Electric Transport in Hybrid Carbon Nanotube-Graphene Devices

A DISSERTATION PRESENTED
BY
AUSTIN KCON CHENG
TO
THE SCHOOL OF ENGINEERING AND APPLIED SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
APPLIED PHYSICS

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
APRIL 2019

Thesis advisor: Professor Philip Kim Austin Kcon Cheng

Electric Transport in Hybrid Carbon Nanotube-Graphene Devices

Abstract

Two remarkable properties exhibited by graphene electrons are their pseudo-relativistic behavior and their ballistic motion that mimics the propagation of electromagnetic waves in a refractive medium. Despite heavy investigation, limitations of existing fabrication and measurement techniques have prevented certain aspects of these phenomena from being thoroughly explored or realized. Two in particular are Landau level collapse and the single-mode waveguide. Observing these phenomena requires a sharp modulation of the potential landscape of graphene as well as a probe to detect the resulting local electronic behavior. Incorporating a carbon nanotube into a graphene heterostructure allows us to do both simultaneously. This thesis describes the fabrication procedure of the hybrid carbon nanotube-graphene device and explains its operational principle. We demonstrate the carbon nanotube can act as a local probe and use it to measure the electronic properties of graphene. With an external magnetic field and with the carbon nanotube as an electric field generator, we perform a detailed study of Landau level collapse and draw connections to relativistic dynamics. When the potential well in graphene due to the carbon nanotube is made deep enough, we observe a clear signature of a single guided mode. Lastly, we present preliminary results on Coulomb drag between a carbon nanotube and graphene, and discuss the next steps toward reaching the goal of observing hydrodynamic behavior in graphene.

Contents

1 Graphene as a Rich Playground for Condensed Matter Physics 2
- 1.1 The Bare Minimum: Single Particle Picture in the Tight Binding Model 5
- 1.2 Analogue to Relativistic Particles 9
- 1.3 External Electrostatic Potential 10
- 1.4 Quantum Hall Effect 12
- 1.5 Hydrodynamics 13
- 1.6 The Small but Mighty Carbon Nanotube 14
- 1.7 Final Foreword: Why the Hybrid CNT−Gr Device? 16

2 The Carbon Nanotube as a Local Probe 19
- 2.1 Using the Carbon Nanotube to Detect Screened Electric Fields 20
- 2.2 Clarification: Chemical Potential vs Electrochemical Potential 22
- 2.3 Quantum Capacitance 23
- 2.4 The Simple Backgated Graphene Device Examined 26
- 2.5 Quantum Capacitance Revisited 31
- 2.6 Quantum Capacitance of Graphene 32
- 2.7 Quantum Capacitance of the Carbon Nanotube 33
- 2.8 Local Compressibility Measurement of Graphene with a Carbon Nanotube 39
- 2.9 Extraction of Quantum Capacitance: Method of χ^2 Test 46
- 2.10 Local Compressibility Measurement Of the Quantum Hall Effect of Graphene . . 48

3 Fabrication of the Carbon Nanotube and Graphene Hybrid Device 54
- 3.1 Challenges Regarding Obtaining a Carbon Nanotube 55
- 3.2 The Road to Deterministic Carbon Nanotube Selection 56
- 3.3 Synthesis of Carbon Nanotube 57
- 3.4 Optical Characterization of the Carbon Nanotube 63
- 3.5 The Rayleigh Setup 74
- 3.6 The Assembly of the Carbon Nanotube-Graphene Hybrid Device 79

4 Relativistic Analogue in Graphene 86
- 4.1 Relativistic Electromagnetism in a Nutshell 87
- 4.2 Brief Timeline Leading Up to the Full-Fledged Relativistic Quantum Theory . . . 94
- 4.3 Klein Tunneling 101
- 4.4 Atomic Collapse 111

4.5 Atomic Collapse-Like Resonances in Graphene . . . 117
4.6 1D Confinement in Graphene and its Connection with Atomic Collapse . . . 118

5 Landau Level Collapse 124
5.1 Connection with Relativity . . . 126
5.2 Landau Levels with a Linear Electrostatic Potential . . . 127
5.3 Device Configuration, Measurement Scheme and Approximations . . . 132
5.4 Electric Field Screening by Landau Levels . . . 133
5.5 Local Compressibility Measurements at Different Magnetic Fields . . . 142
5.6 Landau Level Collapse for the $N = 0$ Landau Level . . . 149
5.7 Conclusion . . . 152

6 Single Guided Mode in Graphene 153
6.1 Optical Analogue . . . 155
6.2 A Single Weakly Guided Mode . . . 159
6.3 Creating and Probing the Single Guided Mode with a Carbon Nanotube . . . 160
6.4 Local Compressibility Measurement of the Single Guided Mode . . . 167
6.5 Connection with Atomic Collapse . . . 170
6.6 Multimodes: Absence of Resonances in Wider Potentials . . . 172
6.7 Briefly on Magnetic Field Dependence . . . 174
6.8 Conclusion . . . 175

7 Towards Probing Hydrodynamics via Coulomb Drag between a Carbon Nanotube and Graphene 176
7.1 The Quest to Probe Electrohydrodynamics . . . 177
7.2 Brief Introduction to Coulomb Drag . . . 181
7.3 Coulomb Drag Mechanisms . . . 182
7.4 Onsager Reciprocal Relations . . . 201
7.5 Parasitic Effects and the Drag Measurement Scheme . . . 204
7.6 Preliminary Results . . . 208

8 Conclusion: Applications and Outlook 218

Appendix Epilogue: Prologue to the Appendix 224

Appendix A Overview of the Electronic Properties of Graphene 226
A.1 Graphene Lattice . . . 227
A.2 Graphene Bandstructure . . . 230
A.3 Linear Dispersion of Graphene: Massless Dirac Fermions . . . 237
A.4 Corrections to the Linear Dispersion: Massive Dirac Fermions . . . 242
A.5 Additional Remarks on the Graphene Hamiltonian . . . 243

A.6 Graphene's Wavefunction around the Dirac Point: A Pseudo-Spin 244
A.7 Low-Energy Density of States 249
A.8 Effective Mass 250
A.9 Source Code for the Dispersion Relation of Graphene with 1D Potential and Magnetic Field 251

Appendix B Overview of the Electronic Properties of Carbon Nanotubes 253
B.1 Briefly about the Carbon Nanotube 254
B.2 Carbon Nanotube Lattice 255
B.3 Dispersion Relation of Carbon Nanotubes: Quantized Graphene 257
B.4 Metallic Versus Semiconducting Carbon Nanotubes 260
B.5 The Density of States and Other Important Quantities of Carbon Nanotubes . . 265
B.6 A Little Further into the Specifics of Carbon Nanotubes 269

Appendix C Overview of the Electronic Transport of Carbon Nanotube Quantum Dots 274
C.1 Confining the Electrons in a Carbon Nanotube Even More 275
C.2 Conditions to Form and Detect a Quantum Dot 277
C.3 Coulomb Blockade 283
C.4 Density of States of a Carbon Nanotube Quantum Dot 292
C.5 Conductance of a Carbon Nanotube Quantum Dot 294
C.6 Other Quantum Dot Behaviors: Gate-Defined and Disorder-Defined Quantum Dots as well as Fabry-Perot Interferences 296

Appendix D The Integer Quantum Hall Effect of Graphene 300
D.1 Classical Regime: Weak Magnetic Fields 301
D.2 Quantum Regime: Strong Magnetic Fields 302
D.3 Landau Level Degeneracy 310
D.4 Relativistic vs Non-Relativistic Landau Levels Comparison 312
D.5 Quantum Hall Edge States and Transport Properties 313
D.6 Berry Phase and Graphene's Unconventional Filling Factors 316
D.7 Briefly on Highly Correlated Phases 318

Appendix E The Hydrodynamic Description of Electrons in Graphene 321
E.1 Introduction to Hydrodynamics in Graphene 322
E.2 The Hydrodynamic Window in Monolayer Graphene 330
E.3 From Semiclassical Electron Transport to Hydrodynamics 333

Appendix F Fabrication and Synthesis Recipes 339
F.1 Polymers 340
F.2 Electron Beam Lithography Recipes 344

F.3 Metal Evaporation and Edge Contact Recipes 345
F.4 Atomic Layer Deposition Recipes 346
F.5 Carbon Nanotube Growth Chip Fabrication Recipe 347
F.6 Etch Recipes 350
F.7 Carbon Nanotube Growth Catalyst Recipe: CoMoCat 351
F.8 Carbon Nanotube Growth Procedure 352
F.9 Janis 356

Appendix G Additional Supporting Data 359

References 393

Author List

The following authors contributed to the work shown in Chapter 5: Kenji Watanabe, Takashi Taniguchi, Philip Kim, Jean Damien Pillet.

The following authors contributed to the work shown in Chapter 6: Kenji Watanabe, Takashi Taniguchi, Philip Kim, Jean Damien Pillet.

The following authors contributed to the work shown in Chapter 7: Laurel Anderson, Kenji Watanabe, Takashi Taniguchi, Philip Kim.

Listing of Figures

1.1 Graphene Lattice . . . 4
1.2 Graphene Energy . . . 8
1.3 Landau Levels of Relativistic and Non-Relativistic Systems . . . 12
1.4 Rolled-Up Graphene . . . 15
1.5 Carbon Nanotube Periodic Table . . . 16

2.1 Illustration of Screening . . . 20
2.2 Graphene with a Backgate . . . 27
2.3 Quantum Capacitance of Graphene . . . 34
2.4 Quantum Capacitance of the Carbon Nanotube . . . 36
2.5 Carbon Nanotube with a Backgate . . . 38
2.6 Hybrid Carbon Nanotube-Graphene Device with a Global Backgate . . . 40
2.7 Coulomb Blockade Periodicity Dependence on Graphene Density of States . . . 42
2.8 Carbon Nanotube-Graphene Hybrid Device Measurement . . . 43
2.9 Local Compressibility Measurement of Graphene . . . 44
2.10 Quantum Capacitance and Fermi Energy Extraction . . . 46
2.11 Slope Extraction Method . . . 47
2.12 Local Compressibility Measurement of Quantum Hall Effect of Graphene . . . 50
2.13 Demonstration of Energy Behavior in the Compressibility Measurement 1 . . . 51
2.14 Demonstration of Energy Behavior in the Compressibility Measurement 2 . . . 52
2.15 Demonstration of Energy Behavior in the Compressibility Measurement 3 . . . 53

3.1 Carbon Nanotube Growth Chip . . . 58
3.2 Carbon Nanotube Growth Furnace . . . 59
3.3 Carbon Nanotube Growth Mechanisms . . . 63
3.4 Suspended Carbon Nanotube Growth . . . 64
3.5 Carbon Nanotube Raman Spectra . . . 67
3.6 Carbon Nanotube Rayleigh Spectra . . . 68
3.7 Metal vs Semiconducting Carbon nanotubes Density of States . . . 71
3.8 Asymmetry of Peak in Rayleigh Spectrum . . . 72
3.9 Demonstration of Trigonal Warping . . . 73
3.10 Types of Carbon Nanotube Seen through Rayleigh Scattering . . . 74
3.11 The Kataura Plot . . . 75
3.12 The Rayleigh Transfer Stage Setup . . . 77
3.13 Carbon Nanotube Raman Spectroscopy Configuration . . . 78

3.14 Carbon Nanotube Rayleigh Spectroscopy Configuration . . . 78
3.15 Carbon Nanotube Transfer Process . . . 82
3.16 The Target Device for Carbon Nanotube Transfer . . . 84
3.17 The Graphene-Carbon Nanotube Hybrid Device Fabrication Scheme. . . . 85

4.1 Clarification of Electron and hole . . . 103
4.2 Klein Tunneling through a Step Potential . . . 109
4.3 Klein Tunneling through a npn Potential Barrier . . . 110
4.4 Non-relativistic vs Relativistic Energy Comparison . . . 113
4.5 Atomic Collapse Theory Graphs and Illustrations . . . 115
4.6 Electric Potential Simulation of 1D Wire over 2D Metal . . . 119
4.7 Potential Depth due to Carbon Nanotube Charging . . . 120
4.8 Simulation of Graphene Density of States under the Influence of a 1D Charge String 123

5.1 Semiclassical and Quantum Mechanical Description of Landau Level Collapse . . 125
5.2 Landau Levels with a Crossed Electric Field . . . 130
5.3 Landau Level Collapse Measurement Scheme . . . 132
5.4 Electric Field Screening by Landau Levels . . . 135
5.5 Comparison with Simulations of the Fan Diagram Electric Field Dependence . . . 138
5.6 Measurements of the Electric Field Dependence of the Fan Diagram . . . 139
5.7 Magnetic Field Dependence of the Carbon Nanotube Quantum Dot . . . 143
5.8 Local Compressibility Measurement and Quantum Capacitance Extraction at 9T . 144
5.9 Local Compressibility Measurement and Quantum Capacitance Extraction at Other Magnetic Fields . . . 145
5.10 Quantum Capacitance Extraction Line Cuts . . . 146
5.11 Junction Polarity and Landau Level Collapse Behavior Explanation . . . 148
5.12 $N = 0$ Landau Level Collapse Fits . . . 150

6.1 Snell's Law in Graphene . . . 156
6.2 Guided Mode Mechanisms . . . 158
6.3 Potential Illustration and Criterion for the Single Guided Mode . . . 161
6.4 Disorder Potential on Graphene . . . 162
6.5 Creating a Potential Well in Charge Puddles . . . 163
6.6 Density of States of the Guided Modes . . . 165
6.7 Local Compressibility Measurement with a Changing 1D Potential . . . 166
6.8 Quantum Capacitance Extraction and Simulation of the Single Guided Mode . . . 168
6.9 Comparison between Experiment and Theory regarding the Quantum Capacitance of a Singe Guided Mode . . . 169
6.10 The Case of a Wide Potential . . . 173
6.11 Compressibility Measurement and Extracted Quantum Capacitance at 1 T . . . 175

7.1 Non-hydrodynamic vs Hydrodynamic Flow Regimes in Graphene . . . 177

7.2 Electronic Phase Diagram of Graphene . . . 178
7.3 The Hydrodynamic Signature in Graphene . . . 179
7.4 Carbon Nanotube for the Creation of a 1D Moving Boundary Condition . . . 180
7.5 Illustration of Momentum Drag . . . 182
7.6 Momentum Drag in a Double Graphene Layer System . . . 183
7.7 1D 2D Electron-Hole Excitations . . . 188
7.8 Coulomb Drag Comparison between Electron-Hole Symmetric and Asymmetric Systems . . . 190
7.9 Illustration of Energy Drag . . . 194
7.10 Heat characteristic of Graphene . . . 197
7.11 Energy Drag Transport Characteristic . . . 198
7.12 Comparison between Momentum and Energy Drag at Non-zero Magnetic Field . . . 200
7.13 Temperature Dependence of Energy Drag . . . 201
7.14 Parasitic Signal and Classic Ways of Avoiding it . . . 206
7.15 Drag Device Measurement Scheme . . . 207
7.16 Drag Device Fabrication . . . 209
7.17 Coulomb Drag in Large Hybrid CNT-Monolayer Device . . . 210
7.18 Layer Asymmetry in Drag Measurements . . . 211
7.19 Coulomb Drag in Small Hybrid CNT-Monolayer Gr Device 2 . . . 212
7.20 Layer Symmetry in Drag Measurements . . . 213
7.21 Drag Measurements on Hybrid CNT-Bilayer Gr Devices . . . 216

A.1 Graphene Lattice . . . 227
A.2 Graphene Energy . . . 238
A.3 Graphene Pseudo-spin . . . 248

B.1 Carbon Nanotube Periodic Table . . . 256
B.2 Structural Types of CNT . . . 258
B.3 Examples of Zig-Zag Carbon Nanotubes and their Bandstructures . . . 261
B.4 Examples of Armchair Carbon Nanotubes and their Bandstructures . . . 262
B.5 Examples of Chiral Carbon Nanotubes and their Bandstructures . . . 263
B.6 Transport of Metallic vs Semiconducting Carbon Nanotubes . . . 264
B.7 Density of States of Different Carbon Nanotubes . . . 270
B.8 The Kataura Plot . . . 271
B.9 Metal vs Semiconducting Carbon nanotubes Density of States . . . 272
B.10 Metallic vs Semiconducting Carbon Nanotubes . . . 273

C.1 Quantum Dot Device . . . 278
C.2 Energy Definitions in a Quantum Dot . . . 286
C.3 Transport Mechanism of a Quantum Dot . . . 287
C.4 Illustration of Coulomb Diamonds . . . 288

C.5 Charge Stability Diagram: Measurement of Coulomb Diamonds 289
C.6 Coulomb Blockade 290
C.7 Confinement in a Bipolar Junction 297
C.8 Other Quantum Dot Behaviors 298

D.1 Landau Levels of Monolayer Graphene 310
D.2 Landau Levels of a Massive System 312
D.3 Quantum Hall Measurement 317
D.4 Illustration of Quantum Hall Effect 319

E.1 Boundary Conditions for a System with a Moving Edge 327
E.2 The Hydrodynamic Simulation of the Electronic Liquid in Graphene with a Moving Boundary Condition 329
E.3 Summary of Scattering Rates in Graphene 331

F.1 Bubble Pushing 340
F.2 Inverted Stack 342
F.3 Put Down Comparison 343
F.4 Carbon Nanotube Growth Furnace 353
F.5 The Janis Setup 357
F.6 Grounding of Janis 357
F.7 Variational Temperature Insert and Magnet 358

G.1 Additional Magnetic Field Dependences of a Carbon Nanotube Quantum Dot on Graphene 360
G.2 Bound State in 6nm Device A 361
G.3 Bound State in 6nm Device B 361
G.4 Coulomb Diamonds of 4nm Device 362
G.5 Additional Fan Diagrams 362
G.6 Kink Zoom of Compressibility Measurement in Quantum Hall Regime 362
G.7 Additional Fan Diagrams showing Landau Level Collapse in Device A 363
G.8 Additional Fan Diagrams showing Landau Level Collapse in Device B 363
G.9 Carbon Nanotube Conductance as functions of Different Parameters 364
G.10 Quantum Capacitance Line Cuts at Different Magnetic Fields 365
G.11 Negative Compressibility at 9T 365
G.12 Carbon Nanotube Conductance as a Function of B and V_{Gr} 366
G.13 Local Compressibility Measurements of 70 and 100nm thick hBN devices 366
G.14 Local Compressibility Measurement and Quantum Capacitance Extraction of the 10 nm hBN Device 367
G.15 Aharonov Bohm Effect in Graphene caused by Carbon Nanotube Gating 368

This thesis is dedicated to my loving family,
to my friends and teachers who shaped me into who I am,
and to those who stumble upon this in their search for knowledge.

Acknowledgments

As I reflect on my journey in graduate school research, I can't help but think of how big an undertaking the whole process has been. At the same time, I find myself trembling with joy realizing how privileged and blessed I was the whole time. Research is definitely not for the faint-hearted and it is not something that I could have done without the people I know around me. The trials and tribulations that is inherent to research can be a big mental and physical toll but pulling through and surviving it has made me grow into a better professional, a wiser problem-solver and most importantly, a more joyful endurer. All this was only possible because of so many people and to whom I am greatly indebted.

First and foremost, eternal gratitude goes to my advisor Philip. To start off let me first say that being a graduate student is a once in a lifetime deal and probably the most awesome one ever. We get paid to learn and explore uncharted territory while being in a very protective and nurturing environment. But Philip, you have made it even better for me. I'm very thankful for your patience and your support over the years. You let me create so freely with no boundaries. I was truly given the opportunity to adopt the free spirit mindset which I do not take for granted. Thank you for being so forgiving and generous. My relentless use of ridiculously expensive helium have very likely dug a big hole in your wallet. Despite knowing so, you have continued to provide me with resources unconditionally. You have made my time here too good to be true. Despite all the failures I had, time and time again, be it the faith you had in me or your desire to ease my sorrows, you still let me reign my PhD. One of my most transformative moments was when you summoned me to your office and then beat into my head that this is *my* PhD. You have given me the world and I had it real easy. It is sad to know what I am parting with once I leave your benevolent jurisdiction. It will be very difficult to come by hereafter. So truly, thank you Philip.

In addition to the amazing lab culture and unbounded means for creativity, Philip also gave me two amazing mentors: Jayakanth and JD. Jayakanth, you were upbeat, lively and critical. You endowed me with your energy that really carried me through graduate school. The times we spent day in and night out, with my constant bugging you about science and your ability to answer everything, ignited the research passion in me that never ever died. JD, you were the most consummate scientist I have ever met and will ever work with. You were always professional and calm, which is hard to do with me around constantly breaking things and being stupid. You taught me everything I know now about

physics and you inspired me to do not just science but anything the right way. You showed me how to approach problems with a clear and systematic mind, and most importantly, you taught me how to become a true experimentalist for not just physics, but for all walks of life. Jayakanth and JD, both of you propelled me to greater heights as an individual and made graduate school a huge success for me. Thank you.

Here is an additional paragraph for you, JD. I'm sure you know that you have been extremely critical to the success of my graduate school career. The projects I was involved in and their successes were all because of you. I am very thankful for your commitment to nurturing me. I appreciate every single one of your emails and Skype sessions. You have always explained things with extreme precision and clarity and you have pushed me to do so as well. You have ingrained in me the importance of preparation, preparation and preparation as well as double, triple, quadruple checking everything. I've learnt so many invaluable lessons from you. I just hope that I can also pass them on to others.

For the most part, graduate school deserves to be glorified but there are of course plenty of downs. Graduate school in experimental condensed matter physics is hard. Over time, you slowly and painfully realize that you really are simply constantly frantically trying to dig yourself out of the ocean. It was a huge challenge especially when you are going through fruitless months of non-stop work. A sane mind is not enough. You need unyielding perseverance and frequent hysterical imagination that what you are doing is going to work. Most critically, you need people who can understand and who you can vent to. For this, Frank and Xiaomeng, I am sorry. Sorry for rotting your brains with the non-stop cursing. And thank you, for always being there. Frank, you are an upstanding individual and have strong positive principles. I am always inspired by your optimism, selflessness, motivation and discipline. The lab would be in a horrible state without you. Those bike rides and hiking trips that you are so willing to partake and navigate were very rejuvenating and important for me to regain inner strength for research. Your exclamation of, "HOOORAAH, onwards to victory!" will stay with me forever. When times are tough, I will remind myself of you. Xiaomeng, you are silently cheeky and humorous and we share a lot of the same taste and perspectives. To have somebody whose thoughts resonate with mine was a great sanity check. Thank you for keeping me company for all the countless all-nighters in lab. The midnight snacks, the blasting of music, the movie critiquing, and the morning walks back home kept my life light and manageable. One thing that sets us apart is that you are one brilliant bastard. Your insight for every problem, scientific or not, was a key ingredient to many of my progresses. Frank and Xiaomeng, the two of you defined my PhD experience. I'm very honored to have been able to work with you guys. You guys make the best labmates one can ever have and I am so happy to have gained two greatest buddies.

To Laurel, you are an extremely hardworking warrior. The project we were working on was tough.

It brought out the best in us to be able to tame it. It is particularly inspiring to see you show so much resolve and conquer so many difficulties and grow on your own over the years. It was a real pleasure working with you. Let me just remind you one last time, it's never going to be easy. So good luck and have fun.

Thank you Professor Loncar for so generously lending your lab space and introducing me to Linbo. Thank you Linbo for introducing me to optics and guiding my way to building the Rayleigh setup. This thesis would not have been possible without both of you.

To Gilho, thank you for filling the void in our lab and my life during the transition period at Harvard. You brought the whole lab together and gave it life. You were always so sincere in helping others and me and the positive energy you brought was so contagious. I expect my fully furnished bedroom when I come visit you.

To all my fellow lab mates: Carlos, Patrick, Fereshte, Chul-Ho, Dmitri, Amelia, KC, Mitsu, Giselle, Daeyong, Clevin, Wei, Ke, Hiroshi, Jonah, Kwabena, Hyobin, Luis, Siyoung, Youngjae, Nicola, Jesse, Jing, Andy, Artem, Rebecca, Mehdi, Zeyu, Tom, Zhongying, Andres, Kate, Sebastian, Onder, and Yuval, it's been a real pleasure being in your presence. You guys all make the lab so lively and made it so worthy to be called my second home. Being able to learn about your diverse backgrounds and witness the different stories unfold for all of you has been eye-opening. You guys have all touched my life in ways you wouldn't know and it was because of every single one of you that I had so much difficulty deciding to graduate. The number of talented, selfless, fun and humble people I have met has truly made this PhD an amazing experience. Thank you all.

To the CNS and LISE staff, you all make the most exemplary staff. When I got to Harvard I was simply amazed by how well everything was maintained and organized. You are all so diligent and always kept our needs a priority. You are the ones who worked behind the scenes and made our research possible. Thank you all so much. "Coach" Mac thanks for keeping the clean room lively and for being such a great fan of basketball. It's a shame basketball isn't a big part of the culture here compared to New York. After I moved to Harvard, your constant request for Friday basketball is what kept my lifelong passion alive. It's always a joy for me to see you show them young bloods that unblockable hookshot.

To my basketball crew (if you ever happen to read this you know who you are!), I extremely appreciate the rampant teasing and jokes. It is unfortunate over the past several years my basketball skills and athleticism have been on a fast decline, but you guys never hesitated to invite me to your tournaments. My Columbia/New York days have been filled with so much fond memories because of you guys. I really appreciate the time when you guys came all the way to Harvard just to ball with me, and I treasure every second we got to be on the court together. And thanks for the "silent" support and just

having me in your thoughts. I promise I'll be hitting the gym more often from now on and I won't hold you guys back in participating in tournaments.

To my first year office mates and partners in crime back at Columbia, it was hell of a ride our time together. To Diego and Olgun, the office dynamic was just so fun. I cannot believe how long ago this was yet the experience is still fresh. Aside from fun there was the non-stop studying for classes and qualifications, the teaching and grading, and the constant worrying of what research to do... what an ordeal. But we powered through, or rather, I powered through thanks to you guys. To Dennis, thank you so much for being just so accommodating and kind, especially during the year when I was "homeless" in New York. Thank you for teaching me how to play the piano and listening to my terrible tunes on repeat. The music pieces you taught me have been my much-needed meditation throughout the years. To Diego, thank you for looking out for me during as well as after grad school. I always admired your drive and truthfully it's something I try to emulate and it has taken me farther than I ever could on my own. Thank you for the workout sessions and the humor we shared. Hopefully, we can have that again if I ever move to California.

To Mr. Sando, the one who started this all. You were the one who captivated me with your teaching and made me pursue physics. Your creative and energetic methods of teaching still inspire me today to be a good teacher for others. Your enthusiasm for teaching will be a permanent reminder of how influential mine is to younger ones.

To Uncle SL, thank you for having such big and deep interest in what I do since the beginning and it means a lot to me. I have always been very touched by the care and encouragement you have given me. Unfortunately, I was never able to fully prepare or have ample time to coherently explain my research. This thesis is written with you in mind so now I hope you will have a full, detailed record of my work— you can go ahead and throw away that notebook. And I think we can finally eat good sushi with no concerns.

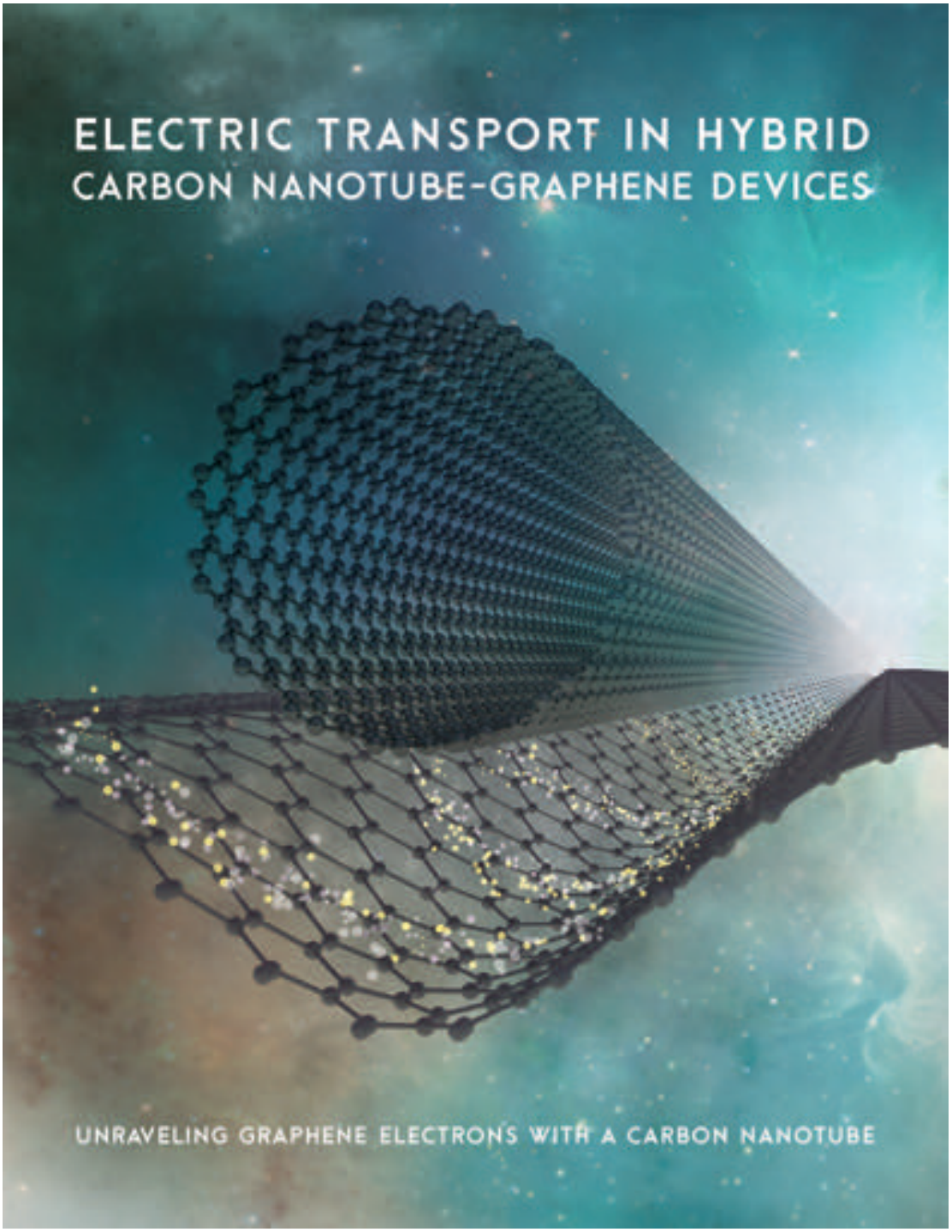
To Mom, Dad, Popo, Kuku, Jessica and Jane, you have been looking forward to my graduation for a long time now and I am happy to finally let you witness it. This thesis is a proud product of mine as well as yours. All the effort I put into it, I did it armed with all the love and support you all together have given me. Thank you Mom and Dad for being the greatest parents. You strike the perfect balance in my life and are responsible for all the life fulfilling emotions I have experienced. The life lessons you have taught me, I carried with me throughout my entire graduate school career. I would not have approached it and tackled it the same fulfilling way had it not been for you. You never pressured me to do anything, and never had daunting expectations. You respected my choices and constantly cheered for me from the other side of the world. Thank you. Popo and Kuku, the hardships both of you had to go through will never be paralleled. This PhD I know is not even a fraction of what

Deepest gratitude to my family! My family picture fabricated with nanotechnology.

you have had to go through. And yet, to this day, you smile like life has never weighed down on you. You are my living inspiration and the only source I need to believe that the world can be so free and wonderful. To my amazing sisters, we are surprisingly functional together. This is largely because of you guys obviously. You guys have always been the selfless ones, and have always made my life easier. Thanks for being so proactive and thanks for planning all those events that I should have helped out with but never did. If it weren't for you, I easily would've taken an extra year or two to finish my PhD. Also, thank you for the artwork. I'm very glad we got to make it a family effort, because this whole thing really is. Here I proudly and shamelessly dedicate a slot in this thesis for you guys!

I am extremely lucky to have met Katie and being with you have shown me another dimension in life that I would otherwise have never seen. You have been my greatest friend with love, support, humor and lots of sweetness. May I add that it is extremely humbling for me to meet a person with so much talent yet with so much modesty. Thinking through all the years we have been together, I am just very amazed and thankful by how things fell into place for us. I cannot see my life unfold any other way and I can't imagine what a disaster it would have been for me had I not met you. I am very excited about what is in store for us and I look forward to our next chapter together!

Graduate school was a huge success simply for all the amazing friends and human connections I have made. All of you have endowed me with wisdom, have been my punchbag, have unhesitatingly lent me a hand and have shared joys of tears and laughter with me. Although my time at the Kim lab will finally come to a close, my passion for science and research will never fade and I will forever hold on to and cherish the moments with you all. Thank you.


ELECTRIC TRANSPORT IN HYBRID
CARBON NANOTUBE-GRAPHENE DEVICES
UNRAVELING GRAPHENE ELECTRONS WITH A CARBON NANOTUBE

Science is beautiful when it makes simple explanations of phenomena or connections between different observations.

Stephen Hawking

1

Graphene as a Rich Playground for Condensed Matter Physics

Graphene is called "the wonder material" for many reasons.[1] From an application standpoint, graphene is highly conductive, has exceptional tensile strength and at the same time is atomically thin, flexible and transparent. Although its lack of a significant bandgap makes it less appealing as a silicon alternative, it can find many uses in other areas such as flexible displays, energy harvesting windows, clothing fabric, water purification filters, super-capacitive batteries and lightweight robust equipment just to

name several. From a condensed matter research standpoint, it is an electronic platform that hosts a plethora of exotic physical phenomena. First, its electronic behavior can roughly be broken into two basic regimes: the ordinary Fermi liquid and the more unique Dirac fluid composed of an electron-hole plasma. Under these two regimes, the charge carriers can behave as ballistic, massless, relativistic Dirac fermions, or as a highly collective hydrodynamic fluid. Coupled with the quantum mechanical properties such as spin and valley physics, graphene provides a highly complex electronic platform with which plenty of physics can be studied. The myriad of behaviors is unparalleled and this richness is made even more valuable because of its easy electronic-phase tunability. Like many of its van der Waals material counterparts, graphene has an intrinsically low density of states and hence sensitive chemical potential. This is due to graphene being a naturally-occurring two-dimensional (2D) material unlike the 2D electron gas (2DEG) created in Gallium Arsenide (GaAs) heterostructures via modulation doping. This brings us to perhaps the most important point about graphene: it can be easily isolated and allows easy electrical access. In contrast to the GaAs days when only few had access to the luxurious 2DEG and which, even for them, only had limited experimental knobs, the advent of graphene has enabled every condensed matter physicist to perform a variety of high level quantum mechanical experiments. Moreover, nowadays nanofabrication facilities are readily available and nanotechnology has advanced to an unprecedented level of sophistication— studying physics in graphene has become mundane. But the complexity that resides in graphene is still far from being untapped completely as we shall see shortly. Graphene research has made the science community as a whole much richer, more capable and frankly speaking, has put us physicists into a frantic play mode.

Graphene was first physically isolated and identified in the year 2004.[2] As mentioned previously, its single atomic layer structure automatically confines electrons to a plane, resulting in a low intrinsic carrier density. This reduced carrier density means less electronic screening, and because graphene is

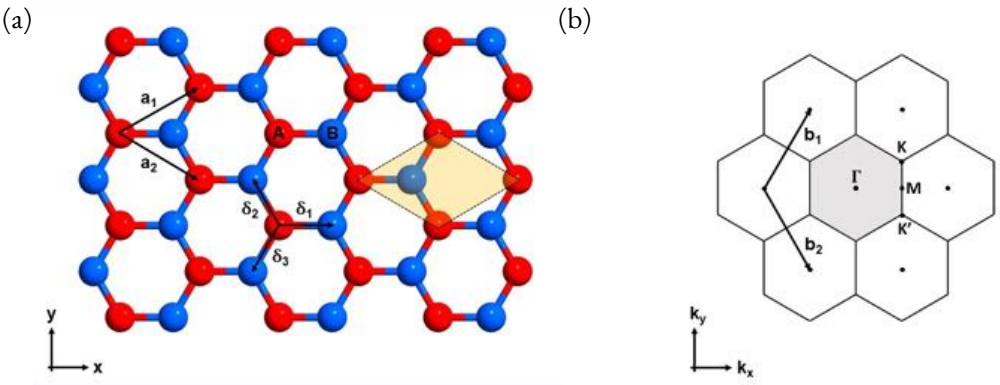


Figure 1.1: Illustration of the direct and reciprocal lattices of graphene, both of which have a hexagonal Bravais lattice. (a) The direct lattice in real space. The coloring scheme shows how it can be broken down into two sublattices. Sublattice A is colored red and sublattice B is colored blue. The yellow shaded parallelogram is the primitive unit cell defined by the real space Bravais vectors $\mathbf{a_1}$ and $\mathbf{a_2}$. The nearest neighbour vectors are given by $\boldsymbol{\delta_1}$, $\boldsymbol{\delta_2}$ and $\boldsymbol{\delta_3}$. (b) The reciprocal lattice in momentum space. The gray shaded hexagon in the middle is the first Brillouin Zone. $\mathbf{b_1}$ and $\mathbf{b_2}$ are the two reciprocal lattice vectors. The high symmetry points are labeled as $\mathbf{\Gamma}$, $\mathbf{K}$, $\mathbf{K}'$ and $\mathbf{M}$ where $\mathbf{K}$ and $\mathbf{K}'$ are two inequivalent positions that have Dirac cones with opposite helicity.

directly exposed, its electrons are much more susceptible to external influence. This is simultaneously a bane and benefit. It is a bane because of how sensitive it is to the unintended and ubiquitous disorder in the substrates; it is a benefit because of how easy it is to induce user-defined external potentials. In this thesis I will demonstrate, with the use of an individual carbon nanotube, how we can reap the benefit and circumvent the bane by performing local measurements. More specifically, I will show how the carbon nanotube can create and detect Landau level collapse, which is a pseudo-relativistic phenomenon, and how it can be used as a narrow electronic waveguide to guide a single electronic mode. In the following sections, I will briefly introduce some of the outstanding qualities of graphene and will highlight the ones that will be extensively utilized and studied in this thesis.

1.1 The Bare Minimum: Single Particle Picture in the Tight Binding Model

As with most problems, it is best to start simple. In the case of electrons in graphene, we begin with the single particle picture where we neglect any interactions that the electron can feel except for the periodic background potential emanating from the carbon lattice. This picture is commonly addressed through the tight-binding approach which will be briefly covered in the language of second quantization. For a more complete and detailed coverage on graphene, the reader is referred to appendix A.

Graphene (Gr) is comprised of carbon atoms arranged in a hexagonal Bravais lattice with two basis atoms. The two basis atoms separately form two hexagonal Bravais lattice and are conventionally referred to as sublattice A and B. The Bravais lattice vectors are:

$$\mathbf{a_1} = \left(\frac{\sqrt{3}a}{2}, \frac{a}{2}\right), \quad \mathbf{a_2} = \left(\frac{\sqrt{3}a}{2}, -\frac{a}{2}\right), \tag{1.1}$$

where $a = |\mathbf{a_1}| = |\mathbf{a_2}|$ is the lattice constant of Gr and has a value of about 2.46 Å. Each carbon has three nearest neighbors, all of which are occupants of the other sublattice. These nearest-neighbor vectors are:

$$\boldsymbol{\delta_1} = \left(\frac{a}{\sqrt{3}}, 0\right), \quad \boldsymbol{\delta_2} = \left(-\frac{a}{2\sqrt{3}}, -\frac{a}{2}\right), \quad \boldsymbol{\delta_3} = \left(-\frac{a}{2\sqrt{3}}, \frac{a}{2}\right), \tag{1.2}$$

where the carbon to carbon bond length or the nearest neighbor distance, $a_{\text{cc}} = |\boldsymbol{\delta_1}| = |\boldsymbol{\delta_2}| = |\boldsymbol{\delta_3}|$, is about 1.42Å. In reciprocal space, the Brillouin Zone (BZ) is also a hexagonal Bravais lattice with the

following Bravias lattice vectors:

$$\mathbf{b_1} = \left(\frac{2\pi}{\sqrt{3}a}, \frac{2\pi}{a}\right), \quad \mathbf{b_2} = \left(\frac{2\pi}{\sqrt{3}a}, -\frac{2\pi}{a}\right). \tag{1.3}$$

Figure 1.1b depicts the momentum space. The hexagon that is shaded in grey is the first BZ. $\mathbf{\Gamma}$ is the origin of momentum space where momentum is zero. $\mathbf{M}$ denotes the BZ edge and is thus a high symmetry point. The $\mathbf{K}$ and $\mathbf{K}'$ points are the inequivalent corners of the BZ and are described by the following vector with respect to $\mathbf{\Gamma}$:

$$\mathbf{\Gamma K} = \left(\frac{2\pi}{\sqrt{3}a}, \frac{2\pi}{3a}\right), \quad \mathbf{\Gamma K'} = \left(\frac{2\pi}{\sqrt{3}a}, -\frac{2\pi}{3a}\right). \tag{1.4}$$

Starting from the tight-binding model for nearest neighbor hopping only, the Hamiltonian is given by:

$$\mathbf{H} = t_{\text{nn}} \sum_{n,m} \mathbf{c}^{\dagger}_{n,A} \mathbf{c}_{m,B} + \text{h.c.} \tag{1.5}$$

$\mathbf{c}^{\dagger}_{n,\alpha}(\mathbf{c}_{n,\alpha})$ is the annihilation (creation) operator of an electron in a p_z orbital situated in the n-th unit cell for sublattice $\alpha = A, B$; $t_{\text{nn}} \approx -2.7$ to -3.1 eV is the nearest neighbor hopping energy; the summation over n, m is implicitly assumed to be only for nearest neighbors; h.c. stands for the Hermitian conjugate and this is to allow the reversed hopping process. In momentum space, the Hamiltonian becomes:

$$\mathbf{H}(\mathbf{k}) = \sum_{n=1}^{3} t_{\text{nn}} \begin{bmatrix} 0 & e^{i\mathbf{k}\cdot\boldsymbol{\delta_n}} \\ e^{-i\mathbf{k}\cdot\boldsymbol{\delta_n}} & 0 \end{bmatrix}, \tag{1.6}$$

written in the basis of $\{\mathbf{c}_A(\mathbf{k}), \mathbf{c}_B(\mathbf{k})\}$— the respective Fourier coefficients of $\mathbf{c}_{n,\alpha}$. Taylor expansion of the Hamiltonian around $\mathbf{K}$ such that $\mathbf{q} = \mathbf{k} - \mathbf{K}$ and retaining only the linear term yields the low

energy Hamiltonian around the $\mathbf{K}$ valley:

$$\mathbf{H}(\mathbf{q}) = \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q}. \tag{1.7}$$

Here, $\boldsymbol{\sigma}$ are the Pauli matrices and $v_F \approx c/300\,(\mathrm{m/s})$ is the Fermi velocity. The low energy Hamiltonian at the $\mathbf{K}'$ valley is related to the $\mathbf{K}$ valley through time reversal symmetry, i.e., $\mathbf{H}(\mathbf{q}') = \mathbf{H}^*(\mathbf{q})$, where $\mathbf{q}' = \mathbf{k} - \mathbf{K}'$. In this thesis, we operate under the assumption that the two valleys are decoupled. Focusing on the $\mathbf{K}$ valley hereinafter, the Hamiltonian has the following eigenvectors:

$$\boldsymbol{\Psi}(\mathbf{q}) = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{-i\theta_{\mathbf{q}}/2} \\ \pm e^{i\theta_{\mathbf{q}}/2} \end{bmatrix}, \tag{1.8}$$

Note these are the wavefunction coefficients. The full wavefunction is accompanied by a spatial contribution given by the basis which is formed by a linear combination of p_z orbitals of sublattice A and B. The eigenenergy is the linear dispersion:

$$\varepsilon(\mathbf{q}) = \pm \hbar v_F |\mathbf{q}|, \tag{1.9}$$

which is isotropic around the Dirac point $\mathbf{q} = 0$ and is electron-hole symmetric. The validity of the linear dispersion coming from the low energy approximation holds up to 0.6 eV, well above our experimentally accessible range.

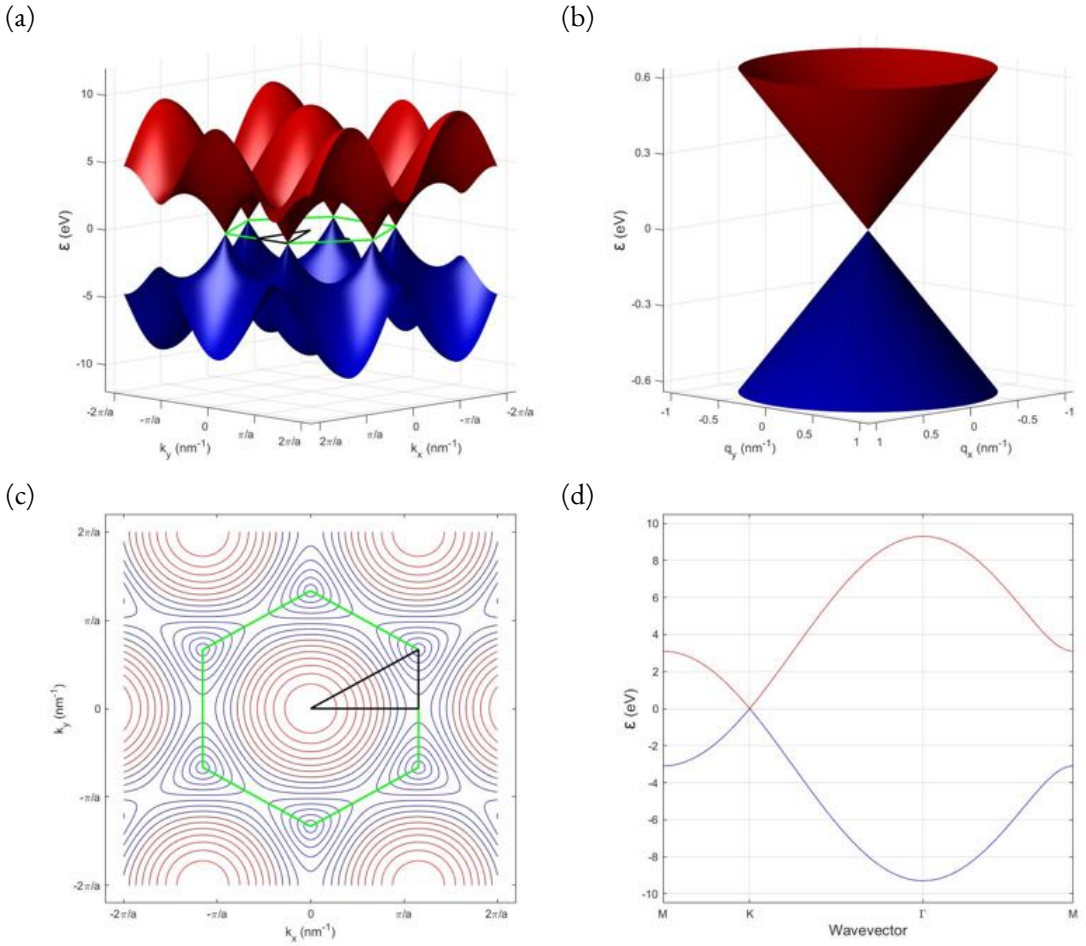


Figure 1.2: Illustration of the dispersion relation (ε vs $\mathbf{k}$) of graphene. (a) The bandstructure of graphene in the first BZ. Red denotes the conduction band and blue denotes the valence band. (b) The Dirac cone, i.e., the cone structure at the $\mathbf{K}$ and $\mathbf{K}'$ points. Here the dispersion relation is linear and symmetric. (c) The projection of the graphene bandstructure into the 2D momentum plane which shows the 2D Fermi surface of graphene. The contour lines represent constant energies. The green line marks the first BZ and the black line traces through the high symmetry points in graphene. $\mathbf{\Gamma}$ defines the origin of the BZ, $\mathbf{K}$ and $\mathbf{K}'$ are any two neighboring corners of the BZ. $\mathbf{M}$ is the midpoint between $\mathbf{K}$ and $\mathbf{K}'$ along the BZ edge. (d) The dispersion relation along the black line as seen in figure (c) which traces through the high symmetry points $\mathbf{\Gamma}$, $\mathbf{K}$, $\mathbf{K}'$ and $\mathbf{M}$.

1.2 Analogue to Relativistic Particles

Shortly after Albert Einstein developed the theory of special relativity, Paul Dirac sought to reconcile quantum mechanics with it. The result was the Dirac equation that describes relativistic particles with half-integer spin known as Dirac fermions. Particles that fall under this category all strictly follow Lorentz transformation laws. This is a direct consequence of the Lorentz invariance of the Dirac equation.

The low-energy Gr Hamiltonian is mathematically identical to the massless 2D Dirac equation except the characteristic speed given by the speed of light is replaced by the Gr Fermi velocity (see equation 1.7). The associated eigenenergy is a linear dispersion (see equation 1.9) which is a signature of a massless relativistic particle reminiscent of light. Because the mathematical treatment for the electrons in Gr and for the relativistic massless Dirac fermions are perfectly identical, safe for the scale factor difference, electrons in Gr are expected to behave as though they were actually massless and relativistic, i.e., in the framework where the speed of light is renormalized to the Gr Fermi velocity, Gr electrons play the role of light. And this has big ramifications: relativistic systems are rare and largely inaccessible. But here, Gr provides us a bench-top system with which we can simulate relativistic experiments.

The Fermi velocity renormalization has an additional perk. The characteristic interaction parameter, which is a measure of how influential particles are with one another, is given by the fine-structure constant. In vacuum it is:

$$\alpha = \frac{1}{4\pi\epsilon_0}\frac{e^2}{\hbar c} \approx \frac{1}{137}, \tag{1.10}$$

which can be interpreted as the ratio between the potential ($e^2/4\pi\epsilon_0 r$) and kinetic ($pc = \hbar c/r$) energy of the particles. ϵ_0 is the permittivity of free space. In Gr, an effective fine-structure can be

similarly defined:

$$\alpha_{\text{Gr}} = \frac{1}{4\pi\epsilon}\frac{e^2}{\hbar v_F} \approx 1. \tag{1.11}$$

$\epsilon = \epsilon_0\epsilon_r$ is the electric permittivity and ϵ_r is the dielectric constant of the system. For typical Gr devices, $\epsilon_r \approx 2$.[3] Here the ratio is drastically increased due to the reduction in the characteristic speed. This means that the interaction barrier among particles is lower and that they are much more susceptible to external influence. This allows us to access the elusive and largely uncharted highly-correlated relativistic regime. It should be emphasized that it is merely a mathematical similarity that connects Gr electrons with the truly relativistic particles. Gr electrons are still traveling at a measly $c/300$, which does not qualify them to be relativistic. The breakdown comes in when performing the pseudo-Lorentz transformation where the scale factor in the time component is adjusted. This type of transformation is nonphysical and is merely a mathematical trick that allows solving the Gr Hamiltonian conveniently. The correct observable result requires that we always return to the original frame of reference after a pseudo-Lorentz boost. Despite this however, we can still bring the insight and intuition developed from special relativity to describe Gr electrons and vice versa. Chapter 5 will report on the phenomenon of Landau level collapse which will substantiate a clear connection between Gr and special relativity.

1.3 External Electrostatic Potential

Under the influence of an external electrostatic potential, the Gr Hamiltonian becomes:

$$\mathbf{H}(\mathbf{q}) = \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q} + V(\mathbf{r}), \tag{1.12}$$

where $V(\mathbf{r})$ is the external electrostatic potential and its sidewalls have a characteristic length scale:

$$l_E = \sqrt{\frac{\hbar v_F}{|\nabla V(\mathbf{r})|}}. \tag{1.13}$$

This quantity defines the sharpness of the potential barrier.[4] In order for the low energy limit Hamiltonian to be valid, $V(\mathbf{r})$ must be smooth varying compared to the carbon-carbon bond length (a_{cc}), that is:

$$l_E \gg a_{cc}. \tag{1.14}$$

This is to avoid inter-valley scattering which allows the assumption of decoupled valleys. If the condition fails, electrons in Gr will no longer be described by the 2D massless Dirac Hamiltonian, which then renders our relativistic analogue invalid.

