (2007).

[137] M. F. Crommie, C. P. Lutz, and D. M. Eigler, "Confinement of electrons to quantum corrals on a metal surface", Science **262**, 218–220 (1993).