Under the condition that $V(\mathbf{r})$ is smooth compared to a_{cc}, there can exist sharp and smooth potentials relative to the Fermi wavelength (λ_F). For sharp potentials, $l_E \ll \lambda_F$, otherwise the potential is considered smooth. In the case of decoupled valleys, $\lambda_F = 2\pi/q_F \approx 10 - 100$ nm for typical experimental values of $\varepsilon_F \approx 0.01 - 0.4$ eV.

Using the carbon nanotube (CNT) as a narrow gating object while using hexagonal boron nitride (hBN) as the variable spacer between Gr and CNT, we can truly access both sharp and smooth regimes. We can confidently create potentials with sharpness that fall under $l_E \ll \lambda_F$ as well as $l_E > \lambda_F$. This is unlike lithographic gates which have at best an effective characteristic width in the high tens of nanometers, making the sharp potential regime unreachable. In chapter 6, we will show for the first time, due to the implementation of CNT as an extremely narrow gate, evidence of a single electronic guided mode in Gr.

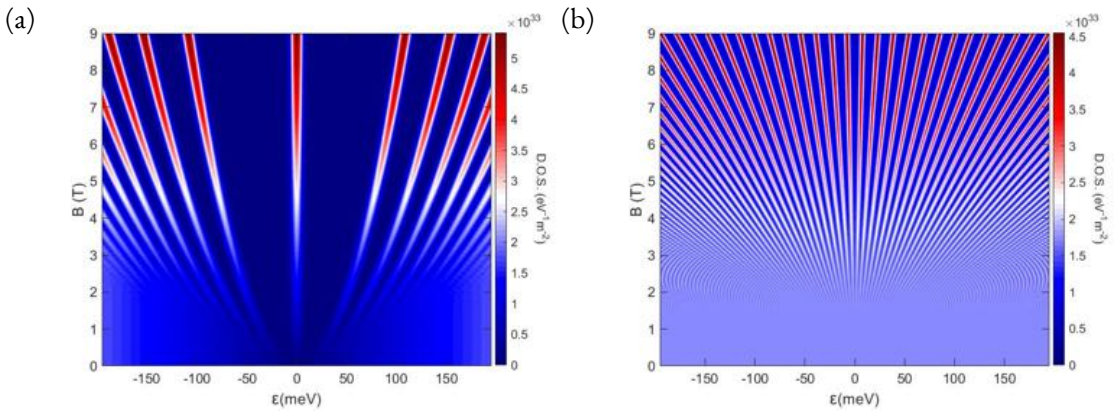


Figure 1.3: Landau levels of a relativistic and a non-relativistic system. (a) shows the continuous evolution of the LLs in the integer quantum hall regime of monolayer Gr, assuming that the LLs retain their four-fold degeneracy. The LLs show its non-linear evolution along the energy axis. Together with a $N = 0$ LL at $\varepsilon = 0$, the two are the trademarks of monolayer graphene's anomalous quantum hall effect. Gr here serves as the relativistic system. (b) shows the quantum hall effect of a non-relativistic parabolic system, also known as the conventional quantum hall effect. The electronic properties used are based on n-doped GaAs. The trademark of this conventional quantum hall effect is the absence of a LL at $\varepsilon = 0$ and that the LLs are closer together and are linearly evolving.

1.4 Quantum Hall Effect

Electrons confined in two dimensions and under the influence of a sufficiently large magnetic field will perform complete circular trajectories called the cyclotron orbit. These orbits have a characteristic length scale called the magnetic length:

$$l_B = \sqrt{\frac{\hbar}{eB}} = \frac{26\,\text{nm}}{\sqrt{B\,[\text{T}]}}. \tag{1.15}$$

If the magnetic length becomes the dominant length scale of the system: $l_B \ll l_m$, where l_m is the mean free path, electrons become localized and as a result they will exhibit discrete energy levels. To

see this, we perform Peierls' substitution, $\hbar\mathbf{q} \to \Pi \equiv \hbar\mathbf{q} + e\mathbf{A}$, and use the following definitions:

$$\begin{aligned} \mathbf{a} &= \frac{l_B}{\sqrt{2}\hbar}(\Pi_x - i\Pi_y), \\ \mathbf{a}^\dagger &= \frac{l_B}{\sqrt{2}\hbar}(\Pi_x + i\Pi_y). \end{aligned} \tag{1.16}$$

The low-energy Gr Hamiltonian for the $\mathbf{K}$ valley will then become:

$$\mathbf{H}(\mathbf{q}) = \sqrt{2}\frac{\hbar v_F}{l_B}\begin{bmatrix} 0 & \mathbf{a} \\ \mathbf{a}^\dagger & 0 \end{bmatrix}. \tag{1.17}$$

The solution to this Hamiltonian is:

$$\varepsilon_{\mathrm{LL}} = \mathrm{sgn}(N)\frac{\hbar v_F}{l_B}\sqrt{2N}. \tag{1.18}$$

These discrete levels are called Landau Levels (LL). Note that there is non-trivial dependence on the different valleys and sublattices, the details of which are addressed in appendix D. The behavior of LLs are plotted in figure 1.3. The behavior of Gr LLs is in stark contrast to the ones from non-relativistic or parabolic band systems such as GaAs. Because LLs are robust features that depend on the presence of a magnetic field, they can be used as a gauge for magnetic field strength. In chapter 5, we will rely on the quantum Hall effect as a means to detect the pseudo-Lorentz invariance in Gr.

1.5 Hydrodynamics

Complex things can be made simple. Such is the case for the highly correlated electrons in Gr. Like the N-body problem, it is an extreme challenge to analytically describe the collective behavior of many

electrons while taking into account their interactions with each other. Typically, one has to resort to symmetries or assume approximations in order to calculate their behavior. It turns out that under a certain condition, macroscopic quantities that characterize a highly correlated electronic system can be described by the classical laws of hydrodynamics. The condition is:

$$\tau_{ee} \ll \tau_p, \tag{1.19}$$

where τ_{ee} is the electron-electron (e-e) scattering time and τ_p is the momentum-relaxing scattering time. If electrons interact with each other much more frequently than with anything else, the electrons together can behave as a single entity with macroscopic properties. This condition bridges the electrons to the constituents of classical objects like water or honey. Inside these liquids are microscopic molecules that are constantly bombarding one another. The microscopic behavior is chaotic and messy; however, over sufficiently large distance and time scale, they form one coherent object that behaves simply and predictably. Electrons can be no different. This hydrodynamic regime has been highly sought after in electronic systems. To date there are only a few observations of hydrodynamic behavior of electrons, including in Gr.[5,6,7,8] But unlike the few other systems, Gr provides a unique and novel venue for studying hydrodynamics due to its exotic physical properties, such as those described above. One prediction is its ability to host a perfect fluid.[9] In chapter 7, we will show our attempt in studying this phenomenon with a carbon nanotube.

1.6 The Small but Mighty Carbon Nanotube

The carbon nanotube (CNT) is a rolled-up graphene (Gr) and was first discovered in 1991— over a decade prior to the isolation of Gr.[10] Despite having a longer history, there is arguably more to unravel

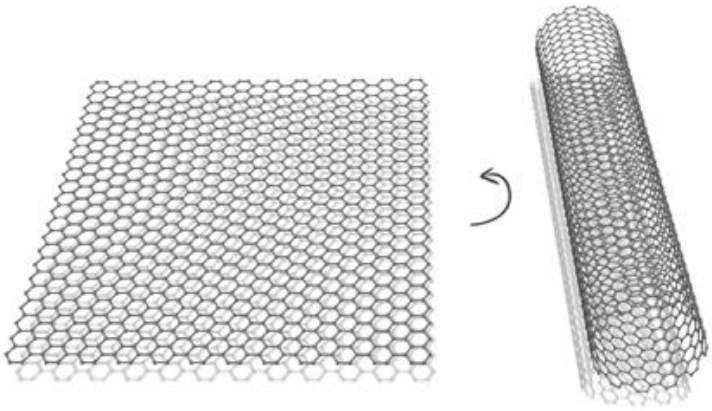


Figure 1.4: Graphene rolling up to be become a carbon nanotube.

still for the CNT. Compared to Gr, it exhibits a wider range of novel physical phenomena due to its one-dimensional (1D) nature. Electrons in the CNT are in a more confined environment which serves to amplify the many-body effects. However, the CNT is inherently more challenging to deal with due to its small size, high sensitivity to external perturbations and statistical acquisition of its kind. In chapter 3 we will delve into what is meant by the statistical acquisition of CNTs, along with the overall fabrication process in incorporating the CNT into our study.

Despite the CNT being a superbly interesting system to study on its own right, here instead we focus on using it as a single electron transistor (SET). The CNT is a very small object— its size is on the same scale as our DNA (a nanometer wide and potentially meters long). It is as close as possible to a perfect 1D system and can be shaped to behave as an atomic-like object, i.e. point-like with discrete energy levels, where electrons are further confined along the tubular axis. The CNT therefore naturally lends itself as a SET with which we can use to sensitively probe the dynamics of nearby charge. Moreover, the CNT is mechanically robust and comes in many different electronic flavors, all of which

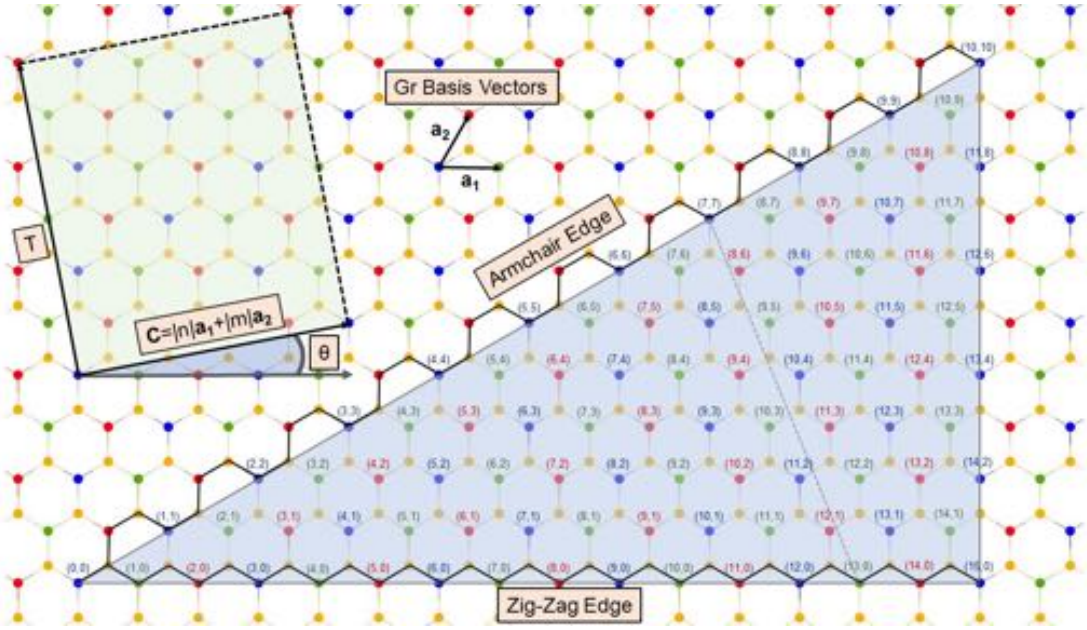


Figure 1.5: The periodic table of carbon nanotubes. As illustrated in figure 1.4, a carbon nanotube is merely a rolled-up graphene. There are, however, a myriad of ways to roll up the graphene and they all end up with different properties. Shown here is a graphene lattice. To make a carbon nanotube, the origin $(0, 0)$ is folded upon another labeled atom. Which atom it folds on to determines the chirality index of the carbon nanotube.

can be determined a priori and be integrated as a property of a nano-scale probe. Coupled with our ability to physically manipulate it, the CNT is an excellent choice for interacting with the nano-scale environment and is definitely a valuable selection among the existing arsenal of local probes. The operation of the CNT SET will be the basis for the majority of the work and will be covered in detail in chapter 2.

1.7 Final Foreword: Why the Hybrid CNT–Gr Device?

Conventional transport measurement primarily relies on fabricating electrical leads directly onto samples whereby a current (or voltage) is applied onto the material and the resulting potential drop (or current) is measured. Doing this typically only measures the global average behavior of conducting

carriers. If we want to study quantum materials like Gr, samples have to be pristine globally so that the quantum mechanical behaviors aren't hidden by disorder. This can be a big fabrication challenge. Phenomena like fractional quantum Hall effect, Hofstadter butterfly physics or hydrodynamics, all require extremely high mobility devices and fabricating them is not easy.

Techniques such as the scanning single electron transistor (SET) and scanning tunneling microscopy (STM) provide an advantage because they directly probe a local area of the sample that can be much smaller than the disorder length scale. Also, because of the locality of these techniques, both can uncover something that conventional transport cannot such as local density of states and the spatial profile of electrochemical potential.

The carbon nanotube (CNT) is a very natural candidate for a local probe. It's only about nanometer wide and it can behave like a quantum dot and so we can use it as an SET to perform sensitive charge detection. There is a very distinct advantage for using the CNT. Because it is so small it has a very small density of states which means that if we use it as an SET and bring it very close to a material for probing, the CNT itself won't screen out electron interactions, which is potentially a problem for metallic scanning SETs. Another advantage of the CNT is it can be used as a very narrow gate. There has been an ongoing development in trying to make electrical gates smaller; presently we are limited by the specifications of electron beam lithographic instruments. As of now, making nanometer wide gates that are smooth on the edges is still very challenging. But here, the CNT provides an immediate solution. Being able to physically manipulate the CNT allows us to noninvasively place it on materials to either perform local probing or local gating.

Gr is an obvious and interesting choice as the material under study. As mentioned in the above sections, Gr provides exotic phenomena many of which still remain to be unraveled, and at this point it is largely due to the limited techniques and tools we have for examining them. Conventional transport

measurements with a wide range of measurement schemes and configurations have been performed on Gr already. Arguably, it can be said that the transport results on a single Gr layer have more or less become saturated. The measurement of the hybrid CNT-Gr device allows us to showcase the additional functionality of the local CNT probe by revealing something new. Gr then serves as a good first benchmark for the utilization of the CNT local probe. The main reason though is Gr is an electronic platform that can be instrumental in pushing different fields in physics forward. It has been called CERN on one's desk[11] for the high energy physics opportunities it holds, has recently been shown to exhibit the novel hydrodynamic behavior, and has been considered to be the main system for electron optics. With the CNT, we look to further advance these respective areas in ways that conventional transport could not.

The structure of the thesis is as follows: chapter 2 will explain the operational principle behind the carbon nanotube local probe; chapter 3 goes into detail on how to fabricate the hybrid CNT-Gr devices; chapter 4 will motivate and introduce the concepts relevant to Landau level collapse and atomic collapse which are presented in chapter 5 and in parts of chapter 6 respectively; the majority of chapter 6 will be on the single guided mode in graphene where a short introduction about electron optics in graphene will be given. Chapter 7 will briefly introduce the concept of hydrodynamics and then present results on Coulomb drag between CNT and Gr; chapter 8 will be on the outlook of the hybrid CNT-Gr system and will also touch upon potential applications and additional experiments.

There's plenty of room at the bottom.

Richard Feynman

2

The Carbon Nanotube as a Local Probe

The carbon nanotube (CNT) has the structure of a graphene (Gr) sheet that is rolled up into a cylindrical tube. The diameter of these tubes are typically on the order of a few nanometers and the length can extend up to centimeters long, making them the materials with one of the highest length-to-diameter ratios. Their aspect ratios qualify them as quasi-one-dimensional and they can be considered as single molecules. Because the underlying lattice of the CNT is identical to Gr's, CNT and Gr share many properties such as exceptional mechanical strength and high mobility electrical transport. On top of these properties, however, CNT's one-dimensionality gives rise to many more intriguing ones that Gr lacks. For instance, a CNT can be either a semiconducting or metallic wire; it can be understood as a

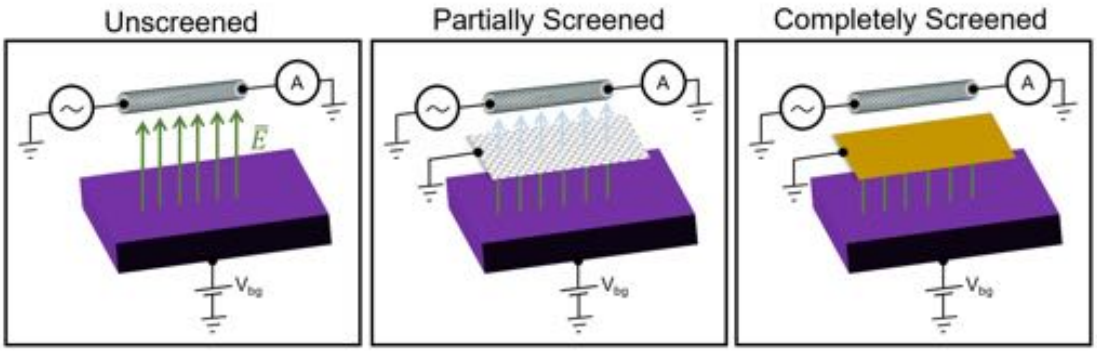


Figure 2.1: Illustration of screening. On the left there is nothing to screen the electric field from the backgate and the CNT sees all of it. In the middle is Gr which is partially screening the electric field because of its non-zero but small DOS. The amount screened depends on its DOS which can be tuned by doping the Gr. On the right is a piece of bulk metal such as gold that is screening everything from the backgate.

highly correlated Luttinger liquid[12,13] or a Wigner crystal[14]; it can exhibit a spin-orbit effect,[15,16] spin-charge separation[17,18] as well as a large Zeeman-splitting effect[19,20]. The most important property that we will use is its quantum dot (QD) behavior. At low enough temperatures, the CNT will exhibit Coulomb blockade due to the discrete nature of charge (a detailed description of CNT QD can be found in appendix C). In this chapter, I will show how this behavior serves as the backbone of the operational principle of the CNT probe.

2.1 Using the Carbon Nanotube to Detect Screened Electric Fields

Electric field screening means the damping of electric fields due to the rearragenment of mobile charge carriers. For instance, a lone charged particle with charge q in a static background medium with dielectric constant ϵ_r has the following potential:

$$\phi(r) = \frac{1}{4\pi\epsilon_0\epsilon_r}\frac{q}{r}. \tag{2.1}$$

If the charged particle were instead in an environment filled with mobile charged carriers of number density n, due to the dynamics of charge repulsion and attraction, a configuration will be reached in which carriers with opposite charges will surround the lone charged particle. The electric field from the charged particle will at least partially be canceled by the surrounding charges and produce a reduced net potential. In the linear screening regime the reduced potential is:

$$\phi(r) = \frac{1}{4\pi\epsilon_0\epsilon_r}\frac{q}{r}e^{-k_{\mathrm{TF}}r}. \tag{2.2}$$

Here, k_{TF} is the Thomas-Fermi screening wavevector and is given by:[21,22]

$$k_{\mathrm{TF}}^{\mathrm{3D}} = \sqrt{\frac{4\pi e^2}{\epsilon}\frac{\partial n}{\partial \mu}}, \tag{2.3}$$

$$k_{\mathrm{TF}}^{\mathrm{2D}} = \frac{2\pi e^2}{\epsilon}\frac{\partial n}{\partial \mu}, \tag{2.4}$$

for 3D and 2D confined electrons respectively. $\epsilon = \epsilon_0\epsilon_r$ is the material's permittivity. Note that the screening length (k_{TF}^{-1}) is reduced in lower dimensions. The equations tell us that the strength of the screened electric field depends on $\partial n/\partial\mu = g(\mu)$, the density of states (DOS) at the Fermi level, which is to say by measuring how effective a material screens an electric field, one can deduce its DOS. Bear in mind that this is just a simple description of screening. The screening of electric field in solid state systems is a complicated phenomenon but the general idea is still true— the screening of electric fields contains information regarding the DOS of the material. In this thesis, the main goal is to uncover the electronic properties of Gr and one of the key quantity is the DOS, or almost equivalently, quantum capacitance. The approach is to use the CNT as an electric field sensor to detect how the Gr is locally screening an electric field, thereby finding the local DOS of Gr. At the same time, the

CNT can be used as a tunable line charge to finely modulate the electronic potential landscape of Gr. As we shall see, the two functions together open up a whole new venue to explore within Gr and the uncovered truths reverberate to other paradigms in the realm of physics.

2.2 Clarification: Chemical Potential vs Electrochemical Potential

In this work, chemical potential (μ) strictly refers to the doping level with respect to the charge neutral state.* In the case of Gr, the chemical potential is in reference to its charge neutrality point (CNP, a.k.a. the Dirac point). It is an energy level that corresponds to the number of mobile free carriers in the system. The Fermi level (ε_F), sometimes used interchangeably with chemical potential, is equal to the chemical potential when $T = 0$, which is a condition we will generally assume unless specified otherwise. And like Gr, the chemical potential of the degenerately doped silicon backgate is also in reference to its charge neutral state.

The electrostatic potential ($e\phi$) is the energy coming strictly from electrostatics. Its role is to bend or shift energy bands in real space because it is changing the potential energy of all the electrons at a specific position. For example, in a double layer system, an interlayer bias creates an electrostatic potential difference between the two layers and causes the band of one layer to become offset entirely from the other (e.g. see figure 2.13). As one can see, a shift of an entire band relative to another will cause the chemical potentials of the two layers to become farther apart in the global energy scale even when there is no active doping taking place, i.e., even when the chemical potential is not moving farther

*Chemical potential is defined as $\partial\varepsilon(k_F)/\partial n = \mu$. It is the amount of work needed to add or remove an electron from the system. It is more informative to write it as: $\varepsilon(N) - \varepsilon(N-1)$, which is the total energy with N electrons minus the total energy with $N-1$ electrons (ignoring electrostatic energy). This happens to be the same as μ. On the other hand, the doping level is defined to be the energy level of the highest occupied state: $\varepsilon(k_F) = \mu$. The difference in definition is subtle; nonetheless, the quantities from the two definitions are equal.

away from its charge neutral state. To address this type of change, the term electrochemical potential can be handy. Electrochemical potential (μ_E) is defined to be the sum of the chemical potential (μ) and electrostatic potential ($e\phi$). There is a natural need for the use of this term especially in devices with gates and interlayer bias. However, in this thesis, the term electrochemical potential will generally be avoided to prevent confusion. The contributions from electrostatics and chemical doping will be pointed out separately whenever possible. There is a clear exception when dealing with CNTs where because of convention and convenience, electrochemical potential has better usage. More details on this will come when necessary. To see the role of electrochemical potential in CNTs, the reader is referred to appendix C.

2.3 Quantum Capacitance

Capacitance in the electrical sense is a measure of how easy it is to store electrical charge in a system, as seen in the definition:

$$Q = CV, \tag{2.5}$$

where Q is the total electric charge (Coulombs, C), C is the total capacitance (Farads, F) and V is the electrical potential or voltage (Volts, V). The larger the capacitance for a given voltage, the larger number of charges is stored in the system. This total capacitance comes from two contributions: geometric capacitance (C_E) and quantum capacitance (C_Q). Geometric or classical capacitance is dictated by the geometry and size of the system. It corresponds to the electric field between the system and its surroundings and can be calculated by Maxwell's equations from classical electrostatics. Linking back to the general idea of capacitance, this electric field represents the force that the charges feel and translates into the energy needed to maintain the charges in their static configuration. Quantum capacitance is

the capacitance that accounts for the intrinsic ability of a material to hold charge, and hence is a material dependent property— as one puts more charges into the material, inherently quantized and finite number of energy levels have to be filled according to Pauli's exclusion principle. This implies that if individual energy levels are spaced far apart, charges will experience a high energy barrier to enter the material which effectively reduces the material's capacity to hold charge. As it turns out, this point is completely captured by the density of states (DOS) of the material and so we can expect the quantum capacitance to be a reflection of the DOS.

Formally, the quantum capacitance is just a re-expression of the general definition of capacitance (equation 2.5) that highlights the energy cost of adding more electrons coming specifically from the quantum nature of energy levels:

$$C_Q = \frac{dQ}{dV} = e\frac{d(-eN)}{d(eV)} = e^2\frac{dN}{d\mu}. \tag{2.6}$$

Here, N is the number of electrons, e is the elementary charge ($-e$ for electrons), V is the voltage and $\mu = -eV$ is the chemical potential that strictly comes from doping (in contrast to electrochemical or electric potential that contain electrostatic contributions). The substitution comes from (while purposefully neglecting any electrostatic contributions):

$$V_1 - V_2 = \frac{\mu_1 - \mu_2}{-e}, \tag{2.7}$$

where we have set the reference voltage (V_2) and chemical potential (μ_2) to zero and V_1 is the system's voltage and μ_1 is the system's chemical potential. Since the quantum capacitance is normally reported in terms of per unit length, area or volume, following this convention, the quantum capacitance (per

unit length, area, or volume) is:

$$\begin{aligned} C_Q &= e^2 \frac{dn}{d\mu} \\ &= e^2 \frac{d}{d\mu} \int_{-\infty}^{\infty} d\varepsilon g(\varepsilon) f(\varepsilon) \\ &\approx e^2 \frac{d}{d\mu} \int_{-\infty}^{\mu} d\varepsilon g(\varepsilon) + \sum_{m=1}^{\infty} [\mathcal{O}((T/\mu)^{2m})] \\ &\approx e^2 g(\mu), \quad \text{for } T \to 0, \end{aligned} \tag{2.8}$$

where n is the number density of charges, T is temperature, $g(\varepsilon)$ is the density of states per unit volume and $f(\varepsilon)$ is the Fermi function. In the above sequence we have employed the Sommerfeld expansion and ignored higher order terms at the end assuming temperature is small. Note that $dn/d\mu$ is also commonly referred to as the electronic compressibility.

Typically, quantum capacitance is not measurable because it is very large (high DOS) in bulk materials. The reason is in a measurement circuit, quantum capacitance is almost always in series with the geometric capacitance (see for example figure 2.2) in which case the total capacitance is given by:

$$C = \left(\frac{1}{C_Q} + \frac{1}{C_E} \right)^{-1} = \frac{C_Q C_E}{C_Q + C_E}. \tag{2.9}$$

From this we can see that if $C_Q \gg C_E$, then $C \approx C_E$. However, when materials becomes small such as the carbon nanotube or graphene, the DOS also becomes small, in which case, the total capacitance will have significant contributions from both the geometric and quantum capacitance. By pin-pointing the exact contribution from the geometric capacitance, the quantum capacitance and hence DOS of a system can be precisely extracted.

2.4 The Simple Backgated Graphene Device Examined

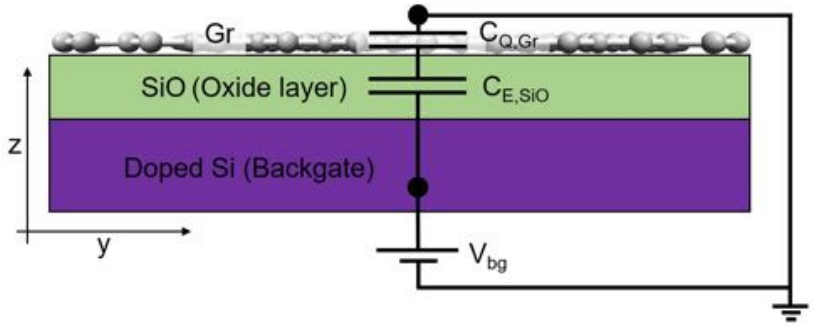
To eventually understand the carbon nanotube (CNT) and graphene (Gr) hybrid device that is the core of this thesis, it is beneficial to first look at how a lone Gr behaves with a global backgate as shown in figure 2.2.

By applying a fixed voltage to the backgate, charges (Q_{bg}) will populate the backgate to reflect the change in potential. These charges emanate electric field lines that go across the dielectric and will have to terminate at the Gr plane (otherwise it will go on forever and this is not allowed because it will take an infinite amount of energy). These electric field lines can only terminate if they attract opposite charges to the Gr. In order to attract a charge in Gr, however, an amount of energy associated with raising the chemical potential has to be paid. The electric field will have to weaken accordingly to accommodate new charges in the Gr. Once with the weakened electric field (by a reduction in number of charges in the backgate), the conditions for attracting charges to the Gr plane would have changed. As one can see, this becomes a circularly dynamical system. Nature, of course, adjusts perfectly and will settle into a stable configuration instantaneously. As for us, to calculate the quantities describing the system, one way is to solve a system of electrostatic equations iteratively and self-consistently. Fortunately, there exists an analytical method. We begin by looking at the total energy of the system:

$$\varepsilon_{\text{total}} = \underbrace{\frac{\epsilon}{2}\int_{\text{all space}} dr^3|\mathbf{E}|^2}_{\varepsilon_E} + \underbrace{\int_0^{\mu_{\text{Gr}}} d\varepsilon\, \varepsilon g_{\text{Gr}}(\varepsilon) + \int_0^{\mu_{\text{bg}}} d\varepsilon\, \varepsilon g_{\text{bg}}(\varepsilon)}_{\varepsilon_Q} + \underbrace{\int_0^{Q_{\text{Gr}}} dq V_{\text{bg}}}_{\varepsilon_S} \tag{2.10}$$

The first term (ε_E) comes from the electric field in between the backgate and graphene and it corresponds to the electrostatic energy. The second and third terms (ε_Q) are due to the filling of the bandstructures in graphene and backgate respectively that give rise to the changes in their Fermi lev-

(a)



(b)

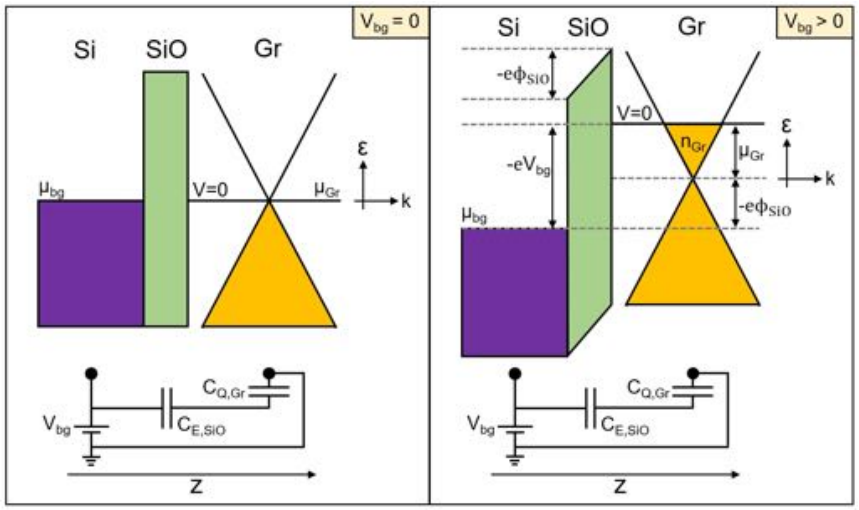


Figure 2.2: Graphene on a SiO_2/Si substrate. The Si is degenerately doped and acts as the global backgate. (a) shows the device from the side and overlain are the circuit elements associated with the device. V_{bg} is the backgate potential, $C_{E,SiO}$ is the geometrical capacitance of the SiO_2 and $C_{Q,Gr}$ is the quantum capacitance of Gr. The capacitance is in series. The top of (b) shows the energy vs real space diagram of the device with the left schematic showing when $V_{bg} = 0$ and the right showing when $V_{bg} > 0$. The bottom shows the circuit elements associated with the device. When a potential is applied to the backgate with respect to Gr that is grounded, an electric field is generated that shifts the energy bands in real space. The energy associated with this electric field is $e\phi$ is linked with the geometric capacitance ($C_{E,SiO}$). On top of the electric field generation, part of the energy goes into raising the chemical potential of Gr (μ_{Gr}) and this is linked with the quantum capacitance $C_{Q,Gr}$. The potential applied on the backgate dictates the entire energy scale: $e|V_{bg}| = |e\phi| + \mu_{Gr}$.

els. Here $g_{Gr(bg)}$ is the density of states of graphene (backgate). The last term (ε_S) comes from the external source, i.e. a battery connected to the backgate. Let us cast the equation into a more transpar-

ent form and rewrite it with experimental parameters. Firstly, the geometry of the device resembles a parallel-plate capacitor and so the geometric capacitance is given by:

$$C_{E,\mathrm{SiO}} = \epsilon_0 \epsilon_r \frac{A}{d} = \epsilon \frac{A}{d}, \tag{2.11}$$

where A is the area of the graphene, d is the thickness of the silicon oxide dielectric and ϵ_r is its dielectric constant. The electric field in the dielectric in between the backgate and Gr can be simply computed by Gauss's Law and is:

$$|\mathbf{E}| = \frac{Q_{\mathrm{bg}}}{\epsilon A} = \frac{N_{\mathrm{bg}}|e|}{\epsilon}, \tag{2.12}$$

where N_{bg} is the number of charge carriers in the backgate, $|e|$ is the elementary charge (here the charge type is irrelevant), and the field lines point in the direction perpendicular to the Gr plane. The electrostatic energy of the system is then:

$$\varepsilon_E = \frac{(N_{\mathrm{bg}} e)^2}{2C_{E,\mathrm{SiO}}}. \tag{2.13}$$

We will hold off in evaluating the expressions for the second and third terms (ε_Q) for now. As for the source term (ε_S) it simply becomes:

$$\varepsilon_S = Q_{\mathrm{bg}} V_{\mathrm{bg}} = -|e| N_{\mathrm{bg}} V_{\mathrm{bg}}. \tag{2.14}$$

Note the negative sign which means that for a positive backgate voltage, energy of the backgate decreases with respect to the electrical ground, which in turn implies the backgate's attraction of electrons since electrons want to migrate toward lower energies. To see the behavior of the Gr with an

applied backgate voltage, we proceed to find the minimum of the total energy:

$$\frac{\partial \varepsilon_{\text{total}}}{\partial N_{\text{bg}}} = 0 = \frac{N_{\text{bg}} e^2}{C_{E,\text{SiO}}} + \frac{\partial}{\partial N_{\text{bg}}} \left(\int_0^{\mu_{\text{Gr}}} d\varepsilon \, \varepsilon g_{\text{Gr}}(\varepsilon) + \int_0^{\mu_{\text{bg}}} d\varepsilon \, \varepsilon g_{\text{bg}}(\varepsilon) \right) - |e| V_{\text{bg}}. \tag{2.15}$$

To evaluate the terms with density of states, we can first take a look at the term regarding the backgate. The density of states of the degenerately doped silicon is really large and has an effective density of states $> 10^{19}$ cm^{-3}*— the energy levels are effectively squeezed into a single energy level.[23] This means that the chemical potential is extremely insensitive to additional doping (our samples are typically doped to $\sim 10^{12}$ cm^{-2}) and that the density of states and energy levels are basically constants and independent of number of carriers. Because of this, the term related to the backgate, once differentiated with respect to number of carriers (N_{bg}) would yield zero. As for the term regarding graphene, we can substitute the following for the density of states of graphene (see equation 2.8):

$$g_{\text{Gr}}(\varepsilon) = \left[\frac{\partial N}{\partial \varepsilon} \right]_{\text{Gr}}, \tag{2.16}$$

or alternatively, one can plug in the actual graphene density of states (see appendix A.6). Either way, with the additional condition of charge conservation:

$$N_{\text{bg}} + N_{\text{Gr}} = 0, \tag{2.17}$$

*The effective density of states is a quantity that describes carrier density and thus has units of cm^{-3}. It appears as a multiplier to a distribution function when calculating the carrier density and so it can be viewed as playing the role of density of states. This is why it is called "effective density of states" and can be interpreted as the actual density of states. To a good approximation, the density of states can be assumed to be equal to the effective density of states and so it would not be imprecise to say the effective density of states has the units of cm^{-3}/eV

it is easy enough to see what the term becomes. Simplifying and rearranging equation 2.15 then yields:

$$|e|V_{\text{bg}} = \frac{n_{\text{bg}}e^2}{C_{E,\text{SiO}}} + \mu_{\text{Gr}}, \tag{2.18}$$

where $n_{\text{bg}} = N_{\text{bg}}/A$ and, by following convention, we have switched to reporting $C_{E,\text{SiO}}$ in units of Farads per unit area (F/m^2). Translated into words, the above just means the external voltage source on the backgate creates an energy window (eV_{bg}) that contains the electrostatic potential energy ($e\phi$) and the chemical potential energy (μ). To be clear, the electrostatic potential in this case is:

$$e\phi = \frac{n_{\text{bg}}e^2}{C_{E,\text{SiO}}}. \tag{2.19}$$

In this thesis, eV will always be associated with potentials (hence energy windows) created by applied external sources whereas $e\phi$ will be reserved for the electrostatic potentials induced within the measurement device.

And so now, after finally establishing the energy picture describing our measurement of a backgated graphene device, let us calculate the expected induced carrier density in graphene with a given backgate voltage. Let us assume the following conditions: $V_{\text{bg}} = 10$ V, $d = 300$ nm, $\varepsilon_r = 4$ (for SiO_2), and Gr is perfectly undoped. Using Gr's dispersion relation: $\mu_{\text{Gr}} = \hbar v_F \sqrt{\pi n}$, one can calculate through equation 2.18 that $n = 7.2950... \times 10^{11}$ cm^{-2} (note that we did not use n_{bg} here because the subscripts indicate the density induced purely by the backgate which is not appropriate). Now, ignoring μ_{Gr} and simply calculating the carrier density induced through electrostatics (so $V_{\text{bg}} = n_{\text{bg}}e/C_{E,\text{SiO}}$) using the same parameters, one will get $n_{\text{bg}} = 7.3685... \times 10^{11}$ cm^{-2}, which is only a small deviation from when taking into account μ_{Gr}.This leads us back to the central topic of quantum capacitance.

2.5 Quantum Capacitance Revisited

The reason we get a smaller doping carrier concentration when taking into account μ_{Gr} is that part of the work coming from the external source is put into raising the chemical potential of Gr. It takes extra work to add electrons when DOS is small as alluded to in sections 2.3 and 2.4. This feature pertains to only quantum materials that have small DOS. Classical, or bulk, materials such as the degnereately doped silicon backgate itself would not exhibit this behavior. We can then attribute the deviation from the classical electrostatic prediction to a quantum mechanical effect and so for convenience, we can rewrite equation 2.18 as:

$$|e|V_{\text{bg}} = n_{\text{bg}}e^2\left(\frac{1}{C_{E,\text{SiO}}} + \frac{1}{C_{Q,\text{Gr}}}\right) \tag{2.20}$$

which is just expressing the total energy as the electrostatic potential plus something more. And it turns out this something more is consistent with and captured by the behavior of an extra capacitor in series as shown in equation 2.9. Such is the reason for the name: quantum capacitance. Bear in mind, however, quantum capacitance is unlike geometric capacitance in that it is not a constant but a function of the material doping level and hence a function of external voltages. The reader is reminded to be wary when seeing quantum capacitance represented as a circuit element— it is not a physical object but a quantum effect within a material. Although generally quantum capacitance can be treated as an external standalone capacitor, it must not be used naively in a device circuity as it may not manifest itself as expected. There will be one scenario where the quantum capacitance needs to be treated specially in which case a gray dotted box will be drawn around the "quantum capacitor".

2.6 Quantum Capacitance of Graphene

Obtaining the theoretical quantum capacitance is straightforward. It simply amounts to evaluating equation 2.8 for Gr:[24]

$$C_{Q,\mathrm{Gr}} = \frac{2e^2 k_B T}{\pi(\hbar v_F)^2} \ln\left[2\left(1+\cosh\frac{\mu}{k_B T}\right)\right], \tag{2.21}$$

with units of F/m^2, where $\hbar$ is the reduced Planck's constant, e the elementary charge, k_B the Boltzmann constant, v_F the Fermi velocity of Gr, and μ the chemical potential (or Fermi energy) of Gr. In the degenerate limit where $\mu \gg k_B T$, we obtain:

$$C_{Q,\mathrm{Gr}} \approx \frac{2e^2}{\hbar v_F \sqrt{\pi}} \sqrt{n}, \tag{2.22}$$

where n is the carrier concentration. Here, we recover the approximated result as seen in equation 2.8: the quantum capacitance is proportional to the DOS (see A.49). This theoretical model, however, assumes a perfect and pristine Gr. In reality, the Gr system always contains some amount of disorder due to charge impurities either in the substrate or within Gr itself. These will ultimately cause the formation of electron-hole puddles, which are long range charge disorder that causes spatial fluctuation in the Gr chemical potential.[25,26] Another type of disorder is short range atomic defects, which include atomic vacancies and grain boundaries. Although Gr is less susceptible to long range disorder it is not invincible. A manifestation of disorder is broadening of the Dirac peak meaning that the bandstructure of Gr gets modified. These factors hence play a role in the quantum capacitance of Gr and must be accounted for. To analytically incorporate the information regarding disorder is tricky. The best way to account for it is empirically by adding a phenomenological correction term and it turns out the

following form describes the observed behavior quite well:[27]

$$C_{Q,\mathrm{Gr}} = \frac{2e^2}{\hbar v_F \sqrt{\pi}} \sqrt{n_G + n_{\mathrm{imp}}}, \tag{2.23}$$

where n_G stands for the carrier concentration caused by the gating of Gr and n_{imp} stands for the intrinsic charged impurities of Gr. To give the reader some sense of how large the quantum capacitance is for Gr, for an applied backgate voltage of $V_{\mathrm{bg}} = 10$ to 90 V on a 300 nm thick SiO_2, which roughly translates to $\varepsilon_F = 0.1$ to 0.3 eV, $C_{Q,\mathrm{Gr}} \approx 20$ to 70 fF/μm^{-2}. Figure 2.3 shows the Gr quantum capacitance calculated by both equation 2.21 and 2.23. From this figure we can see that impurity and temperature both serve to smear out the Dirac cone.

2.7 Quantum Capacitance of the Carbon Nanotube

Having examined the lone Gr device extensively, what do we expect from the carbon nanotube (CNT)? CNTs are even smaller than Gr, so in some sense it is even more quantum mechanical. But before we get to the quantum capacitance of the CNT, it is worth mentioning its geometric capacitance. In our work, the CNT is always capacitively coupled to a piece of Gr. This geometry is described by a wire over an infinite plane and from classical electrostatics, one can calculate the geometric capacitance to be:

$$C_E = \frac{2\pi l \epsilon_0 \epsilon_r}{\cosh^{-1}(d/r)} \xrightarrow{d \gg r} \frac{2\pi l \epsilon_0 \epsilon_r}{\ln(2d/r)}, \tag{2.24}$$

in units of F, where l is the length of the CNT, ϵ_0 the permittivity of free space, ϵ_r is the relative permittivity or dielectric constant of the medium the system is in, d is the distance between the center of the CNT and the Gr plane, and r is the radius of the CNT. To give one an idea of the magnitude

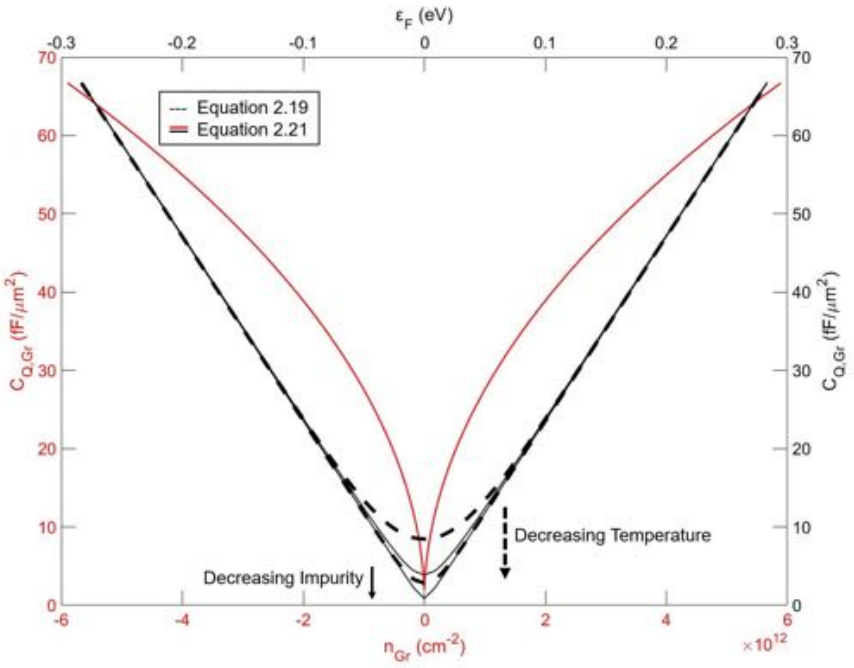


Figure 2.3: Quantum capacitance of graphene. The red line is plotted with respect to carrier density and the black lines are plotted with respect to energy. The line shapes change due to the energy and carrier density nonlinear relation. The dotted black line corresponds to the ideal theoretical model that only accounts for temperature and the full black and red lines corresponds to the semi-empirical model where impurity is used as a phenomenological fitting parameter. Temperature and impurity have a common effect which is to smear our the Dirac cone.The red line corresponds to 300K while the higher and lower black line corresponds to 300 and 100K respectively. The higher and lower dotted line corresponds to Gr having an intrinsic impurity level of 20×10^9 and $1 \times 10^9 \text{cm}^{-2}$ impurity concentration.

of the geometric capacitance, assume that $\epsilon_r = 4$ (for hBN, the material used for separating the CNT from the Gr), $r = 1$ nm and $d = 5$ nm, then in capacitance per unit length, $C_E \approx 100\,\text{aF} \cdot \mu\text{m}^{-1}$.

As for the quantum capacitance of CNT, following the same treatment as Gr (i.e., evaluate equation 2.8 but using CNT's DOS), if it is metallic then:

$$C^M_{Q,\text{CNT}} = \frac{8e^2}{hv_F} \approx 310\,\text{aF} \cdot \mu\text{m}^{-1}, \tag{2.25}$$

where we have used metallic CNT's DOS (see B.16). If it is a semiconducting CNT, one can only carry out the calculation numerically. Below we cite the semi-empirical analytical formula from the reference:[28]

$$C_{Q,\mathrm{CNT}}^{SC,n} = \sum_{q=1}^{2} \frac{e^2 n_e(q)}{k_B T}\left[1 - \frac{(2N_0 e^{x_n} - n(q))(\alpha + 2\beta x_n + 3\gamma x_n^2)}{2N_0 e^{x_n}}\right], \tag{2.26}$$

$$C_{Q,\mathrm{CNT}}^{SC,p} = \sum_{q=1}^{2} \frac{e^2 n_h(q)}{k_B T}\left[1 - \frac{(2N_0 e^{x_p} - n(q))(\alpha + 2\beta x_p + 3\gamma x_p^2)}{2N_0 e^{x_p}}\right], \tag{2.27}$$

where for equation brevity we have defined the following:

$$\begin{aligned} x_{\mathrm{n}} &= \frac{2\varepsilon_{\mathrm{F}} - \varepsilon_{\mathrm{g}}(q)}{2k_B T}, \\ x_{\mathrm{p}} &= \frac{-2\varepsilon_{\mathrm{F}} - \varepsilon_{\mathrm{g}}(q)}{2k_B T}. \end{aligned} \tag{2.28}$$

Here, $C_{Q,\mathrm{CNT}}^{SC,n}$ and $C_{Q,\mathrm{CNT}}^{SC,p}$ are the quantum capacitances for the conduction and valence bands respectively. $n_{e(h)}(q)$ are the carrier densities of electrons (holes) of subband q and are given by equations B.23 and B.24; N_0 is the effective DOS given by equation B.25; $\varepsilon_g(q)$ is the bandgap of subband q; α, β, γ are the fitting parameters. q refers to the subband index starting from closest to the bandgap. The reader is referred to appendix B.5 for more information on these quantities. Figure 2.4 shows the quantum capacitance of CNTs with different chiralities and hence diameters. From the plot we can see that semiconducting CNTs with larger diameters have a smaller bandgap, also, semiconducting CNTs have a larger quantum capacitance compared to the metallic CNT which has a non-varying and smaller quantum capacitance. Finally, here we see that unlike Gr, the quantum and geometric capacitance of the CNT is very similar in order of magnitude and so generally when considering the overall capacitance regarding the CNT, one needs to account for both types of capacitances by adding them

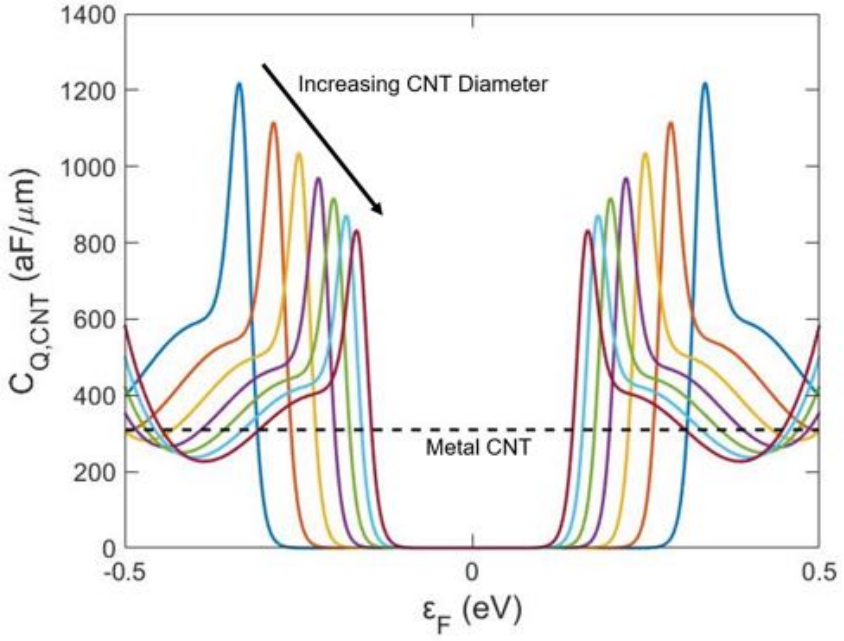


Figure 2.4: Quantum capacitance of the carbon nanotube. Plotted are the quantum capacitances of semiconducting CNTs with different diameters as well as a metallic CNT which is denoted by the dashed black line. The CNT diameters included are: 17, 20, 23, 26, 29, 32 and 35nm. Semiconducting CNTs with larger diameters have a smaller bandgap and they also generally have larger quantum capacitances compared to metallic CNTs.

in series as in equation 2.9.

Thus far we have assumed that the CNT behaves as a perfect 1D object, allowing itself to retain its 1D DOS. When temperature is low enough ($\varepsilon_C \gg k_B T$), the CNT will behave as a quantum dot (QD). In this thesis, however, the CNT QD falls in the classical Coulomb blockade regime where the confinement energy is not significant compared to the charging energy ($\Delta\varepsilon_{\text{conf}} \sim k_B T \ll \varepsilon_C$). Roughly, for our $\sim$ 500 nm long CNT QD devices at $\sim$ 1.6 K, the characteristic energies roughly

are:[29,30,31]

$$\Delta\varepsilon_{\text{conf}} = \frac{\hbar v_F}{2L} \approx 600\,\mu\text{eV},$$
$$k_B T \approx 200\,\mu\text{eV}, \tag{2.29}$$
$$\varepsilon_C = \frac{e^2}{C_{00}} \approx 5\,\text{meV}.$$

For the self-capacitance we used the following:

$$C_{00} = \frac{C^M_{Q,\text{CNT}} C_E}{C^M_{Q,\text{CNT}} + C_E} \approx 50\,\text{aF}, \tag{2.30}$$

where we assumed $C_E \approx 100\,\text{aF} \cdot \mu\text{m}^{-1}$ as before. In this classical Coulomb blockade regime, the DOS is not discretized and consequently, the quantum capacitance is simply given by equation 2.25 to 2.28. The reader is referred to appendix C.4 for more information on CNT QD DOS. To be more convincing, the experimental measurements of the CNT Coulomb diamonds in this work (see figure C.5) also show a high level of regularity as well as an even larger charging energy ($\sim$ 20 meV). These imply that confinement energy is negligible and substantiate that we are truly in the classical Coulomb blockade regime.

Much like the case for Gr, figure 2.5 shows how a gated CNT QD similarly behaves. When a potential is applied to the backgate with respect to CNT that is grounded, an electric field is generated that shifts the energy bands in real space. The energy associated with this electric field is $e\phi$ and is linked to the geometric capacitance ($C_{E,\text{SiO}}$). The rest of the energy goes into raising the chemical potential of CNT (μ_{CNT}) and this is linked to the quantum capacitance $C_{\text{Q,CNT}}$. The potential applied on the backgate dictates the entire energy scale: $|eV_{\text{bg}}| = |e\phi| + |\mu_{\text{CNT}}|$. Appendix C addresses more in-depth the transport behavior of a QD, and the reader is referred there for more details.

All this being said, however, the quantum capacitance of the CNT is not a critical quantity when

(a)

(b)

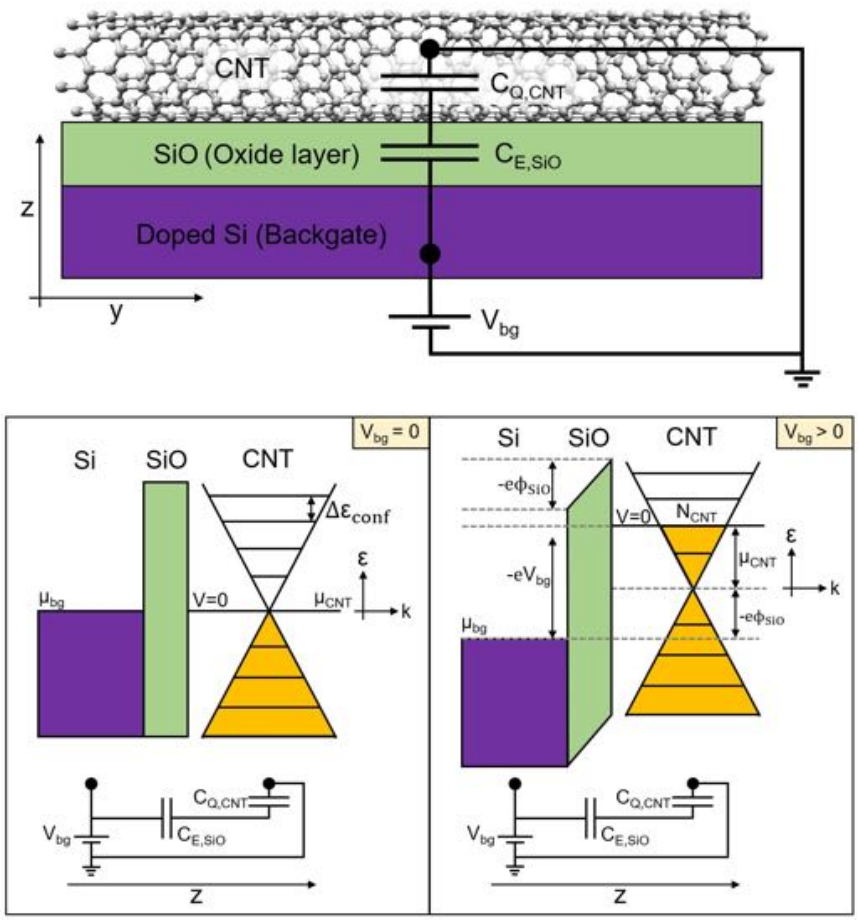


Figure 2.5: Carbon nanotube on a SiO_2/Si substrate. The Si is the global backgate. (a) shows the device from the side and overlain are the circuit elements associated with the device. V_{bg} is the backgate potential, $C_{E,SiO}$ is the geometrical capacitance of the SiO_2 and $C_{Q,CNT}$ is the quantum capacitance of CNT. The top of (b) shows the energy vs real space diagram of the device with the left schematic showing when $V_{bg} = 0$ and the right showing when $V_{bg} > 0$. The bottom shows the circuit elements associated with the device. The operation principle is the same as Gr. The only difference is the behavior of the quantum capacitance of the CNT.

determining the quantum capacitance of Gr. In the determination of the Gr quantum capacitance, we are only concerned with tracking a constant number of electrons in the CNT QD. We do not care about the energy level spacing or how the electrochemical potential is distributed between the geometric and quantum capacitances. All we care about is being able to point at a CNT energy level

and observe how it evolves with other experimental knobs. In the following section we will see how this all works.

2.8 Local Compressibility Measurement of Graphene with a Carbon Nanotube

Armed with all the information regarding the different capacitances, we are now ready to evaluate the CNT−Gr hybrid device. As a reminder to the reader, the motivation behind creating and studying this device is to use the CNT as a probe to measure the electronic properties of Gr. Specifically, we would like to measure the local density of states (LDOS) or quantum capacitance of Gr by using the CNT as a single electron transistor (SET). To do this, we need to observe carefully how the CNT behaves as Gr is changing, illustrated in figure 2.6. The CNT is separated from Gr by a thin hexagonal-Boron Nitride (hBN). The hBN thicknesses used in this work range from $4 - 100$ nm. In order to understand how the system behaves, we need to breakdown the energetics of the system and see how different elements are affected by one another. From section 2.4 we saw that the external voltage bias between layers defines an energy window that encompasses all the movements of energy levels happening in those layers. Following the same principle,[32] examining the V_{bg} effect on Gr (left of figure 2.6b), we have:

$$-e(V_{\text{bg}} - V_{\text{Gr}}) = -\mu_{\text{Gr}} - e\phi_{\text{SiO}}. \tag{2.31}$$

Here we have taken into account the potential of Gr (V_{Gr}). Keep in mind that the ground of the measurement is set to be the CNT, meaning the potential of the CNT is the reference potential of the entire system. The potential of Gr is not necessarily always zero in which case the potential difference between the Si backgate and Gr is given by $(V_{\text{bg}} - V_{\text{Gr}})$. ϕ_{SiO} is given by the electric field across the

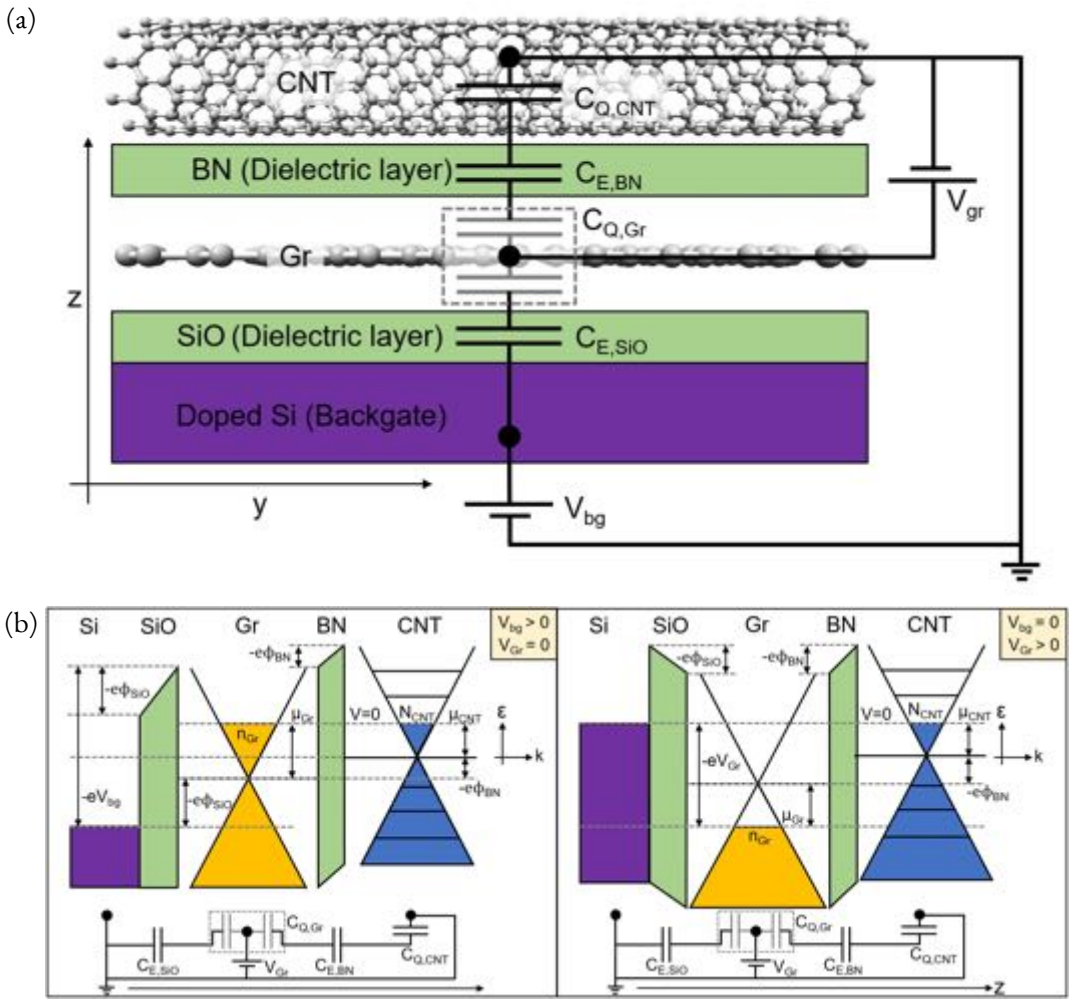


Figure 2.6: A CNT and Gr separated by a hBN sitting on a SiO_2/Si substrate. (a) shows the device from the side and overlain are the circuit elements associated with the device. V_{bg} and V_{Gr} are the backgate and Gr potential; $C_{E,SiO}$ and $C_{E,BN}$ are the geometrical capacitances of the SiO_2 and hBN; $C_{Q,Gr}$ and $C_{Q,CNT}$ are the quantum capacitance of Gr and CNT. The capacitance are all in series. The CNT is the reference ground. (b) shows the energy vs real space diagram of the device with the left schematic showing when $V_{bg} > 0$ and the right showing when $V_{Gr} > 0$. The bottom shows the circuit elements associated with the device. Note that $C_{Q,Gr}$ is bounded by a gray box. When considering the total capacitance between any two layers, only one $C_{Q,Gr}$ should be counted. Between V_{bg} and the CNT layer, there should only be one count of $C_{Q,Gr}$, not two. Separately, between V_{Gr} and V_{bg} or between V_{Gr} and the CNT layer, both should include one instance of $C_{Q,Gr}$.

SiO_2 layer (SiO in short). Following the same procedure shown in equations 2.11 to 2.13, by using

Gauss's law as well as charge conservation ($N_{\text{bg}} + N_{\text{gr}} + N_{\text{CNT}} = 0$), the equation turns into :

$$e(V_{\text{bg}} - V_{\text{Gr}}) = e^2 \frac{(n_{\text{Gr}} + n_{\text{CNT}})}{C_{E,\text{SiO}}} + \mu_{\text{Gr}}(n_{\text{Gr}}), \tag{2.32}$$

where we have now written the carrier number density dependence explicitly as this becomes an important quantity to keep track of. Examining the effect of V_{Gr} on the CNT (right of figure 2.6b), we have:

$$-eV_{\text{Gr}} = -\mu_{\text{CNT}} - \mu_{\text{Gr}} - e\phi_{\text{BN}}. \tag{2.33}$$

Once again, performing Gauss's law we get:

$$eV_{\text{Gr}} = \mu_{\text{CNT}}(n_{\text{CNT}}) + \mu_{\text{Gr}}(n_{\text{Gr}}) - e^2 \frac{n_{\text{CNT}}}{C_{E,\text{BN}}}. \tag{2.34}$$

Equations 2.32 and 2.34 are the core equations that contain all the measurement parameters and quantities of interest. This set of equations tells us how the two knobs (V_{bg} and V_{Gr}) are linked to all the other parameters of the system. Let us now discuss the measurement. The mode of operation is to change Gr's electronic properties by tuning the global Si backgate, and while doing so, we constantly monitor the conductance of the CNT and see how the CNT conductance changes. The idea is as the backgate is changing, Gr would screen the electric field coming from the backgate differently. This is due to Gr being a semi-metal with vanishing DOS near the charge neutrality point (CNP) and with large DOS away from the CNP (see appendix A for the theoretical description of Gr electronic properties). As the DOS of Gr increases away from the CNP, more electric field will get screened and the CNT will consequently become less sensitive to the backgate. This effect is demonstrated in figure 2.7.

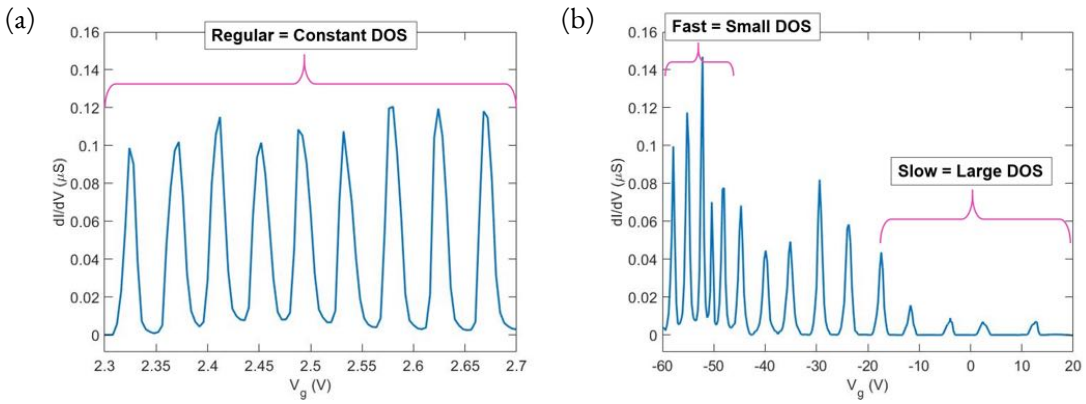


Figure 2.7: Coulomb blockade periodicity dependence on graphene density of states. The CNT conductance is measured as a function of the global gate (V_g). In (a), the device measured is like in figure 2.5. The CNT feels the electric field from the backgate completely and in this case the CNT behaves like a QD exhibiting regular Coulomb blockade peaks. In (b), the device measured is like in figure 2.6 where a Gr is in between the CNT and global gate. As the V_g is swept, the Gr gets doped and its DOS changes. On the left is when Gr is at the Dirac point where it has small DOS and so the CNT just sees much of the electric field from the global gate. As the gate is swept, Gr gets more doped and its DOS increases. The electric field from the gate in turn gets screened more and the CNT becomes less sensitive to the global gate and this results in the Coulomb blockade peaks becoming spaced farther apart.

The actual measurement, which we call compressibility measurement, is done as shown in figure 2.8. A small voltage is applied to the CNT and the resulting current is measured. The conductance is given by the dI/dV where we operate at zero DC bias. The two measurement knobs are V_{bg} and V_{Gr}. V_{bg} serves to dope the Gr. The unscreened electric field that goes through the Gr will dope the CNT, and this remnant electric field is what we want to detect. V_{Gr} serves to electrostatically energize the Gr, i.e., increase the potential energy of the electrons in Gr. The measurement is done at 1.6K where the CNT behaves as a QD which exhibits Coulomb blockade peaks. Each valley signifies a certain constant number of charges in the CNT (see appendix C for details on CNT QD). Sweeping both axes simultaneously and measuring the conductance of the CNT, we obtain figure 2.9. The lines here correspond to the regularly spaced Coulomb blockade peaks of the CNT QD. We see the lines follow an S-like trajectory and this is because of the changing DOS of Gr. One can gauge the relative DOS

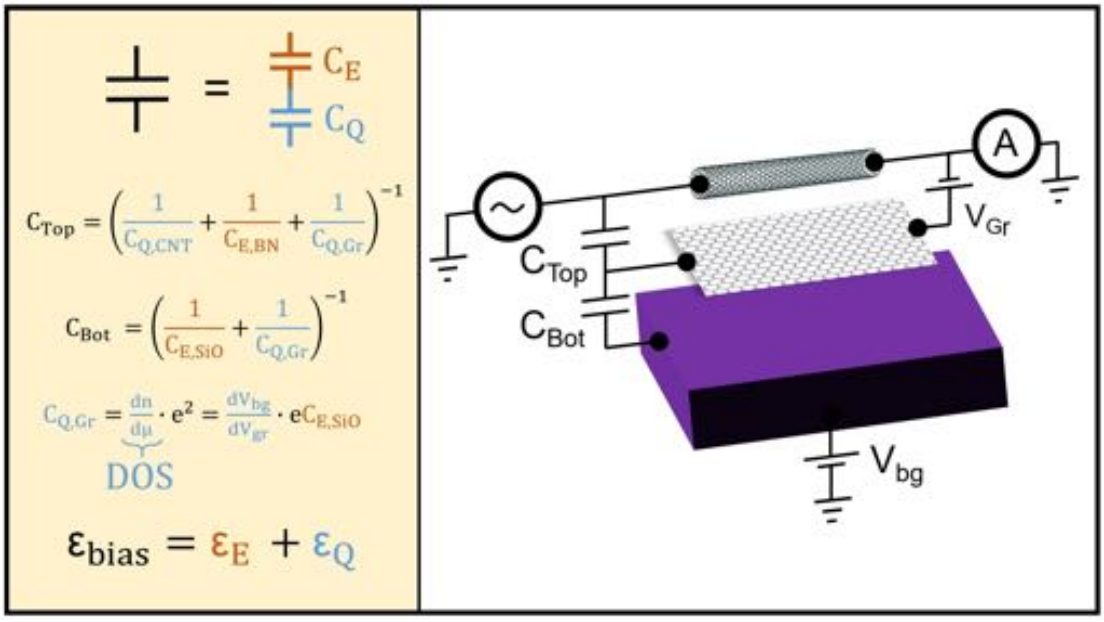


Figure 2.8: The measurement scheme for the carbon nanotube-graphene hybrid device. A small AC voltage excitation (typically 0.1 to 1mV) is applied onto the CNT and the resulting AC current is read through an ammeter (typically through a current pre-amplifier. The CNT is the ground reference for the system. Other biases such as the interlayer bias (V_{Gr}) or backgate bias (V_{bg}) is applied with respect to the CNT ground. On the left are notes on how the device can be broken down into circuit elements. The electrostatic contribution is in orange and the quantum mechanical contribution is in blue.

of Gr by looking at the slope of these lines. The reasoning is lines that are more perpendicular to one of the axes are more sensitive to that axis. A line that is closely parallel to an axis would mean that the signal has no dependence on that axis. To be more quantitative, let us start with equations 2.32 and 2.34. We will operate under the premise of keeping n_{CNT} constant. This entails only looking at one line in figure 2.9, since each line corresponds to a constant number of charges in the CNT. The idea is as before: to see how a particular line behaves with respect to V_{bg} and V_{Gr}. Grouping all constants together into Const. (including n_{CNT}), starting with equation 2.34 we have:

$$eV_{Gr} = \mu_{Gr} + \text{Const.}, \tag{2.35}$$

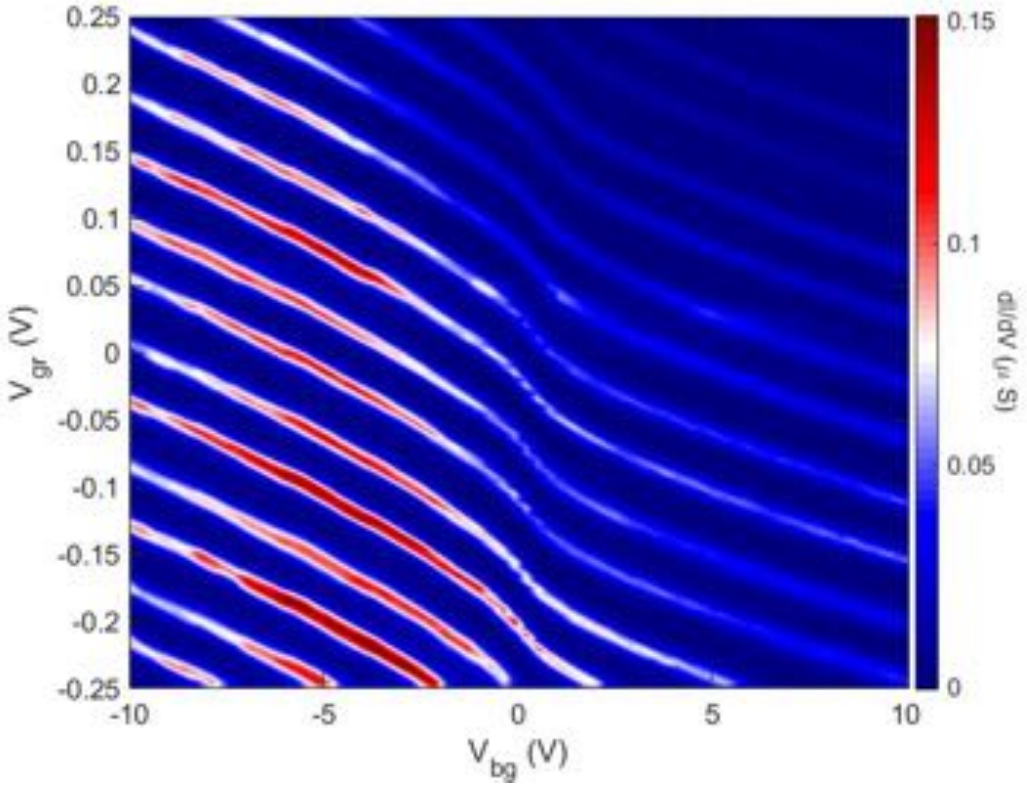


Figure 2.9: The compressibility measurement of graphene. This electronic compressibility is analogous to the mechanical compressibility. It is a measure of how the density of electrons changes with respect to their energy change, and this is equivalent to both density of states and quantum capacitance. The lines above correspond to Coulomb blockade peaks and they track the same electronic state of the CNT. As the backgate (V_{bg}) and Gr potential (V_{Gr}) vary, the CNT levels will move according to how sensitive they are to the screened electric field from the backgate. The CNT level sensitivity is marked by its slope and is a reflection of the screening strength and hence the DOS of Gr.

which is our first all-important equation. It says that lines of constant n_{CNT} tells us directly the chemical potential (or Fermi energy) of Gr. Similarly, for equation 2.32 we have:

$$V_{\text{bg}} = \frac{en_{\text{Gr}}}{C_{E,\text{SiO}}} + V_{\text{Gr}} + \text{Const.}, \tag{2.36}$$

then by differentiating the whole equation with respect to V_{Gr} and substituting in equation 2.35 we

obtain:

$$\frac{dV_{\text{bg}}}{dV_{\text{Gr}}} = \frac{e^2}{C_{E,\text{SiO}}}\frac{dn_{\text{Gr}}}{d\mu_{\text{Gr}}} + 1. \tag{2.37}$$

With some rearrangement:

$$g_{\text{Gr}}(V_{\text{bg}}, V_{\text{Gr}}) = \frac{dn_{\text{Gr}}}{d\mu_{\text{Gr}}} = \frac{C_{E,\text{SiO}}}{e^2}\left(\frac{dV_{\text{bg}}}{dV_{\text{Gr}}} - 1\right) \approx \frac{C_{E,\text{SiO}}}{e^2}\left(\frac{dV_{\text{bg}}}{dV_{\text{Gr}}}\right), \tag{2.38}$$

where $g_{\text{Gr}}(V_{\text{bg}}, V_{\text{Gr}})$ is the DOS of Gr as a function of V_{bg}. The approximation at the end is because the value one was taken out arbitrarily mainly for simplicity, but this is highly justifiable because $dV_{\text{bg}}/dV_{\text{Gr}} \gg 1$ always. To summarize, the important equations written in terms of reported quantities are:

$$\left.\begin{aligned} \mu_{\text{Gr}} &= eV_{\text{Gr}}, \\ C_{Q,\text{Gr}} &= C_{E,\text{SiO}}\left(\frac{dV_{\text{bg}}}{dV_{\text{Gr}}}\right), \end{aligned}\right\} \text{for } n_{\text{CNT}} \text{ constant.} \tag{2.39}$$

The above set of equations is the basis for analyzing all of our compressibility measurements. Compressibility measurement being the tracking of CNT energy levels as a function of V_{Gr} and V_{bg}. By performing these calculations on figure 2.9, we obtain figure 2.10a and 2.10b, which are plots for quantum capacitance ($C_{\text{Q,Gr}}$) and Fermi energy ($\varepsilon_F = \mu$) respectively. Despite going through all the algebra to arrive the equations, the beauty of this measurement is how it can be understood simply by examining the energetics of the system graphically. Figure 2.13 to 2.15 show in detail how the compressibility measurements can be examined.

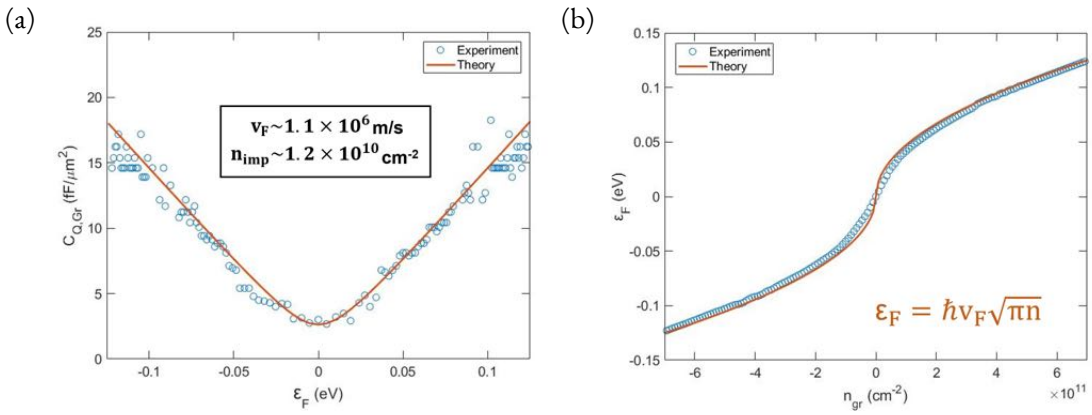


Figure 2.10: The extraction of quantum capacitance and Fermi energy of graphene. The extractions are based on equation 2.39. (a) shows the extracted quantum capacitance of Gr. The fit is from equation 2.23. (b) shows the extracted Fermi energy of Gr. The fit here assumes the linear dispersion of perfect Gr. The fit (orange lines) for figure (a) uses an intrinsic impurity level of $n_{\text{imp}} = 1.2 \times 10^{10} \text{cm}^{-2}$ and a Fermi velocity of $v_F = 1.1 \times 10^6 \text{m/s}$. The fit for the Fermi energy (figure (b)) deviates from experimental values near the charge neutrality point and the cause for it is most likely disorder.

2.9 Extraction of Quantum Capacitance: Method of χ^2 Test

Looking more closely into figure 2.10a, each blue circle is an extracted data point that reflects the slope of the CNT levels around the same V_{bg} and around $V_{\text{Gr}} = 0$ in the compressibility measurement plot (figure 2.9). The reason we say "around" is we need a finite sampling window to obtain the slope. In principle, only two points along the line defined by constant n_{CNT} are needed, but this severely limits the number of places we can extract slopes to just the peaks or valleys in the compressibility measurement, i.e. to places with identifiable features. To extract slopes everywhere, including in between peaks and valleys where there are no identifiable features, we assume that the profile of each sweep continuously evolves. We then define a sampling window around the point where we want to extract the slope. This sampling window must have at least two rows of data. In this window we pick for instance the top row of data points to be the reference. This row of data points traces out a local profile of the compressibility measurement along the x axis (because we are looking at rows). The neighboring row

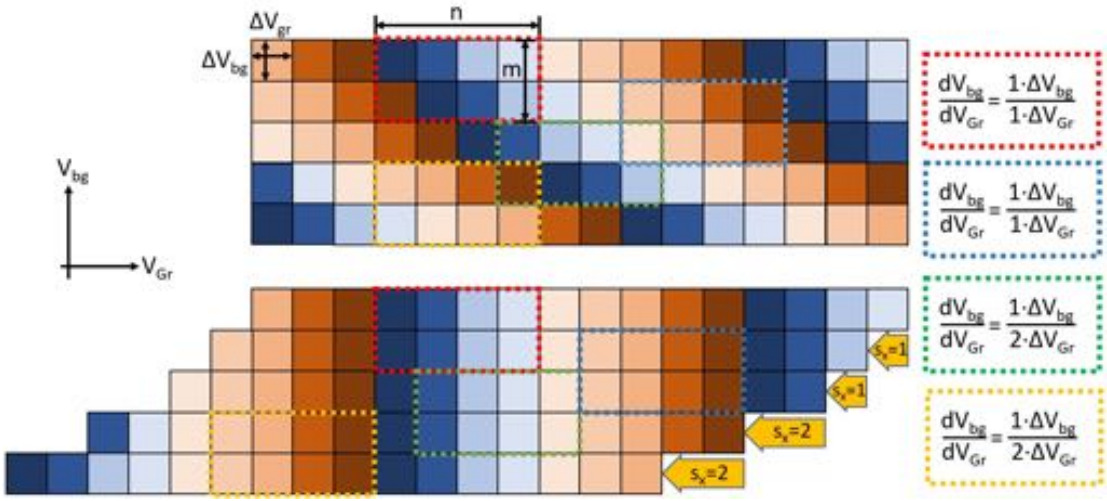


Figure 2.11: Slope extraction method. In the top picture are squares that represent individual data points taken in a measurement. Take for instace the deep blue to be a peak and deep red to be a dip in the signal. In the top picture there is a clear trend of the peak. It is slanted with a slight curve. To find the slope at each local position, we need to first define a sampling window as shown by the dotted square boxes ($n \times m$ number of data points). In each box we define the first row to be the reference row. The operation is to shift the neighboring row until the shifted row matches the reference row. This shift (denoted s_x) tells us the x component of the slope in the sampling window in terms of units along the x axis. To know whether the rows match best a χ^2 test is performed. If we look at the region defined by red dotted box, using the top row in the box as reference, we see that one unit of shift ($s_x = 1$) is required to make the neighboring row match. The slope in the red dotted box is then simply $\Delta V_{\text{bg}}/\Delta V_{\text{Gr}}$. Similarly for the yellow dotted box, the slope is $\Delta V_{\text{bg}}/2\Delta V_{\text{Gr}}$. The values of ΔV_{bg} and ΔV_{Gr} are step sizes of the compressibility measurement.

below the reference row will also trace out the same profile but it will be slightly shifted in one direction. To extract the slope, we digitally shift the second row relative to the reference row until the two rows match everywhere in the sampling window along the y axis. See figure 2.11 for a depiction. The shift along the x axis required to make the two rows match is the x component of the slope in units of the x axis unit. This is denoted by s_x. The y component of the slope in this sampling window is the y unit difference between the reference row and the shifted neighboring row. The y units shift can be more general than just immediate neighbors and is denoted by s_y. In figure 2.11, the x axis unit is ΔV_{Gr} and the y axis unit is ΔV_{bg}. Because digitally data points are discretized, shifts are discretized

movements. To calculate the actual slope inside the $n \times m$ sampling window at $\{i, j\}$, it is:

$$\left.\frac{dV_{\text{bg}}}{dV_{\text{Gr}}}\right|_{ij} = \left.\frac{s_y \cdot \Delta V_{\text{bg}}}{s_x \cdot \Delta V_{\text{Gr}}}\right|_{ij}. \tag{2.40}$$

s_x and s_y are the unit shifts along the x and y axes respectively and $\{i, j\}$ are the indices of the position of the sampling window. To decide when two rows are best matched after the shifts, we employ the χ^2 test. In the sampling window of $n \times m$ size (see figure 2.11) located at position $\{i, j\}$, we subtract the elements (denoted by z) in the same column of the two rows, square the difference, add them all together and divide it by the sum of the squares of each element in the reference row. Mathematically it can be expressed as:

$$\chi^2|_{ij} = \frac{\sum_n \left\{ z\big(\Delta V_{\text{bg}}[i+n+s_x], \Delta V_{\text{Gr}}[j+s_y]\big) - z\big(\Delta V_{\text{bg}}[i+n], \Delta V_{\text{Gr}}[j]\big) \right\}^2}{\sum_n z\big(\Delta V_{\text{bg}}[i+n], \Delta V_{\text{Gr}}[j]\big)}, \tag{2.41}$$

where $z\big(\Delta V_{\text{bg}}[i], \Delta V_{\text{Gr}}[j]\big)$ is the data point that is a function of V_{bg} and V_{Gr} located at index $\{i, j\}$ and n is the sampling window size along the x axis. We then seek to minimize the χ^2 quantity by varying s_x. The s_x that yields the lowest χ^2 will tell us the best match for shifted rows and thereby the slope. We typically incorporate three to five signal peaks in our sampling window size.

2.10 Local Compressibility Measurement Of the Quantum Hall Effect of Graphene

The compressibility measurement can be extended to detect the quantum effect of Gr. With the same measurement scheme but with the application of a magnetic field, and with the same method of data analysis detailed above, we can extract the DOS of the Landau levels (LLs) formed in Gr. Figure 2.12

shows the measurement and extraction results. When reading the compressibility measurement plots, the same logic as before applies. Each kink corresponds to a change in DOS of Gr. When the CNT level dips towards the V_{bg} axis, it means CNT is sensitive to the backgate which means Gr DOS is low. If CNT level plateaus with respect to V_{bg}, Gr DOS is high. What's happening here is we see the development of LLs and each kink is associated with one level. The LLs begin to appear around 2 T, then at 5 T we begin to see hints of degeneracy breaking. By 8T, all the low filling factors from $\nu = 1$ to 4 can be seen. For more information of the Quantum Hall effect of Gr, the reader is referred to Appendix D. This is yet another demonstration of the capability of the CNT local probe.

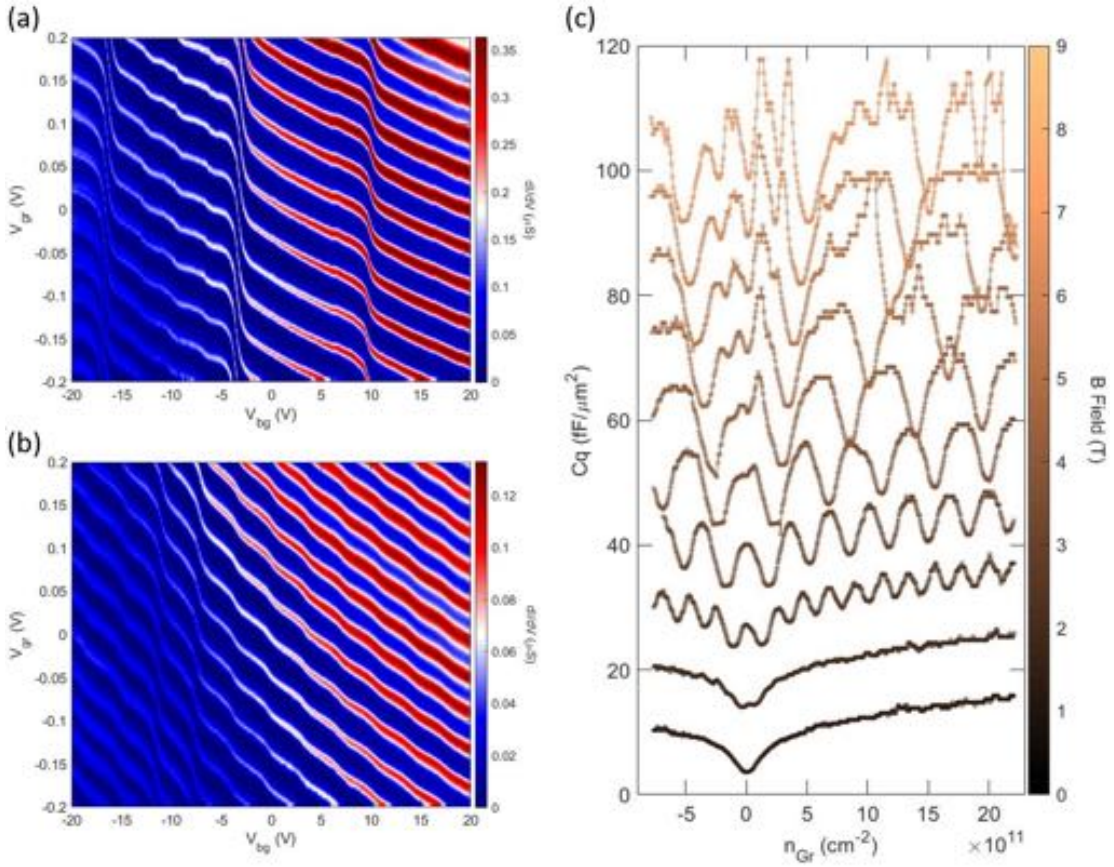


Figure 2.12: The local compressibility measurement of Quantum Hall effect in graphene. The measurement is done the same way as detailed in figure 2.8 except with the application of a magnetic field. (a) is measured at 9 T (b) is measured at 3 T. Compared to 0 T, the measurements show additional kinks in the CNT levels. Because V_{bg} is the x-axis, each kink corresponds to a sudden change in Gr DOS. As it dips, Gr DOS is low; as it plateaus, Gr DOS is high. At 3 T (figure (b)), the kinks are not as prominent compared to at 9 T (figure (a)) and this is because the LLs are not as big at lower magnetic fields. At 9 T, the kinks are obvious because the LLs are well formed. In addition to that, in between the large kinks, one can see the formation of smaller subtler kinks. These are due to the breaking of LL degeneracy. Figure (c) is the C_q extraction from 0 (lowest curve) to 9 T (highest curve) and increasing by 1 T for each offset. We can see that LLs begin to form completely by 2 T and we begin to see LL degeneracy breaking at around 5 T. At 7 T we see all filling factors from $\nu = 1$ to 4 completely develop.

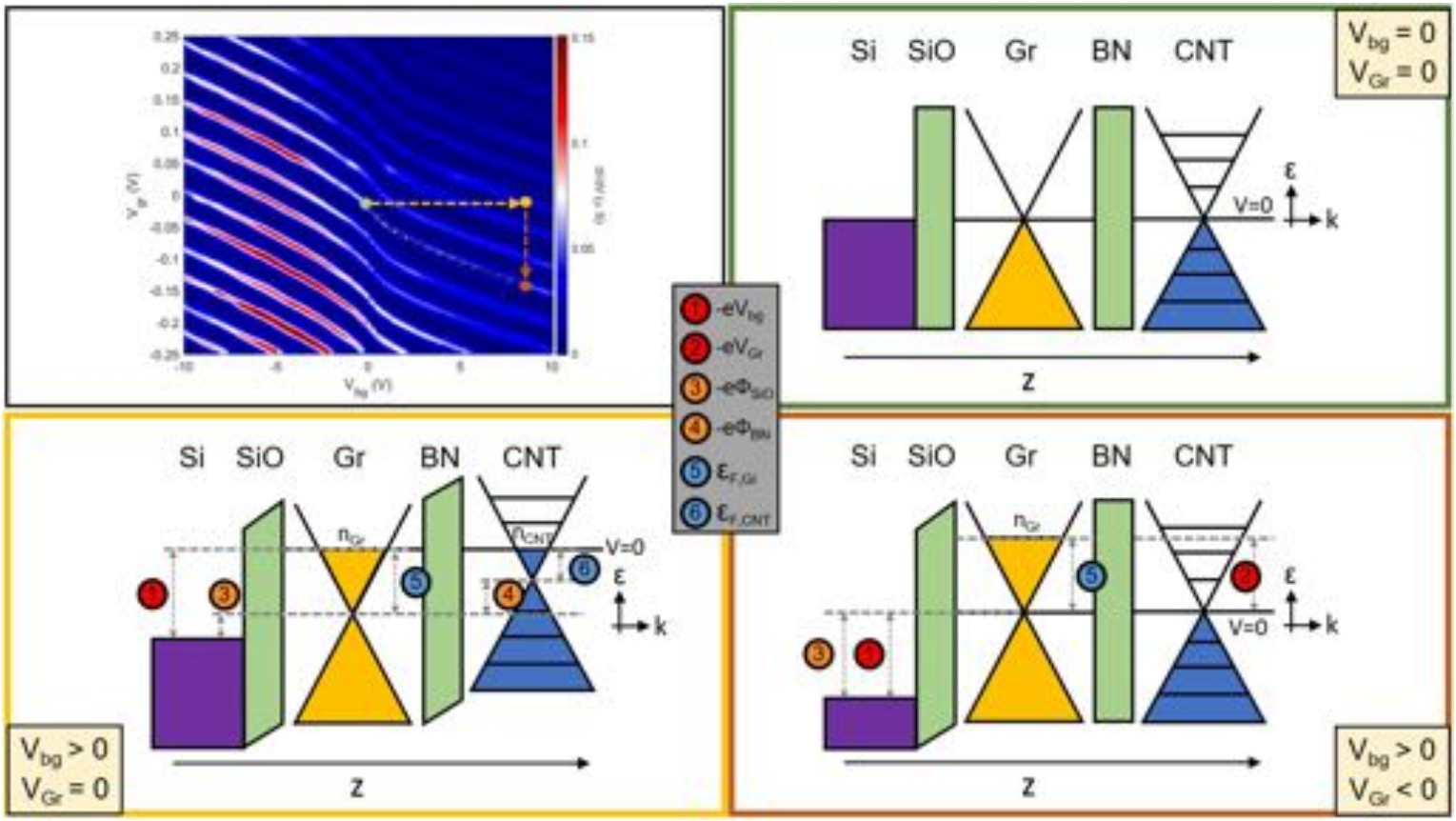


Figure 2.13: Demonstration of how the energy window provided by the voltage bias (red labels) are decomposed into an electrostatic (orange labels) and quantum mechanical (blue labels) components. By tracking the behavior of the CNT levels as a function of V_{gr} and V_{bg}, the quantum capacitance or density of states of Gr can be deduced. Energy shifts not drawn to scale. See main text for details.

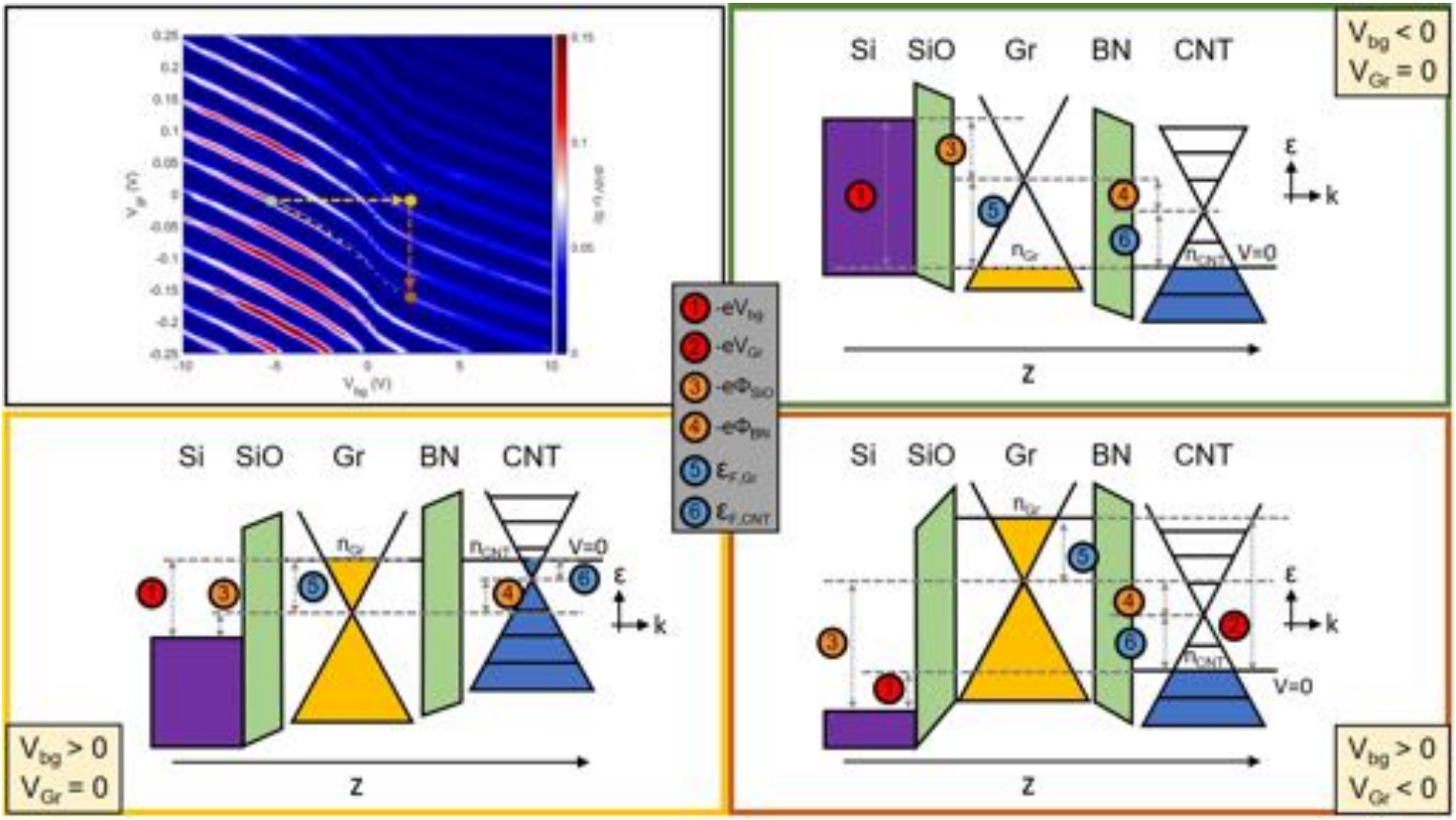


Figure 2.14: Another demonstration of how energies are distributed between the electrostatic and quantum mechanical components. Energy shifts not drawn to scale.

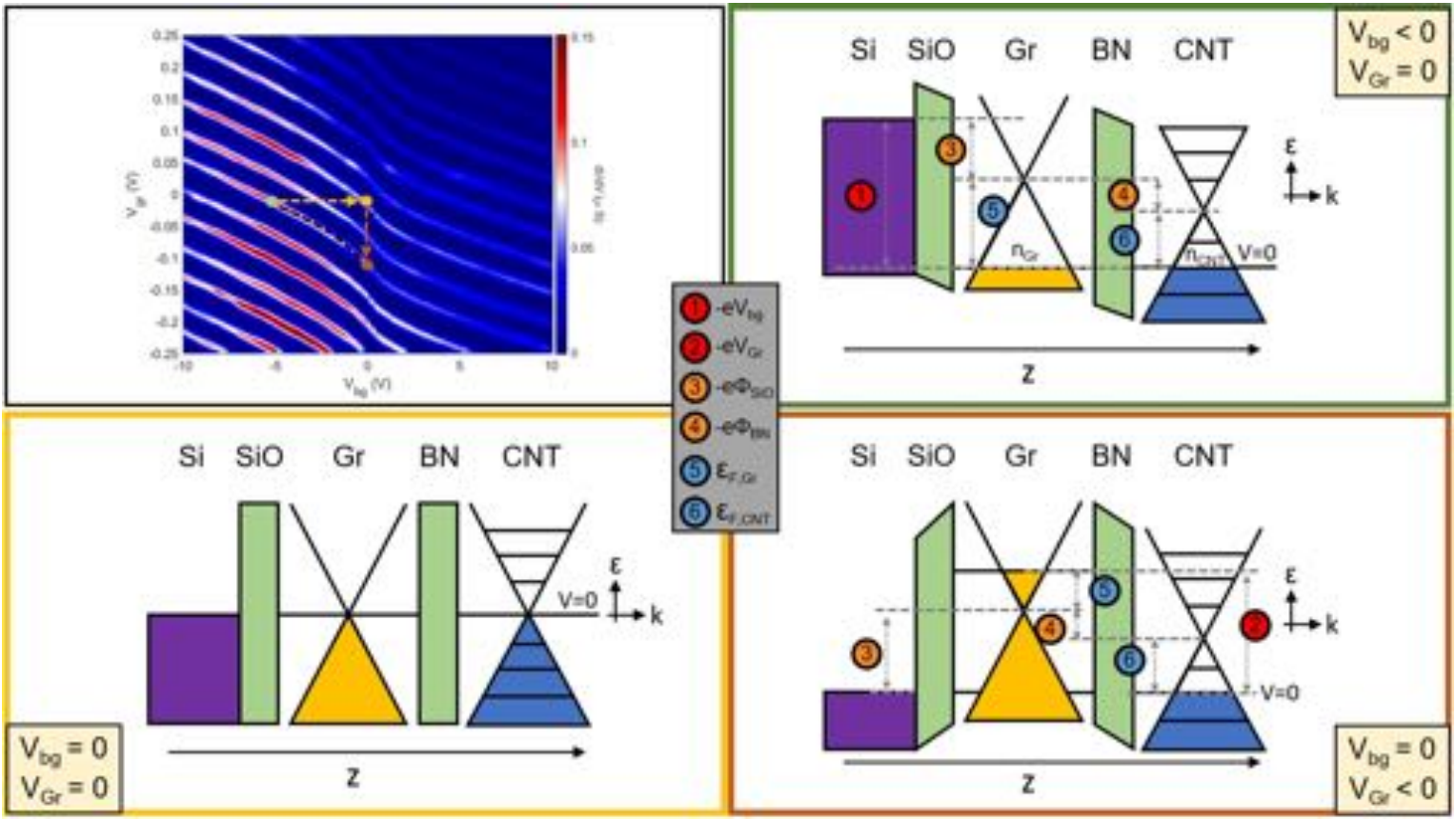


Figure 2.15: Another demonstration of how energies are distributed between the electrostatic and quantum mechanical components. Energy shifts not drawn to scale.

I have not failed. I've just found 10,000 ways that won't work.

Thomas Edison

3

Fabrication of the Carbon Nanotube and Graphene Hybrid Device

The experiments detailed in this thesis all involve the carbon nanotube (CNT)-graphene (Gr) hybrid device and the realization of such a device hinges on the capability of manipulating a single well-characterized CNT with a high level of controllability. This particular step poses great difficulty, but on top of that, merging it with material synthesis as well as the ensuing nanofabrication steps makes the whole process extremely grueling. Undoubtedly, the fabrication of this hybrid device is one of

the biggest challenges and tremendous amount of time, creativity and perseverance were put forth to make it work. In this chapter, I will lead the reader step by step through the whole fabrication process but will first begin with the challenge associated with CNTs.

3.1 Challenges Regarding Obtaining a Carbon Nanotube

Carbon nanotubes (CNTs) come in all types and sizes that are marked by their unique chirality. A single CNT can be anywhere from 0.3nm to tens of nanometers in diameter, and it can be metallic or semiconducting. There are hundreds of CNTs, each with its own unique property. The problem is up to this date there is no reliable way to obtain a certain type of CNT deterministically. The most popular method of acquiring CNTs is through chemical vapor deposition (CVD) and with this method, or any other method for this matter, CNTs are grown completely random. There are ongoing endeavors and some successes in chiral-selective growth of CNTs, however, the grown CNTs are either specific to one chirality or limited to a very small range of chirality.[33,34] There is still a need to deterministically obtain any type of CNT.

In this work, knowing what CNT we are dealing with is of utmost importance. The mesoscopic behavior of our devices are completely dependent on the nature of the CNT. To circumvent the problem that CVD CNT growth yields random CNTs, we accept this fact but we develop a method to characterize the grown CNTs and then controllably pick out the desired CNT from the mix of randomly grown CNTs. With this approach, we put ourselves into the best situation we are aware of, i.e., guaranteeing that we can have access to any type of CNTs. Acquiring CNTs with a certain property becomes a statistic game for us. As it turns out though, the statistics is not bad at all. The remaining challenge is that in developing this method, we need to ascertain the CNT be kept clean, long and straight.

3.2 The Road to Deterministic Carbon Nanotube Selection

Our carbon nanotubes (CNTs) are acquired through the popular method of chemical vapor deposition (CVD). In our synthesis, we grow suspended CNTs over a slit in a raised platform. This is to fulfill the purpose of being able to mechanically stamp a CNT onto a target area. This raised platform on our growth chip is shown in figure 3.1. The reason we want to grow CNTs over a slit is so that we can perform dark field Rayleigh spectroscopy to carefully characterize the grown CNTs. This subject will be covered in the following section (see figure 3.14 for illustration of dark field spectroscopy). As for the reason we want to grow CNTs on a raised platform, it is to guarantee a high success rate of stamping and transferring a CNT onto a target area. CNTs are nanometers in size; when attempting to transfer a CNT, the CNT must make sufficient physical contact with the target area. The physical contact is what provides the Van der Waals force that attracts the CNT down to the target area. If the target area or its vicinity has anything that is larger than the size of the CNT, there will be no hope of the CNT ever making contact with the target area. So to improve the conditions of making contact, we engineer growth chips that have a raised platform which will effectively reduce the area critical for establishing contact. The size of this platform was picked based on several factors: 1) the platform should be wide enough such that it can accommodate many CNTs— the wider the platform, the wider the slit, the larger the number of CNTs that can grow over the slit, which finally means we can have a larger selection of grown CNTs to choose from; 2) the platform should be small so that when it comes to transferring CNTs, only a small area (the raised platform) needs to be flush against the target area. If the platform becomes too big, we risk having more unwanted particulates preventing the platform from making flush contact with the target area; 3) the height of the platform should be high enough so that we don't need to worry about most dirt when the growth chip comes into contact with the

target area; 4) the height of the platform should not be too high to prevent CNTs from growing over the platform (the height of the platform affects the gas flow in the reactor furnace and can have an adverse effect on CNT growth). These four factors had to be juggled with and through trial and error we ended up with these parameters: slit width is 3 mm while the platform is 3.2 mm wide and $50-55$ μm tall. Finally, regarding the length of the slit, it can range from 30 μm to over 150 μm depending on the need. If longer CNTs are needed, then we will use the longer slits. The downside of using longer slits is that fewer CNTs will make it over the slit. Oppositely, shorter slits will yield extremely crowded growths which will make characterization and singling out an individual CNT more difficult. The most popular slit length we use is $60-80$ μm long.

3.3 Synthesis of Carbon Nanotube

Chemical vapor deposition (CVD) is the most widely-used growth method and is also our method of choice when it comes to producing carbon nanotubes (CNT); we choose this because it yields all types of CNTs and because it can produce extremely long, aligned and suspended CNTs. In our method we use argon as the main carrier gas. For the carbon source, we flow argon through ethanol. Hydrogen, which acts as the process gas, is used to initiate the reaction between the hydro-carbon and the catalyst. On top of all this, small amounts of water vapor is also injected to improve overall yield and cleanliness. All together, the gaseous ingredients are flown slowly into the hot zone of the furnace where the reactor gases meet the catalyst. During the encounter, carbon-containing gas is broken apart with the help of the process gas. The growth then follows the root-type growth (according to our observation) in which the CNT begins to form and extend out of the metallic nanoparticles in the catalyst. I shall not expound the growth mechanisms (for the chemistry detail please refer to the cited reference [35]) but instead simply state the guidelines formulated after the performance of many rounds of tests. Please

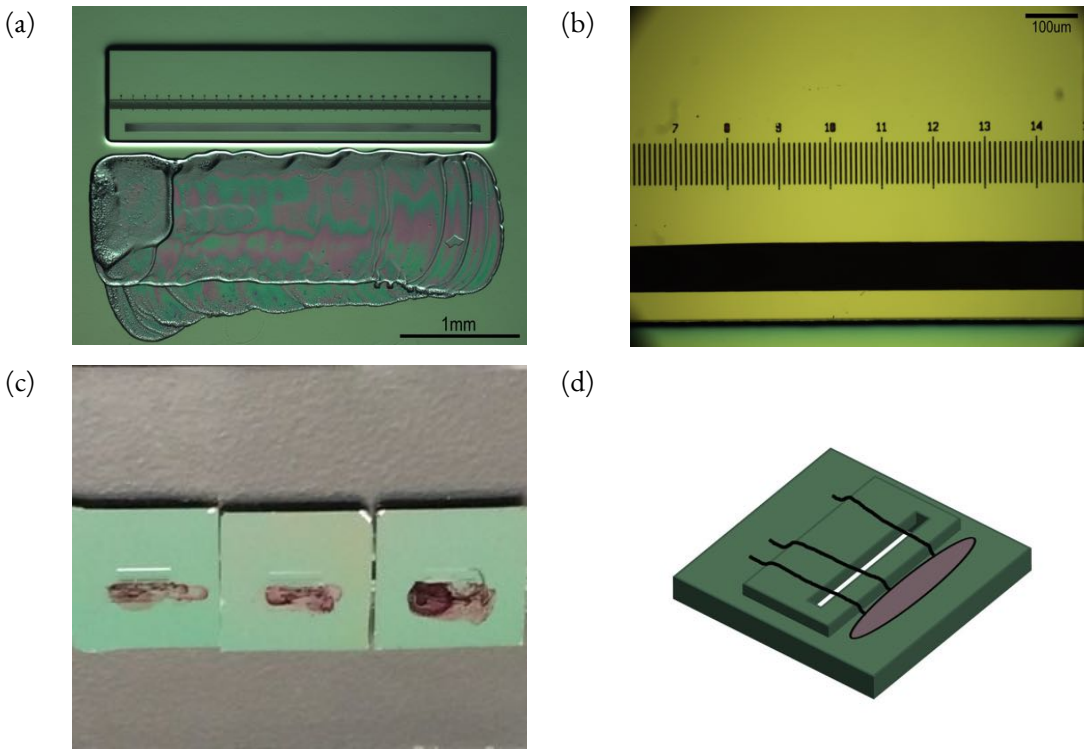


Figure 3.1: The Carbon Nanotube Growth Chip. (a) and (b) are optical images of the CNT growth chip with 1 mm and 100 μm scale bars respectively. (c) are three CNT growth chips as seen by eye. In the middle of the chip is the raised platform with a slit (not visible) on it. The dark solution below the platform is the hand-applied catalyst. (d) shows a 3D rendering of the CNT growth chip with CNTs growing out of the catalyst (gray patch) and over the slit.

refer to appendix F for recipes (and see figure 3.3 for the picture of the furnace). To achieve ultra-long single walled suspended carbon nanotubes in CVD growth, the following are important factors to consider:[36]

— Carbon feedstock: ethanol has been demonstrated to cause extremely high speed and low temperature carbon nanotube growth relative to other choices such as methane gas. Shorter growth time and lower growth temperature means less chances of catalyst to sinter (see below) as well as less opportunities for unwanted amorphous carbon deposition. It also increases the yield of single walled carbon

Figure 3.2: The carbon nanotube growth furnace. The furnace is composed of three sections highlighted in red. Left: outlet for vacuum pumping to evacuate the tube furnace and for growth gas exhaust. Middle: moving furnace for fast heating and cooling. Right: inlet that includes four mass/gas flow meters for hydrogen and argon, two bubblers for ethanol and distilled water, and a hygrometer.

nanotubes as well as reducing amorphous carbon deposition due to its provision of OH (hydroxl) radicals that absorb C (carbon) radicals.[37]

— Hydrocarbon flux: the carbon source comes from the ethanol in the form of hydrocarbon groups. It is carried by the Argon flowing through the bubbler holding the ethanol. This parameter can be changed by changing the temperature of the ethanol (icing or heating it to change the vapor pressure) and by changing the amount of ethanol in the bubbler (more ethanol means the Argon gas will travel longer through the ethanol, hence increasing the chances of the hydrocarbon being carried out of the bubbler into the reactor). A high influx of hydrocarbon means denser CNT growth as well as more bundles and amorphous carbon deposition. Too low an influx means little CNT growth.

— Catalyst type, size and density: the size of the nanoparticles in the catalyst dictates the diameter of carbon nanotubes. Large nanoparticles will cause multiwall or bundles to grow. We target catalyst made of small nanoparticles to increase the yield of single walled and small diameter CNTs. The density of catalyst, if it is too high, will cause the nanoparticles in the catalyst to sinter and form bigger nanoparticles which result in large, multiwalled or bundled CNT growth. The other effect of high

nanoparticle density is it will prevent CNTs from growing out of the catalyst area. Carbon nanotubes attract each other through Van der Waals interactions. If CNTs are grown too dense, the CNTs will entangle and conglomerate into a forest or network of dense CNTs. CNTs that try to grow out will get sucked back into the nanotube forest. To address all these factors, the catalyst of choice is the bimetallic cobalt-molybdenum catalyst (CoMoCat) that is embedded in a mesoporous silica substrate. Cobalt and molybdenum interacts to prevent cobalt from sintering to form larger nanoparticles whose size dictate the diameter of carbon nanotubes. The mesoporous silica substrate that holds the CoMoCat is a solid object that is created upon calcination. This will hold the catalyst in place and hence prevents them from moving together and sinter. This achieves the same effect of guaranteeing smaller size cobalt nanoparticles.[38]

— Pressure: our growth is performed at a very slight over-pressure (immeasurable with our analog pressure gauge). As mentioned in the growth recipe, gas is flown into the tube furnace which is blocked by pump oil at the exhaust. This pump oil serves to separate the exhaust which is in slight vacuum from the tube furnace (it also acts as a gas flow indicator as it will bubble when gas is flowing through it). A certain pressure build up is required in order for gas to flow through the pump oil. This slight over-pressure is to ensure that gas and vapor outside the furnace do not flow into the tube furnace during growth. The growth takes place in essentially atmospheric pressure. At the time of this writing, the study of pressure was not heavily investigated; according to literature, however, both the structure and yield of CNT depends on the pressure. Further improvements on our CNT growth can possibly be made.[39]

— Gas flow: High buoyancy (large Richardson number) is desired so that CNTs can stay afloat above the substrate. When afloat, the CNTs are susceptible to the gas flow around the growth chip. Here, laminar flow (low Reynolds number) is helpful to gently carry the CNTs in the direction of the gas

flow. Gas flows that are on the contrary will cause the CNT to stick to the substrate surface (due to low buoyancy) which diminishes the effect from the gas flow. And even if CNTs make it out of the CNT forest and float above the substrate, the convective flow (due to high Reynolds number) will disrupt the CNTs flow and cause them to bundle up prevent them from extending far from the catalyst region.[40]

— Anneal temperature and time: the highest temperature the growth is held at is 910°C for about 5mins (in the growth recipe, the timer is started the moment the furnace is moved over to the growth chips. According to the temperature controller, it takes about 2 mins for the temperature to reach 910°C, which leaves us 5 mins of annealing time). There is an optimal annealing temperature for particular carbon feedstocks and for particular types of CNTs. For ethanol and single-walled CNTs, $800-900^{\circ}$C is shown to be the most optimal.[37,41] Our furnace is set to a slightly higher temperature to account for the fact that the substrates do not get as hot as the furnace temperature. As mentioned above, the anneal time affects the density of growth as well as whether we get single-walled or multi-walled CNTs. 5 mins gives us the best collection of CNTs per growth.

— Ramp time: fast heating of the growth chip is performed in our CNT growth and this is done by moving the pre-heated furnace from away the growth chips to directly over the them. Studies suggest that CNTs grow via the "kite-mechanism" in which case the fast heating is beneficial because it helps levitate the floating CNT more by the preserving a steep thermal gradient surrounding the growth chips. This is because in the initial stages of growth, the cool gas will not be able to thermalize with the growth chips that are heated much faster. The temperature difference between the growth chips and the cool gas will create convection flow that helps lift the open-end of the CNT up while the end touching the catalyst stays stuck to the substrate due to Van der Waals interaction. The lifted end of the CNT is then carried by the laminar flow of the reactor gas where it will eventually lay down onto

the surface across the slit.[42]

— Reactor geometry: the quartz tube of the growth furnace is chosen to be small so that less gas is consumed per growth. The tube diameter of our CVD furnace is 1in. However, it is likely more beneficial to use a larger quartz tube because more gas means higher heat capacity which means it will be harder for the overall temperature of the gas to increase (for CNT growth we want to keep the gas cold).

— Water flux: Injection of right amounts of water vapor into the reactor chamber has been shown to both significantly improve yield and cleanliness. The water serves to prolong the life time of the catalyst (by preventing them from oxidizing which is what limits the reaction time available for the catalyst) and by reacting with the amorphous carbon to create gaseous hydrocarbon that will then flow away. Too much water, however, will cause the etching of the carbon nanotubes which will then decrease yield.[43,44,45]

— Hydrogen concentration: Hydrogen is important because it's a catalyst activator, a catalyst reducing agent and also protects the growth environment from unwanted substances such as oxygen and water. However, hydrogen can also etch carbon nanotubes, hence, the amount of hydrogen must be carefully chosen.[46] The above guidelines are most likely just a part of everything that actually matters. The truth is growth mechanisms are so microscopic that at this point of time it is beyond our complete control. Growth results will vary from setup to setup no matter how identical one tries to make them and as any grower would say, when it comes to material synthesis nothing can ever be truly reproduced. It would not be surprising to see that the types of gas-lines used (we use stainless steel throughout to prevent hydrogen from escaping and water from permeating into the system and we use VCR connections whenever possible to create the most leak-tight system) or the moisture level during summer time would have an effect on the growth. The blackbox nature of growth in general makes this type of work daunting and the one formula for success is to perform an extensive and systematic trial and

(a)

(b)

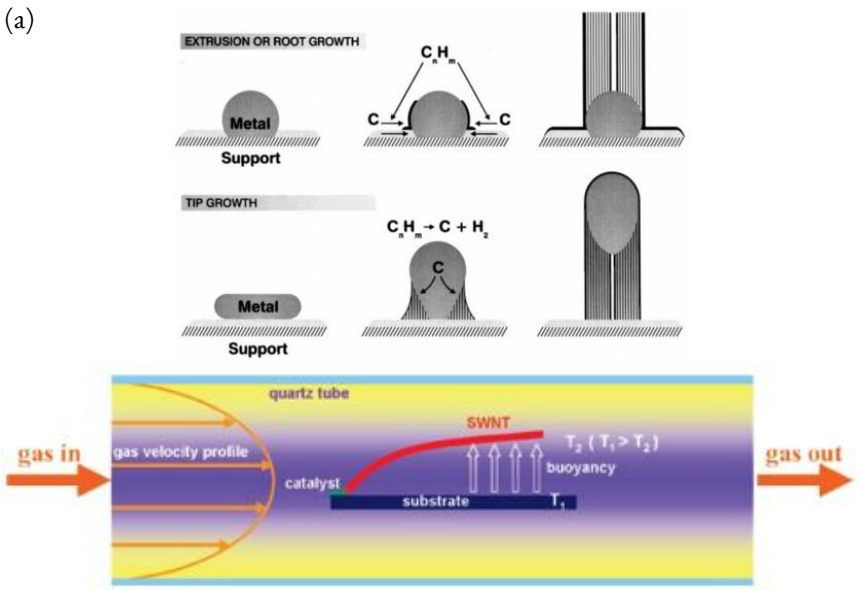


Figure 3.3: Carbon nanotube growth mechanisms. (a) The two proposed types of carbon nanotube growths. The picture is adapted from the work by Sinnott et al.[47] (b) During growth the carbon nanotube is like a kite and gets caught by the reactor gas flow. At the same time there is an attraction towards the substrate due to van der Waals force. A sufficient buoyancy force is required to keep the carbon nanotube afloat and continue its growth until either the catalyst lifetime has ended or gas flow has stopped. The picture is taken from the work by Peng et al.[40]

error. The recipe used for growth and one that has stood the test of time is shown in detail in appendix F. Figure 3.4 shows examples of SEM images of the CNT growths.

3.4 Optical Characterization of the Carbon Nanotube

One big challenge in working with an individual CNT lies in being able to identify it. The CNTs we work with are nanometer(s) in diameter and tens of microns in length. Apart from the small size, the optical absorption of an individual CNT is very weak (less than 0.01%). Together, it becomes hopeless to see an individual CNT under the optical microscope in bright-field mode. Tools such as the atomic force micoscopy (AFM) and its variants, scanning electron microscope (SEM) or transmission

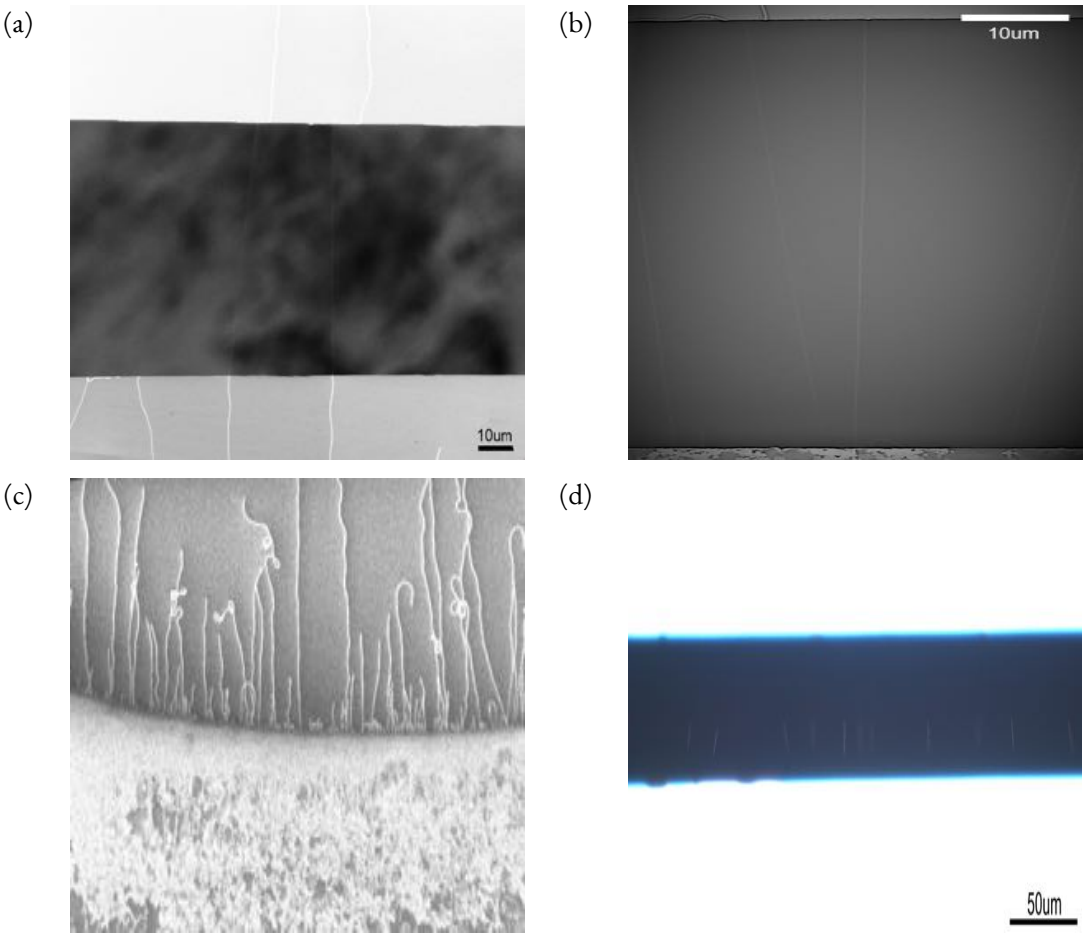


Figure 3.4: Suspended carbon nanotube growth. (a) SEM image of CNT growing over the slit of the growth chip (scale bar: 10 μm). (b) Zoomed in SEM image on the suspended CNTs (scale bar: 10 μm). (c) SEM image of aligned CNT growing out of the catalyst (area with high density of CNT networks). (d) Optical image of suspended CNTs that are locally metallized to make them visible under optical microscope (scale bar:50 μm).

electron microscope (TEM) are popular choices for imaging nanoscale objects. However, AFM, SEM and TEM all have drawbacks. Troublesome preparation, slow imaging speed, restricted access whilst imaging and sample contamination or even destruction are all legitimate. To circumvent these annoyances, an alternative optical method for seeing carbon nanotubes live was developed. The basic idea is to shine a very intense light directly onto the CNT in a scattering or dark field geometry. The light

that hits the CNT will scatter in different manners and can be categorized into two obvious categories: elastic and inelastic scattering.

3.4.1 Raman Spectroscopy

Raman spectroscopy is a widely used and industrialized technique for the optical characterization of materials. It is a measure of the inelastic scattering of light due to the interaction with optical phonons in the material. Overall it is an invaluable tool when studying CNTs and is most helpful when trying to identify metallic CNTs from semiconducting ones. This is because the Raman signature of a metallic CNT is very different from that of a semiconducting CNT. However, because of the rate at which light scatters inelastically, to obtain a substantial enough signal requires a significantly long period of time. For a suspended CNT, compared to the seconds (typically 1 to 3 secs) of exposure time in a Rayleigh spectroscopy setting, many minutes (in our setup it takes over 10 mins) of exposure time is required for the Raman spectroscopy. Adding the fact that one cannot see CNTs live (and hence we are unable to determine the optimal laser focus) and the fact that we need to go through a high number of CNTs for our experiments, Raman spectroscopy is an inefficient tool and is rarely used in comparison to the high-throughput Rayleigh spectroscopy. Without going into much detail, we simply state the key features we look out for when identifying metallic vs semiconducting CNTs. Representative CNT raman signals are presented in figure 3.5. A comparison be the Raman and Rayleigh setup is shown in figures 3.13 and 3.14.

One of the main Raman peaks of the CNT (the G-band) is positioned around $1590\,\mathrm{cm}^{-1}$ (the other main Raman peak is the radial breaking mode which is largely inaccessible in our setup). This particular peak is associated with the LO (longitudinal optical) phonons and TO (transverse optical) phonons. LO phonons are associated with vibrations that occur along the CNT longitudinal axis,

while TO phonons are associated with vibrations that occur circumferentially around the CNT. Depending on the underlying lattice and hence wavefunction of the CNT, only particular modes are allowed and this is reflected in the detected Raman signal. The following summarizes in bullet points the features used for spotting a metallic CNT:[48,49]

— Zig-zag CNT only has a LO mode

— Armchair CNT only has a TO mode

— TO peak is positioned at around $1590\,\mathrm{cm}^{-1}$

— Single TO peak always has a FWHM (full-width-half-max) of $10\,\mathrm{cm}^{-1}$

— Metallic CNT: LO peak shifts below TO peak

— Metallic CNT: LO peak may broaden

3.4.2 Rayleigh Scattering

Rayleigh scattering describes the elastic scattering of light. It is the predominant scattering mechanism of any dipole object and accounts for over 99.99999% of scattered light. The remaining 0.00001% of scattered light constitutes the inelastic Raman scattering.[50] The elastic scattering cross-section of the CNT is given by:

$$\sigma(\omega) = \frac{\pi^2}{64}\frac{\omega d}{c}^3 |\epsilon(\omega) - 1|^2 d \tag{3.1}$$

where c denotes speed of light, d the diameter of the CNT, $\epsilon(\omega)$ the dielectric function of the CNT, and ω denotes the angular frequency of the incident light. If the light is polarized parallel to the longitudinal axis of the CNT, the scattered light is isotropic with respect to the azimuthal direction. If one were using a laser of wavelength 500 nm, and assuming a measly $\epsilon = 6$ (ϵ of a CNT is shown to have peaks that go to $\epsilon > 100$), the resulting cross-section is $\sigma = 1.2 \times 10^{-7}\,\mu\mathrm{m}$. With a

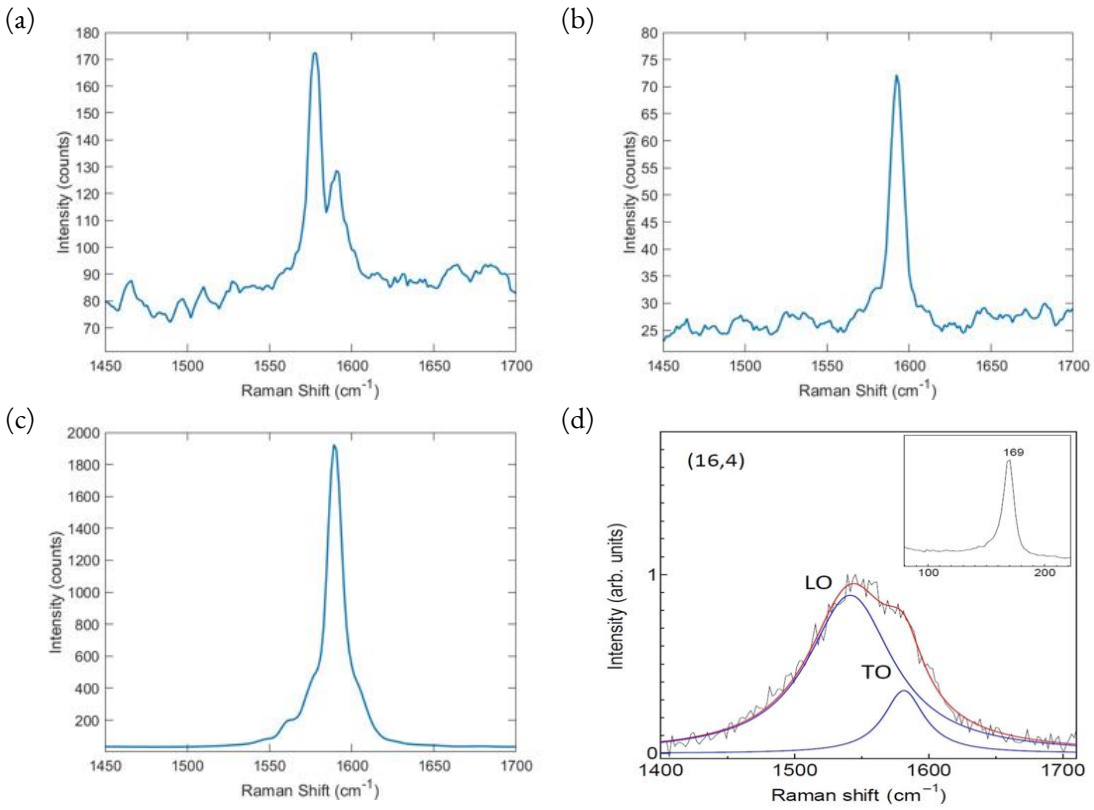


Figure 3.5: Raman spectra of different types of carbon nanotubes. Raman spectrum of (a) a metallic CNT, (b) a semiconducting CNT, and (c) a CNT bundle. (d) shows the signature Raman spectrum peak of a CNT decomposed into its two constituents: the LO (longitudinal optical) and TO (transverse optical) modes. Figure (d) was adapted from the work by Zhang et al.[49] Single-walled CNT have defined peak(s) with low signal counts, which is in contrast to a CNT bundle which have a broad peak with very high signal counts.

100 μW laser focused to a spot size of 1 μm, photons are then scattered at a rate of 100 μW$\times 1.2 \times 10^{-7}/(hc/500\,\text{nm}) \approx 3 \times 10^{7}\ \text{s}^{-1}$. With an overall signal collection rate of 10^{-2}, the final count rate is on the order of 10^{5} counts/sec— more than enough for live detection.[51] And so if one were to shine a broadband laser onto the CNT, the Rayleigh scattered light would show peaks at certain wavelengths as shown in figure 3.6. This is due to the peaks in the CNTs' dielectric function which is associated with its joint density of states. By examining the details of the peaks, one can deduce the chirality and many other properties of the CNT. This becomes a very powerful tool for CNT charac-

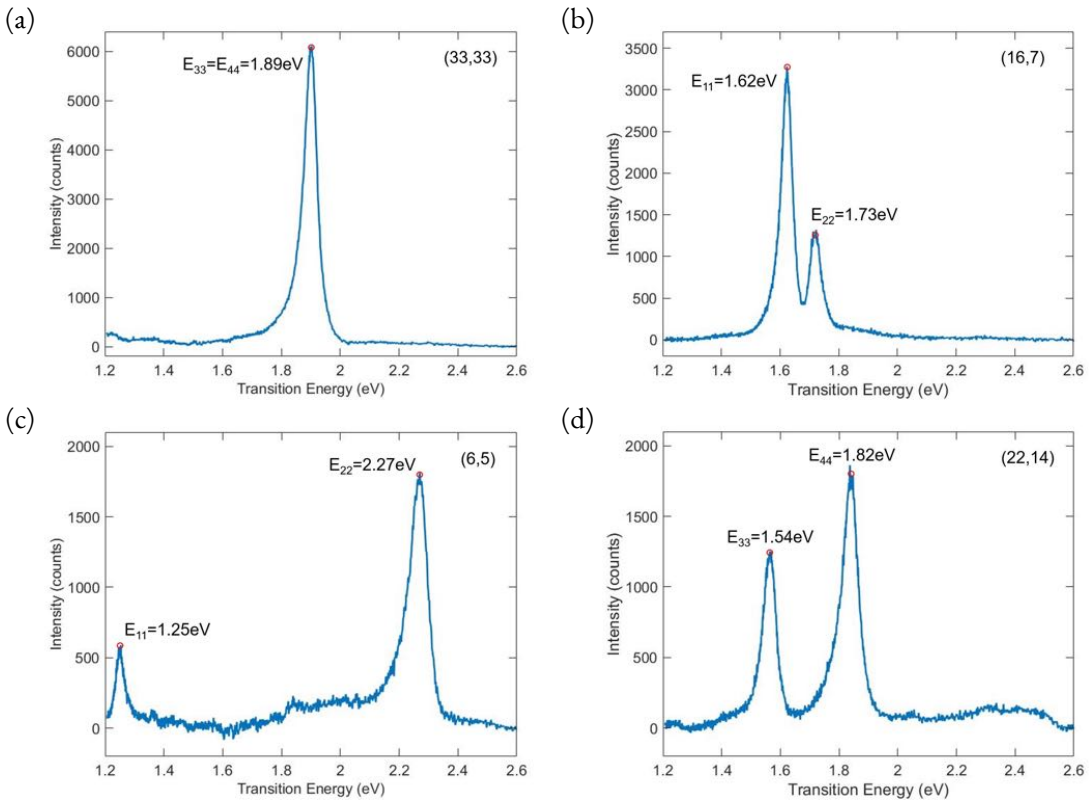


Figure 3.6: Rayleigh spectra of different types of carbon nanotubes. (a) and (b) are Rayleigh spectra of metallic CNTs, (c) and (d) are of semiconducting CNTs. The assigned chirality indices going from (a) to (d) are: $(33, 33)$, $(16, 7)$, $(6, 5)$, $(22, 14)$. The distinguishing feature of a metallic CNT is a double-split peak with the right peak having roughly half the intensity.

terization. The details of the Rayleigh spectroscopy and the setup will be covered shortly. Let us first take a look at how the optical transition energies arise in the following section.

3.4.3 Carbon Nanotube: A Dipole Moment

The CNT is nanometer in size in diameter but many microns long in length. This effectively restricts the electron movement to one direction, making it behave as though it were a dipole. So starting

generally, the Hamiltonian of a charged particle in an electromagnetic field is:

$$H = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m}. \tag{3.2}$$

In the Coulomb gauge where we set $\nabla \cdot \mathbf{A} = 0$, expanding and retaining only the lowest order term as warranted by the dipole approximation due to the anisotropy of the CNT, we get:

$$\begin{aligned} H &= H_0 + V(t) \\ &= \frac{p^2}{2m} - \frac{q}{m}\mathbf{A} \cdot \mathbf{p}, \end{aligned} \tag{3.3}$$

where the first term is single particle Hamiltonian and the second term is the time dependence interaction term. Substituting the ansatz $\mathbf{A}(\mathbf{r}, t) = A_0 \mathbf{n} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} + \text{c.c.}$, where $\mathbf{n}$ denotes the polarization direction, and noting that in the Coulomb gauge, $\mathbf{E} = -\partial\mathbf{A}/\partial t$, and assuming the electric dipole size is much smaller than the wavelength of the incident light (hence allowing us to drop the spatial dependence), the time dependent interaction term becomes:

$$V(t) = -\frac{q}{m}(\mathbf{n} \cdot \mathbf{p})\frac{E_0}{\omega}\sin\omega t = V_0 \sin\omega t. \tag{3.4}$$

By inserting the interaction term (V_0) into Fermi's Golden Rule which is:

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}\left| \langle f | V_0 | i \rangle \right|^2 \delta(\varepsilon_f - \varepsilon_i + \hbar\omega), \tag{3.5}$$

we get for the transition matrix element:

$$\langle f | V_0 | i \rangle = -\frac{qE_0}{m\omega} \langle f | \mathbf{n} \cdot \mathbf{p} | i \rangle \tag{3.6}$$

The rates are given by the magnitude of the transition matrix element where the initial and final states are given by the CNT wavefunctions. Calculation shows that the associated selection rules are:[52]

$$\begin{aligned} &\text{always} : \Delta k_{||} = 0, \\ &\text{for } E_{||} : \Delta J = 0 \,(\text{or } \Delta k_{\perp} = 0), \\ &\text{for } E_{\perp} : \Delta J = \pm 1 \,(\text{or } \Delta k_{\perp} = \pm 2/d), \end{aligned} \tag{3.7}$$

where $E_{||(\perp)}$ means the incident light is polarized parallel (perpendicular) to the CNT longitudinal axis. The CNT energy subbands of index j has an angular momentum number of $J = |j|$ (see appendix B.3), where we ignore spin since $\Delta s = 0$ for optical transitions. And so by choosing to polarize the incident light to be parallel with the CNT longitudinal axis (so $E_{\perp} = 0$), we are restricted by the selection rule: $\Delta J = 0$. This just means that optical transitions can only occur between subbands defined by the same linecut associated with j, which further means transitions happen symmetrically about the charge neutrality point as shown in figure 3.7.

3.4.4 Rayleigh Spectrum and Line Shapes of Optical Transitions

Hinted in the previous section, the scattering signal given by Fermi's Golden Rule will depend on the joint density of states between the initial and final states. This is the easiest way to see that peaks in the Rayleigh spectrum correspond to the transitions between the van Hove singularities. Looking at the scattering cross-section ($\sigma(\omega)$), we see that it depends on the dielectric function ($\epsilon(\omega)$) which is similar to an optical absorption. However, unlike optical absorption which depends purely on the imaginary part of the dielectric function (or electric susceptibility which is $\chi_e(\omega) = \epsilon(\omega) - 1$),[53] $\sigma(\omega)$ also obtains a contribution from the real part of $\epsilon(\omega)$. Typically, the optical absorption curve is given

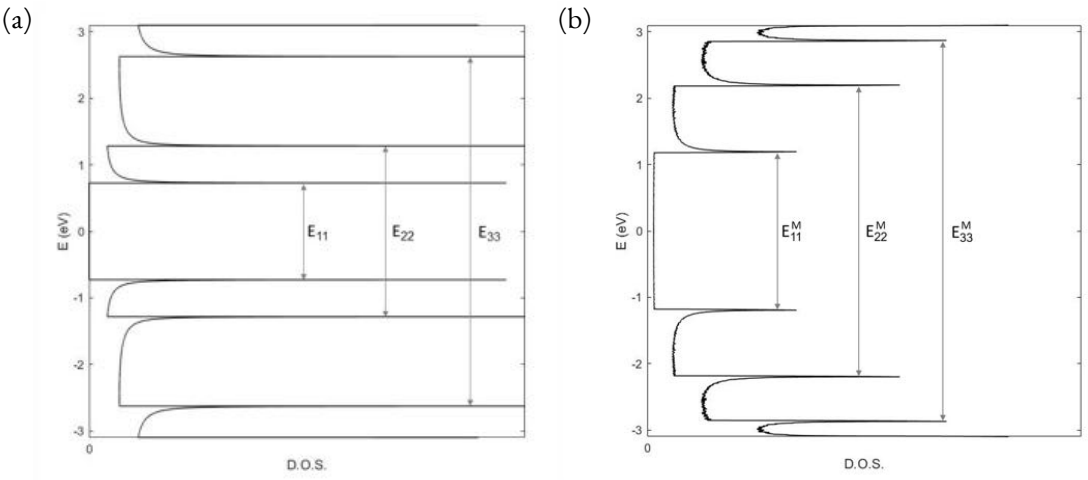


Figure 3.7: Zoomed-in density of states plots of (a) a $(8, 0)$ semiconducting zig-zag carbon nanotube and of (b) a $(8, 8)$ metallic armchair carbon nanotube. E_{11}, E_{22}, E_{33} are the first three allowed optical transitions (where superscript M denotes the optical transitions for metallic CNT). These are the energies that are required to excite an electron from the valence band to the conduction band while obeying the selection rule. Note that the metallic carbon nanotube has a constant non-zero density of states around the charge neutrality point as opposed to the vanishing density of states in the bandgap of a semiconducting nanotube.

by a symmetric Lorentzian, however, the inclusion of $\mathrm{Re}|\epsilon(\omega)|$ which has an intensity tail towards the low energy side will cause an asymmetry in the Rayleigh peaks (see figure 3.8).[54]

Distinguishing metallic CNTs from semiconducting ones can be quick and easy. Majority of metallic CNT will exhibit a a double-split peak due to trigonal warping of the bandstructure of Gr.[55] Trigonal warping is the warping of the energy contours of Gr and it is due to the higher order energy terms in the Gr Hamiltonian. If trigonal warping were not present, energy contours around the Dirac point of Gr would be perfect circles. Because the allowed linecuts for a metallic CNT are placed symmetrically about the Dirac point, in this case the linecuts would yield identical subbands and hence would yield overlapping optical transitions. As a result we would only see a single Rayleigh peak. But this is not the case as trigonal warping exists with varying levels of strength. As one can see in figure 3.9, the two linecuts immediately left and right of the Dirac point would yield different subbands if the

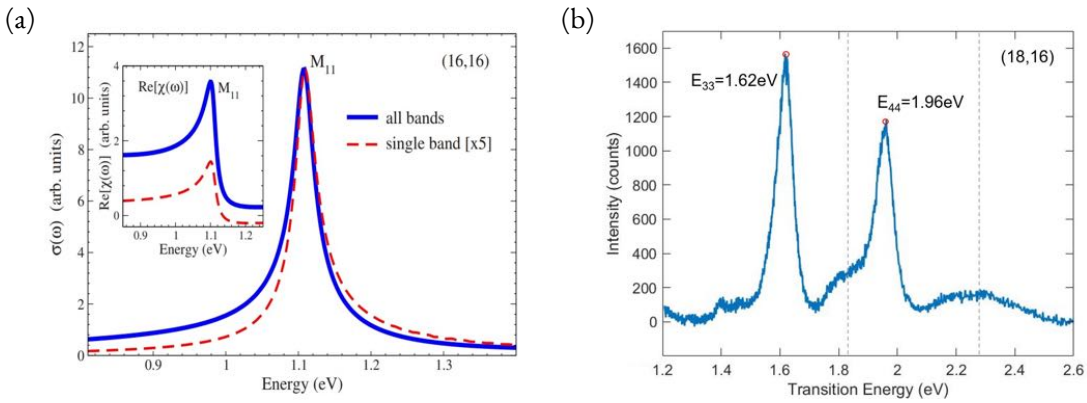


Figure 3.8: Contributions to the asymmetry observed in the peaks in the Rayleigh spectrum. (a) The underlying asymmetry in the peaks is due to $\mathrm{Re}|\epsilon(\omega)|$ as shown in the inset. Figure (a) is adapted from the work by Malic et al.[54] This is in contrast to purely optical absorption which is only dependent on $\mathrm{Im}|\epsilon(\omega)|$ and yields a symmetric Lorentzian instead. (b) The phonon side bands that arises due to exciton-phonon interactions further contributes to the asymmetry in a peak. In the plot the gray dashed lines indicate the center of the phonon side bands. This effect is more severe for semiconducting CNTs and mostly undetectable for metallic ones.

energy contours were warped. This then causes the optical transitions to be slightly different between the two linecuts. By drawing out the allowed linecuts, one can see that the splitting distance between the two peaks increases with decreasing chiral angle and diameter. Because of this, the armchair CNT with zero chiral angle (see appendix B) will have no double split peak but a much taller single peak. Furthermore, because of the direction of warping, the higher energy peak ($E_{22} = M_{11}^{-}$) will roughly be half the height of the lower energy peak ($E_{11} = M_{11}^{+}$).[55] Figure 3.6a and 3.6b show examples of an armchair and chiral metallic CNT respectively.

Another distinguishing characteristic between metallic and semiconducting CNTs is that semiconducting CNTs exhibit a phonon side band that causes further asymmetry in the Rayleigh peaks (see figure 3.8b). These features have been predicted to arise from exciton-phonon bound states that is especially strong in semiconducting species as compared to metallic ones which have reduced strength in excitonic interactions.[56] Finally, to wrap up the analysis of the Rayleigh peak line shapes, lower optical

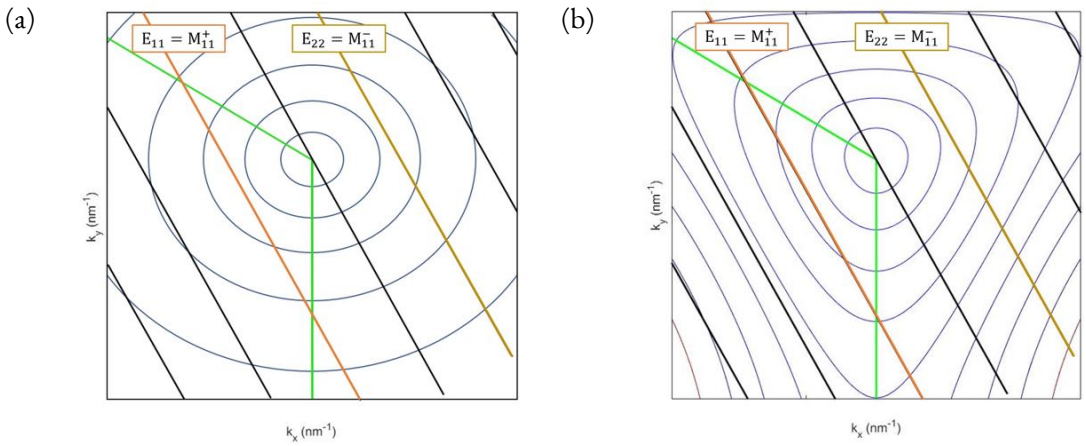


Figure 3.9: Demonstration of the effect of trigonal warping. The plots show the energy contours in momentum space of Gr. The lines are the allowed linecuts for a metallic carbon nanotube. The apex of the green line is the Dirac point of graphene. In this thesis, E_{ii} is used as a generic label for optical transitions. In literature, a common label for transition energies of metallic carbon nanotubes is M_{ii}. The plots show how the two labeling schemes are related. (a) shows circular energy contours due to lack of trigonal warping. In this case, the E_{11} and E_{22} would yield the same optical transition energies as the associated linecuts are symmetric. (b) shows the energy contours that gradually transform into a more triangular shape as they go farther away from the Dirac point. In this case, E_{11} and E_{22} have associated linecuts that cut through a different energy contour and this results in two slighty different optical energies. The difference between the energies depends on the severity of the trigonal warping. Higher energies show more trigonal warping and hence one would expect a bigger energy shift.

transitions such as E_{11} and E_{22} generally have peaks widths that are $10-50$ meV wide, whereas higher optical transitions like E_{33} and E_{44} are typically $80 - 90$ meVs wide.[56] Because we typically see the E_{33} and E_{44} transitions for semiconducting CNTs while we typically see the E_{11} and E_{22} transitions for the metallic CNTs, in the identification of CNTs, apart from the double split peak and asymmetry considerations, we also expect peaks associated with metallic CNTs to be sharper. Finally, bundles and multi-walled CNTs are quite easy to distinguish. We typically categorize any Rayleigh spectra with four or more peaks to be associated with non-single CNTs. Moreover, it is shown that bundles or multi-walls will generally include interacting CNTs which will exhibit an overall background in the Rayleigh spectrum.[57] Figure 3.10 show examples of the live-images and Rayleigh spectra of different

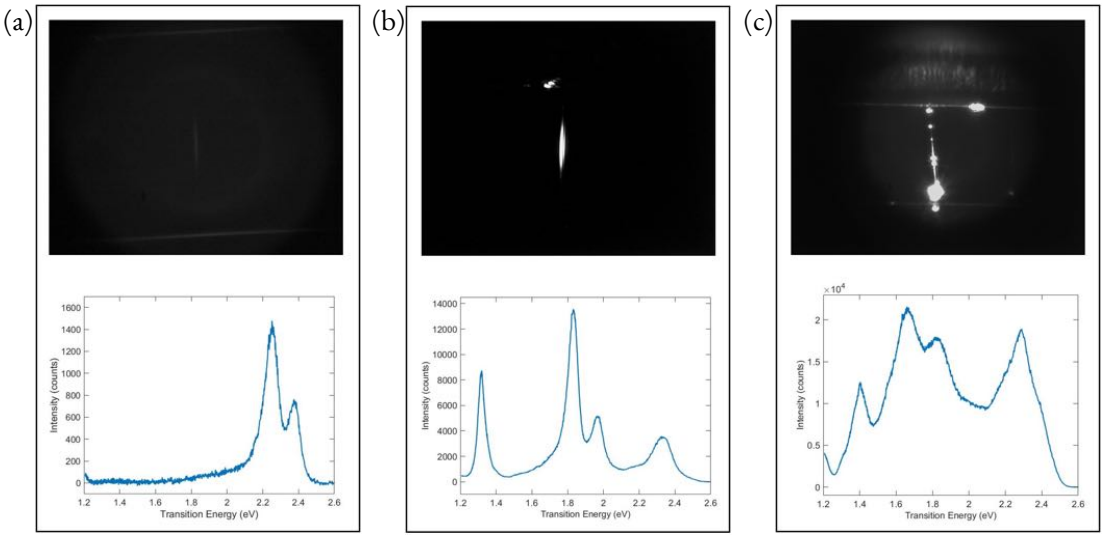


Figure 3.10: Different carbon nanotubes seen through Rayleigh scattering. The top row shows the Rayleigh scattered signal of different CNTs collected through the infrared camera. The bottom row shows the Rayeligh spectra of the corresponding CNTs collected through the spectrometer. A single CNT (a) typically shows up fainter in the live image and the spectrum is clean with one to sometimes three peaks (seeing two peaks is the most common). A CNT bundle (b) is bright and the spectrum has many peaks that can only be fitted by using multiple CNTs' transition energies. A dirty CNT (c) has blobs all over and the image is bright. Dirty CNTs are common if water was not introduced during CVD growth. The spectrum of a dirty CNT has a large background signal.

types of CNTs.

3.5 The Rayleigh Setup

The Rayleigh setup (shown in figure 3.12) was built based on the concepts conceived by Professor Tony Heinz.[58,51] He was the pioneer behind the method of seeing individual CNTs via Rayleigh scattering. What we have done was adapt his Rayleigh spectroscopy setup into a system that also includes an optical microscope, transfer stage and Raman spectroscopy setup. Having all these functionalities together gives us the ultimate tool to simultaneously characterize and physically manipulate individual CNTs. The operational principle behind Rayleigh spectroscopy is to shine a very intense polarized broadband

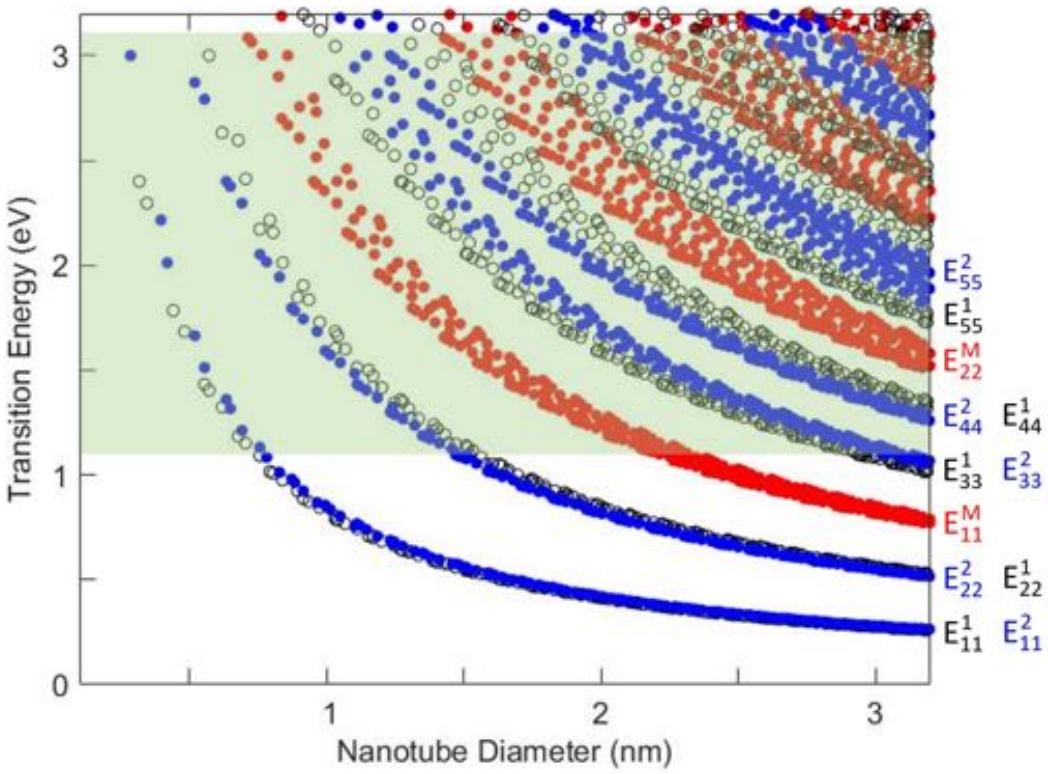


Figure 3.11: The Kataura plot and depiction of the CNT transition energies. The Kataura plot is generated from the calculation of tight-binding theory. The parameters assumed are: $t_{nn} = 2.9\,\text{eV}$ (the nearest neighbor hopping energy) and $a_{cc} = 0.144\,\text{nm}$ (the carbon to carbon bond length).[59] The black circles correspond to type 1 semiconducting CNTs, the blue points correspond to type 2 semiconducting CNTs and the red points correspond to metallic CNTs. These points are the transition energies of taking an electron from the valence band to the conduction band as shown in figure 3.7. The first five optical transitions are labeled. The shaded green area corresponds to the CNTs that our Rayleigh setup can see.

light onto the CNT, and then detect the scattered light in a dark-field geometry. The scattered light of a suspended tube is isotropic azimuthally and will exhibit peaks in intensity at wavelengths that correspond to the energy difference between CNT's von Hove singularities but with additional corrections. In addition to the slight deviation due to the contribution from $\text{Re}|\epsilon(\omega)|$ (see previous section), other possible causes for corrections are electron-electron interactions and spin-orbit coupling due to curvature effects. Up to the time of this writing, there are still ongoing efforts in describing the optical

transitions more accurately. So far, revised descriptions are all empirical models and here we stay away from going into the explicit equations. The interested reader is referred to the cited references.[60,61,62] But to illustrate the general behavior of the optical transition energies, figure 3.11 shows the energies that are calculated theoretically from the simple tight-binding model. From this plot, we can see that $E_{ii} \propto 1/d$, where d is the CNT diameter, and this is an important characteristic. But as mentioned previously, the real values of the optical transition energies deviates slightly from the ones predicted by tight-binding theory. The experimental values are tabulated in the cited online webpages,[59,63] and these are the values that we refer to when matching CNTs' chirality indices. For more information the reader is asked to refer to appendix B.6.

In our setup, we use the Super K Extreme EXW-12 as our supercontinuum (or broadband or white light) laser capable of delivering a total of 4W across the entire bandwidth of 400 to 2400 nm. The wavelengths in the near-infrared region is critical as it turns out many of the CNTs scatter in that range. The objectives used all belong to the Mitutoyo NIR Infinity Corrected series. All of the optics used are compatible with the aforementioned bandwidth (except for the cameras). The other specifications of the laser are irrelevant as the only feature we require is the broad bandwidth and high power. Typically, we set the power to output around 500 mW as measured in an optical power meter with silicon detector. We estimate that the power density after the polarizer (with no further attenuation) at the CNT is about $10 - 50$ mW/μm. The spectrometer we use is the Princeton SpectraPro HRS-300 with a PIX-400BRX camera (with a sensitivity range from $400 - 1100$ nm). The infrared camera used for the live-imaging of the CNT is the DCC3240N from Thorlabs which also only has a sensitivity range from $400 - 1100$ nm. The quantum efficiency towards the near-infrared is boosted, however, making it more sensitive to seeing CNTs scattering in that range. The incident light is directed through an attenuator and polarizer. The laser spot size in the setup reaches down to about 5 μm. We use the 150

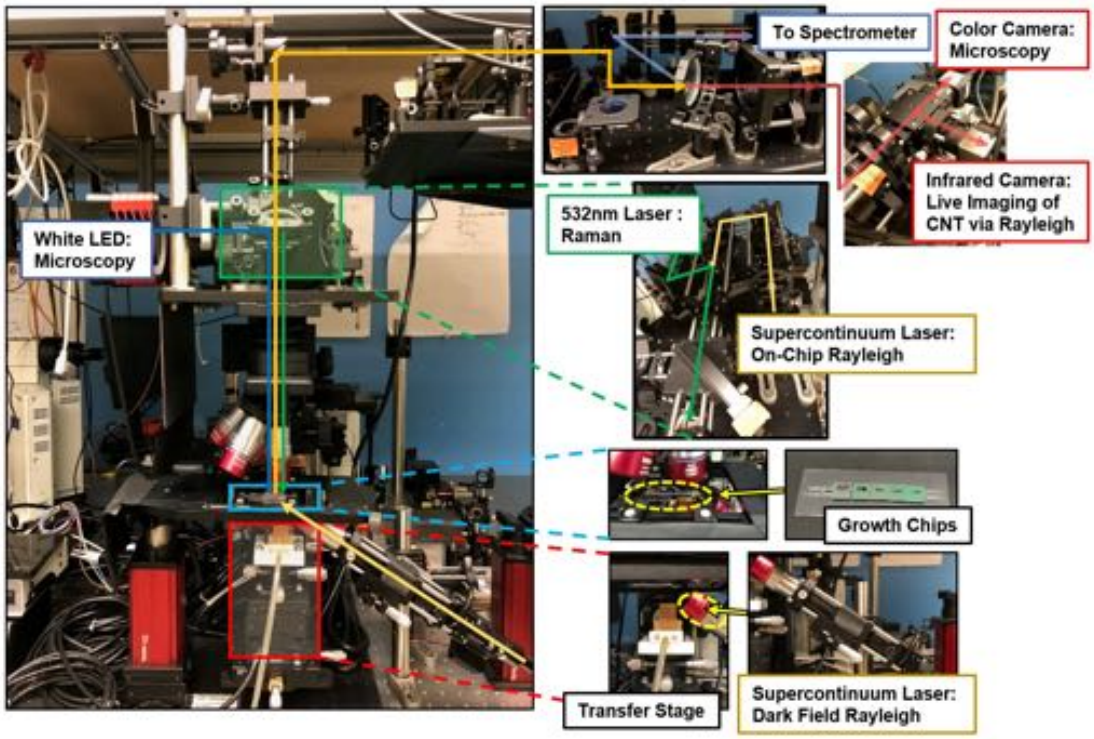


Figure 3.12: The Rayleigh transfer stage setup. The setup is equipped with a broadband white light laser and a 532 nm wavelength laser used for Rayleigh and Raman spectroscopy respectively. There is a color and near-infrared sensitive camera used for normal optical microscopy and the live imaging of CNTs during Rayleigh spectroscopy respectively. For normal optical microscopy, the white LED is used. Along with the optical components is the motorized transfer stage capable of heating, cooling (through thermoelectric or fan cooling), and tilt, rotation and translation in all axes. This setup is used for CNT imaging as well as 2D flakes and CNT transfers.

g/mm grating centered around 750 nm to acquire a one-shot spectrum with the range 450 to 1050 nm (which converts to roughly 1.2 to 2.6 eV).

Figure 3.12 is the Rayleigh setup with the main components illustrated in the figure. Figures 3.13 and 3.14 show the Raman and Rayleigh configurations respectively. During the Rayleigh operation, the signal is divided among three cameras. 50% goes to the camera of the spectrometer, 25% goes to the infrared and color camera individually. To see the CNT live, we use the infrared camera (Thorlabs DCC3240N) with a gain boost and together with the 50x objective. The CNT growth chips are

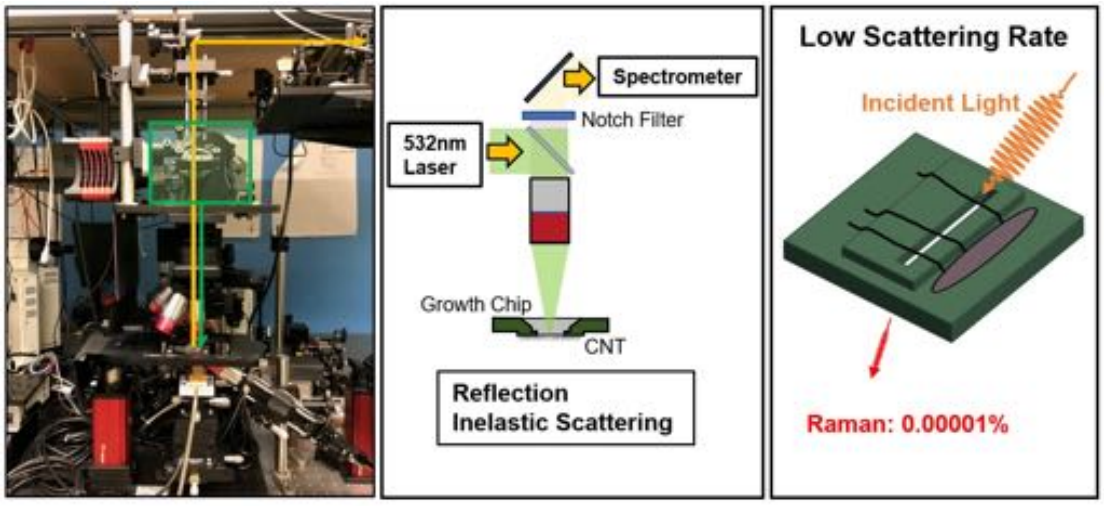


Figure 3.13: The Raman spectroscopy configuration. The 532nm wavelength laser is used. The laser enters from the back of the setup and is directed onto the sample through the objective lens. The inelastically scattered light is detected in the reflection and brightfield geometry. A notch filter is installed in the detection line and is used to filter out the excitation wavelength. Only wavelengths away from the excitation is detected. Raman scattering happens at a very low rate and so collection time has to be very long.

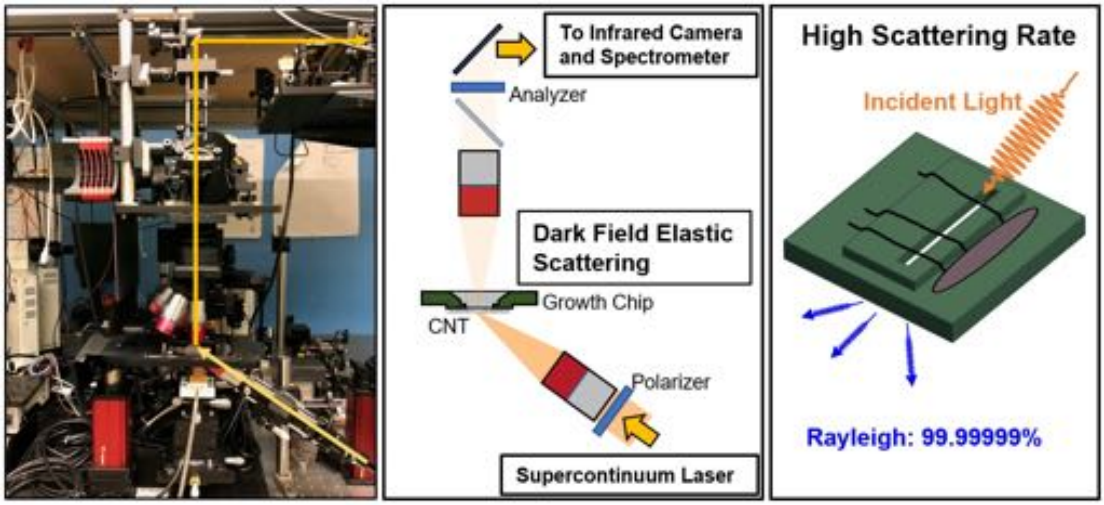


Figure 3.14: The Rayleigh spectroscopy configuration. The broadband white light laser is directed from underneath the sample in a dark field/off axis geometry. The elastically scattered light at a particular range of angle goes into the objective lens and is fed into both the near-infrared light sensitive camera and the spectrometer camera. One is used for the live imaging of CNTs and the other is for the collection of the Rayleigh spectrum. Because Rayeligh scattering happens at a very high rate, immediate imaging of CNT is possible.

loaded onto a glass slide that has a slit. This slit is aligned with the slit on the growth chip and this together creates an opening such that the incident light from the supercontinuum laser can go through. Because the incident light is off-axis, meaning that we are in dark-field geometry mode, none of the incident light is detected by the cameras. The default state of the cameras is dark. To aid finding a good focal position for the growth chip slit with respect to the objective, the white LED is turned on so that we can acquire optical imaging of the growth chip slit. Once the slit is in focus, the white LED is turned off and the laser is directed into the slit. The laser position is also free to move in all directions and we tune it such that the laser focal spot is at the slit. This is done by looking at the profile of the transmitted laser beam. The transmitted laser beam will be perfectly circular and will show densest amount of light when its focus is at the slit. Following these initial instructions, scanning the growth chip slit across a stationary laser beam will eventually show suspended CNTs along the slit. Finer tuning can be done on the focus of the laser and objective afterward. Figure 3.10 shows examples of the live imaging of different types of CNTs.

3.6 The Assembly of the Carbon Nanotube-Graphene Hybrid Device

The purpose of the Rayleigh setup is to allow the user to see and characterize CNTs with live feedback, and then to transfer a desirable CNT to a target area immediately. In the previous section I described how we locate and characterize CNTs suspended along the slit in a growth chip. To be able to then transfer the CNT onto a target area, this target area needs to be coated with a substance that can expand, flow, conform and retract so that a suspended CNT can be grabbed away from the growth chip. This role is fulfilled by Poly(methyl methacrylate) (PMMA). Without it the transfer of CNT will not be possible and the reasoning is as follows: The general idea of transferring CNT is to bring the growth chip with the suspended CNTs and the target chip into contact and hope that the van-der Waals force

will attract the suspended CNT and grab hold onto it. But this assumes that the CNT and target chip can be in sufficiently close such that the attraction can happen but this is unlikely due to the local roughness of the two chips. The CNT is nanometers wide which is on the same order of typical substrate roughness. It is unlikely that the CNT sitting on a substrate to stand out. Even for the suspended CNT, it is very difficult for the target substrate to come close to it because the growth and target chips away from the suspended CNT region will have a roughness that can prevent the two chips from becoming flush against each other. Despite this, we still want the chips to be able to come as close to a perfect contact as possible so that the following step will have the highest possible chance of happening (see section 3.2): Once the chips are as close to each other as possible, heating the PMMA will force it to go into the glass phase, which is a viscous liquid that will then expand and bridge the remaining distance between the CNT and the target area. The viscous PMMA will begin to conform around the CNT, the van-der Waals force will attract it, and then finally upon cooling it will shrink, retract and solidify again, and in the process grabbing the CNT away from the growth chip.

The role of PMMA can be fulfilled by many other substances (mainly any thermoplastics), however, PMMA provides an advantage and that is it is a high resolution electron beam (ebeam) resist. This means that fine selective areas coated in PMMA can be removed and the substrate underneath can be exposed. Being able to "open a window in the PMMA" is critical because eventually there may be a need to get rid of the PMMA through either immersing the entire substrate into a solvent (like Acetone) or through vacuum annealing. The lift-off of the PMMA will cause the CNT that was originally stuck onto the PMMA to be lifted and float away from the substrate. The vacuum annealing of the PMMA will cause the CNT on it to twist and bend because the PMMA evaporates unevenly and possibly violently. By opening a window in the PMMA and making this area the target area, during a CNT transfer, part of the CNT will be grabbed by the PMMA and part of the CNT will settle onto

the exposed substrate. The part settled onto the exposed substrate will be strongly anchored down due to the attractive van-der Waals force and this is enough to hold the entire CNT down. So despite part of the CNT possibly floating away during liftoff, the CNT stuck to the substrate remains stuck to the substrate. The CNT also remains completely straight inside the PMMA window regardless of the processing afterwards.

3.6.1 Carbon Nanotube Transfer Procedure

Figure 3.15 shows the whole CNT transfer process. With the target chip onto the transfer stage, we use the Rayleigh spectroscopy to find and choose the candidate CNT along the slit in the growth chip after which we park the CNT in place. The target chip is brought up closer to the growth chip but still not touching. We focus onto the target chip and align the target area with the parked and stationary CNT. Once aligned, we bring the two chips together and heat up to at least 140°C and wait for at least 5 mins. The point here is to allow the PMMA to flow and conform around the CNT. After, we cool the stage to at least below 90°C and then lower the stage.

The hybrid CNT-Gr device is made by first assembling the BN-encapsulated Gr stack as inspired by the original pickup technique. [64] For the devices in this thesis, because a thin hBN is typically used to separate the Gr from the CNT, and because thin hBN are hard to be picked up alone by Polypropylene carbonate (PPC), we employ the inverted stacking method (see appendix F). In this method, we pick up the layers in the inverted order so we begin by first picking up the thick hBN and attempt to flip the entire stack at the end such that the thin hBN is facing upwards. Figure 3.16 shows the steps to the stage where the device is ready for a CNT transfer. First, the stack is etched so that flakes or unwanted parts on the substrate are removed. Second, Gr contacts are made (see appendix F for more details). An in-situ contact etch and metal evaporation with a tilted rotating stage are employed for best contact.

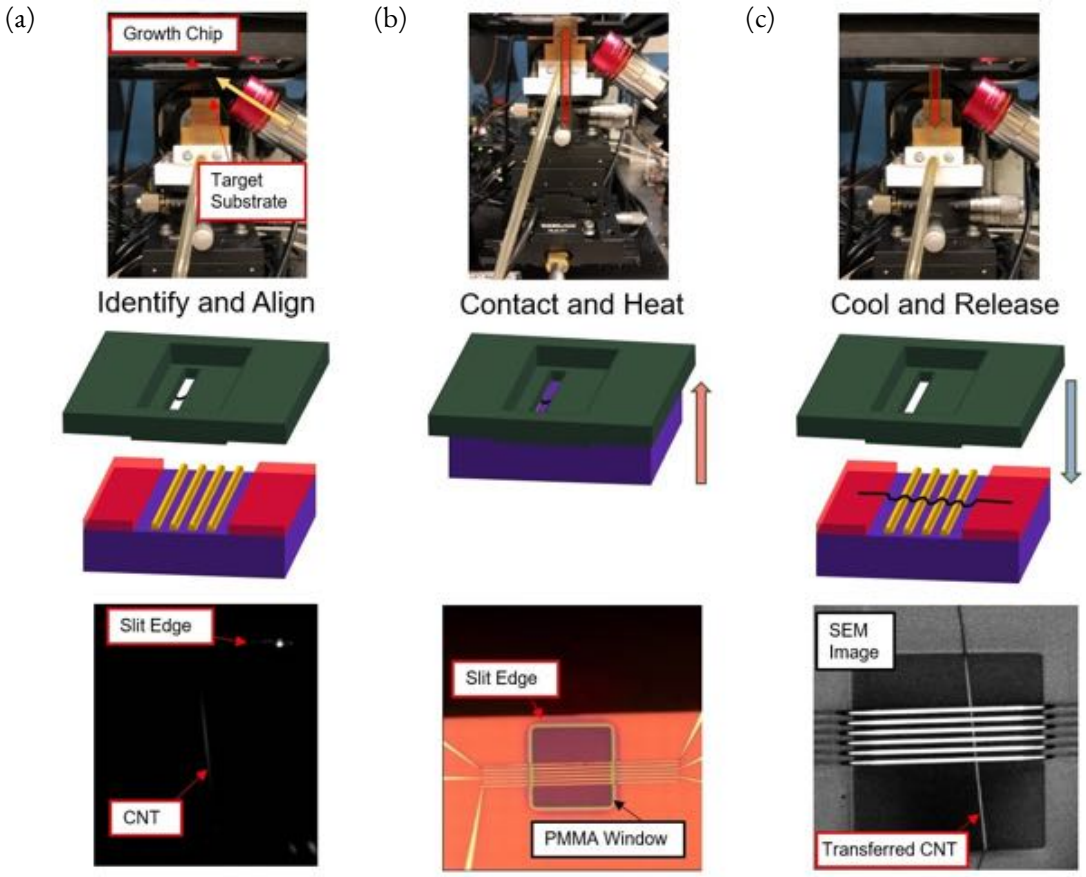


Figure 3.15: The transfer process of a carbon nanotube. The images from (a) to (c) show the transfer process in chronological order. The top row are images of the transfer stage position, the middle row are graphics of the growth chip (with a suspended CNT) and target substrate (with an array of gold electrodes and a red-purple PMMA window), and the bottom row of (a) and (b) are the optical images as seen through the microscope during the transfer process while that of (c) is an SEM image of the target substrate after the transfer process. Beginning with (a), we perform Rayleigh spectroscopy to hunt for a candidate CNT and then we park the growth chip in place and mark the CNT position as seen at the bottom image. We then focus onto the target substrate and align the target area with the marked position of the CNT. Once the CNT and target area are aligned, we get to (b) where we slowly bring the growth chip and target substrate into contact. The bottom image shows the moment when the growth chip and target substrate are in contact. Upon contact, we heat the stage up to at least 140°C such that the PMMA melts and conforms around the CNT to grab it. We wait at least 5mins and then we proceed to (c) where we cool and slowly detach the target substrate from the growth chip. The bottom image shows an SEM image of a CNT transferred onto an array of electrodes.

Third, pre-made electrodes for CNT contacts are deposited on top of the device. These electrodes are made thin (3/7/20 nm Cr/Pd/Au) to avoid the formation of bunny ears along the evaporated metal. Bilayer resist (see appendix F for more details) is used to further prevent bunny ears. Bunny ears are bad for CNT contacts as they can cut it or create a big kink on the CNT that increases contact resistance. Finally, a PMMA window is opened on the target area. PMMA A4 or A6 spun at 3000 rpm are both used with no noticeable difference. This window is typically 40 μm long, a little smaller than the slit of the growth chip. This is to guarantee that the PMMA and the substrate can both touch the suspended CNT on the growth chip. They are both important in grabbing and anchoring the CNT down onto the substrate. With this, the device is ready for a CNT transfer by following the steps detailed above. Figure 3.17 shows the general outline of the device fabrication. Typically without removing the PMMA, the device is wirebonded and directly loaded into the cryostat for measurements.

3.6.2 Quick Tip on Carbon Nanotube Contacts

A single-walled CNT is on the order of a nanometer in diameter which is smaller than typical grain sizes of evaporated metals (grain size of evaporated gold (Au) is about tens to hundreds of nanometers, whereas for chromium (Cr) and palladium-gold (Pd-Au), which are regarded as having some of the smallest grain sizes, they are on the order of a few nanometers). As a result, wetting the CNT with metal becomes very difficult and thus establishing consistent transparent electrical contacts is a big challenge. To improve the contact yield, the contact method of choice is to lay the CNT on top of premade gold electrodes. Gold is used because it does not oxidize (as opposed to palladium or chromium, a common contact metal except it oxidizes). This way, a part of the CNT is guaranteed to always be making physical contact with some of the metal, unlike top contacts where the evaporated metal can

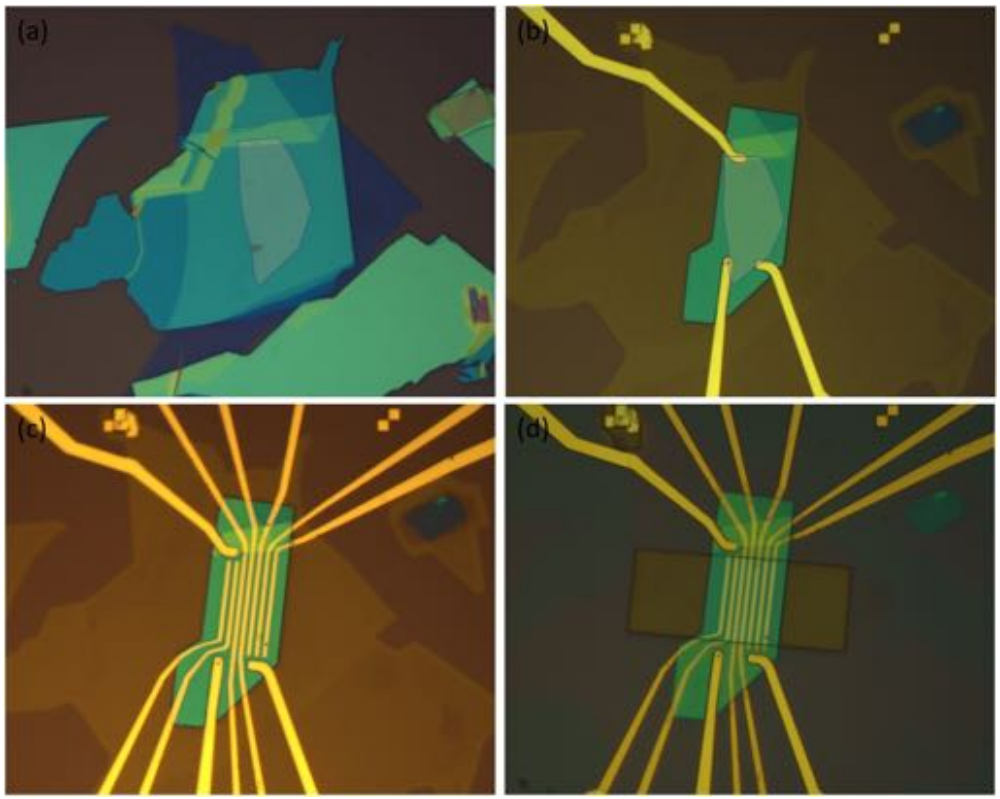


Figure 3.16: The steps to making a target device for carbon nanotube transfer. (a) shows the assembled hBN-encapsulated Gr stack. The outline shows the Gr flake. (b) shows the step after the etching of the flake and depositing Gr contacts after an in-situ contact etch. (c) shows the CNT electrodes deposited on top of the stack. The CNT leads and Gr are electrically isolated. (d) shows the PMMA transfer window that is 40 μm long written on top of the target area. This device is now ready for a CNT transfer.

potentially form a hollow tunnel structure around the CNT. Another advantage of laying down CNT onto pre-made electrodes is that the CNT contact area is still exposed. This allows the contact area of the CNT to be cleaned further by current anneal. Current annealing is performed simply by sending a large current through the CNT (can be done in ambient conditions). The typical procedure is to slowly ramp the bias voltage while detecting the current with an ammeter. The ramp rate followed is about 1 μA/min and current is usually ramped to 10 μA (the canonical current capacity of a CNT on SiO_2 is about 25 μA per shell[65]). The heat generated by the current is thought to be responsible

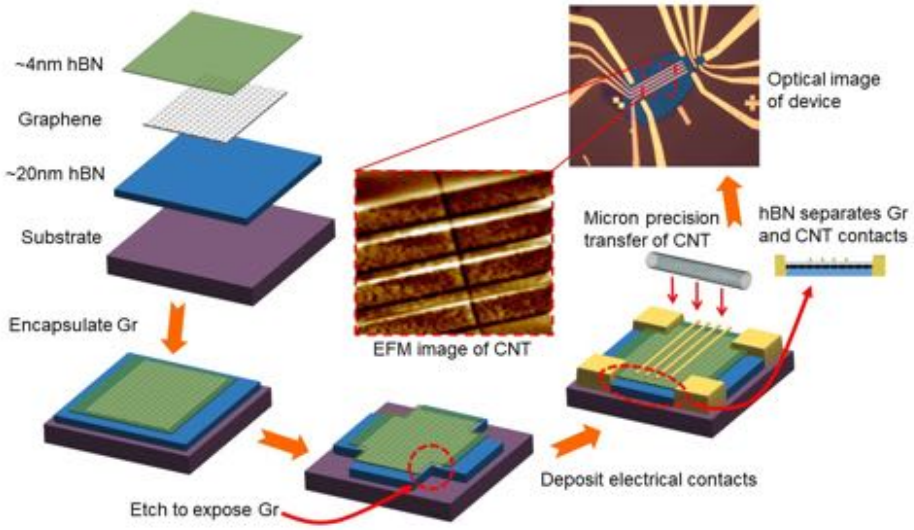


Figure 3.17: The graphene-carbon nanotube hybrid device fabrication scheme. We first assemble a standard BN encapsulated Gr stack. Because thin BN ($4-10$nm) is typically used at the top, the inverted stack method is used. Once the stack sits on the substrate of choice, contacts to the Gr are made via edge contact with in-situ etching. After, thin pre-made electrodes topped with gold is fabricated on the stack on which a CNT of choice will be placed with the CNT transfer setup. The end product is a device where the CNT is electrically isolated from but still in close proximity to the Gr.

for removing the nearby impurities such as polymer residues. The observation during current anneal is that while holding the voltage bias constant, the current along the CNT will increase on its own. CNTs with initial contact resistances on the order of a hundred of MOhms have been reduced to low hundreds of kOhms.

Everything we hear is an opinion, not a fact. Everything we see is a perspective, not the truth.

Marcus Aurelius

4

Relativistic Analogue in Graphene

Near the charge neutrality point of graphene, the energy dispersion of electrons is linear and hence light-like. This is a result of the electrons' equation of motion being mathematically analogous to the massless Dirac equation which effectively renders the electrons as relativistic and massless Dirac fermions. This electronic property allows the unprecedented study of relativistic particles in a condensed matter system. In this chapter, the background and theoretical components of two relativistic phenomena will be addressed. The first is the Lorentz invariance of the Dirac equation, which is the principle behind the observation of Landau level collapse detailed in chapter 5. The second is the atomic collapse, which is a phenomenon analogously reproduced in graphene and a brief connection

will be shown in chapter 6.

4.1 Relativistic Electromagnetism in a Nutshell

We are all accustomed to imagining things in three spatial dimensions which is described by 3-vectors. In this formalism the governing equations of electrodynamics— Maxwell's equations, are written as:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \qquad \text{(Gauss's law)}, \tag{4.1}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \qquad \text{(Gauss's law for magnetism)}, \tag{4.2}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{(Faraday's law)}, \tag{4.3}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \qquad \text{(Maxwell-Ampere's law)}. \tag{4.4}$$

In 1905, Einstein published the theory of special relativity that embodied and unified many of the theoretical and experimental results on light obtained by various notable scientists. In his work, Einstein postulates the following:

1. The laws of nature are identical in all inertial frames of reference.
2. The speed of light c is the same to all observers at rest in inertial frames of reference.

These two statements are the postulations of special relativity and a direct consequence of maintaining the validity of the postulations in our physical laws is that the spatial and temporal dimensions must be interwoven. This four dimensional space where time is merely another dimension just like space is called spacetime. In the special case where spacetime is flat it is called Minkowski spacetime. Generically, spacetime can be curved but one can safely assume that locally (especially here on earth) spacetime is flat and this will be the assumption throughout the thesis.

The validity of special relativity has been demonstrated numerous times and it is inarguable that the world we live in is relativistic. With this thought in mind, Maxwell's equations can possibly be cast with a different and perhaps a more accurate perspective. Make no mistake, Maxwell's equations are correct as they stand but concealed in these equations is a deeper manifestation of the relativistic reality we live in. As shown by Lorentz (and a few others), Maxwell's equations are invariant under the Lorentz transformation. This is to say, the equations, when Lorentz boosted to another moving inertial frame, retain the exact same mathematical form. Looking at the equations directly, it is hard to see this Lorentz invariance. Therefore, there is an advantage in being able to write out the equations that immediately show Lorentz invariance, and this is simply done by writing out the equations in a Lorentz covariant form, i.e. in terms of four-tensors. If this is possible, we say that the equations are Lorentz invariant.

Quantities that are expressed in 4-vector form automatically satisfy Einstein's postulates of relativity because by definition a 4-vector is a set of four numbers that transform according to the Lorentz transformation. The interval defined by the 4-vector is invariant, i.e., they are the same across different inertial frames of reference. The magnitude of the 4-vector is the interval and is called the Lorentz invariant, or Lorentz scalar. The 4-vectors are formulated based on the foundation of the Minkowski spacetime where space and time are treated at the same level. Below we list several relevant ones in

their contravariant forms:*

$$r^\alpha = (ct, \mathbf{r}) \qquad \text{(four-position)}, \tag{4.5}$$

$$\partial^\alpha = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\boldsymbol{\nabla}\right) \qquad \text{(four-gradient)}, \tag{4.6}$$

$$p^\alpha = \left(\frac{\varepsilon}{c}, \mathbf{p}\right) \qquad \text{(four-momentum)}, \tag{4.7}$$

$$J^\alpha = (\rho c, \mathbf{J}) \qquad \text{(four-current)}, \tag{4.8}$$

$$k^\alpha = \left(\frac{\omega}{c}, \mathbf{k}\right) \qquad \text{(four-wavevector)}, \tag{4.9}$$

$$A^\alpha = \left(\frac{\phi}{c}, \mathbf{A}\right) \qquad \text{(four-potential)}, \tag{4.10}$$

where the indices correspondence goes as: $\{t, x, y, z\} = \{0, 1, 2, 3\}$, which is to say, index 0 corresponds to the temporal dimension and indices $1, 2, 3$ are the spatial dimensions (coordinates of the 3-vector). Regarding describing electromagnetism in Minkowski spacetime, or in any model for that matter, there is a need to express the four-potential in terms of the electromagnetic fields because fields are the more physically enlightening quantities. Using the following fields' relations with the poten-

*Contravariant and covariant vectors are projected onto bases that are orthnormal to one another. As a result, a contravariant vector has coordinates (denoted by superscripts) that scale inversely to the basis' length while a covariant vector has coordinates (denoted by subscripts) that scale together with the basis' length. There is a need to distinguish contravariant and covariant 4-tensors because together they contain information about the geometry of spacetime. In the case of curved spacetime, where the metric is no longer as simple as equation 4.21, the use of contravariant and covariant 4-tensors become critical. In this thesis, however, we will only deal with a flat spacetime, in which case there is limited reason to use contravariant and covariant vectors except as a bookkeeping device (which can be useful because they follow different Lorentz transformations), as well as to follow convention in the field of relativity. Ultimately, what we want to show is that the Dirac equation is Lorentz invariant and following convention will be important as one will see. For the mystified reader, the vectors we usually deal with is contravariant. The covariant vector is the one that can contain extra information. Regarding the 4-vectors, in flat spacetime, going from contravariant to covariant the spatial indices acquire a negative sign relative to the temporal index. For more information on the application of this mathematical tool in relativistic electrodynamics, the reader is referred to the cited reference.[66]

tials:

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi, \tag{4.11}$$

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}, \tag{4.12}$$

and writing out the fields explicitly in terms of the four-potential coordinates, one can see a symmetry arise in the way both electric and magnetic fields can be written out. It turns out that both fields can be encompassed in one entity which is a four-tensor conventionally called the electromagnetic field tensor:

$$F^{\alpha\beta} = \partial^{\alpha} A^{\beta} - \partial^{\beta} A^{\alpha} = \begin{bmatrix} 0 & -\frac{E_x}{c} & -\frac{E_y}{c} & -\frac{E_z}{c} \\ \frac{E_x}{c} & 0 & -B_z & B_y \\ \frac{E_y}{c} & B_z & 0 & -B_x \\ \frac{E_z}{c} & -B_y & B_x & 0 \end{bmatrix}. \tag{4.13}$$

Another 4-tensor can be concocted to compliment the electromagnetic field tensor and this other 4-tensor is called the dual electromagnetic field tensor:

$$G_{\alpha\beta} = \frac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}F^{\gamma\delta} = \begin{bmatrix} 0 & B_x & B_y & B_z \\ -B_x & 0 & \frac{E_z}{c} & -\frac{E_y}{c} \\ -B_y & -\frac{E_z}{c} & 0 & \frac{E_x}{c} \\ -B_z & \frac{E_y}{c} & -\frac{E_x}{c} & 0 \end{bmatrix}, \tag{4.14}$$

which is written in covariant form and where $\varepsilon_{\alpha\beta\gamma\delta}$ is the Levi-Civita symbol given by:

$$\varepsilon_{\alpha\beta\gamma\delta} = \begin{cases} +1 & \text{if } \{\alpha\beta\gamma\delta\} \text{ is an even permutation of } \{1,2,3,4\} \\ -1 & \text{if } \{\alpha\beta\gamma\delta\} \text{ is an odd permutation of } \{1,2,3,4\} \\ 0 & \text{otherwise} \end{cases}, \tag{4.15}$$

To wrap up the discussion on 4-tensors: 4-tensors of any rank (where a 4-vector is a 4-tensor of rank 1) all transform from an inertial frame to another according to the Lorentz transformation, i.e., they are all Lorentz invariant. In the contravariant case:

$$S'^{\alpha\beta\ldots} = L^{\alpha}_{\ \kappa} L^{\beta}_{\ \lambda} \ldots S^{\kappa\lambda\ldots}, \tag{4.16}$$

and in the covariant case:

$$S'_{\alpha\beta\ldots} = (L^{-1})^{\kappa}_{\ \alpha} (L^{-1})^{\lambda}_{\ \beta} \ldots S_{\kappa\lambda\ldots}, \tag{4.17}$$

where S and S' are any four-tensor in the lab and moving inertial frames respectively. The Lorentz transformation matrices for a boost in the x direction are given by:

$$L^{\alpha}_{\ \beta} = \begin{bmatrix} \gamma & -\frac{v}{c}\gamma & 0 & 0 \\ -\frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (L^{-1})^{\alpha}_{\ \beta} = \begin{bmatrix} \gamma & \frac{v}{c}\gamma & 0 & 0 \\ \frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.18}$$

where v is the speed of the moving inertial frame and the Lorentz factor is given by:

$$\gamma = \frac{1}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}, \tag{4.19}$$

which should not be confused with the gamma matrices γ^i and γ_i (contravariant and convariant forms) which are identified by the presence of superscripts or subscripts. Note also that it is common to define $\beta = v/c$ but we will refrain from doing this here to avoid confusion with the Greek letter indices. To go from contravariant to covariant, the following operation is followed:

$$S_{\alpha\beta\ldots} = g_{\alpha\gamma} g_{\beta\delta} \ldots S^{\gamma\delta\ldots}, \tag{4.20}$$

where in the flat Minkowski spacetime (assumed in the entirety of this thesis), the metric is given by:

$$g_{\alpha\beta} = g^{\alpha\beta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{4.21}$$

The reason behind coming up with the two electromagnetic field 4-tensors is three-fold. First, the electromagnetic field tensor tells us how electric and magnetic fields transform from one frame to another and here we see that the two fields are completely intermingled. When we move from frame to frame, even when we start off with an electric field and no magnetic field in one frame, we may end up instead with only a magnetic field and no electric field in the other. The electric and magnetic field really is just one entity with different manifestations. To see how the fields transform, we perform the

Lorentz transformation operation on the electromagnetic field tensor as in equation 4.16. With some tedious math, for a Lorentz boost from the lab frame (S) to the moving frame (S') traveling at velocity $\mathbf{v}$, we have:

$$\begin{aligned} \mathbf{E}'_{||} &= \mathbf{E}_{||}, \qquad & \mathbf{E}'_{\perp} &= \gamma(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}), \\ \mathbf{B}'_{||} &= \mathbf{B}_{||}, \qquad & \mathbf{B}'_{\perp} &= \gamma\left(\mathbf{B}_{\perp} - \frac{1}{c^2}\mathbf{v} \times \mathbf{E}\right). \end{aligned} \tag{4.22}$$

For the second reason, Maxwell's equations can now be transcribed in terms of the electromagnetic field tensors as follows:

$$\partial_\alpha F^{\alpha\beta} = \mu_0 J^\beta, \tag{4.23}$$

$$\partial_\alpha G^{\alpha\beta} = 0, \tag{4.24}$$

This set of two equations encompasses all four of Maxwell's equations and here we see how Maxwell's equations belong in the four dimensional spacetime. The significance of this is, since Maxwell's equations can be expressed in a covariant formulation, Maxwell's equations must be Lorentz invariant. For the third and final reason, the electromagnetic field tensors are used to determine the field-associated Lorentz invariants. Using the two tensors, we arrive at the following Lorentz invariants:

$$F_{\alpha\beta}F^{\alpha\beta} = \text{invariant} = 2\left(B^2 - \frac{E^2}{c^2}\right), \tag{4.25}$$

$$F_{\alpha\beta}G^{\alpha\beta} = \text{invariant} = -\frac{4}{c}\mathbf{E} \cdot \mathbf{B}. \tag{4.26}$$

The first invariant tells us that if the magnetic field energy is greater than the electric field energy in one frame, then this will hold true in all other reference frames. The same applies for the opposite case. If

$B < E/c$ in one frame, then $B < E/c$ for all frames. For the case of light where the magnetic field energy is equal to the electric field energy, the invariant is zero which means $B = E/c$ for all frames. The second invariant tells us that if $\mathbf{E}$ and $\mathbf{B}$ form an acute (or obtuse angle), then they will make an acute (obtuse) angle for all frames. In the particular case of $\mathbf{E} \cdot \mathbf{B} = 0$, not only will the fields be perpepndicular, but one can always find a reference frame in which $E = 0$ or $B = 0$ depending on $B^2 - E^2/c^2 <$ or > 0, i.e., the field will become either purely magnetic or purely electric.[66] The implications of these two Lorentz invariants will serve as the benchmark tests for verifying whether Gr electrons fall under the relativistic category. This topic will be investigated in chapter 5.

4.2 Brief Timeline Leading Up to the Full-Fledged Relativistic Quantum Theory

The dawn of quantum theory took place throughout the 1800's when many opposing propositions regarding the theory were made. Along with the propositions were discoveries that only gradually hinted at the existence of the quantum world (or more correctly, the atomic world. The idea of quantization did not exist till later). It is the turn of the twentieth century that the majority of significant findings occurred, allowing the quantum theory to truly take form.

In 1900, Max Planck postulated that electromagnetic energy could only be emitted in quantized form (later termed photon) and in multiples of h— Planck's constant. This constant put the development of quantum theory onto a road-map and serves as the very foundation that all of quantum theory rests upon. In the ensuing short period of time the world would see a flurry of experimental and theoretical discoveries but it was 1924 when Louis de Broglie proposed the idea of wave-particle

duality through his most famous equation:

$$\lambda = \frac{h}{p}, \tag{4.27}$$

that further catapulted the quantum theory into prominence. Here, λ is the de Broglie wavelength, p is momentum and h is Planck's constant as mentioned. This equation, which states that anything that has a momentum will also have a determinable wavelength, is arguably one of the most influential in physics because it inspired Erwin Schrödinger to come up with one of the most important equation in the history of man-kind— Schrödinger's equation:

$$i\hbar\frac{\partial}{\partial t}\Psi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t), \tag{4.28}$$

where $\hbar = h/2\pi$ is the reduced Planck's constant, m is the mass of the particle of interest, $V(\mathbf{r},t)$ is the energy potential the particle experiences at a given position and time, and finally, $\Psi(\mathbf{r},t)$ is the wavefunction— a quantity whose squared amplitude tells us directly the probability of finding the particle at a certain position and time. This equation can be simply viewed as a restatement of energy conservation but in the language of quantum mechanics. The left-most term yields the total energy, the first term on the right hand side yields the kinetic energy and the last term yields the potential energy. Published in 1926, this equation holds tremendous importance as it completely revolutionized our understanding of the universe. Our world, as we know it, is not deterministic but probabilistic! The equation, however, is not completely general as it only describes non-relativistic particles.

During the rise of quantum theory, Albert Einstein introduced the theory of relativity in which one of the main takeaway points is that time is merely another physical dimension like space and they all belong to a greater common entity called space-time. There is no reason why one dimension should

be treated differently then any other. But if we look at Schrödinger's equation we can see that spatial derivatives are second order whereas the time derivative is of first. This suggested that space and time were treated differently in the quantum theory and that some sort of reconciliation between space and time was required.

In 1928, Paul Dirac successfully combined the concepts from the theory of relativity with the ones from quantum theory and the result was the Dirac equation. This equation is essentially a modified Schrödinger's equation and is able to describe accurately the effects a particle experiences due to relativity which becomes more prominent as the particle travels closer to the speed of light. It is interesting to note that in his attempt to find the Dirac equation, Paul Dirac simply treated the problem at hand as a mathematical one and solved it via mathematical manipulation, much like his predecessor Erwin Schrödinger who also arrived at his equation through mathematical trial and error. The physical interpretation only came after.

The root of the Dirac equation is Einstein's energy-momentum relation formula which was derived in Einstein's special theory of relativity and is given by:

$$\varepsilon^2 = p^2c^2 + m^2c^4, \tag{4.29}$$

where ε is the energy of the particle, m is the rest mass of the particle, c is the speed of light and p is the momentum of the particle. By employing the concept of first quantization, which indirectly means transforming observables into operators that can act on a wavefunction, where

$$\varepsilon = i\hbar\frac{\partial}{\partial t}, \quad \mathbf{p} = -i\hbar\boldsymbol{\nabla}, \tag{4.30}$$

we acquire the following transformed energy-momentum relation:

$$-\hbar^2 \frac{\partial^2 \Psi(\mathbf{r},t)}{\partial t^2} = \big(-\hbar^2 c^2 \nabla^2 + m^2 c^4\big)\Psi(\mathbf{r},t). \tag{4.31}$$

This equation is known as the Klein-Gordon equation and is fundamental to the understanding of modern quantum theory (such as quantum field theory) where the treatment of Bosons (integer spin) is more prevalent and where this equation is more applicable. The use of this equation, when it comes to condensed matter systems which are dominated by fermions (half spin), is rather limited however. At the time when Paul Dirac was facing this equation, his intention was to transform the spatial and time derivatives to first order. The reason behind this was that the Klein-Gordon equation yielded seemingly unphysical solutions due to it being second order, i.e., it allowed negative probabilities (although the interpretation behind the seemingly unphysical solutions was later on resolved by Richard Feynman[67]). His solution to transforming the equation into a first order equation (by taking square root of both sides of the equation perfectly) was to convert the equation into a four component matrix problem. He introduced a four-component wavefunction and 4×4 matrices that would allow him to properly take the square root of the operators. The final form he arrived to was:

$$i\hbar \frac{\partial \Psi(\mathbf{r},t)}{\partial t} = \big(-i\hbar c\, \boldsymbol{\alpha} \cdot \boldsymbol{\nabla} + \beta m c^2\big)\Psi(\mathbf{r},t). \tag{4.32}$$

This is the Dirac equation written out in full. The matrices he introduced were:

$$\begin{aligned} \beta &= \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}, \quad \alpha_1 = \begin{bmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{bmatrix}, \\ \alpha_2 &= \begin{bmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{bmatrix}, \quad \alpha_3 = \begin{bmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{bmatrix}, \end{aligned} \tag{4.33}$$

where I_2 is the 2×2 identity matrix and $\boldsymbol{\sigma}$ are the Pauli matrices given by:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{4.34}$$

It is sometimes more useful to cast the original Dirac equation into the Lorentz covariant form (which implies Lorentz invariance):

$$(ic\hbar\gamma^\mu \partial_\mu - mc^2)\Psi(\mathbf{r}, t) = 0. \tag{4.35}$$

This is the Dirac equation in covariant form written in Einstein's notation, meaning that there is an implied summation over the twice-repeated index μ which takes on the value of $0, 1, 2, 3$ and they correspond to the t, x, y, z dimensions respectively. $\partial_\mu = (c^{-1}\partial/\partial t, \boldsymbol{\nabla})$ is the covariant four-gradient and γ^μ are the 4×4 gamma (or Dirac) matrices. The matrices (in chiral representation and in con-

travariant form) are as follows:

$$\gamma^0 = \begin{bmatrix} 0 & I_2 \\ I_2 & 0 \end{bmatrix}, \quad \gamma^1 = \begin{bmatrix} 0 & \sigma_x \\ -\sigma_x & 0 \end{bmatrix}, \\ \gamma^2 = \begin{bmatrix} 0 & \sigma_y \\ -\sigma_y & 0 \end{bmatrix}, \quad \gamma^3 = \begin{bmatrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{bmatrix}. \tag{4.36}$$

Note that there are many other gamma matrices representations. The choices will work as long they obey the following anticommutation relations:

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4 \tag{4.37}$$

where $g^{\mu\nu}$ is the metric tensor and I_4 is the 4×4 identity matrix. This anticommutation ensures that the square root of the operators is possible. The eigenvector for the Dirac equation is the four component wavefunction:

$$\Psi(\mathbf{r}, t) = \begin{bmatrix} \Psi_{+,\uparrow}(\mathbf{r}, t) \\ \Psi_{+,\downarrow}(\mathbf{r}, t) \\ \Psi_{-,\uparrow}(\mathbf{r}, t) \\ \Psi_{-,\downarrow}(\mathbf{r}, t) \end{bmatrix}, \tag{4.38}$$

where the entries from top to bottom correspond to: particle with up spin, particle with down spin, anti-particle with up spin, anti-particle with down spin; and the eigenvalue is simply Einstein's energy-momentum relation given by equation 4.29. The discovery of this equation revealed tremendous amount of untapped information about our world at the time. Its prediction of the existence of particles and anti-particles as well as spin were hysterical notions that only proved that we knew basically

nothing. The birth, rise, crystallization of the quantum theory and leading up to now the utilization of it in commercial products such as the quantum computer is an amazing non-fiction. And all this is happening in just a century. To be able to learn and now witness the development of such an abstract field has been extremely awe-inspiring.

So far the written equations are for $3+1$ dimensions (three spatial and one temporal). In anticipation of dealing with graphene, we shall move on to its $2+1$ dimensional and massless counterpart. The equation becomes:

$$(ic\hbar\sigma^{\mu}\partial_{\mu})\psi(\mathbf{r}) = 0, \tag{4.39}$$

with $\sigma_{\mu} = \sigma^{\mu}$ for $\mu = \{0,1,2\} \Leftrightarrow \{t,x,y\}$ and $\sigma_0 = i\sigma_z$. The eigenfunction becomes a two-component spinor with each component corresponding to particle and anti-particle (see appendix A for details on the wavefunction and how it relates to Gr's lattice). The gamma matrices have essentially been replaced by the Pauli matrices which follow the same anticommutation relations:

$$\{\sigma_{\mu},\sigma_{\nu}\} = \sigma_{\mu}\sigma_{\nu} + \sigma_{\nu}\sigma_{\mu} = 2I_2. \tag{4.40}$$

Like the 3D case, this anticommutation allows the square root of the operators to be taken but in 2D and like the gamma matrices, there is more than one way of designating the Pauli matrices. The important point here is that the Dirac equation retains its Lorentz covariant form meaning that the equation is still Lorentz invariant in 2D. In the steady-state case we are forced to relinquish the covariant formulation and the time independent 2D massless Dirac equation becomes:

$$c\boldsymbol{\sigma}\cdot\mathbf{p}\Psi(\mathbf{r}) = \varepsilon\Psi(\mathbf{r}), \tag{4.41}$$

which we can see has a Hamiltonian that is identical to Gr's except v_F is replaced by c (see equation 1.7). Here $\mathbf{p} = \hbar\mathbf{q} = \hbar(q_x, q_y) = -i\hbar(\partial_x, \partial_y)$.*

The mathematical similarity here means that the equation of motion of Gr electrons is the same as that of massless Dirac Fermions— apart from their speeds. Bear in mind that Gr electrons are still traveling at only $\sim c/300$ which is nowhere near relativistic speeds. So more accurately speaking, these electrons are pseudo-relativistic but this does not detract the equivalence in their observable behaviors— and this equivalence is one of the most outstanding property of Gr and has been put to the test in numerous occasions. One of the most significant experiments that substantiates the pseudo-relativistic nature of Gr electrons is Klein tunneling.

4.3 Klein Tunneling

A big part of this work is to study the effects of a one dimensional potential on the massless Dirac fermions, which is analogously provided by graphene in the low energy limit. This problem can be transcribed into the following Hamiltonian:

$$\mathbf{H} = \hbar v_F(\boldsymbol{\sigma} \cdot \mathbf{q}) + V(x), \tag{4.42}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$ are the Pauli spin matrices and $\mathbf{q} = (q_x, q_y)$ are the momenta with respect to the Dirac point in the x, y directions only (see appendix A for more information regarding Gr's Hamiltonian). More importantly is $V(x)$, the external one dimensional potential. The profile of this potential must be considered very carefully. The above Hamiltonian is rendered inapplicable to the massless Dirac fermions of Gr when $V(x)$ is atomically sharp. This is because once the potential

*The massless Dirac equation is more accurately known as the Weyl equation.

becomes atomically sharp Gr is no longer blind to the two sublattices of graphene, i.e. sublattice A and B can behave differently under the same potential. In the presence of atomically sharp potentials, one would have to consider intervalley mixing and it will become necessary to consider the full Gr Hamiltonian as opposed to the low energy limit and linear regime of the Hamiltonian. This will cause us to deviate from the original problem we want to pursue.

To retain the Dirac equation of Gr we have to operate under this condition:

$$l_E \equiv \sqrt{\frac{\hbar v_F}{|\nabla V(\mathbf{r})|}} \gg a_{\text{cc}}. \tag{4.43}$$

l_E is a length that defines the sharpness of the potential barrier.[4] $V(\mathbf{r})$ must be smooth varying compared to the carbon-carbon bond length (a_{cc}).

In the following, we will briefly address the mathematical description of Klein tunneling. We will always assume that Gr retains its linear and low energy limit Hamiltonian, hence implicitly assuming the smoothness of the potentials with respect to the atomic length scale. This assumption can aptly be applied to our CNT-Gr hybrid device. Our CNT is on the order of ~ 1 nm wide in diameter. The closest separation between the CNT and Gr is 4 nm. Considering these two parameters, as one shall see in our electrostatic simulation, the shortest span of the potential barrier side wall will roughly be $l_E \gtrsim 4$ nm. This is over hundreds times larger than the carbon-carbon bonding distance ($a_{cc} \approx 1.42$ Å) and so we are safely in the smooth potential regime always. Under this smooth potential regime, there exists a sharp and smooth potential relative to the Fermi wavelength. As we will see, the two yields different behaviors.

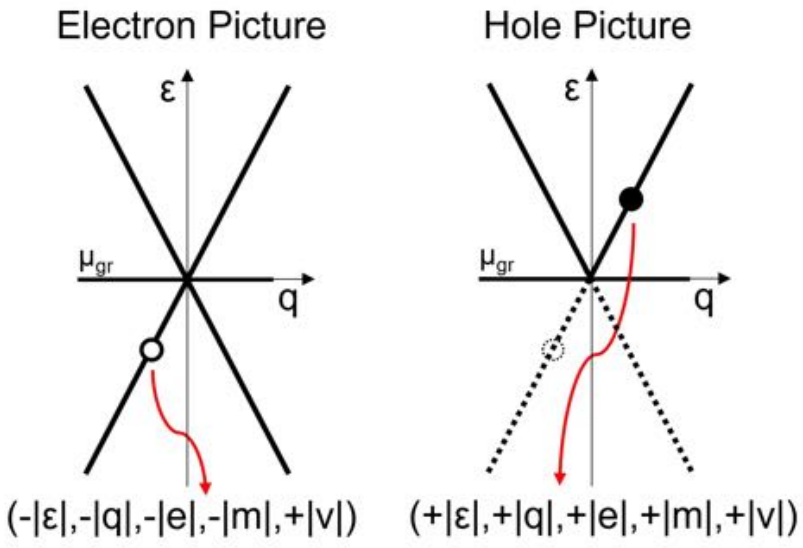


Figure 4.1: Illustrations to show the distinction between electron and hole. The left shows a valence band with a missing electron. The dynamics of the whole valence band in the electron picture are given by all the occupied states. The right shows a hole in a hole band. This hole describes completely the dynamics of the whole partially filled valence band.

4.3.1 Clarification: Electron vs Hole

Whenever there is a need to address both valence and conduction bands (or excitations of electron-hole pairs about the Fermi level), there is a natural need to use the electron and hole concept. Often times we find ourselves confused when describing the state of a missing electron and a hole. The two are different and below we will elucidate the difference.

Unfilled bands contribute to the overall electrical dynamics of a solid. For a valence band with missing electrons, a.k.a. partially filled valence band, the dynamics of this particular band is given by all the electrons occupying the band. To calculate the electrical current from this band would be extremely inconvenient. To facilitate describing the dynamics of a partially filled valence band, we can introduce the concept of a hole.

Looking at figure 4.1, we have in this case a charge neutral graphene where we shall call the energy at the chemical potential $\varepsilon = 0$. The state that the missing electron in the valence band would otherwise occupy has these dynamic properties (stated in terms of polarity): $(-|\varepsilon|, -|q|, -|e|, -|m|, +|v|)$. To speak in terms of holes, we need to first have a constructed hole band in mind which is created through these relations (only valid for symmetric conduction and valence bands, which is the case for graphene):

$$\varepsilon_h = -\varepsilon_e, \tag{4.44}$$

$$\mathbf{k}_h = -\mathbf{k}_e, \tag{4.45}$$

where the subscript h and e refers to the hole of the hole band and the missing electron of the partially filled valence band respectively. This simply means that the hole band is an inversion of the electron valence band in both energy and momentum space. Here, $\mathbf{k}$ refers to the crystal momentum with respect to the origin of the Brillouin zone and the above relation applies to the full graphene BZ. In forming the hole band for graphene, one will see that $\mathbf{q} = \mathbf{k} - \mathbf{K}$ (the electron crystal momentum respect to the Dirac point) rightfully becomes $-\mathbf{q}$ (the hole momentum in the hole band). The remaining properties such as effective mass and velocity is given by the same original equations except applied onto the hole band. Note that the number of holes corresponds to the number of missing electrons in the partially filled band. As for the electric charge of the hole, it is positive and this is so that the equation of motion for the hole stays consistent with the equation of motion for the partially filled band that the hole is replacing.[68] To summarize, the dynamics of the whole valence band with the missing electron behaves exactly as though it were a single hole particle in the hole band occupying the state: $(+|\varepsilon|, +|q|, +|e|, +|m|, +|v|)$. Care must be taken to distinguish electrons in the valence

band and holes representing a valence band with missing electrons.* Finally, keep in mind that there is no real particle associated with a hole. It is a quasi-particle manifested by the combined behavior of all the electrons in a partially filled band.

4.3.2 1D Klein Tunneling through Different Types of Potentials

Using the equation of motion in Heisenberg's picture, the velocity operator for graphene in the low energy limit around the $\mathbf{K}$ valley is given by:

$$\begin{aligned} \mathbf{v} \equiv \frac{d\mathbf{r}}{dt} &= \frac{i}{\hbar}[\mathbf{H}, \mathbf{r}] \\ &= \frac{i}{\hbar}[\hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q}, \mathbf{r}] \\ &= v_F \boldsymbol{\sigma}, \end{aligned} \tag{4.46}$$

where $\boldsymbol{\sigma}$ is the usual Pauli spin matrix operator which commutes with $\mathbf{r}$ and $\mathbf{q}$ is the momentum with respect to the Dirac point and has the usual commutation relation with real space position. From this we can see that the direction of velocity is given by graphene's pseudo-spin and the magnitude is its Fermi velocity (as for the $\mathbf{K}'$ valley, the velocity operator is given by $\boldsymbol{\sigma}^*$ instead). The average velocity of a particle with momentum $\mathbf{q}$ is then:

$$\begin{aligned} \mathbf{v} &= \langle \boldsymbol{\Psi}(\mathbf{q}) | \, v_F \boldsymbol{\sigma} \, | \boldsymbol{\Psi}(\mathbf{q}) \rangle \\ &= \pm v_F \frac{\mathbf{q}}{|\mathbf{q}|}, \end{aligned} \tag{4.47}$$

*In general, one is better off not mixing pictures for a given band; however, in some cases such as Andreev reflection, there is a natural need to incorporate the idea of both electrons and holes within the same band. In this particular case, electrons and holes are similarly defined but with respect to the Fermi level that lies inside a band. This serves to distinguish particles above and below the Fermi level; however, in this thesis, this phenomenon is not relevant and there is no need to worry about mixing electrons and holes in the same band.

where positive and negative are for conduction and valence band electrons respectively. Note that this is the same for the opposite valley. Expectedly, the semiclassical approach for calculating group velocity ($\mathbf{v} = (1/\hbar)\partial\varepsilon/\partial\mathbf{k}$) yields the same information though one needs to note that the semiclassical group velocity is intended for a wavepacket whereas the Heisenberg picture describes a quantum mechanical single particle.

Next up is to see whether the velocity, or equivalently the psuedospin, of the electron is conserved over time as it traverses the potential barrier. In the presence of a 1D potential $V(x)$, the time evolution of the x component velocity operator in Heisenbeg's picture is:

$$\begin{aligned} \frac{dv_x}{dt} &= \frac{i}{\hbar}[\mathbf{H}, v_x] \\ &= \frac{i}{\hbar}[\hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q} + V(x), v_F \sigma_x] \\ &= 2v_F^2 \sigma_z q_y. \end{aligned} \tag{4.48}$$

To find the value of q_y, we note that because of translational symmetry in the y direction, we have:

$$\frac{dq_y}{dt} = \frac{i}{\hbar}[\mathbf{H}, q_y] = 0, \tag{4.49}$$

so q_y is a constant, i.e., $q_y(t) = q_y(0)$. And so if we take $q_y = 0$, which means the incident electron is heading perpendicular to the potential barrier, the velocity of the electron will not change— it will penetrate through the barrier without fail. This is also referred to as conservation of pseudo-spin.

The perfect transmission at angles perpendicular to the potential barrier is a robust feature regardless of the potential type and is a true testament to Gr electron's analogous behavior with relativistic particles. Its general transmission characteristics such as the transmission angle dependence, however, will have to be calculated differently depending on the potential barrier profile. For a sharp potential

where $l_E \ll \lambda_F$, the transmission behavior is solved following the quantum tunneling treatment.[69] By imposing the continuity conditions on Gr wavefunctions spatially across the junction, for a bipolar sharp potential step we have:

$$|T(\theta)|^2 = -\frac{\cos\theta\cos\theta_t}{\sin^2\left(\frac{\theta+\theta_t}{2}\right)}, \tag{4.50}$$

where θ is the incidence angle away from the normal angle and the transmission angle is:

$$\theta_t = \sin^{-1}\left(\frac{\varepsilon}{V_0-\varepsilon}\sin\theta\right) + \pi. \tag{4.51}$$

The details of the sharp potential is inconsequential to the result. The tunneling mechanism is solely reliant on the conservation of Gr's pseudo spin and the existence of matching energy states after and before the potential step. Strictly speaking, this is not a tunneling effect because there is no classically forbidden region and there is no propagation of evanescent waves.[70]

For a bipolar smooth potential step where $l_E \gtrsim \lambda_F$, the scenario was treated via method of least action.[71] The transmission probability for the smooth potential is:

$$|T(\theta)|^2 = e^{-\pi\hbar v_F q_y^2/eE} = e^{-\pi q_F w \sin^2\theta}, \tag{4.52}$$

where in the last step we have assumed the potential drop at the barrier to be linear and goes as $\varepsilon = eEw$, w being the potential drop distance. For the tunneling across a smooth potential, it is shown that at non-head-on angles the tunneling relies on evanescent waves propagation and hence is more in line with actual quantum tunneling. The transmission probabilities of the sharp and smooth potentials are compared and contrasted in figure 4.2. For both sharp and smooth potentials, the transmission is perfect if the angle of incidence is zero as expected. A key difference between them is that

for the sharp potential case, transmission is always quite large whereas for the smooth potential case, the transmission probability dies off sharply away from normal incidence. A key takeaway point here is smooth potentials have a true collimation effect.

For the transmission coefficient for a square potential barrier (sharp npn or pnp potential), we follow the same treatment as the sharp bipolar potential. Imposing the continuity conditions on Gr wavefunctions spatially across all junctions, we get in the low energy limit:

$$|T(\theta)|^2 = \frac{\cos^2\theta}{1 - \cos^2(q_x' d)\sin^2\theta}, \tag{4.53}$$

where $q_x' d = -2\pi l\sqrt{1 - 2(\varepsilon/V_0) + (\varepsilon/V_0)^2\cos^2\theta}$ and $l = V_0 d/\sqrt{2\pi\hbar v_F} = V_0 d/2\pi$.[72] Figure 4.3 shows a case for npn tunneling. As an electron goes from the conduction band to the valence band in the barrier, momentum is flipped but velocity stays pointed in the same direction along x. This is a result of pseudo-spin conservation. Pseudo-spin is flipped going from conduction to the valence band. The reason velocity does not flip but momentum does is that electric current needs to be conserved but momentum in the x direction does not. Keep in mind that this whole Klein tunneling process we are using solely the electron picture.That is, an electron in the conduction band enters a valence band as an electron, occupying an originally empty state.[70] It is helpful to think of the valence band as being empty (though of course it is not) in which case the concept of hole will become out of place. This electron state in the valence band is then transported across the barrier with an opposite momentum until it hits the other junction and reverts to an electron in the conduction band.

Figure 4.3b and 4.3c show the transmission probabilities as functions of the potential width and density. In both cases, we see a petal shape form. This is due to Fabry-Perot resonance caused by the constructive interferences of the waves inside the potential. The condition for a peak in transmission

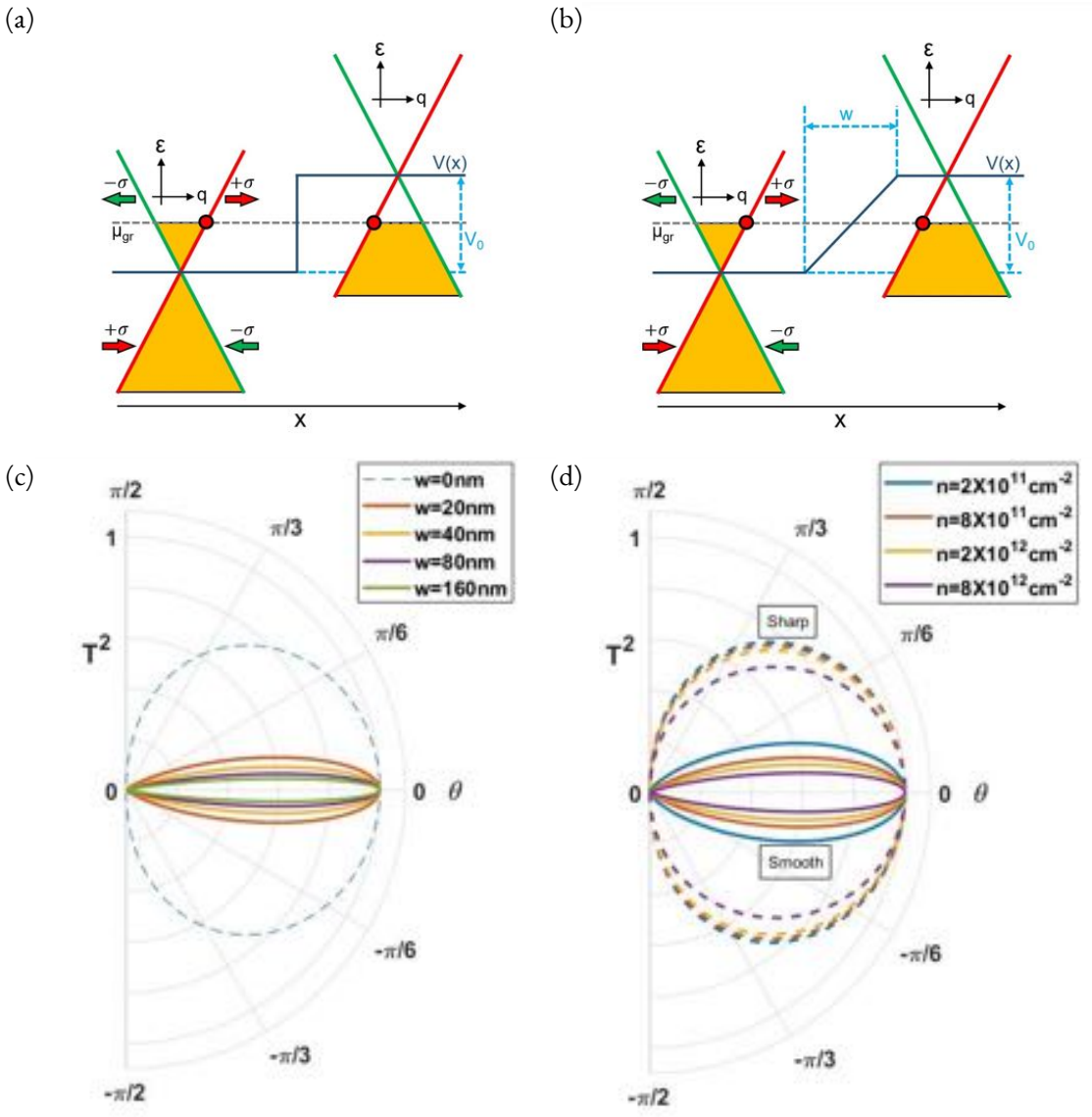


Figure 4.2: Klein Tunneling through a bipolar step potential. (a) Illustration of a sharp step potential (corresponding to $w = 0$). (b) Illustration of a smooth step potential. (c) The transmission probability ($\mathbf{T}^2$) of an electron from the conduction band to the valence band across a step potential of varying widths (w). The step height is set to be $V_0 = 1$ eV and the carrier density of the incident electrons is set to be 1×10^{12} cm^{-2}. (d) The transmission probability ($\mathbf{T}^2$) of an electron from the conduction band to the valence band across a sharp (dotted lines) and smooth (solid lines) step potential that is 100 nm wide with varying incident electron carrier densities.

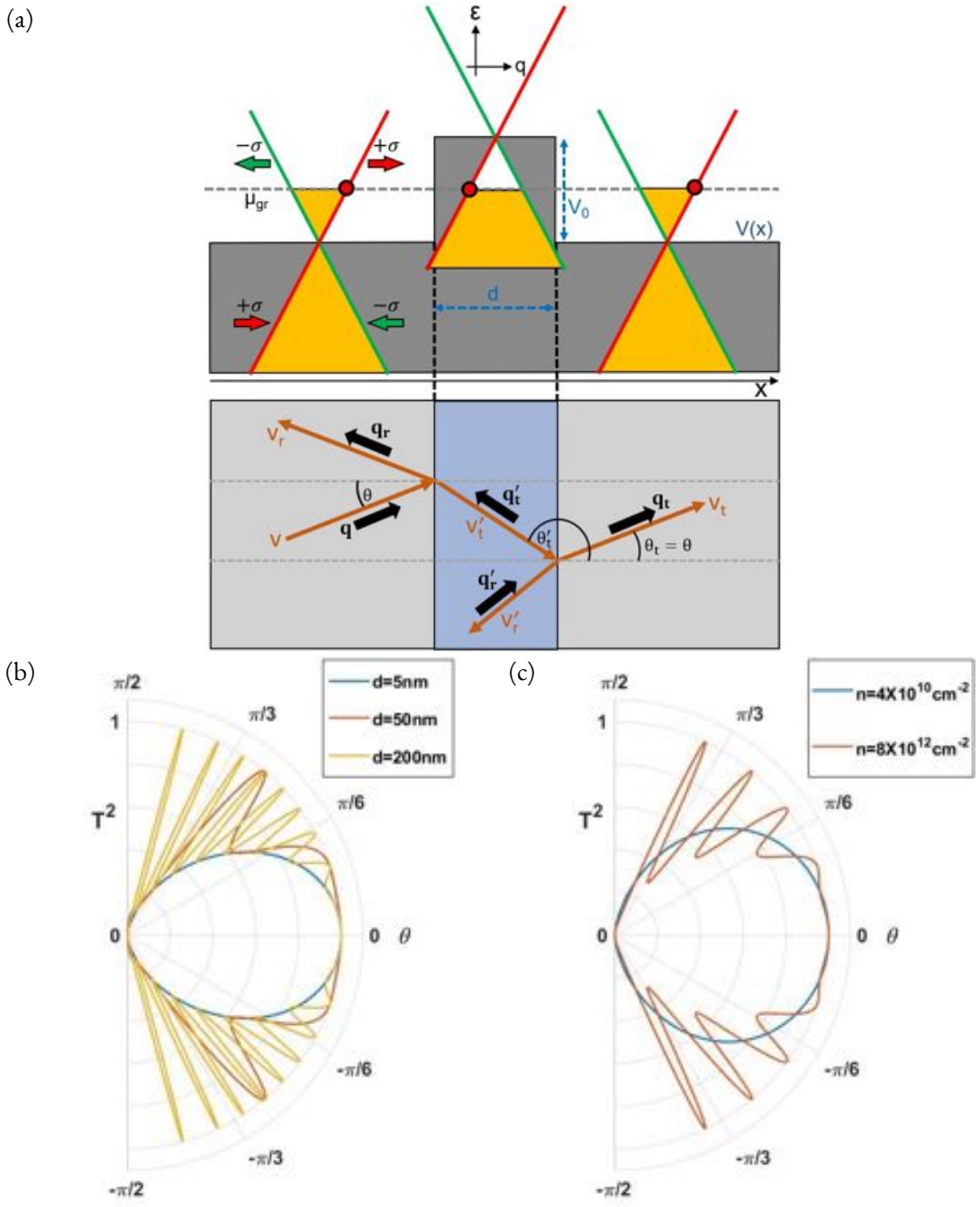


Figure 4.3: Klein Tunneling through a npn potential barrier. (a) Illustration of a sharp npn potential (in the electron picture). (b) The transmission probability as a function of potential width with $n = 8 \times 10^{12}$ cm^{-12}. (c) The transmission probability as a function of carrier density with $d = 100$ nm. The step height is $V_0 = 1$ eV.

probability is when the potential can fit the half period of the wave, so the peaks occur when:

$$q_x' d = \pi N, \tag{4.54}$$

for N being any integer. $\mathbf{q}'$ is the wavevector inside the barrier and d is the barrier width (see figure 4.3a). What remains to be addressed is the smooth bipolar potential barrier. It turns out the treatment for this scenario is not straightforward and the transmission probabilities do not have a clean presentable form. The general behavior, however, can be deduced by putting two smooth pn junctions in sequence and finding the transmission probability across the two junctions. Due to the phase accumulation within the barrier and the non-zero reflection probability at oblique angles, this results in a Fabry-Perot behavior as shown theoretically and experimentally.[73,74]

The theory of Klein tunneling will be a core part of physics in the experiment regarding electron confinement which will be shown in chapter 6. It also has a direct connection to one of the most well-known problems in high energy relativistic physics: atomic collapse.

4.4 Atomic Collapse

The idea of an atom was first postulated in the early 1800's by John Dalton. Fast forward a century, constituents of the atom like the electron and nucleus of an atom were discovered and these discoveries illuminated scientists towards a more microscopic description of the atom. The first decade of the 20$^{\text{th}}$ century saw a flurry of atomic theories that replaced one after another. The general picture finally took a more settled form in Bohr's model which was built upon the initial foundation of quantum mechanics and described electrons as having stationary orbits around a positively charged nucleus. Over time, with further development of quantum theory taking place, Bohr's model became more

refined and mature. Shortly after the rise of Dirac's equation and its application to the atomic theory, the atomic behavior became unprecedentedly consistent with experimental results. However, a glaring problem also arose which initially was termed a "catastrophe".[75] It begins by simply solving Dirac's equation with a Coulomb (or central) potential:

$$\left[(c\boldsymbol{\alpha}\cdot\mathbf{p}+\beta mc^2)-\frac{Ze}{r}\right]\psi(\mathbf{r})=\varepsilon_H\psi(\mathbf{r}). \tag{4.55}$$

The equation describes the hydrogen-like atom with atomic number Z and the solution is:

$$\varepsilon_H=mc^2\left[1+\frac{(Z\alpha)^2}{\left(n-|j|-\frac{1}{2}+\sqrt{\left(j+\frac{1}{2}\right)^2-(Z\alpha)^2}\right)^2}\right]^{-\frac{1}{2}}. \tag{4.56}$$

Examining the $1s^2$ state case where $n=1$ and $j=1/2$ (single spin), the catastrophe was if Z, the atomic number, became larger than $1/\alpha\approx 137$, where α is the fine structure constant, the energy becomes imaginary. This behavior was not well understood and for decades theorist sought for ways around the problem. By taking into account the finite size of the nucleus, they were able to push the condition at which the catastrophe occurs to $Z=172$ but they were unable to get rid of it completely. Eventually, after the discovery and acceptance of the existence of positrons which constituted the Dirac continuum, a physical interpretation came for this catastrophe and the phenomenon was termed "atomic collapse".[76,77]

4.4.1 Semiclassical Picture of Atomic Collapse

Atomic collapse can be best seen through the semiclassical perspective as shown in figure 4.4. In the non-relativistic case, the total energy of the orbiting electron is the sum of the kinetic energy (blue

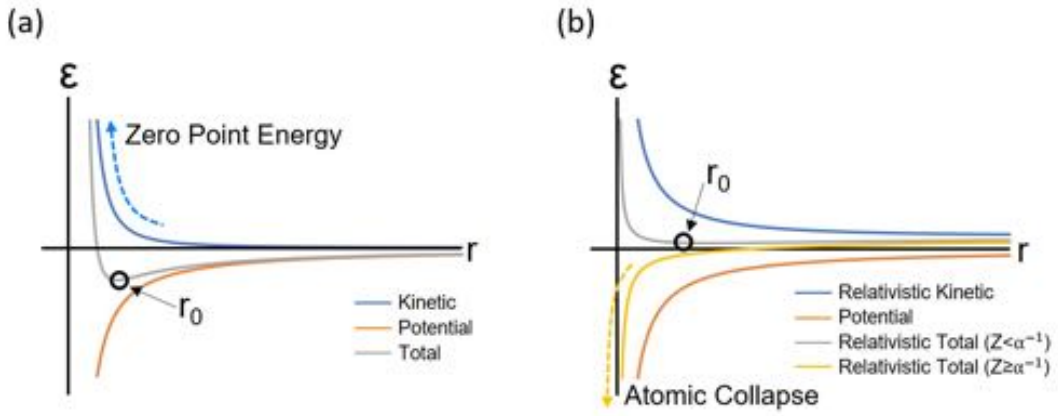


Figure 4.4: Comparison of the hydrogen atom energy in the non-relativistic and relativistic case. (a) shows the non-relativistic case. Summing the kinetic and potential energy results in a energy profile that has a energy minimum at r_0, the Bohr radius. (b) shows the relativistic regime. The kinetic energy is replaced by Einstein's relation and depending on the value of Z, there can either be a stable solution yielding the corrected Bohr radius or one that tells us that the electron is more energetically stable at the nucleus. The latter behavior is the atomic collapse.

curve) which has a $1/r^2$ dependence due to the zero point energy (restatement of Pauli's exclusion principle) and the potential energy coming from the electrostatic attraction to the nucleus which has a $-1/r$ dependence(orange curve). The total energy comes to the following:

$$\varepsilon = \frac{\hbar^2}{2mr^2} - \frac{Ze^2}{4\pi\epsilon_0 r}. \tag{4.57}$$

This total energy has a minimum at a distance called the Bohr radius which dictates the size of the atom and is given by:

$$r_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2}. \tag{4.58}$$

Turning to the relativistic case which is much more relevant when atoms become heavier, the kinetic term is replaced by Einstein's relation and the energy equation becomes:

$$\varepsilon = \sqrt{\frac{\hbar c}{r}^2 + (mc^2)^2} - \frac{Ze^2}{4\pi\epsilon_0 r}. \tag{4.59}$$

This energy equation nets you two behaviors. If the atomic number Z is below $1/\alpha$, where α is the fine structure constant, an energy minimum exists at a finite distance away and this is the corrected Bohr radius (see gray curve in figure 4.4), but if Z is greater than or equal to $1/\alpha$, the minimum exists at nowhere but at the center of the nucleus (yellow curve). This second behavior indicates that if the atomic number becomes higher than ~ 137, the electrons would spiral into the nucleus and the atom would collapse. By taking into account the finite size of the nucleus, the threshold at which atomic collapse would happen is predicted to be when $Z = 172$ and this condition is termed supercriticality Z_{cr}.[76,77]

4.4.2 Signatures of Atomic Collapse

Focusing on the behavior of the $1s^2$ state, (whose behavior can be extended to other higher energy states) as Z increases the orbit of the electrons would begin to shrink (figure 4.4.1b and c). Once the atom reaches supercriticality, the orbit size becomes on the same order as the nucleus size (~ 1.6 fm). The $1s^2$ bound state ceases to exist and this is meant by the state no longer being as localized as before. The originally bound state mixes with the negative energy continuum states and spreads out energetically in width to the order of Γ (see figure 4.4.1a). In quantum electrodynamics, the vacuum can be interpreted as the negative energy continuum called the Dirac sea. In this nomenclature, the vacuum now contains $-2e$ charges hence the notion of a charged vacuum. The charged vacuum only

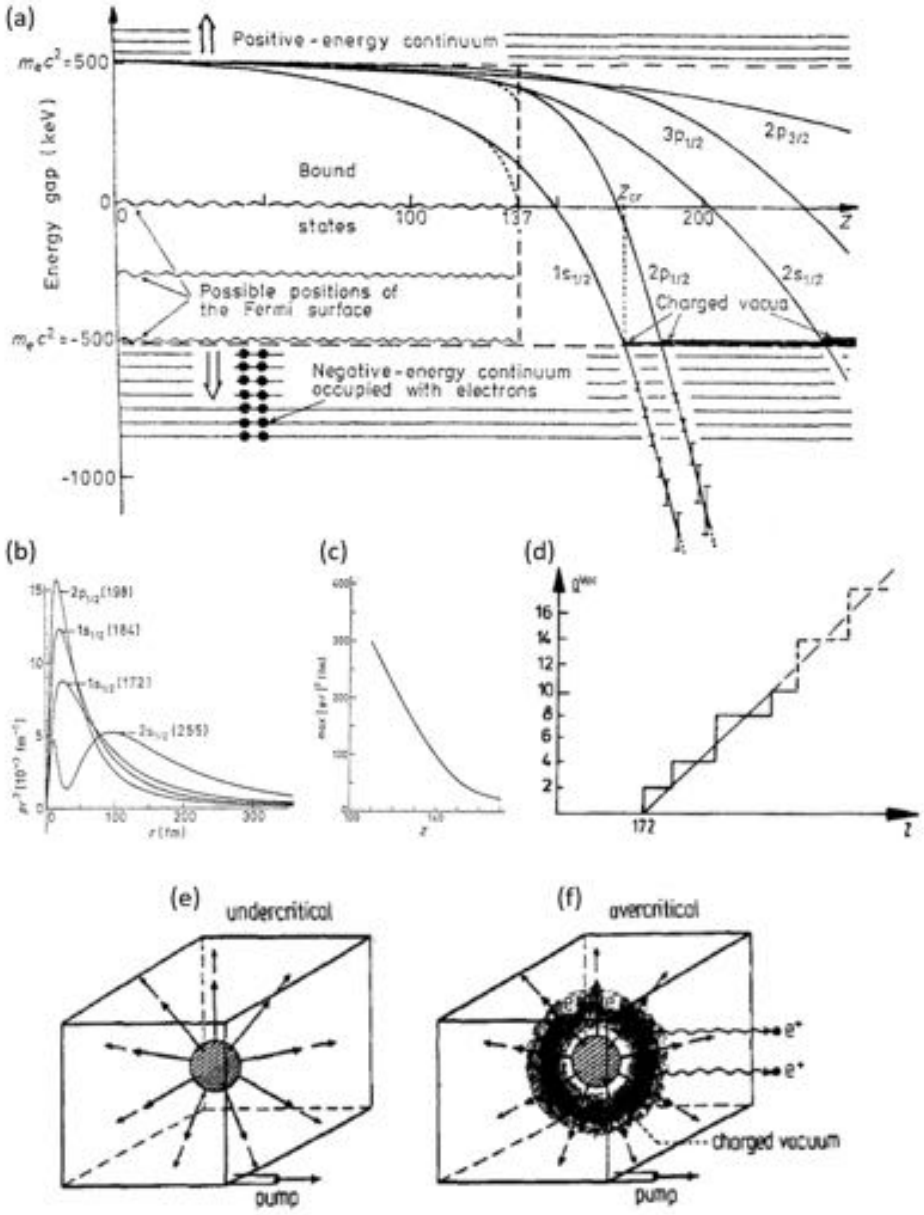


Figure 4.5: Atomic Collapse theory Graphs and illustrations. (a) the single Dirac Fermion energy as a function of the nucleus charge (Z) it is orbiting. As Z increases above the critical threshold marked by Z_{cr}, the energy states begin to dive into the negative energy continuum one by one. (b) the wavefunction of the electron orbit plotted for different Z. As Z goes up, the atomic size, denoted by the maximum of the wavefunction, tends towards the size of the nucleus $\sim$ 1.6fm. (c) plot of the maximum of the wavefunction as a function of Z, which illustrates the rapid shrinking of atomic size. (d) the vacuum around the nucleus gets charged linearly as Z goes above Z_{cr}. (e) and (f) are pictures showing the undercritical and overcritical cases respectively. In the overcritical case, the vacuum immediately outside the nucleus is charged and hence screens the fields from it. All pictures are adapted from the work by Greiner et. al.[78,79]

increases in units of $-2e$ and this is because each state is spin degenerate and can hold two charges. A spin degenerate empty state that dives down into the negative energy continuum will also charge the vacuum with $-2e$. The interpretation is as follows: as soon as an empty bound state mixes with the negative energy continuum, electrons from the occupied continuum levels will spontaneously fill the empty bound state and leave holes in the continuum. These holes are positively charged positrons that get repelled by the positively charged nucleus and escape into infinity. The originally charge neutral vacuum is now two negative charges more than neutral. With this, the vacuum charge only increases in units of $-2e$. The precise increase of vacuum charge depends on which state dives into the negative continuum. This is better shown in figure 4.4.1d. The formation of charged vacuum then serves to screen the nucleus and so looking from away the nucleus, we would only be able to detect a maximum nucleus charge Z_{cr}. This effectively means that in the theory of quantum electrodynamics the point charge cannot be larger than $Z_{cr} = 172$.[80]

The signature behaviors of atomic collapse can all be deduced from the atomic solution (equation 4.56) and these are summarized in figure 4.4.1a. From the hydrogen atom solution, in the case when $Z\alpha > j$, i.e. past the supercriticality threshold, the energy can be recast generically into two terms:

$$\varepsilon_H = -\varepsilon_0 + i\Gamma. \tag{4.60}$$

ε_0 is the position of the energy levels of the bound state and Γ is the imaginary component of the solution and refers to the lifetime of the collapsed state. When the bound state dives into the negative continuum, the state does not decay away completely but maintains a resonant behavior with the negative continuum states. This is represented by the spreading of the energy level across the continuum. This spreading can be detected physically by optical absorption spectroscopy.[79] The inverse of

this energy width gives the timescale for the unstable neutral vacuum to decay into the stable charged vacuum. This time scale is on the order of 10^{-19} s.[80]

The energy behavior of the collapsed state is:

$$\delta\varepsilon_0 \sim (Z - Z_{cr}). \tag{4.61}$$

This is known as the linear diving of states, i.e. states once collapsed will continuously dive down in a linear fashion towards negative infinity. The lifetime behavior is:

$$\Gamma \sim (Z - Z_{cr})^2, \tag{4.62}$$

which is known as the quadratic scaling of lifetime.[79]

In attempt to observe atomic collapse, scientists have gone to particle accelerators to bombard atoms to artificially engineer heavier atoms. To date, the heaviest atom ever created is Oganesson with $Z = 118$.[81] This is far from supercriticality and the possibility of ever observing it seemed grim. Gr, however, can possibly bring light to the seemingly futile prospect of observing atomic collapse.

4.5 Atomic Collapse-Like Resonances in Graphene

Graphene provides particles that behave analogously with relativistic particles. This fulfills the one main requirement for observing atomic collapse. In atomic collapse a limiting factor is the fine structure constant because it dictates the threshold of supercriticality. In vacuum $\alpha \sim 1/137$. In Gr, the effective fine structure constant is $\alpha_{Gr} \sim 1$. This means that the condition of $Z\alpha > 1$ to reach supercriticality becomes roughly 137 times easier to attain if atomic collapse were to happen in Gr. Theoret-

ically, it means that if Gr carriers are subject to a Coulomb type potential with a charge $Z = 1e$, the carriers would collapse into the center of this potential. The difficulty here is, how do we observe this phenomenon in a condensed matter system? The one signature that can be probed is the resonance in the negative energy continuum. This resonance translates to a peak in density of states in the hole band of Gr if electrons were collapsing. An experiment along this line of thought was performed by Wang et. al.[82] In their experiment, they placed Ca ions on top of Gr. These ions acted as impurities that they can move around with their scanning tunneling microscope (STM) tip. After moving three Ca together, they saw that the local density of states (LDOS) around the cluster of Ca developed a peak in the hole band. This provided the first evidence for atomic collapse-like resonances in Gr.

4.6 1D Confinement in Graphene and its Connection with Atomic Collapse

A similar experiment can be performed in our hybrid CNT-Gr device but the initial premise is completely different. In our hybrid devices, the CNT can serve as a 1D gate and the Gr as a high mobility 2DEG whereby this hybrid system as a whole can be used to study Gr as a quantum electron optics platform. The CNT can create a potential that confines electronic waves in Gr and possibly guide them throughout the 2D plane like a nano-photonics waveguide. This is to further the progress in the field of Gr plasmonics and will be the central topic in chapter 6. However, in the vein of Gr electrons being analogous to relativistic and massless Dirac Fermions, we can consider the system as a means to address the 1D atomic collapse problem. The governing equation of this hybrid system is:

$$\Big(\hbar v_F \boldsymbol{\sigma} \cdot \mathbf{p} + V_{\text{CNT}}\Big)\psi(x) = \varepsilon\psi(x). \tag{4.63}$$

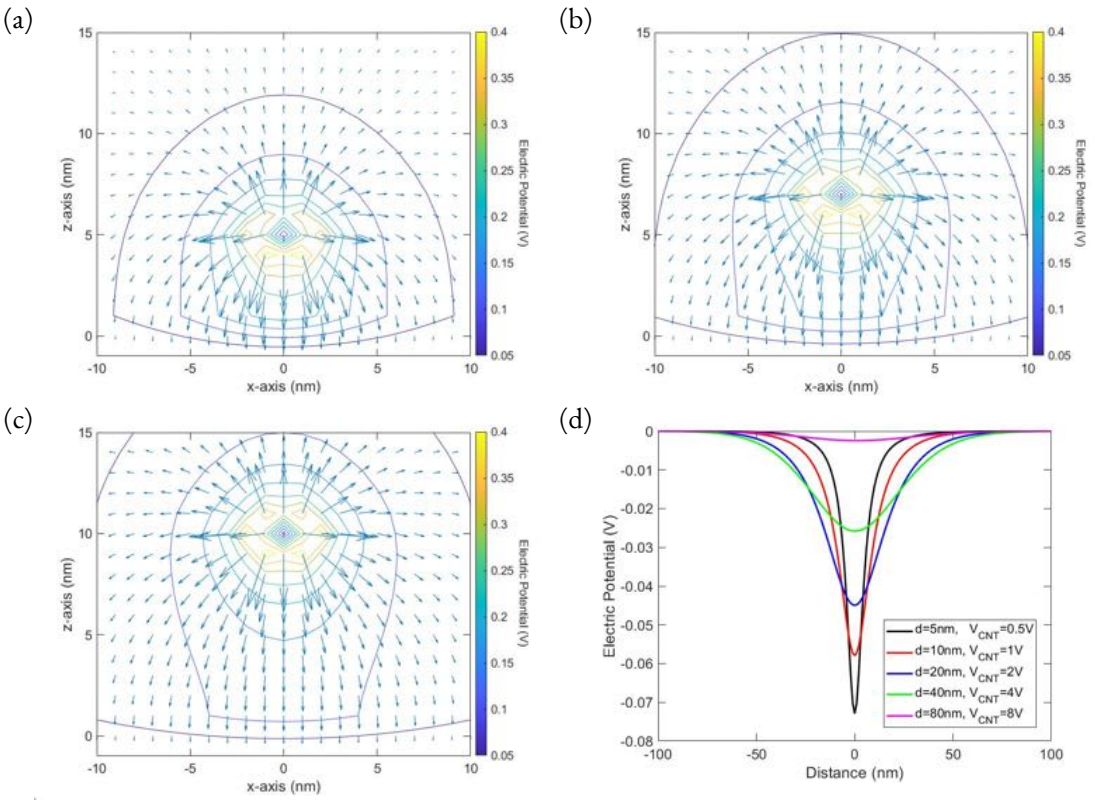


Figure 4.6: Electric potential simulation of 1D wire over 2D metal. (a) wire is at 1.5V and 5 nm above metal plane. (b) wire is at 1.75V and 7 nm above metal plane. (c) wire is at 2V and 10 nm above metal plane. Comparing the potential contour in (a) to (c) we see the potential at the 2D plane becomes wider as the wire is farther away from the 2D plane. (d) shows direct comparison of potential profiles for different physical configurations.

Here we can choose the electrostatic potential generated by the CNT to be the same as one given by the potential created by a charged wire:

$$V_{\text{CNT}} = \frac{Ne^2}{2\pi L \epsilon_0 \epsilon_r} \ln \frac{L}{\sqrt{x^2 + d^2}}, \tag{4.64}$$

where L is the length of the CNT, N is the number of charges in the CNT, d is the distance between the CNT and Gr and ε_r is the dielectric constant of hBN— the spacer between CNT and Gr. Note for the line potential we chose arbitrarily for $x = L$ to be the reference potential (i.e. $V(x = L) = 0$)

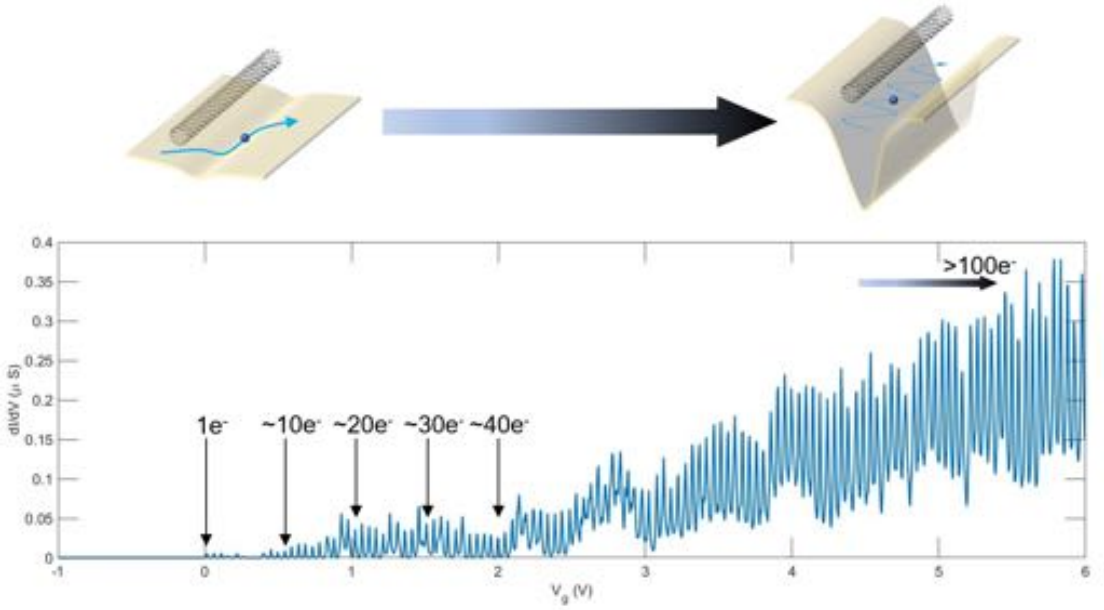


Figure 4.7: Potential depth due to carbon nanotube charging. The carbon nanotube can be charged with a known number of charges. The plot shows a conductance of a carbon nanotube as a function of gate voltage. Increasing the gate voltage will cause charges to pile up in the carbon nanotube one by one. Using this knob, we can controllably tune the potential imposed on graphene. The expectation is once the CNT becomes charged sufficiently, graphene carriers will become confined within the potential created by the CNT.

and this is a suitable reference point because L is typically $\sim$ 500 nm which is much greater than the region the CNT is sensitive to as well as the region of interest. We estimate this region of sensitivity to be on the order of the full width half maximum (FWHM) of the potential coming from the 1D wire. Our electrostatic simulations show that the potential due to a 1D wire over 2D plane geometry dies off quickly beyond the spacer distance (see figure 4.6). This spacer distance is given by the hBN thickness and in our devices it typically ranges from 4 to 10 nm, which is much smaller than L. This justifies the chosen form for the 1D potential.

The idea is to then use the CNT as a tunable charge string and as a local probe at the same time. By gradually increasing the number of charges in the CNT, the potential in Gr created by the CNT can become deep enough to trap electrons (as depicted in figure4.6). This may result in peaks in density

of states that the CNT can detect through the local compressibility measurement. The nature of this peaks can be interpreted as atomic-collapse like resonances and below we show how this may be the case.

4.6.1 Simulation of Graphene Density of States with 1D Potential

Figure 4.8 show the simulation of the Gr DOS due to the 1D potential given by equation 4.64. The calculation is done by numerically solving equation 4.63. In the calculation, the y direction is assumed to be translationally invariant and the problem is reduced into a 1D problem that extends along the x axis. With this in mind, the wavefunction takes the form:

$$\psi(x) = Ce^{iq_y y} \begin{bmatrix} u_A(x) \\ u_B(x) \end{bmatrix}, \tag{4.65}$$

where q_y is the momentum along the y direction, $u_{A(B)}$ is the weight of the wavefunction on sublattice $A(B)$, and C is the normalization constant. Plugging the wavefunction into 4.63 we get two coupled equations:

$$\begin{aligned} \partial_x u_A &= \left(-\frac{\varepsilon}{\hbar v_F} + V_{\text{CNT}} \right) u_B + q_y u_A, \\ \partial_x u_B &= \left(\frac{\varepsilon}{\hbar v_F} - V_{\text{CNT}} \right) u_A - q_y u_A. \end{aligned} \tag{4.66}$$

The x variable in the equations are first discretized into units of $d/10$. The coupled equations are solved numerically with a boundary condition of $u_A(x = 200 \cdot d) = u_B(x = 200 \cdot d) = 1$. The wavefunction is then normalized within the same boundaries. The DOS is given by $|\psi(x)|^2$. To reflect the fact that the CNT is only sensitive to the DOS just below it, we define an effective DOS that

is the convolution of $|\psi(x)|^2$ with a Gaussian function, similarly done by Peres et. al.[83] The 2D plots shown in figure 4.8 are acquired by solving for all ε and q_y where the two are discretized into units of $\hbar v_F \pi/L$ and π/L respectively.

The simulation shows signatures that are reminiscent of atomic collapse. First, as the charge in the CNT is increased, a spike-like feature begins to form on the surface of the Dirac cone. The spike corresponds to a peak in DOS. For the case when the CNT is charged positively (positive N), the spike forms at the top of the valence (hole) band. As the CNT charge increases further, the spike begins to dive down towards negative energies while becoming wider. At the same time new spikes begin to form above and dive as well. This is reminiscent of an electron collapsing into a positively charged nucleus, after which, the electronic states form a resonance in the hole band that gets wider and begins to dive down towawrds negative infinity with additional resonances to follow. In chapter 6, we will continue to draw connections between the signatures of atomic collapse and the behavior of guided modes due to 1D confinement. As a concluding remark, despite the similarities that one shall see, the hybrid CNT−Gr device only serves as a relativistic analogue that addresses the 1D atomc collapse problem. Atomic collapse is not actually happening and the 1D atomic collapse analogy we are drawing is highly nonphysical. It remains to be seen how closely related the behavior observed in Gr is to the actual atomic collapse.

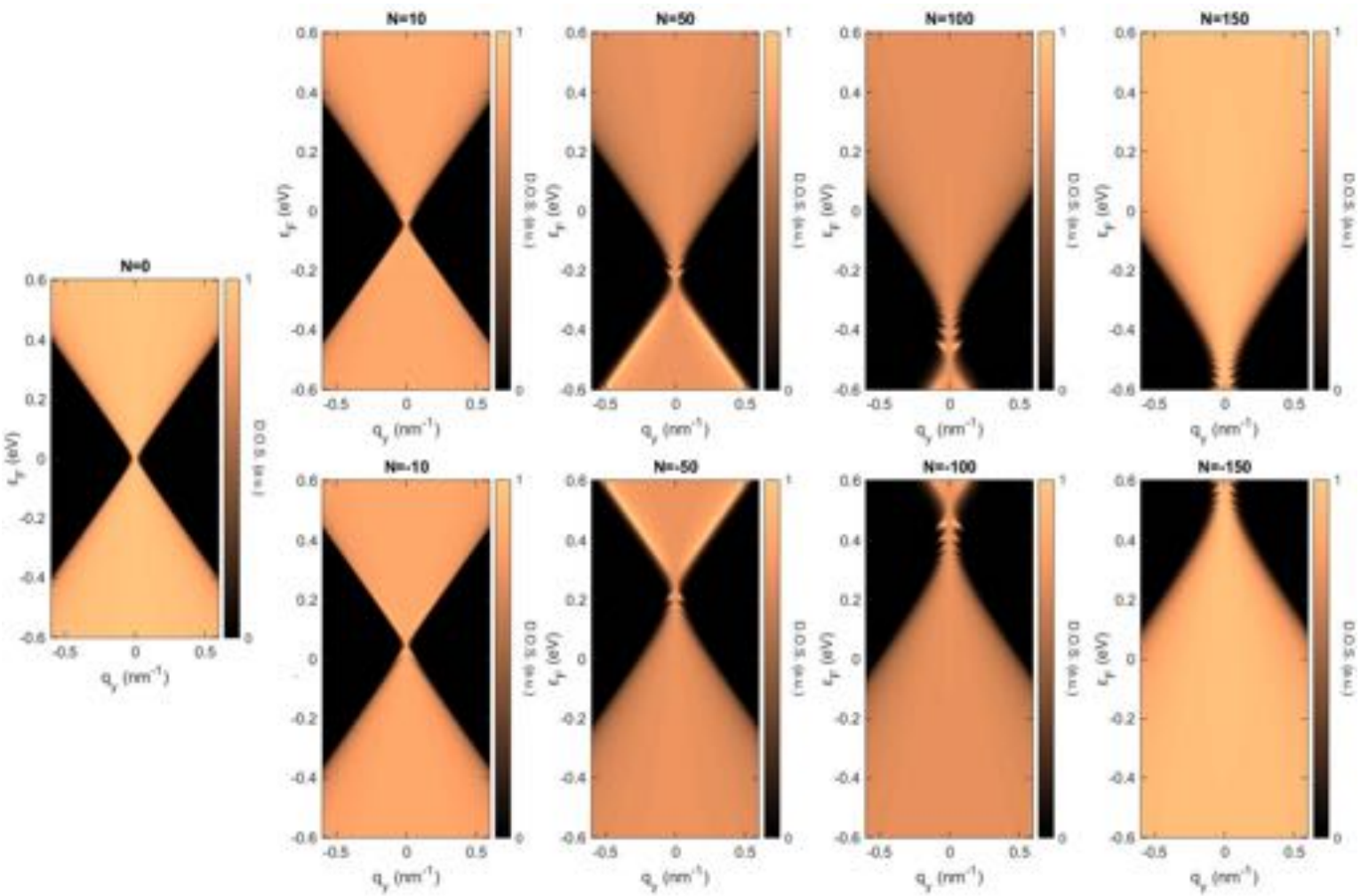


Figure 4.8: Simulation of graphene density of states under the influence of a 1D charge string. The density of states are calculated by solving equation 4.63. A more detailed discussion to be followed in Chapter 6

Главное, делайте всё с увлечением, это страшно украшает жизнь [It is important to do everything with passion, it embellishes life enormously].

Lev Landau

5

Landau Level Collapse

The experimental realization of graphene (Gr) and its recent improvements in electronic mobility performance have provided new opportunities to explore effects that were previously inaccessible. One such effect is Landau level collapse. The low energy properties of Gr are governed by the two dimensional (2D) massless Dirac equation. Under the influence of a sufficiently strong perpendicular magnetic field, electrons in a 2D plane will perform complete cyclotron orbits that are localized and have discrete energies called Landau levels (LLs). These localized orbits give rise to the signature transport behaviors of the quantum Hall effect such as the insulating bulk and chiral edge states.

Subject to a sufficiently strong in-plane electric field, however, the quantum hall effect can become

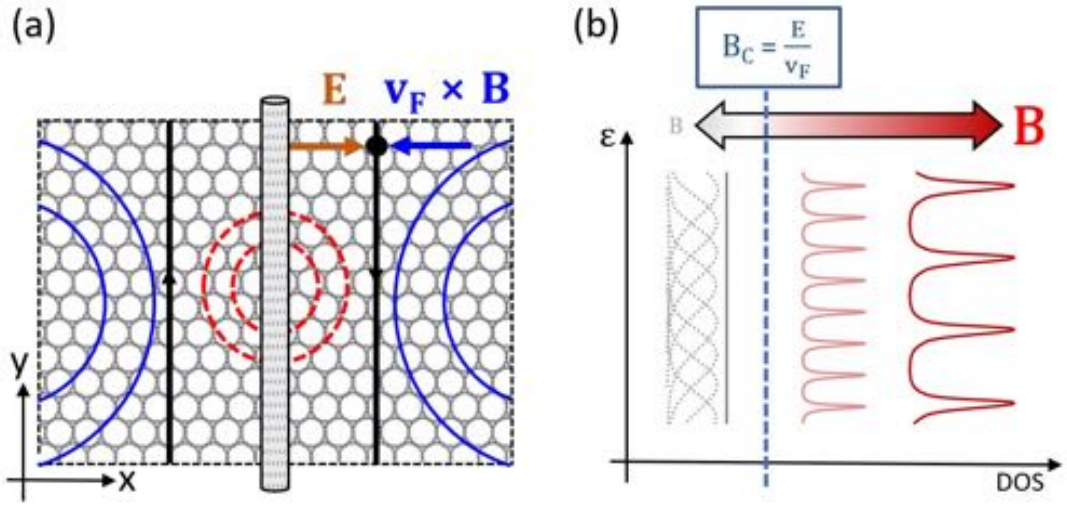


Figure 5.1: Semiclassical and quantum mechanical description of Landau level collapse. In (a), the red dotted and blue line orbits are the cyclotron orbits. At the condition when $E = v_F B$, the electrons following the red dotted orbits will become deconfined and make extended trajectories following the black line. (b) shows the energies of the Landau levels. When $B > E/v_F$, the Landau levels behave normally (red lines) but once $B < E/v_F$, the Landau levels will broaden and begin to merge and become an energy continuum (gray lines).

quenched. From the semiclassical perspective, because the deconfinement effects of an in-plane electric field is in direct competition with the confining magnetic field, once the electric field overpowers the magnetic field, the originally closed orbit will open up and the electron trajectory will become extended. Quantum mechanically, the electric field causes the broadening of LLs until the individual levels begin to completely merge and form an energy continuum. The originally quantized LLs will become indiscernible and this marks the onset of Landau level collapse.

Previous attempts involved the use of the two dimensional electron gas (2DEG) in Gallium Arsenide (GaAs) heterostructures as the testing platform. This 2DEG is typically buried about 100nm below the surface however. To create in-plane electric field effects, local gates can be deposited on top of the heterostructure but this meant the gates will have to be far from the 2DEG, therefore limiting the strength and sharpness of the induced potential barrier. This severely diminishes the ability

to create the necessary electric field strength to observe Landau level collapse. Gr on the other hand is a naturally exposed 2D electronic platform. With the availability of thin hexagonal-boron nitride (hBN) as a dielectric separating the local gate, one can create potential barriers that are much narrower and deeper than before.

In this work, we employ the carbon nanotube (CNT) as a local gate and this even further enhances the functionality of the Gr test device. The CNT is a 1D wire with a diameter on the order of $\sim$ 1 nm. Coupled with the use of a 4 nm thick hBN (the thinnest hBN used in this work), we can create a potential barrier that is unprecedentedly sharp and deep. Compared to typical electron beam patterned gates which at best is about 10 nm wide, the CNT is an order of magnitude narrower. More importantly, the edges of the CNT is guaranteed to be smooth as opposed to gates fabricated through metal deposition. The electric field emanating from gates with rough edges may not be homogeneous at the nanoscale and this can complicate the phenomenon.

The main advantage of using the CNT is it can simultaneously be used as a local probe to detect the formation and disappearance of LLs. This is an extremely useful functionality because with it we can probe directly at the site where the electric field is strongest. In contrast, in conventional transport measurements where current detection happens far from the locally gated region, the signature of LL collapse may become obscured due to the spatial remoteness of the phenomenon and the analysis may not be as precise.

5.1 Connection with Relativity

The electric field causes the broadening of the discrete Landau levels and the peak-like levels begin to overlap until the spectrum of DOS loses its fast oscillating structure. The rate at which this happens as the field approaches the critical field is alarmingly fast (as we shall shortly see). Effectively, it is as

though the magnetic or electric field has suddenly vanished and transformed into the other. This is reminiscent of the vanishing field effect due to the Lorentz invariance of the equation of motions of relativistic particles (see equation 4.25). It is worth it to note that in non-relativistic systems where Galilean transformations take place, the behavior is completely different. In the presence of perpendicular electric and magnetic fields, a Galilean transformation can be used to eliminate the electric field by performing a boost to a frame with velocity $v = E/B$ in the direction perpendicular to both E and B. It is not possible, however, to eliminate magnetic field by any Galilean transformation. The vanishing of magnetic field is specifically a relativistic effect.

In this experiment, we will show Landau level collapse and draw a connection between this phenomenon with special relativity. We will see that in the framework where we replace the role of the speed of light c with Gr's Fermi velocity v_F, there exists a critical field $B_c = E/v_F$. Above this critical field, electrons in Gr feel predominantly a magnetic field, exhibiting behaviors due to the formation of discrete Landau levels; below this critical field, the effect from magnetic field vanishes and the electrons become insensitive to the magnetic field.

Although Landau level collapse is not a consequence of Lorentz invariance, simply based on the fast collapse of the levels as well as the condition at which it occurs, the resemblance to the vanishing of magnetic field due to the Lorentz invariance of relativistic systems is uncanny and here a strikingly fit connection can be made. This adds to the novel set of existing relativistic simulations performed on Gr, showing that Gr can indeed be a relativistic test-bed.

5.2 Landau Levels with a Linear Electrostatic Potential

We first begin by examining theoretically the behavior of electrons in crossed electric and magnetic fields. We will assume that the electrostatic potential imposed on Gr is linear, hence producing a ho-

mogeneous in-plane electric field.

The Gr Hamiltonian in the presence of an external magnetic field is:

$$\mathbf{H} = v_F \boldsymbol{\sigma} \cdot \boldsymbol{\Pi}, \tag{5.1}$$

where $\boldsymbol{\Pi} \equiv \hbar\mathbf{q} + e\mathbf{A}$ and $\mathbf{A} = (A_x, A_y, A_z)$ for $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$. Choosing the magnetic field to be perpendicular to the Gr plane (so pointing in the z direction), the eigenenergies are Landau levels (LLs) and they are:

$$\varepsilon_{\mathrm{LL}}(N) = \mathrm{sgn}(N)\frac{\hbar v_F}{l_B}\sqrt{2N}, \tag{5.2}$$

where $l_B = \sqrt{\hbar/eB}$ is the magnetic length and N is the Landau level index (see appendix D for details on the Quantum Hall effect). Consider now the Gr to be under the influence of crossed magnetic and electric fields. The Hamiltonian is given by:

$$\mathbf{H} = v_F \boldsymbol{\sigma} \cdot \boldsymbol{\Pi} + eEx, \tag{5.3}$$

Here we choose the electric field, E, to be pointing in the x direction, perpendicular to the magnetic field. To solve this, because we know that the Hamiltonian can be written in a Lorentz covariant form with v_F playing the role of light, we can copy the treatment used in relativistic kinematics (see chapter 4 regarding background information on this topic). In Lorentz covariant form and in terms of Einstein's notation, Gr's time dependent Dirac equation is:

$$-iv_F\hbar\sigma^{\mu}\left(\partial_{\mu} + i\frac{e}{\hbar}A_{\mu}\right)\psi(x^{\mu\prime}) = 0, \tag{5.4}$$

where $\mu = \{0, 1, 2\} \Leftrightarrow \{t, x, y\}$, σ^μ are the Pauli matrices with $\sigma_0 = i\sigma_z$. $A^\mu = (\phi/v_F, A^1, A^2)$, where ϕ is the scalar potential defined by equation 4.11.

According to the consequence of $\mathbf{E} \cdot \mathbf{B} = 0$, which applies in our case of crossed fields, if $B > E/v_F$ then we can always boost the frame to one where $\mathbf{E} = 0$. To achieve this in our scenario, from the stationary frame (S) we boost in the y direction (perpendicular to $\mathbf{E}$) to a moving frame (S') with velocity $v = E/B$. Following the transformation rules of fields shown in equation 4.22, one will find $E' = 0$ and $B' = B\sqrt{1-\beta^2}$ where $\beta = v/v_F$. The Dirac equation in the moving frame coordinate system is then:

$$-iv_F\hbar\sigma^\mu\left(\partial'_\mu + i\frac{e}{\hbar}A'_\mu\right)\psi(x^{\mu\prime}) = 0. \tag{5.5}$$

In the Landau gauge, where we choose $A^1 = -Bx^2$ (in Einstein's notation, $x^2 = y$), it becomes:

$$-iv_F\hbar\left(\sigma^\mu\partial'_\mu + i\sigma^1\frac{e}{\hbar}B'x^{2\prime}\right)\psi(x^{\mu\prime}) = 0. \tag{5.6}$$

This is simply the time dependent Dirac equation under a perpendicular magnetic field and the eigenenergy in the moving frame is:

$$\varepsilon'_{\mathrm{LL}}(N) = \mathrm{sgn}(N)\frac{\hbar v_F}{l_B}\sqrt{2N}(1-\beta^2)^{\frac{1}{4}}, \tag{5.7}$$

which is simply the Landau level energies in the stationary frame but with a reduced magnetic field, $B' = B\sqrt{1-\beta^2}$. Note here l_B is the unchanged magnetic length. Working now in $2+1$D, in the 3-vector formalism with v_F replacing the role of c, energy is the time component of the 3-momentum which is: $p^\mu = (\varepsilon/v_F, p^1, p^2)$. Bear in mind that all the 3-vectors in the boosted frames are not

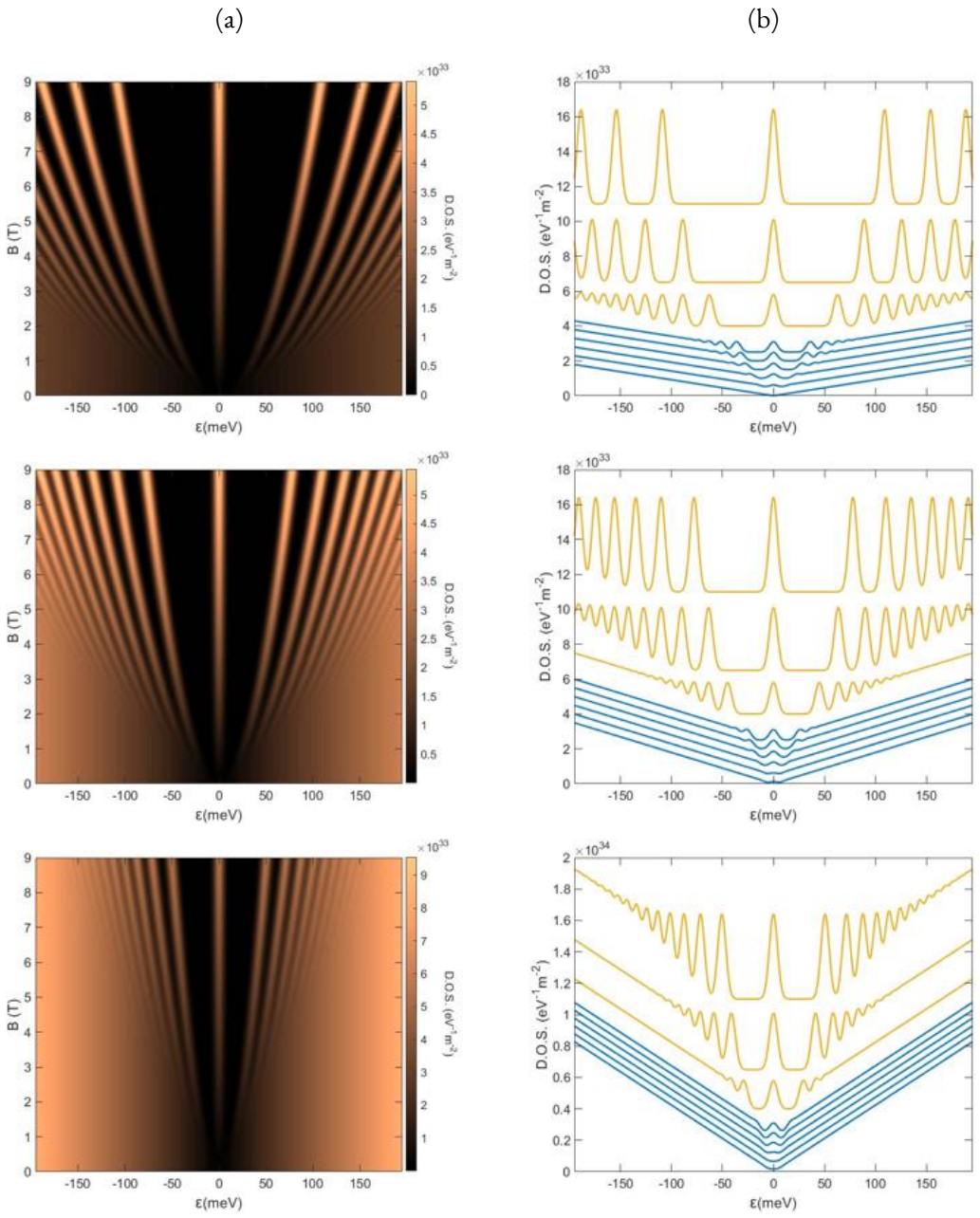


Figure 5.2: Landau level energies with different crossed electric field strengths. Column (a) shows the Landau level DOS as a function of B Field and energy. The LLs get squeezed more closely together as well as merge together at a higher magnetic field for larger electric field strength. Top to bottom rows correspond to $\beta = 0, 0.6, 0.8$ respectively. Column (b) shows the line cuts of column (a) at different magnetic fields. Blue lines are cuts at $0, 0.2, 0.4 \ldots 1$ T, and orange lines are cuts at $3, 6, 9$ T.

physical objects and this is due to v_F having replaced c. Consequently, equation 5.7 are not the real LL energies. The real physical energies are obtained by transforming the boosted frame results back to the stationary frame.[84] Hence, performing the Lorentz transformation on the psuedo-3-momentum, we acquire:

$$\varepsilon_{\mathrm{LL}}(N) = \mathrm{sgn}(N)\frac{\hbar v_F}{l_B}\sqrt{2N}(1-\beta^2)^{\frac{3}{4}} - \hbar v_F \beta q_y, \tag{5.8}$$

where $(p^1, p^2) = \hbar(q^1, q^2) = \hbar(q_x, q_y)$, $\mathbf{q}$ being the Gr wavevector. Equation 5.8 shows the real LL energies in the presence of crossed electric and magnetic fields. When the electric field vanishes, we recover the standard LL energies. An important contrast to point out is, the LL energies of a non-relativistic, parabolic band system in crossed magnetic and electric fields is given by:

$$\varepsilon_{\text{Non-Rel,LL}} = (N+1/2)\hbar\omega_c - \hbar k_y \frac{E}{B} - \frac{m}{2}\left(\frac{E}{B}\right)^2. \tag{5.9}$$

The main difference between the two is that the Landau level spacing for Gr scales as $(1-\beta^2)^{(3/4)}$, whereas, the Landau level spacing for parabolic band 2DEGS is independent of the electric field (see figure 1.3b and simulations will show this fan diagram to be unchanging with respect to an external homogeneous electric field). This difference is one of the key signatures that separates relativistic vs non-relativistic systems. In Gr, as dictated by equation 5.8, an external electric field causes Landau levels to mix and eventually merge into a continuum (see figure 5.2). With a large enough electric field, the Landau levels will become indiscernible and this is referred to as Landau level collapse. Even though the cause of Landau level collapse is strictly speaking not due to the vanishing of magnetic field as a result of Lorentz invariance, the overall coarse-grained behavior is strikingly similar.

It is important to keep in mind that the results only hold for homogeneous fields. The quantitative results derived thus far become invalid as soon as either field becomes inhomogeneous but qualitatively

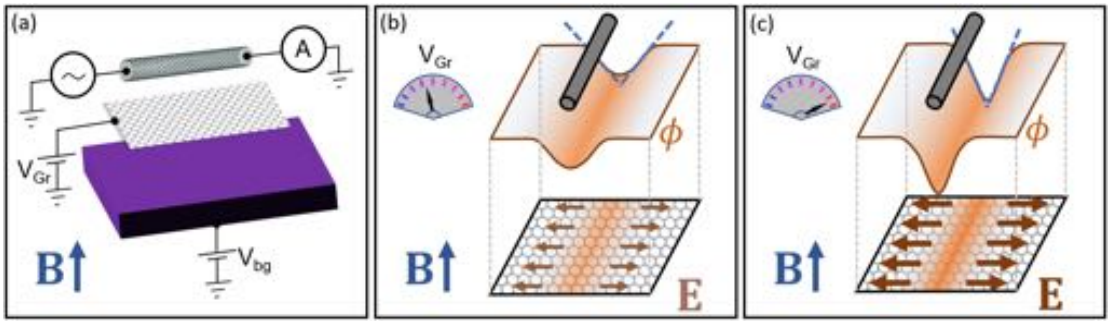


Figure 5.3: Landau level collapse measurement scheme. (a) shows the measurement scheme which is the same as the local compressibility measurement covered in chapter 2. The role of V_{Gr} serves to charge the CNT. Charging the CNT in turn modulates the electrostatic landscape in Gr. (b) and (c) are illustrations showing the effect of V_{Gr}. (b) is at a lower V_{Gr} compared to (c) and hence creates a gentler electrostatic modulation (ϕ) in Gr. Consequently, the electric field is smaller in (b). The blue dotted line shows the linear part of the potential that is used to calculate the homogeneous in-plane electric field.

we expect the behavior to hold.

5.3 Device Configuration, Measurement Scheme and Approximations

Figure 5.3 shows a cartoon of the device and the measurement scheme for detecting LL collapse. The device used is an electrically contacted CNT placed on top of a hBN encapsulated Gr. The hBN that separates the CNT and Gr used ranges from 4 to 20 nm. We see the strongest effect in our 4 nm hBN device and results presented in this chapter are based on this device unless noted otherwise. Additional data from other devices are included in the appendix G. The fabrication of the device and quantum capacitance extraction method is covered in detail in chapter 2 and 3. In this experiment, the Gr potential (V_{Gr}) with respect to the grounded CNT is the knob that tunes the in-plane electric field. What happens is V_{Gr} serves to charge the CNT which in turn modulates the electronic landscape in the Gr plane. In actuality, the electric field emanating from the CNT is not uniform in the in-plane direction. This is reflected by the non-linear potential created in Gr. This is shown through theory (see equation 5.19 and figure 5.12b) as well as simulation (see figure 4.6). However, the potential in the Gr can be ap-

proximated to be linear with a high degree of accuracy. Comparison between the linearized potential and the theoretical non-linear potential show to only differ by less than 1mV from 0 to $\gtrsim 10$ nm away from the center of the potential. This is a negligible variation given our experimental parameters. And because we are effectively probing a very local part of the Gr (up to ~ 4 nm away from the CNT in the Gr plane for the 4 nm hbN device), the linearized potential is a completely justifiable approximation. Nonetheless, the linearization is only to simplify the collapse analysis. With the assumption of a linear potential, the in-plane electric field can be treated to be homogeneous in the area of interest. The occurrence of the phenomenon does not hinge on the linearity of the potential. Relaxing the linearity of the potential may introduce small corrections to the quantitative results but the overall qualitative behavior should remain similar.

5.4 Electric Field Screening by Landau Levels

A measurement that can showcase Landau level collapse is to use the CNT to probe when Landau levels form and disappear. This is done similarly to the local compressibility measurement: while sweeping the magnetic field B and global backgate V_{bg}, monitor the conductance of the CNT (dI/dV). Doing so, at some magnetic field the Gr will be driven into the quantum Hall regime. In the quantum Hall regime, V_{bg} will serve to dope the Gr and cause charges to occupy different LLs. This parameter will allow us to see the screening effect for different LLs. The conductance of the CNT will be a reflection of how heavily screened the electric field from the backgate is. Following the circuit schematic shown in figure 5.3a, while keeping Gr grounded initially ($V_{\text{Gr}} = 0$), we obtain the result shown in figure 5.4a. The lines formed by the conductance peaks are the Coulomb blockade peaks of the CNT. To conceptually understand the behavior of these zig-zag lines, we follow the local compressibility analysis method described in section 2.8: lines that go relatively more horizontally mean that the CNT is more

insensitive to the backgate in those conditions, while lines that go relatively more vertically indicate a greater dependence on the backgate. In the quantum Hall regime, the zig-zag behavior comes from the chemical potential going in and out of the highly degenerate LLs. When the chemical potential is inside a LL, the Gr is populated with a high concentration of mobile carriers which will screen very effectively the electric field from the backgate ($\mathbf{E}_{\text{bg}}$), causing the CNT to be insensitive to the backgate changes. When the chemical potential is outside a LL, the Gr bulk is insulating which leaves $\mathbf{E}_{\text{bg}}$ unscreened. The zig-zag lines are hence a good indicator of when Landau levels are present. Looking more closely at these lines, one will see that below a certain point, the lines will lose its oscillation and just form a straight line down to zero magnetic field. The point at which this transition occurs is the critical magnetic field B_c denoted by the dotted black line. In the following section, we will discuss the nature of this critical field.

5.4.1 Electric Field Screening at Different Graphene Doping Levels

The application of an external electric field across a system with mobile carriers will be screened. The strength of the screening largely depends on the concentration of mobile carriers which is a function of the material's DOS. For a perfect metal, due to the high DOS, external electric fields will get screened nearly perfectly. For quantum materials which have a reduced and variable DOS, electric fields will become relatively less screened and will also see a variation in how it gets screened depending on the material's doping level. In chapter 2 this topic was covered in detail where we showed that as Gr is doped with a global backgate voltage, the CNT sees a varying electric field from this backgate. This response can be quantified through the inclusion of a quantum capacitance added in series to the geometric capacitance of the system (see equation 2.9).

In this experiment, the CNT is responsible for the generation of electric field in the system. The

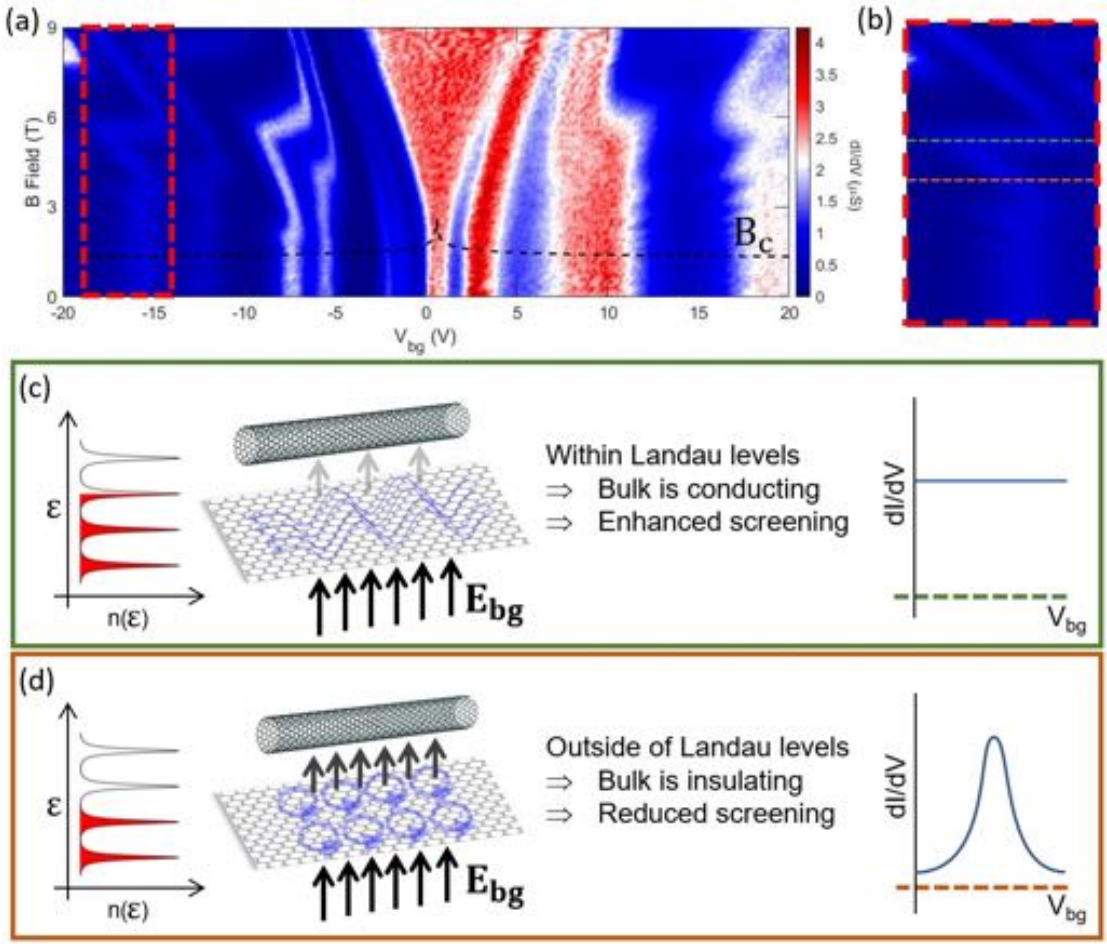


Figure 5.4: Illustrations to explain the effect of screening by Landau levels. (a) shows the CNT conductance as a function of backgate (V_{bg}) and magnetic field (B). The lines formed by the conductance peaks correspond to Coulomb blockade peaks. The modulation of these lines is due to the screening of the Gr. To see this, we look at a specific section demarcated by red dotted lines and this is shown in (b). The black dotted line is the critical field line plotted using equation 5.12. In (b) is one green dotted line and one orange dotted line. The two indicate screening when Gr LLs are partially filled and completely filled respectively. (c) illustrates when LLs are partially filled, in which case, the Gr bulk hosts conducting carriers that screen the electric field from the backgate, thus causing the CNT conductance to be insensitive to V_{bg}. (d) illustrates when LLs are completely filled, in which case, the Gr bulk is insulating and is unable to screen the electric field from the backgate because there are limited mobile carriers. This causes the CNT to efficiently feel the electric field due to V_{bg}.

resulting in-plane electric field will become screened according to the doping level of Gr. We can account for this through the model of Thomas-Fermi screening. Using equation 2.4 for Thomas-Fermi screening wavevector and plugging in Gr's properties we have:

$$q_{\mathrm{TF}} = 4\alpha_{\mathrm{Gr}} q_F = 4\alpha_{\mathrm{Gr}}\sqrt{\pi n}, \tag{5.10}$$

where α_{Gr} is Gr's effective fine structure constant given by equation 1.11, q_F is Gr's wavevector with respect to the Dirac cone and n is the carrier density of Gr. This quantity is the inverse of the screening length of an electric field due to the dielectric response of Gr. In the linear response regime, this screening length can be incorporated into the overall electrostatics of the system by having a variable effective distance in its geometric capacitance. This effective distance in the linear response regime is given by:[22,85]

$$d_{\mathrm{eff}} = d + \frac{1}{2} q_{\mathrm{TF}}^{-1}, \tag{5.11}$$

The dielectric response is accounted for through the Thomas-Fermi screening wavevector. Using equation 2.24 as the geometric capacitance between the CNT and Gr, and using Gauss's law to calculate the approximate electric field, the critical magnetic field becomes:

$$B_c = \frac{E}{v_F} = \frac{n_{\mathrm{CNT}} e}{C' d v_F} = n_{\mathrm{CNT}} e \frac{\cosh^{-1}(d_{\mathrm{eff}}/r)}{2\pi l \epsilon d v_F}, \tag{5.12}$$

where the length of the CNT is $l \approx 500$ nm, the separation distance between the CNT and Gr is $d \approx 6$ nm, the radius of the CNT is $r \approx 1$ nm, $\alpha_{\mathrm{Gr}} \approx 0.5$ and $\epsilon = \epsilon_0 \epsilon_r$ with $\epsilon_r \approx 4$ for our hBN encapsulated Gr device. n_{CNT} is the charge number in the CNT, which can be controlled by the voltage applied to the CNT. This serves as a fitting parameter in our analysis. The value used here is $n_{\mathrm{CNT}} = 2$.

$C' = C'(d_{\text{eff}})$ is the total effective capacitance that includes contributions from both geometric and quantum capacitance. This is done by simply replacing the distance dependence (d) with the effective one (d_{eff}) in the equation for geometric capacitance. It is important to note that only the capacitance distance dependence gets modified. The d that relates the electric field with the capacitance ($E = n_{\text{CNT}}e/C'd$) is the actual physical distance that does not change. All the necessary adjustments to account for the dielectric response are taken within the capacitance term. The resulting electric field due to this effective capacitance ($C'(d_{\text{eff}})$) is the screened electric field. With the parameters above, B_c is the dotted line plotted in figure 5.4a. We see that B_c peaks near the charge neutrality point of Gr and this is expected because there are limited number of mobile carriers and so screening will be the weakest there. This causes the required magnetic field to overpower the electric field to increase and hence the peak.

Because figure 5.4a is performed at $V_{\text{Gr}} = 0$, theoretically we should expect $B_c = 0$ as shown in figure 5.5b. This suggests that there are intrinsic residual charges in the CNT that is already generating an in-plane electric field. Figure 5.5b to 5.5d show simulations of the fan diagrams with different in-plane electric field strengths. We can see that as the electric field strength increases, the LL levels lose their fast oscillating structures at larger magnetic fields. This result is well captured by the existence of B_c. Figure 5.6a through 5.6c are the experimental measurements of the fan diagrams at varying V_{Gr}. We can see that the behavior similar to the simulations are well reproduced. Figure 5.6d summarizes the evolution of the critical magnetic field near the charge neutrality point (CNP) of Gr. In the figure, we see that as the effective V_{Gr} increases, B_c also increases as expected.

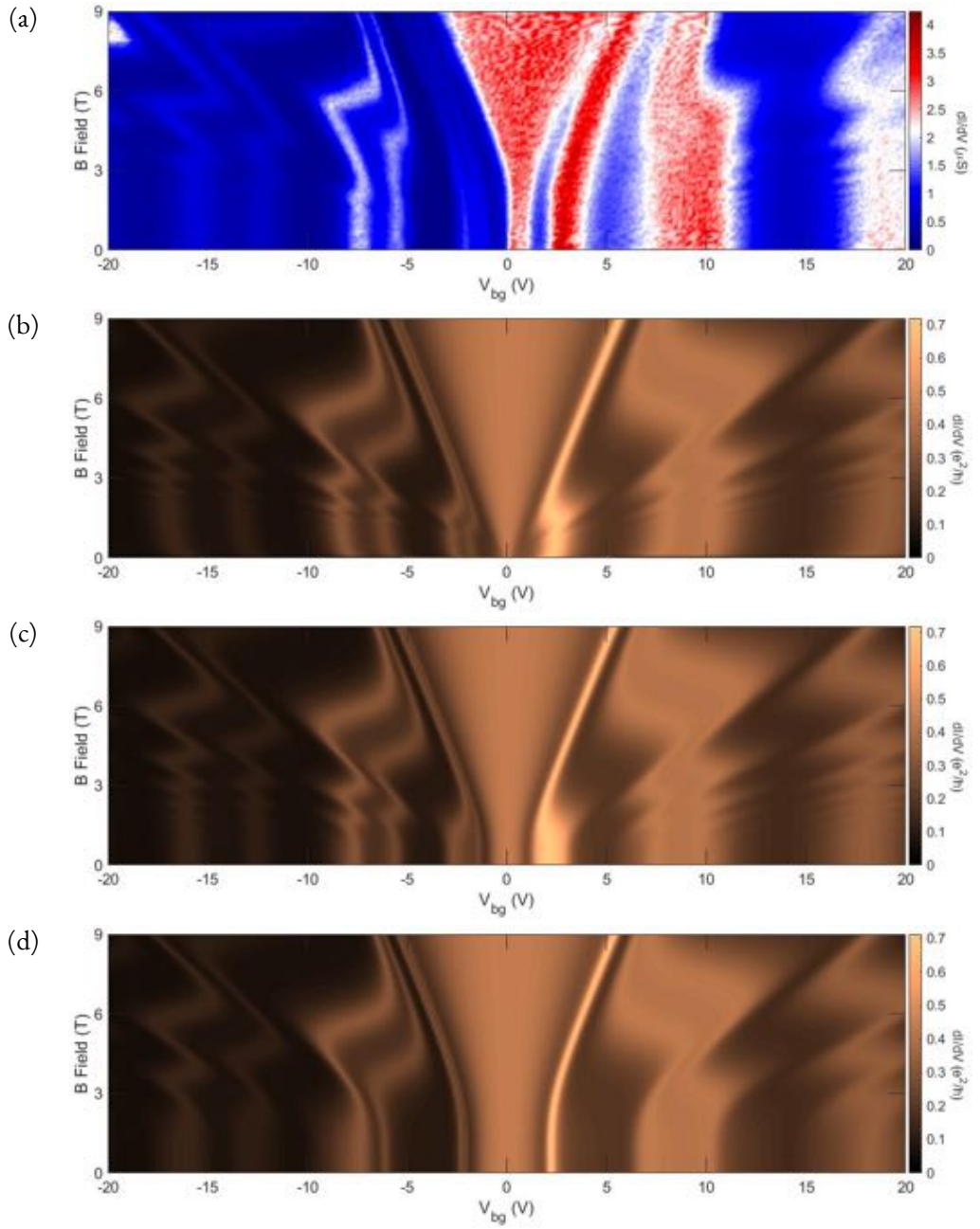


Figure 5.5: Comparing the simulated fan diagrams at different electric field strength with the experimental result. (a) is the experimental result performed at $V_{\text{Gr}} = 0$ V. (b) to (d) show the simulated fan diagrams for different electric field strength going from none to large strengths. Plotted are values of CNT conductance as a function of backgate (V_{bg}) and magnetic field (B). (b) with $E = 0$ V/m the Y shape structure is absent. (c) with $E = 1.25 \times 10^6$ V/m the Y shape structure appears signifying a delayed onset of Landau level formation. (d) with $E = 3 \times 10^6$ V/m the Y shape is prominent and LLs are much smoother overall.

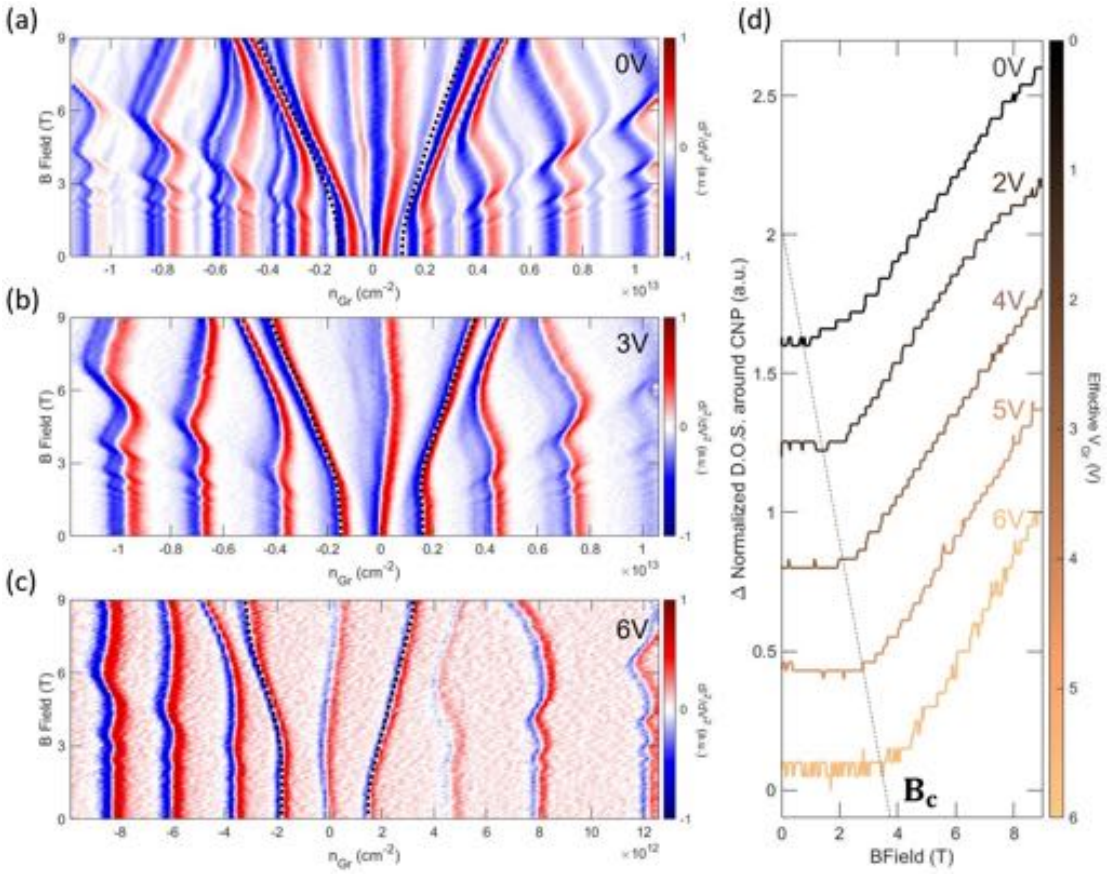


Figure 5.6: Experimental comparisons of the fan diagrams at different V_{Gr} with a 10 nm hBN device. (a) is at $V_{\mathrm{Gr}} = 0$ V, (b) is at $V_{\mathrm{Gr}} = 3$ V and (c) is at $V_{\mathrm{Gr}} = 6$ V. Note that the V_{Gr} here constitutes an effective V_{Gr} because the values taken are with respect to $V_{Gr} = 2$ V which we saw yielded the smallest critical field values. (d) shows the normalized DOS about the CNP for different V_{Gr} (see main text for DOS extraction method). At the top the effective $V_{\mathrm{Gr}} = 0$ V as measured in reference to $V_{\mathrm{Gr}} = 2$ V that yielded the lowest B_c. As the effective V_{Gr} increases, so does the B_c marked by the inflection point after the plateau. The plateau signifies the insensitivity of Gr to an increasing B field, suggesting a non-existent B field. The dotted black line is hand drawn to mark the approximate position of B_c.

5.4.2 Simulation of Carbon Nanotube Conductance

Figure 5.5 shows simulations of the conductance of a CNT on top of Gr as a function of magnetic field (B) and backgate voltage (V_{bg}). The general form of the CNT conductance used is given by the Breit-Wigner formula for CNT in the strong coupling regime (see appendix C):

$$G = \sum_i \frac{4e^2}{h} \frac{\Gamma_L^i \Gamma_R^i}{\left(\Gamma_L^i + \Gamma_R^i\right)^2 + \left(\mu_{\text{NT}} - \varepsilon_i\right)^2}, \tag{5.13}$$

where i represents the counter for the nanotube levels seen in the measurement, ε_i is the nanotube energy level, $\Gamma_{L(R)}^i$ is the energy coupling between the ith nanotube level with the left (right) electrode, and μ_{NT} is the electrochemical potential of the nanotube. To obtain the simulation, we use the concept drawn from equation 2.39: each nanotube level tracks the Fermi energy of Gr, so $\mu_{\text{NT}} \to \mu_{\text{Gr}}$. We treat Γ^i and ε_i as fitting parameters to adjust the profile of the nanotube conductance to match the ones seen from experiment.

To acquire the Fermi energy of Gr under the influence of a magnetic and electric field, we do not use the analytical results from section 5.2 because the result is particular to a single electron and does not take into account screening. Instead, we perform a more phenomenological simulation. We follow the ordinary Gr Landau levels theory. We assume that the DOS of Gr ($g_{\text{Gr}}(\varepsilon)$) can be decomposed into Landau levels where each level is given by a Lorentzian:

$$g_{\text{Gr}}(\varepsilon) = \sum_N \frac{g}{\pi} \frac{\gamma_N}{\gamma_N^2 + \left(\varepsilon - \varepsilon_{\text{LL}}(N)\right)^2} \tag{5.14}$$

where the degeneracy of each Landau level is:

$$g = \frac{4eB}{\hbar 2\pi}, \tag{5.15}$$

the Landau level energy is:

$$\varepsilon_{\mathrm{LL}} = v_F\sqrt{2e\hbar BN}, \tag{5.16}$$

and γ_N is the disorder broadening of the Nth Landau level. The DOS can be used to calculate the carrier density through:

$$n_{\mathrm{Gr}} = \int_0^{\mu_{\mathrm{Gr}}} d\varepsilon g_{\mathrm{Gr}}(\varepsilon). \tag{5.17}$$

The carrier density maps directly to a certain backgate voltage (V_{bg}). Through these relations, the Fermi energy of Gr can be calculated as a function of B and V_{bg}.

To account for the effects of the in-plane eletric field, we replace the ordinary magnetic field (B) with an effective field $B_{\mathrm{eff}} = \alpha(B - B_c)$. $\alpha = \sqrt{1 - (B_c/B)^2}$ and $B_c = E/v_F$ is the screened critical field given by equation 5.12. These phenomenological terms together with the Fermi energy of Gr allows us to simulate the conductance of the nanotube as shown in figure 5.5.

5.4.3 Density of States Extraction

The density of states (DOS) shown in figure 5.6d is acquired by first tracking the evolving CNT conductance peaks. Considering only the DOS about the CNP, we track the two lines that define the Y shape (denoted by the dotted black line). Because each conductance peak line tracks a constant carrier concentration level in Gr (see equation 2.39), the difference between the two lines along V_{bg} tells us the carrier density increase in Gr (Δn_{Gr}) needed to increase the electrochemical potential of the

CNT by one level (for CNT to gain one extra charge). This Δn_{Gr} has an associated energy to it and this can be extracted by looking at the CNT charging energy (ε_C) of the Coulomb diamonds associated with the conductance peak lines in the parameter space V_{sd} vs V_{bg}. The DOS is then given by $\Delta n_{\text{Gr}}/\varepsilon_C$. Figure 5.7 shows examples of these Coulomb diamonds at different magnetic fields. We can see that ε_C is invariant with respect to magnetic fields. Many Coulomb diamonds were collected over the course of the measurements and they all show that ε_C is unchanged with respect to different magnetic fields. Figure G.1 shows many more of these Coulomb diamonds. It is worthwhile to point out that the Coulomb diamonds are virtually identical at 0 T and 1 T. At 3 T the diamonds are only slightly modified. This is another clue showing the absence of a magnetic field effect below B_c.

5.5 Local Compressibility Measurements at Different Magnetic Fields

The previous type of measurements where we swept V_{bg} and B while fixing V_{Gr} allowed us to see a continuity in the magnetic field coming in and out of existence by looking at when LLs formed or disappeared. However, extracting a concrete B_c was difficult because the boundary that divides when LLs disappear is blurry. The local compressibility measurement performed at fixed magnetic fields will give us better clarity in seeing the onset of Landau level collapse. The local compressibility measurement scheme (tracking CNT conductance while sweeping V_{Gr} and V_{bg}) as well as the quantum capacitance extraction method is covered in detail in chapter 2. Figure 5.8 to 5.9 show the results at different magnetic fields. Focusing on the extracted quantum capacitance (C_Q) plot at 9 T (horizontal lines in the compressibility plots yield peaks in quantum capacitance), we see that the C_Q (or DOS) of LLs decrease generally (from blue to red) as V_{Gr} increases in magnitude. Higher LLs decrease at higher values of V_{Gr} and they all eventually merge together into a roughly homogeneous region (indicated by red). This behavior can be generalized to the LLs at other magnetic fields. The only differences are

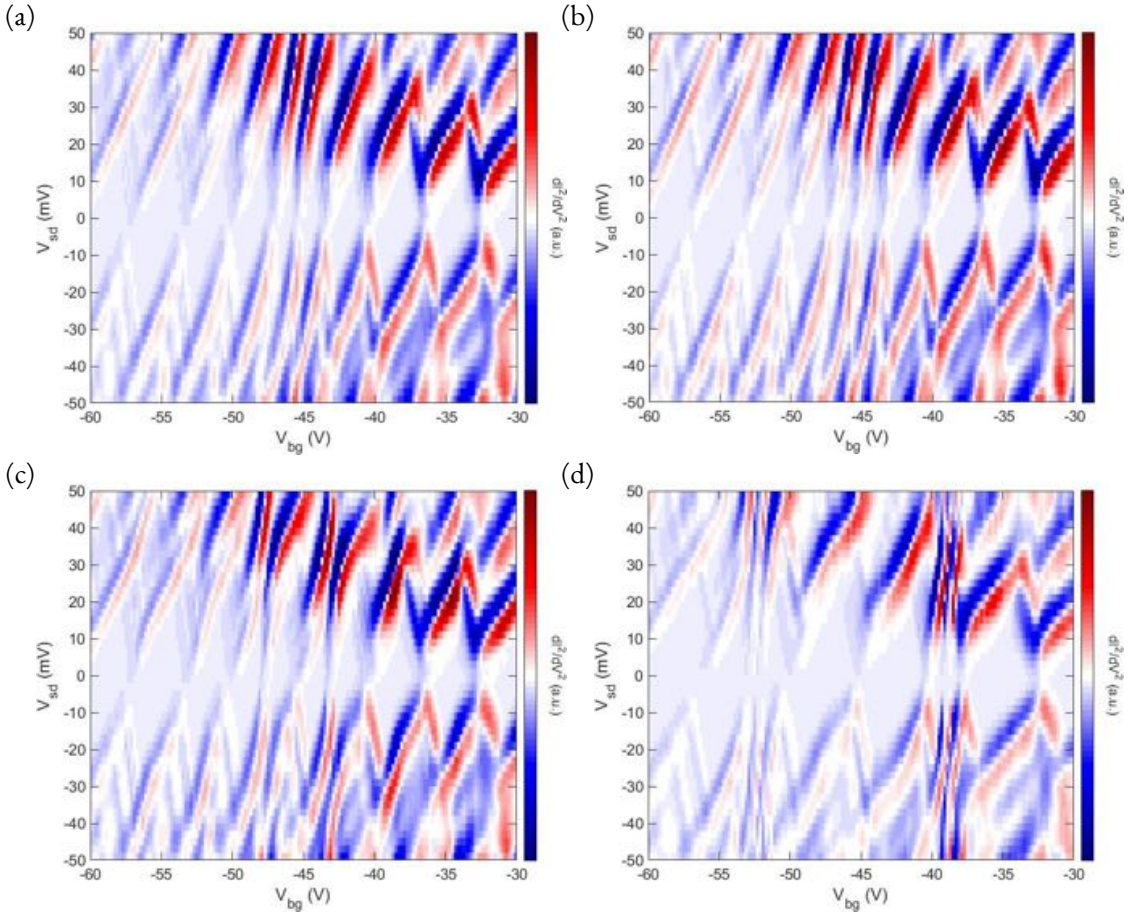
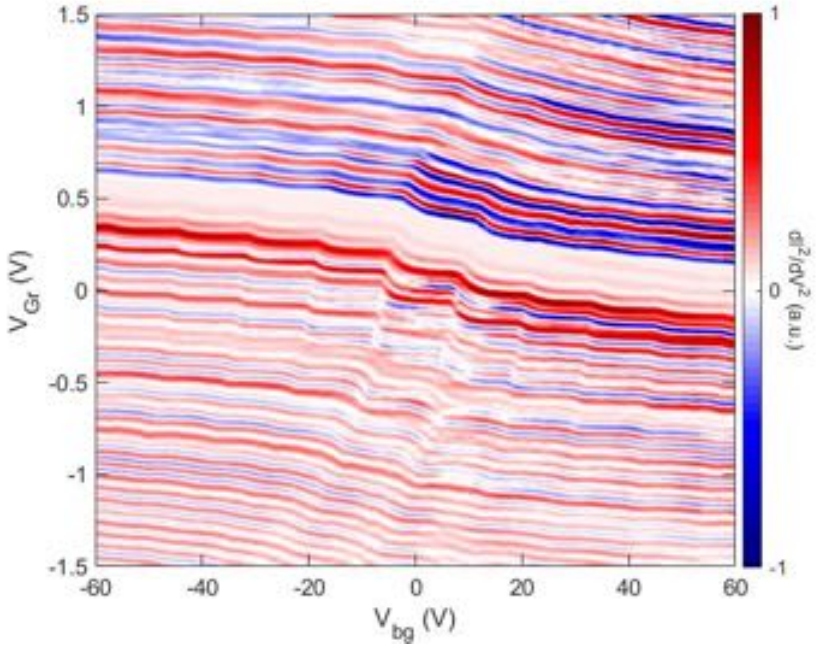


Figure 5.7: Magnetic field dependence of the carbon nanotube quantum dot on graphene. Plotted are the derivatives of differential conductances with respect to the global Si backgate (V_{Gr}) vs the CNT source-drain bias (V_{sd}). (a) is taken at 0 T, (b) at 1 T, (c) at 3 T, (d) at 9 T. For more magnetic field dependences, please refer to appendix G.

that the LLs are denser together as magnetic field is lowered and the V_{Gr} required to observe the substantial diminishing of LLs' C_Q is also lower. At 3 T, the LLs are noticeably not as well resolved and at 1T they are completely washed away (see figure 6.11 for 1 T).

Figure 5.10 shows line cuts of the C_Q spaced across the full V_{Gr} range and at different magnetic fields. We can see from this figure clearly that the LLs collapses as V_{Gr} increases in magnitude. More peculiarly though is that as seen throughout the data, LLs seems to preferentially happen at top right

(a)



(b)

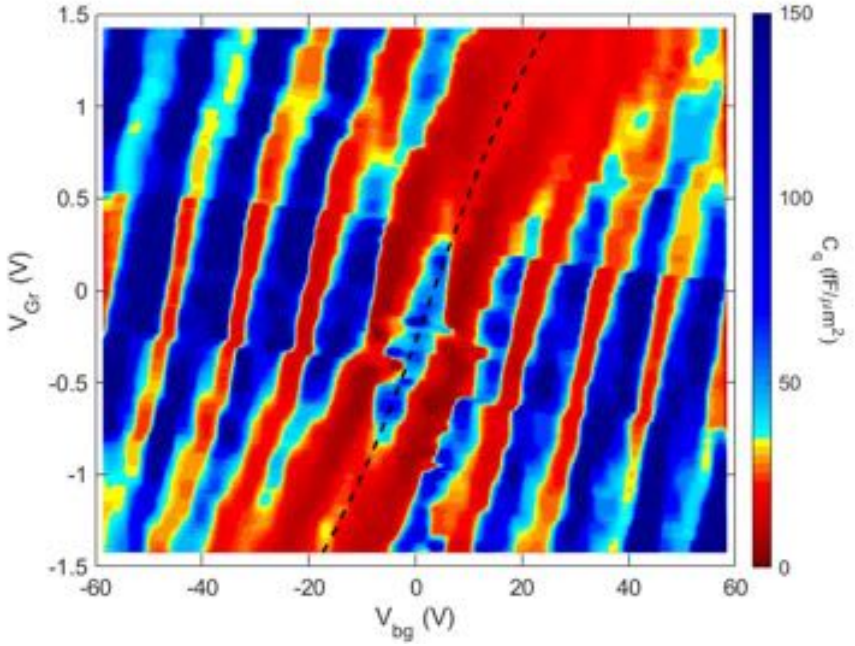


Figure 5.8: Local compressibility measurement and quantum capacitance extraction at 9 T. (a) is the local compressibility measurement and (b) is the quantum capacitance extraction both as functions of the backgate (V_{bg}) and Gr potential (V_{Gr}). The dotted line denotes the $N = 0$ Landau level. We see the Landau levels collapse as V_{Gr} increases. Quantum capacitance data is cut at 150 fF/μm^2 for better feature clarity.

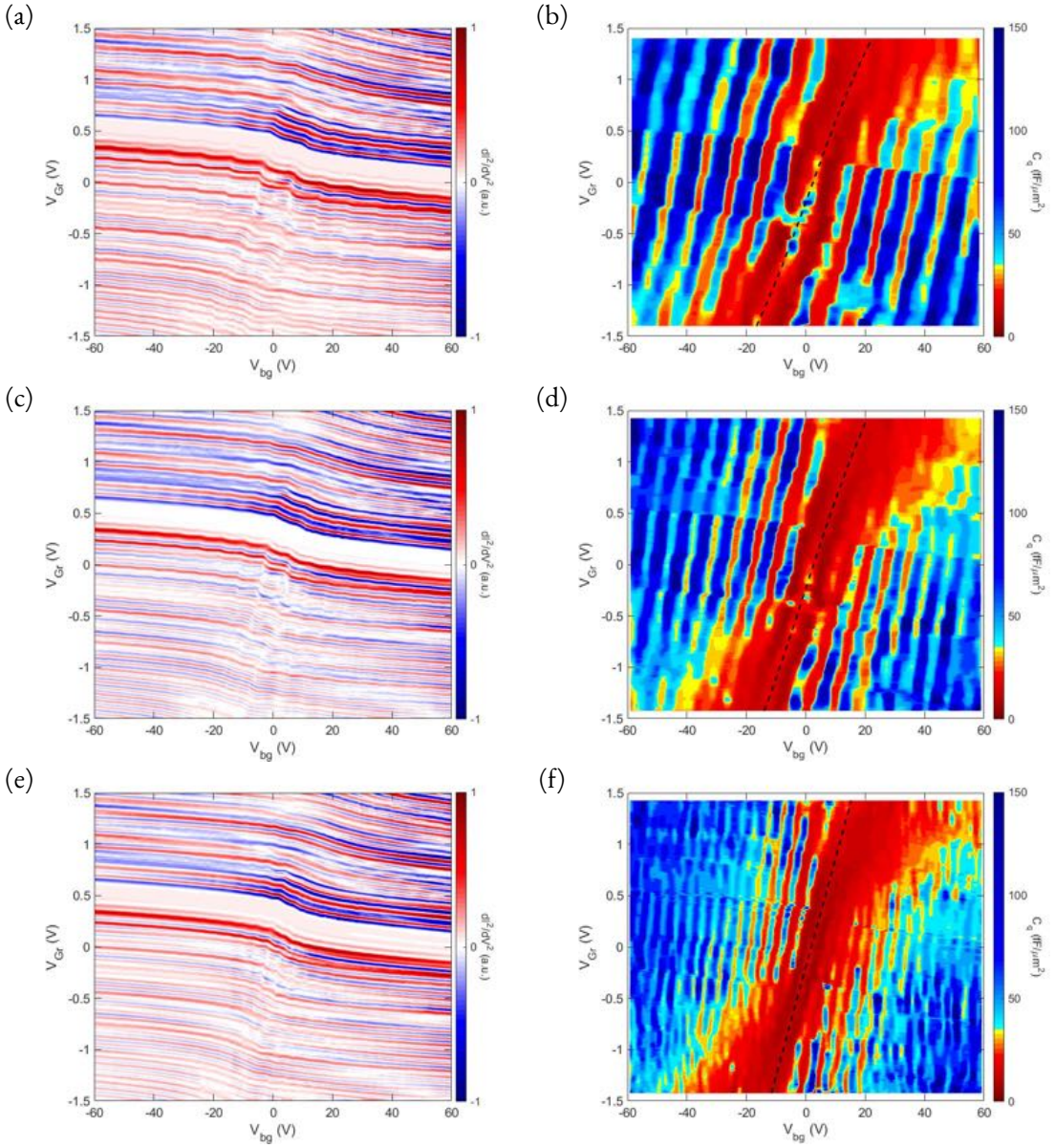


Figure 5.9: Local compressibility measurement and quantum capacitance extraction at other magnetic fields fields. (a),(b) are at 7 T; (c),(d) are at 5 T; (e),(f) are at 3 T. The left column is the local compressibility measurement and the right column is the quantum capacitance extraction. Landau levels collapse at smaller V_{Gr} for smaller magnetic fields. Quantum capacitance data is cut at 150 fF/μm^2 for better feature clarity.

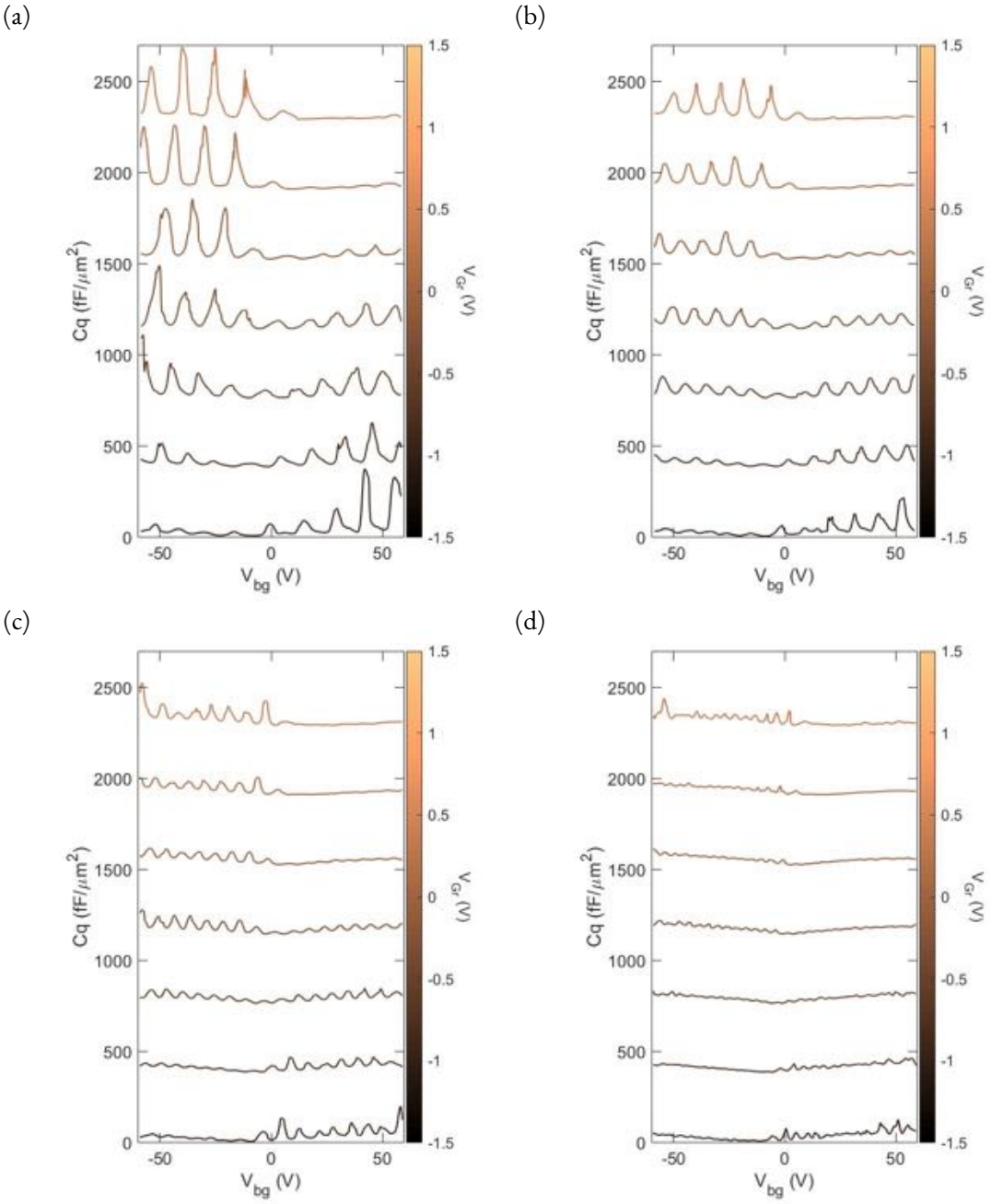


Figure 5.10: Quantum capacitance extraction line cuts at different graphene potentials for different magnetic fields. From top to bottom most lines are cuts taken at $V_{\mathrm{Gr}} = 1.5, 1, 0.5, 0 \ldots -1.5$ V respectively. (a) is at 9 T, (b) is at 7 T, (c) is at 5 T, (d) is at 3 T. Each curve is displaced vertically for clarity.

and bottom left of the measurement plots. This is best seen in lower magnetic fields presented in figure 5.9. The continuous red region in the quantum capacitance graphs (figure 5.9b,d,f) qualitatively represent the collapsed region. In the top right or bottom left, as LLs' indices increase, screening becomes more severe due to the higher number of carriers in the Gr. This explains why LL collapse requires greater V_{Gr} as LL index increase. However, on the top left and bottom right of the measurement plots, we would expect something similar naively but this is not the case. We simply do not see LL collpse happening. An explanation comes from the nature of the potential. The electric field of the potential is given by the slope of the potential profile ($\mathbf{E} = -\boldsymbol{\nabla} V = \boldsymbol{\nabla}\varepsilon/e$). Then, looking at the potential profiles of the potentials in figure 5.11 (given by $V(x)$ denoted by the dark blue lines), one can see that different regions in the V_{Gr} vs V_{bg} map may either have potential wells or barriers and these potentials can be either attractive or repulsive depending on the carrier type in the gated region. A potential barrier is repulsive for electrons but attractive for holes, whereas, a potential well is attractive for electrons but repulsive for holes. Figure 5.11b shows how as the Gr doping under the CNT inverts while maintaining a potential barrier, the potential acting on the Gr carriers would suddenly go from repulsive to attractive or vice versa. When the potential is repulsive, the electric force forces the carrier to escape the gated region and this competes directly with the magnetic confinement that is responsible for pinning the carrier down. In this case if the deconfining electric force overpowers the magnetic confinement, localized cyclotron orbits will not form and LL collapse is possible. When the potential is attractive, both electric force and the Lorentz force will pin the carrier down beneath the gated region. No matter how big the electric field gets, cyclotron orbits when formed will never get destroyed. Hence, LL collapse would not be possible.

As mentioned previously, screening plays a big role in LL collapse. In an attempt to quantitatively analyze LL collpse, in the following we will investigate specifically the behavior of the $N = 0$ LL. This

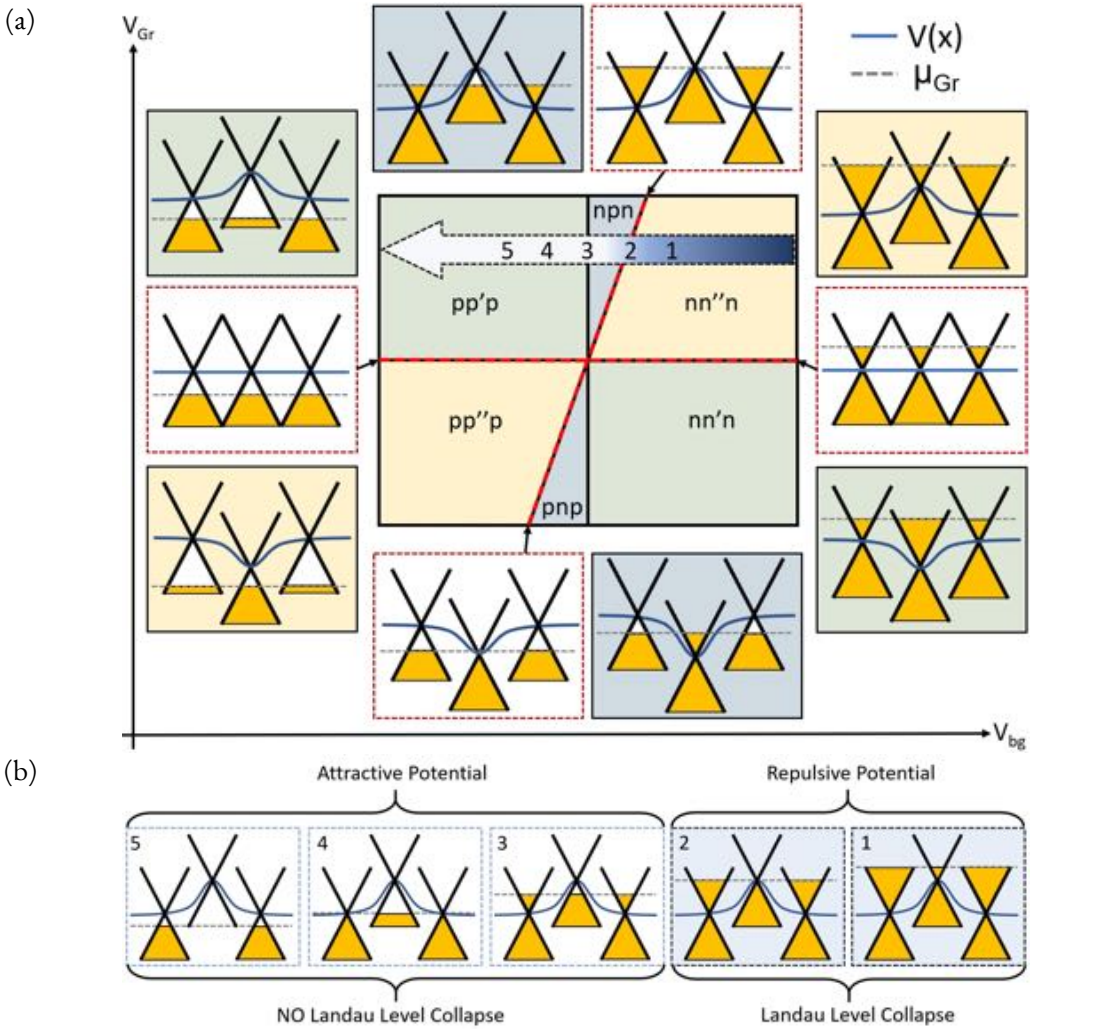


Figure 5.11: Junction polarity and Landau level collapse behavior explanation. (a) The hybrid CNT-Gr device is comprised of a CNT gated region and the bulk region. These regions together constitute either the unipolar or bipolar regime. $pp'p$, $nn'n$,$pp''p$, $nn''n$ are the unipolar regime. npn and pnp are the bipolar regime. The boxed scenarios showing the Dirac cones are colored coded to match the colored regions in the V_{Gr} vs V_{bg} map. The diagonal red dashed line tracks the CNP (DOS minimum) underneath the CNT. The arrow with the dashed outline contains numbers that correspond to the scenarios shown in (b). (b) shows the behavior of the Gr regions corresponding to the numbered positions in the blue arrow shown in (a). Going from right to left (1 to 5, the device exhibits LL collapse and then abruptly ceases to show LL collapse after passing the CNP (at 2). This observation can be explained by whether the carriers under study are feeling an attractive or repulsive potential. LL collapse only occurs with repulsive potentials which compete against magnetic field confinement.

is because the $N = 0$ LL presents the least amount of screening and is thus the most reliable feature to analyze LL collapse accurately. In moving to the next section, we make a note to the reader that the $N = 0$ LL the collapse happens symmetrically about a reference V_{Gr}. This potential reference situates at $V_{\text{Gr}} \approx -0.2$ V due to residual charge in the CNT and Gr. V_{Gr} will be adjusted to reflect the shifted reference.

5.6 Landau Level Collapse for the $N = 0$ Landau Level

Figure 5.12 presents fits for the behavior of the $N = 0$ LL. Figure 5.12a shows the DOS averaged across the full width of the $N = 0$ LL along the V_{bg} axis for $B = 3, 5, 7, 9$ T. A moving average filter along the LL was performed to reduce noise along with a detrending function that gets rid of a small background linear slope. The experimental data is denoted by hollow circles and the solid lines are the fits. A double sigmoid curve of the following form is used:

$$f(x) = \frac{a}{1 + e^{-b\cdot(x+c)}} + \frac{a}{1 + e^{b\cdot(x-c)}} - d. \tag{5.18}$$

This falls under the category of logistic curves. This type of curve best represents behaviors with an activation trigger, which in our case is the onset of LL collapse. This type of curve also offers the best flexibility in fitting steep slopes with wide peaks. Many other peak-like curves (such as Gaussians, Lorentzians and their variations) do not provide as good a fit as equation 5.18.

Examining the curves we see that as magnetic field increases, the curve becomes wider, reflecting that a higher electric field is required to cause LL collapse. The green-blue diamond marks the potential value at which we consider LL collapse to have occurred. We chose the occurrence of LL collapse to be when C_q reaches 0.5 (fF/μm^2) from zero. We note that the value chosen here is in some sense

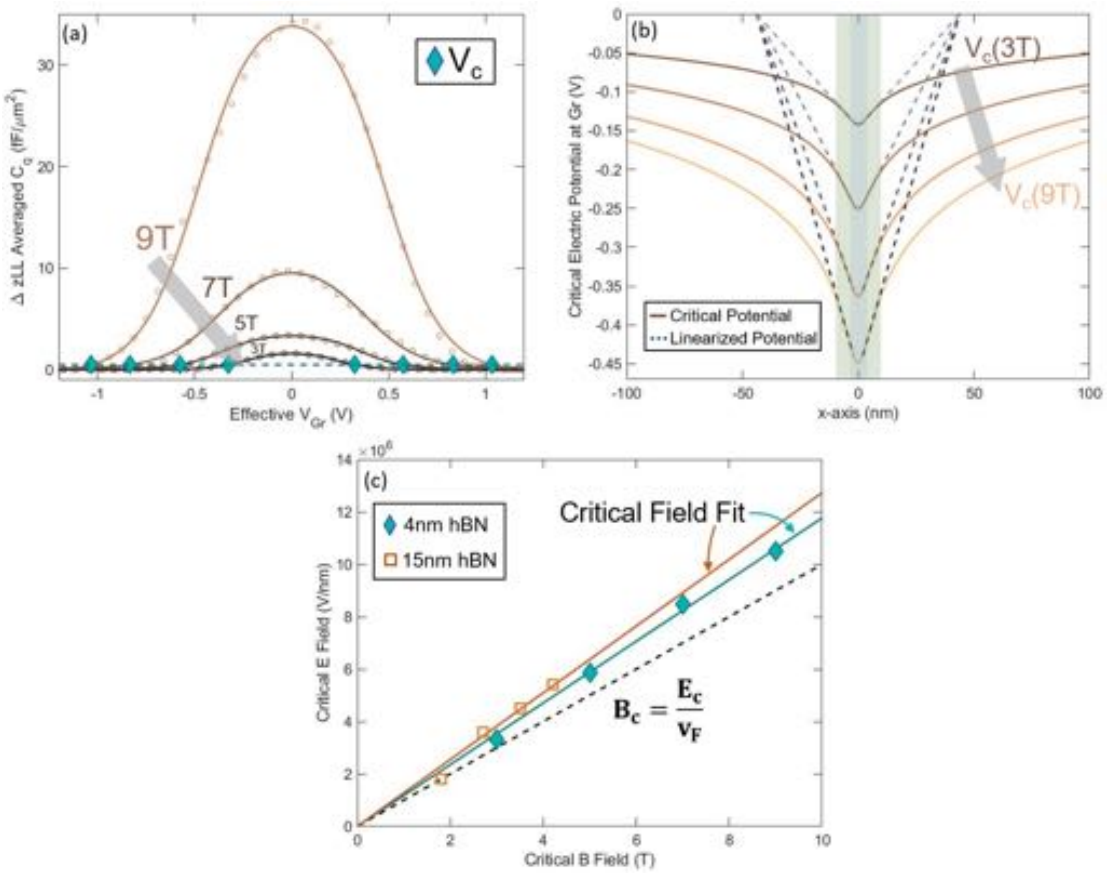


Figure 5.12: $N = 0$ Landau level collapse fits. (a) shows the density of states of $N = 0$ Landau levels for $B = 3, 5, 7, 9$ T denoted from dark to light colors respectively. The blue dotted line demarcates the onset of the collapse. The interception of the dotted line and the LL DOS is denoted by the green-blue diamonds and this interception defines the critical potential (V_c) associated with the critical electrical field at which LL collapse occurs. The hollow circles are the filtered experimental points and the solid lines are the sigmoid (logistic curve) fits. (b) shows the critical potentials plotted with respect to real space. The profile of the potential is based on equation 5.19. The dotted line is the linearized potential. The shaded blue shows the section which the linearization is based on (slope of the $1 - 5$ nm section away from the center of the potential). The shaded green represents the region in which the linearized potential approximates well (< 1 mV variation). This is ~ 14 nm about the center of the potential. (c) shows the theoretical and experimental critical field behavior. The dashed line is according to theory. The green-blue diamonds and orange squares are the experimental points with the solid lines being the linear fits.

arbitrary. This collapse value was chosen such that the critical field fit stays above the $B_c = E/v_F$ as seen in figure 5.12c. By picking a collapse point to be at a higher value such as 1 (fF/μm^2) from zero

would cause the critical electric field at 3 T to dip below the $B_c = E/v_F$ line.

The selected critical potentials each have an associated spatial potential profile. For instance, critical potential for 9T is at an effective $V_{\text{Gr}} \sim 1.1$ V. This critical potential is the potential we chose the CNT to be at with respect to the Gr plane. Gr electrons will then feel the electrostatic forces due to the potential of the CNT and the force felt varies spatially in the Gr plane depending on distance of separation. According to theory, the potential energy of an electron at the Gr plane due to a 1D gate of radius r at a distance $h - d$ from the Gr plane is:[86]

$$V(x) = \frac{e\phi_0'}{2} \ln\left[\frac{x^2 + (h-d)^2}{x^2 + (h+d)^2}\right], \tag{5.19}$$

where

$$\phi_0' = \frac{\phi_0}{\ln\left[(2h - r)/r\right]}, \tag{5.20}$$

and ϕ_0 is the voltage applied between the 1D gate and the doped Si substrate, r is the radius of the 1D wire, h is the distance between the center of the wire and the surface of the Si substrate, and d is the distance between Gr and the surface of the Si substrate. In acquiring this potential, the condition of decoupled Dirac cones is assumed and the Gr dielectric response is neglected. Equation 5.19 provides us the spatial dependence of the potential the Gr electrons feel at a particular critical potential. These potentials as functions of real space is plotted in figure 5.12b. To account for the in-plane dielectric response of Gr, we divided the potential by $\epsilon_{\text{Gr}} = 2$.[3]

To extract the electric field, we linearize potential from equation 5.19 and take the negative of the spatial derivative of it. The linearized potential is based on the linear slope of the $1 - 5$ nm section as outlined by the dashed line. More details and justification regarding the linearization of the potential is covered in section 5.3. The electric field extracted for each critical potential is the critical electric field.

The extracted electric field for each magnetic field is plotted in hollow orange circles in figure 5.12c. The orange line is a linear fit. We see that the extracted critical field values fall quite close to the theoretical $B_c = E/v_F$ line. The extracted E_c values sit above the theoretical line and deviates from it linearly as B_c increases. This is expected due to the screening of the $N = 0$ LL. As magnetic field increases, the DOS of the every Gr LL will increase linearly and this will screen more of the electric field coming from the CNT.

5.7 Conclusion

Using the carbon nanotube as a 1D wire, we were able to induce a 1D potential in graphene that translated into an in-plane electric field. The in-plane electric field is responsible for destroying Landau levels in graphene in the quantum Hall regime. This effect is normally particular to relativistic systems and is rooted in the Lorentz invariance of the equation of motion of relativistic particles. In our result, assuming the Fermi velocity of graphene replaces the role of speed of light, the extracted values for critical field falls impressively well with the theory from relativity, that $B_c = E_c/v_F$ while accounting for screening. Although it must be stressed that graphene is not inherently a relativistic system, the Landau level collapse behavior adds to the novel set of existing relativistic simulations and certainly makes the case for graphene being an excellent relativistic test-bed.

Ideas do not always come in a flash but by diligent trial-and-error experiments that take time and thought.

Charles Kao

6

Single Guided Mode in Graphene

Like photons, electrons can be used as carriers of information. But even better, electrons have the distinct advantage of squeezing into spaces smaller than the diffraction limit of light. At the same time, electrons exhibit different types of interaction effects. All together, it is possible to realize complex logic gates for the exchange or process of information in the nanoscale. However, there is still limited tools to control electrons and the simple fact of guiding them in a solid, like an optical fiber for light, is still a technological challenge. One-dimensional materials such as nanowires or carbon nanotubes naturally provide guidance for electrons, but they cannot be transmitted without loss of information over long distances. High mobility GaAs/AlGaAs heterostructures with engineered narrow gates[87]

or cleaved-edge over growth[88] have shown 1D modes but these require highly involved and delicate fabrication schemes that are difficult to implement. Furthermore, the electrons gas in GaAs/AlGaAs heterostructures are buried roughly a hundred nanometers deep from the surface, which restricts their usefulness for plasmonics applications.

Graphene is a 2DEG that is easy to obtain and is completely exposed. It also has high mobility, allowing electrons to be transmitted ballistically over several microns even at room temperature.[89,64] Additionally, the wavelength of electrons in graphene can be tuned by changing the carrier density. Coupled with the fact that electrons can reflect and refract analogously to electromagnetic waves, the "true colors" of graphene naturally extends to optics. Its potential for plasmonics has generated huge excitement and has made the ambition of creating on-chip communication systems that can interact with light in the nanoscale within reach.[90,91,92] In pursuit of this, there is a need to develop a complete electro-optical toolbox. Some of the essential tools developed include: collimators, beamsplitters, filters, modulators, interferometers, lenses and multimode waveguides.[93,94,95,96,97,98] One of the remaining critical components that have yet to be realized is the single mode waveguide. The single guided mode is important because, like in optical fibers, it allows the transmission of information over long distances without loss. Multi-mode fibers have the advantage of transmitting information more easily and in parallel but is hindered by its high loss and information distortion due to modal dispersion. With the realization of an isolated single guided mode in graphene, the electron-optics toolbox becomes one step closer to completion.

One obvious method for creating guided modes is through the quantum hall effect but it is then necessary to use a large magnetic field, especially for single to few modes,[99] which becomes difficult to be integrated into electronic devices and hence impractical for applications. On the other hand, guided modes in zero magnetic field have been demonstrated in graphene. This was either done by

utilizing the graphene edge or by line gates.[100,101,102] The edge defined guided mode can be complicated by intervalley scattering that is dominant at the edges and because it lacks potential tunability, it is unclear how many modes are being guided. As for the line gate guided modes, the gates used so far are substantially wider than the threshold for a single guided mode; furthermore, the current detection scheme is difficult to single out the existence of distinct guided modes and the overall fabrication scheme is delicate. An isolated and reliable single guided mode have yet to be demonstrated.

6.1 Optical Analogue

The guiding mechanism in graphene waveguides falls into two independent categories. The first is the optical fiber (OF) guiding and the second is the pnp (or npn) guiding (we will refer to both pnp and npn guiding as pnp guiding for convenience). Both of these mechanisms stem from the fact that ballistic electrons in Gr can reflect and refract like light rays in ray optics. In the following we will show why.

6.1.1 Snell's Law in Graphene

The Gr system with a 1D line gate on top exhibits translational invariance along the line gate axis. This requires the momentum along the gate direction (which we pick to be in the y direction, see figure 6.3) to be conserved. Momentum perpendicular to the line gate (x direction), however, need not be. From this momentum conservation reasoning, we obtain the condition:

$$|\mathbf{q}|\sin\theta = |\mathbf{q_t}|\sin\theta_t, \tag{6.1}$$

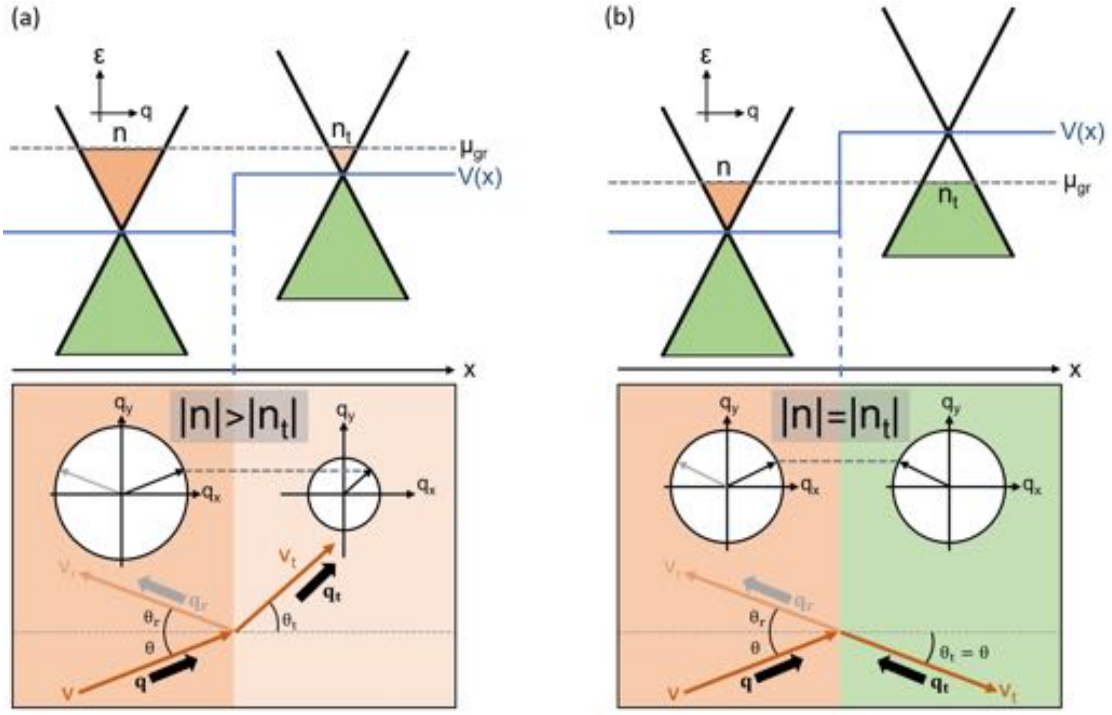


Figure 6.1: Snell's law in graphene. (a) in the unipolar regime, graphene exhibits positive index of refraction. (b) in the bipolar regime, graphene exhibits negative index of refraction where electrons refract away from the norm in the opposite direction. Illustration is adapted from the work by Handschin et al.[103]

where $\mathbf{q}$ is the momentum or wavevector of Gr with respect to the Dirac point, θ is the angle of incidence measured from the norm and θ_t is the angle of transmission (see figure 6.1). This condition is analogous to Snell's law for light and so similarly we define the index of refraction to be the momentum of Gr ($\mathbf{q}$). Consequently, the critical angle is given by:

$$\theta_c = \sin^{-1}\left(\frac{|\mathbf{q_t}|}{|\mathbf{q}|}\right). \tag{6.2}$$

The critical angle defines the threshold for the incidence angle, below which will cause electrons to transmit across the junction and above which will cause cause electrons to reflect completely at the

junction. The critical angle hence characterizes the confinement capability of the waveguide. The reflection or refraction behavior is restricted by one additional condition and that is the conservation of charge number. This means that $\sqrt{\pi n} = \text{Const.} = q_x^2 + q_y^2$. Figure 6.1a shows the behavior when Gr junction is in the unipolar regime and this corresponds to a Gr junction having a positive index of refraction. The index of refraction of the junction itself is defined to be:

$$n_{\text{Gr junc}} = \frac{\sin\theta_t}{\sin\theta}. \tag{6.3}$$

There exists a very unique and highly sought-after scenario and this is illustrated in figure 6.1b. When Gr is in the bipolar regime, the junction behaves as having a negative index of refraction. The transmitted electrons bend away from the norm but at a negative (opposite from positive case) angle. The behavior of the negative index of refraction is what allows Gr to become a superlens and in this case specifically, a Veselago lens.[97,104,98]

6.1.2 Guiding Mechanisms

In accordance to Snell's law, Gr electrons can become confined and hence guided in a gated region if the index of refraction (given by $\mathbf{q}$) is larger in the gated region. They are confined if the angle of incidence towards the junction is greater than the critical angle θ_c. This behavior is analogous to the guiding in optical fibers and is hence called optical fiber (OF) guiding. Transmission across a pn junction decreases exponentially with angle (see section 4.3.2), therefore, when angle of incidence is high, electrons have a larger chance of reflecting back, thus increasing the chances of becoming confined. This is termed pnp guiding and in this type of guiding, the smoother the potential profile, the better the guiding capability (see figure 4.2). Figure 6.2 illustrates the two types of guiding. As shown later,

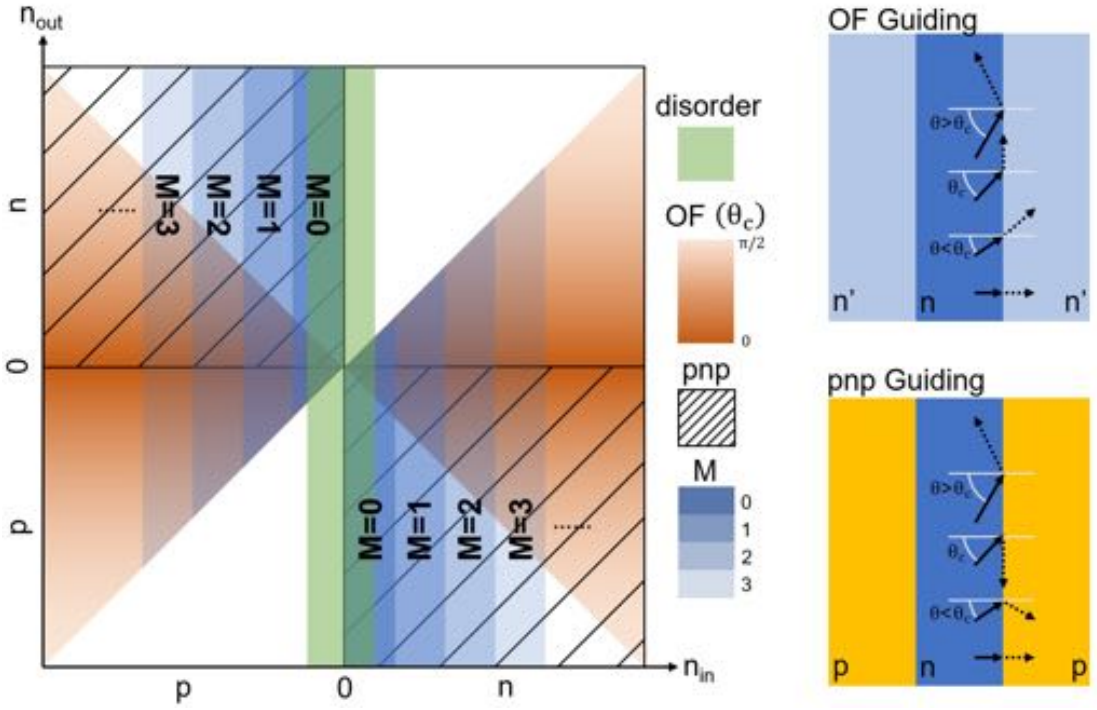


Figure 6.2: Guided mode mechanisms. If carrier density within the gated region (n_{in}) is larger than density outside (n_{out}), electrons reflect if the angle incidence is below θ_c like in an optical fiber. This is called optical fiber (OF) guiding. If n_{in} is opposite in polarity as n_{out}, electrons also reflect if the angle incidence is below θ_c. This is, however, attributed to the *pnp* junction. The shaded blue steps correspond to different modes (M) in the gated region, the hatched quadrants indicate the regions of *pnp* guiding, the orange shaded triangles indicate the regions of OF guiding, and the green region is the disorder dominated zone. The sizes of the blue shaded steps changes as a function of potential well depth and width. On the right shows the trajectories of electrons at the junctions in the OF and *pnp* guiding cases. Note that the relative sizes for the disorder and mode regions are arbitrarily chosen. It is important that the mode of interest not be masked by disorder, i.e. the green zone must not overlap the shaded blue region of interest. Illustrations are adapted from the work by Rickhaus et al.[101]

the guided modes that we detect is when Gr is in the bipolar regime (see figure 5.11 for junction polarity reference). In this figure, it is important to take note of the relative size of the green shaded region (disorder zone) relative to the blue shaded regions (guided modes). The goal here is to observe the single guided mode and it is imperative that the disorder zone not overlap the $M = 1$ guided mode. To ensure this, one can decrease potential width such that the blue step sizes increase (see equation 6.5 where U_0 is directly tied with the width of the blue shaded steps). The effect of disorder will be

addressed in detail in the following sections.

6.2 A Single Weakly Guided Mode

The superiority of the single guided mode is due to its capability of transmitting information unhindered and with high fidelity for long distances. For multi-mode waveguides, because different modes are in close proximity to each other, scattering between modes is possible, and because different modes travel at different velocities, carriers go from one end of a fiber to the other end at different times resulting in signal distortion at the receiving end.

Recent manufactured optical fibers are mostly categorized as weakly guiding.[105,106] These weakly guiding waveguides have a cladding refractive index that is only slightly smaller than the core refractive index, in which case a large portion of the transmitted wave exists as an evanescent wave in the cladding region. This provides an advantage because transmitted wave flows through with a wider width, allowing it to diffract around any propagation-loss-inducing defects along the optical fiber. Likewise, electrons can have a similar property and this is where Gr truly shines.

Gapped semiconducting materials such as GaAs or gapped bilayer Gr, or materials that have hard-wall potentials such as nanoribbons and carbon nanotubes, have one common disadvantage when it comes to electron guiding. The confinement of electrons are caused by a strict boundary condition which limits the possibility of evanescent waves beyond the confined area. Gate-defined potentials in Gr on the other hand do not confine electrons as much because of the conducting bulk. The wavefunction of the electron persists beyond the guided core created by the potential well. This allows the electron wave to diffract around disorder residing within the guided core. The theorized scattering rate in the guided core goes as $\gamma \approx ak^2$ where a is a proportionality constant dependent on disorder strength and potential strength, and $k = 2\pi/\lambda$. For a realistic carrier wavelength of 100 nm (0.01

eV), the weakly guided mode can have a mean free path as large as a few hundred of microns.[107] Such guiding behavior is unique to the gapless Gr.

6.3 Creating and Probing the Single Guided Mode with a Carbon Nanotube

Here we demonstrate the creation and direct probing of a single guided mode in graphene by utilizing a carbon nanotube. A carbon nanotube (CNT) that is carefully placed on a hexagonal-boron bitride (hBN) encapsulated Gr acts as a extremely narrow gate and at the same time as a single electron transistor that can monitor the local density of states (LDOS) directly beneath it. The structure and fabrication scheme of the hybrid CNT-Gr device is shown in chapter 3. All the data presented in this chapter are taken on a device with 4 nm hBN as the dielectric spacer. Similar behavior was seen in two other devices with 6 nm hBN (see appendix G for additional data). We will address the behavior in devices with even thicker hBN towards the end of the chapter. The operational principle and method of DOS extraction are covered in detail in chapter 2.

The essence of the gated guided mode lies in the Gr Hamiltonian subject to an external 1D electrostatic potential:

$$\mathbf{H}(\mathbf{q}) = \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q} + V(x). \tag{6.4}$$

The central question to ask is: what is the necessary $V(x)$ in order to realize the single guided mode? For a given 1D potential, there are two main parameters: the potential depth (U_0) and the potential width (d). Figure 6.3a shows the qualitative depiction of the potential created by a charged CNT. In our hybrid device, the depth of the potential well can be continuously adjusted by a voltage difference applied between the nanotube and the graphene. Its width depends on the diameter of the nanotube and the thickness of hBN between the CNT and Gr. By using a thin hBN ($\sim 4-6$ nm) as the dielectric

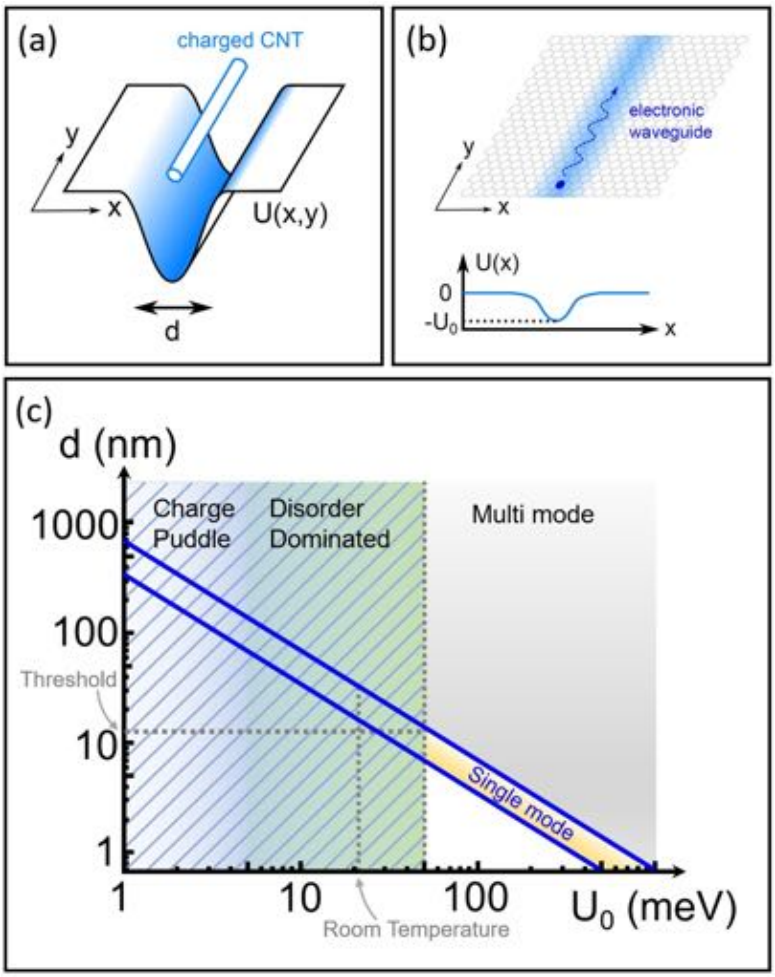


Figure 6.3: Potential illustration and criterion for the single guided mode. (a) the potential caused by the charged CNT and its defining parameters. (b) the potential trench created in the Gr plane once deep enough confines electrons. This potential forms the electronic waveguide. (c) the criterion for the observation of a single guided mode (shaded in yellow). The hatched region is the region where disorder dominates the potential landscape in Gr. To reliably create a single guided mode, the potential trench must be at least an order of magnitude deeper than the potential fluctuation caused by disorder.

spacer, potential variations can be as small as just a few nanometers, which as we shall see is critical to the existence of an isolated single guided mode. Figure 6.3b shows an illustration of the electronic waveguide as a result of the potential trench (shaded in blue) created in the Gr plane. Figure 6.3c shows the criterion to observe the single guided mode. The number of modes (M) is approximately

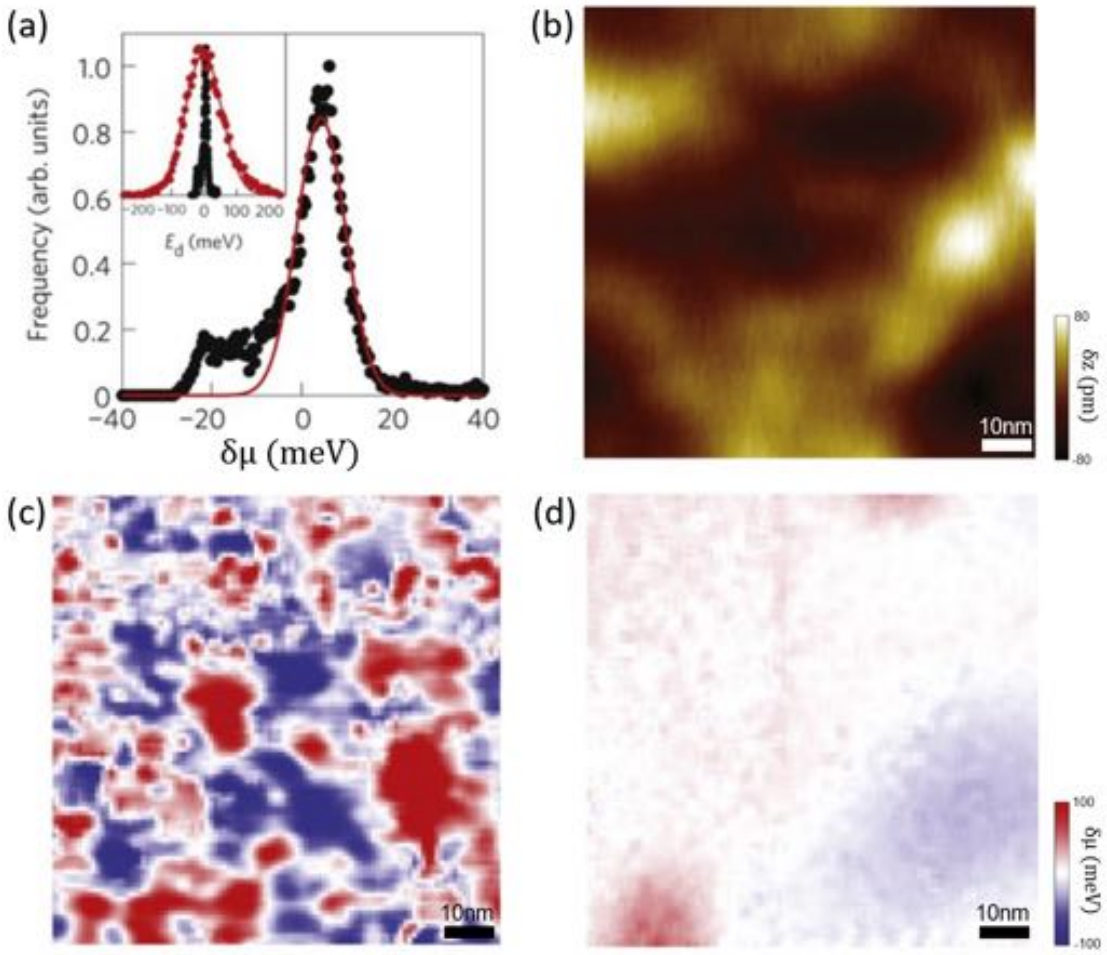


Figure 6.4: Disorder potential on graphene. This disorder is usually caused by charge impurities resulting in a disorder potential that traps puddles of electrons and holes. These are commonly called charge puddles. (a) the distribution of energies of chemical potential fluctuations (or Dirac point energy shifts). The inset shows the same data but also with the distribution for SiO_2 shown in red. (b) topography of Gr on hBN. (c) chemical potential modulation for Gr on SiO_2. (d) chemical potential modulation for Gr on hBN. (c) and (d) share the same colorbar. All plots are adpated from the work by Xue et al.[26]

given by:[109,86,107]

$$\left\lfloor M + \frac{1}{2} \right\rfloor \approx \left\lfloor \frac{d}{\lambda} \right\rfloor \approx \left\lfloor \frac{U_0 d}{\hbar v_F} \right\rfloor , \tag{6.5}$$

for $M = 0, 1, 2...$ and $\lfloor\ \rfloor$ denotes rounding down to nearest integer. v_F is the Fermi velocity of the

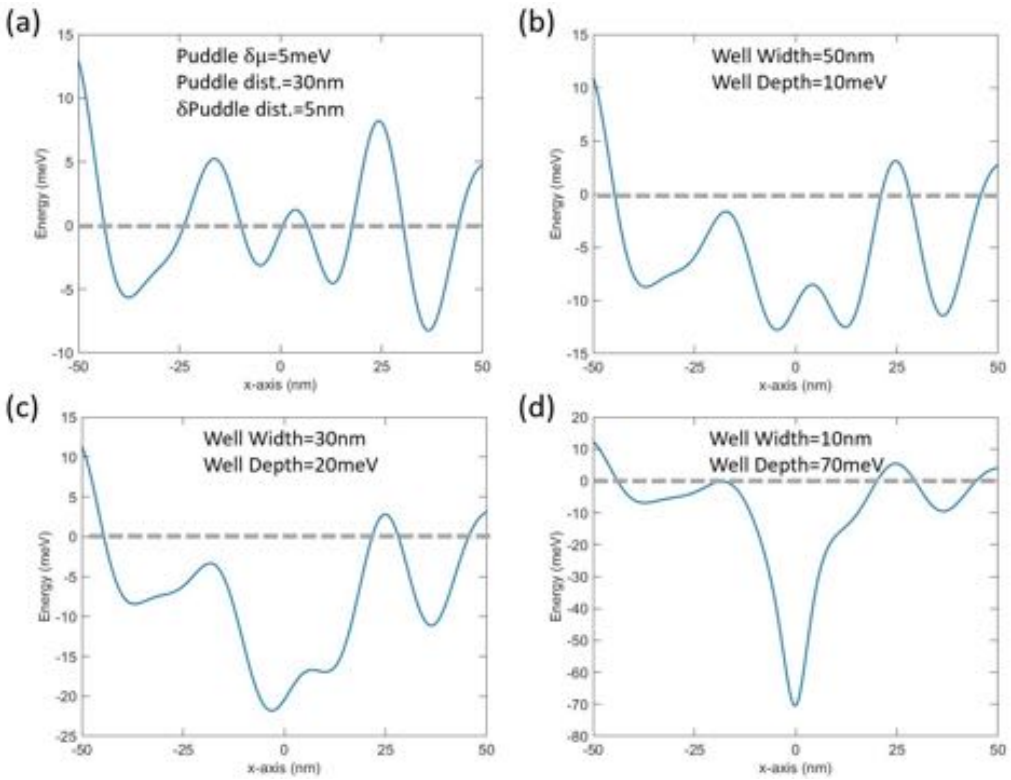


Figure 6.5: Creating a potential well in charge puddles. (a) shows a randomly generated charge puddle potential with a chemical potential fluctuation $\delta\mu = 5$ meV, average puddle distance being 30 nm and the puddle distance variance to be 5 nm. These values are taken to match the experimental findings of a realistic charge puddle in Gr on hBN.[26,25,108] The blue line can be viewed as the chemical doping level with respect to the Dirac point (zero energy). (b) to (d) uses the same randomly generated charge puddle with their respective potential well (Lorentzian form) at $x = 0$ (b) has a well width of 50 nm and depth of 10 meV, (c) has a well width of 30 nm and depth of 20 meV, and (d) has a well width of 10 nm and depth of 70 meV. The value are chosen such that it matches with the single mode criterion. And qualitatively by looking at the graphs one can see that only (d) creates a noticeable enough well and this well has a depth that is an order of magnitude larger than $\delta\mu = 5$ meV. The realistic condition may even be stricter.

electrons and U_0 is the potential well depth. Here U_0 roughly dictates the allowable wavelength of the confined particles. The exact form is dependent on the potential profile. For a square potential for example there is no threshold for observing the first mode and this criterion is in contrast to conventional nonrelativistic systems which always hosts bound modes.[109,86] This relation is analogous to the number of allowed modes in an optical fiber. The single guided mode must therefore require this ratio

to be of the order unity or smaller. In principle this condition can be fulfilled for very wide and shallow potentials, but for the mode to be well defined it is necessary that the depth be much higher than the fluctuations of chemical potential due to the disorder such as charge puddles (see figure 6.4). This explains in particular why it is difficult to guide the electrons in very disordered diffusive graphene and that there has been no convincing report of a single guided mode so far. For Gr encapsulated by hBN, these fluctuations are about $\sim$ 5 meV (at best) and to beat these fluctuations, a potential of an order of magnitude deeper is required (see figure 6.5).[26] This roughly means that to create a single guided mode, the potential will need to be at least well over 50 meV in depth and at most $\sim$ 10 nm wide.

The electronic modes generated by such a potential are manifested as branches appearing outside of the Gr Dirac cone (figure 6.6). These branches are similar to the dispersion relation of optical modes. Being isolated, they are unlikely to mix with other modes and therefore constitute protected carriers of information. These modes are formed locally, at the heart of the potential well, so that they do not affect the overall density of states (DOS) of graphene but appear as resonances in the local state density (LDOS). Therefore, to detect them a local probe must be used. Figure 6.6b shows the evolution of these branches of guided modes with increasing potential well depth. The key in detecting these modes is their von Hove singularity (VHS). Each branch will eventually develop a VHS and each branch is associated with at least one. This VHS becomes more prominent as the potential well gets deeper. In figure 6.6b) we can see such VHS arising. As the potential becomes stronger, the VHS will merge into the continuum and will no longer be seen isolated, however, this does not preclude the existence of a guided mode. The existence of a guided mode in our detection scheme is predicated on a first showing of peak in DOS.

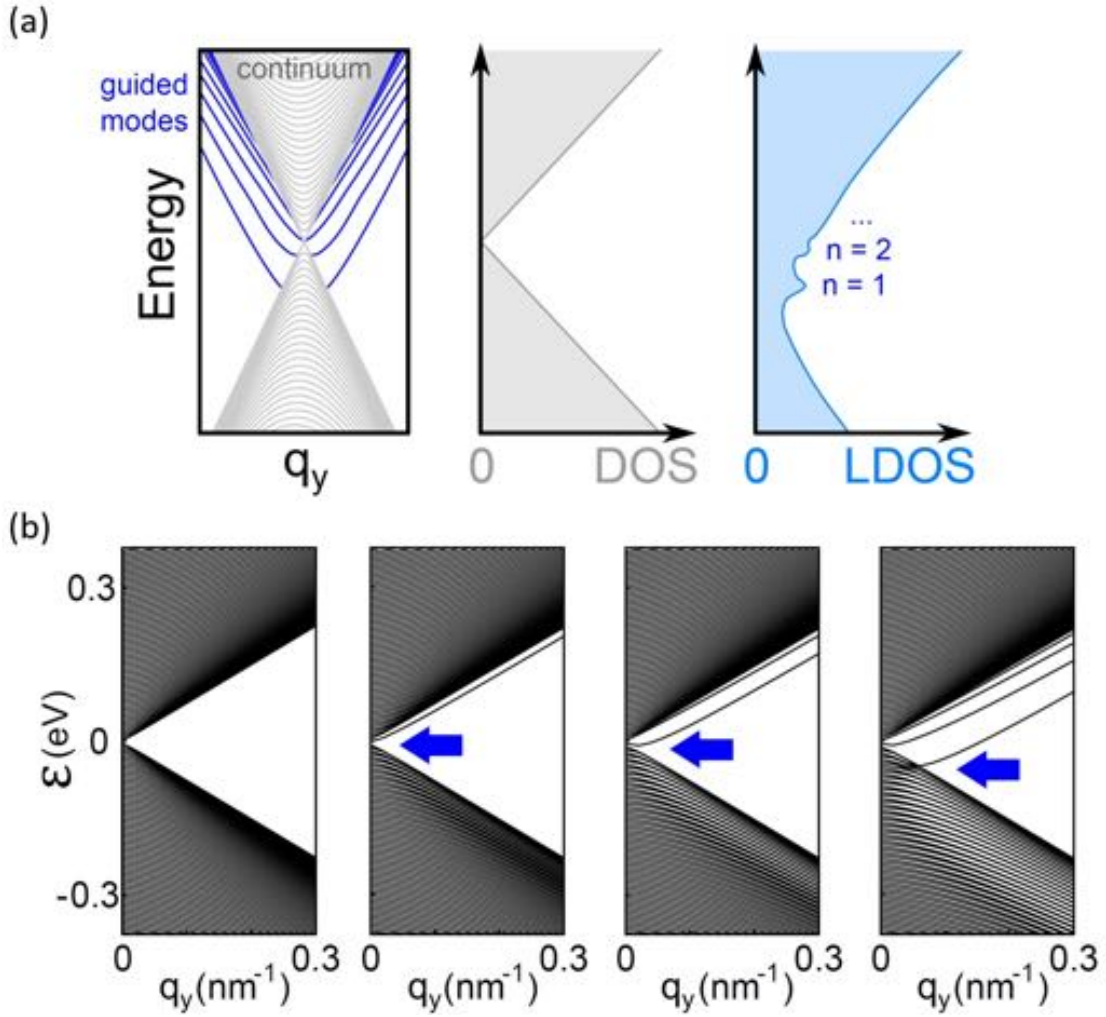
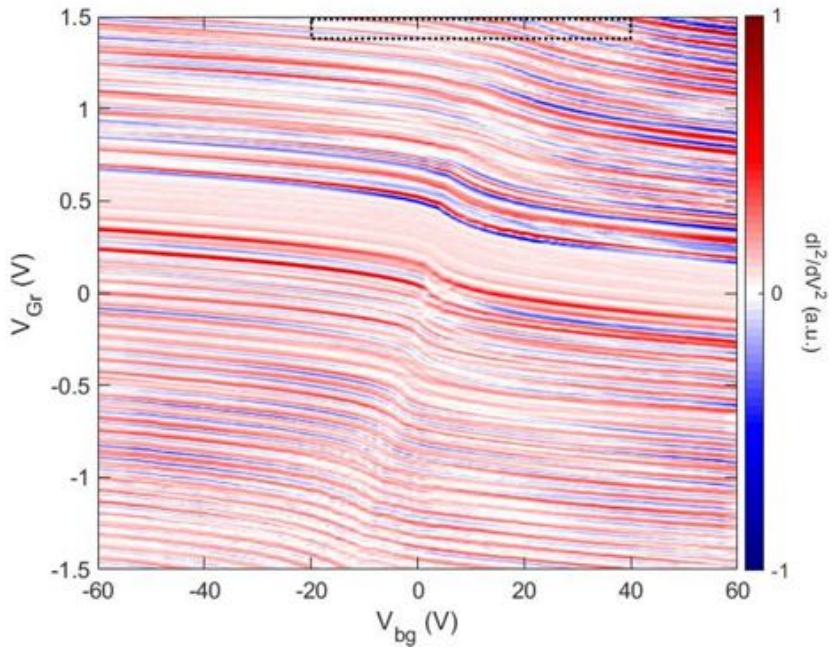


Figure 6.6: Density of states of the guided modes. (a) the left panel shows the dispersion relation of the guided modes (blue lines) which show up as branches sticking out of the continuum (gray lines); the middle panel shows the global DOS of Gr; the right panel shows the LDOS that is immediately beneath the CNT. The local peaks in LDOS is due to the branches of the guided modes (as can be seen in figure 4.8). Each of these guided modes have a von Hove singularity (VHS) that contributes to the LDOS resonance. Note that these VHS can exist inside the continuum. (b) the calculated global dispersion relation of Gr with the 1D potential. Going from left to right, the potential induced (U_{Gr}) are $0, 45, 90, 180$ meV respectively. As the potential increases, more guided modes are allowed. The blue arrow indicates the first guided mode.

(a)



(b)

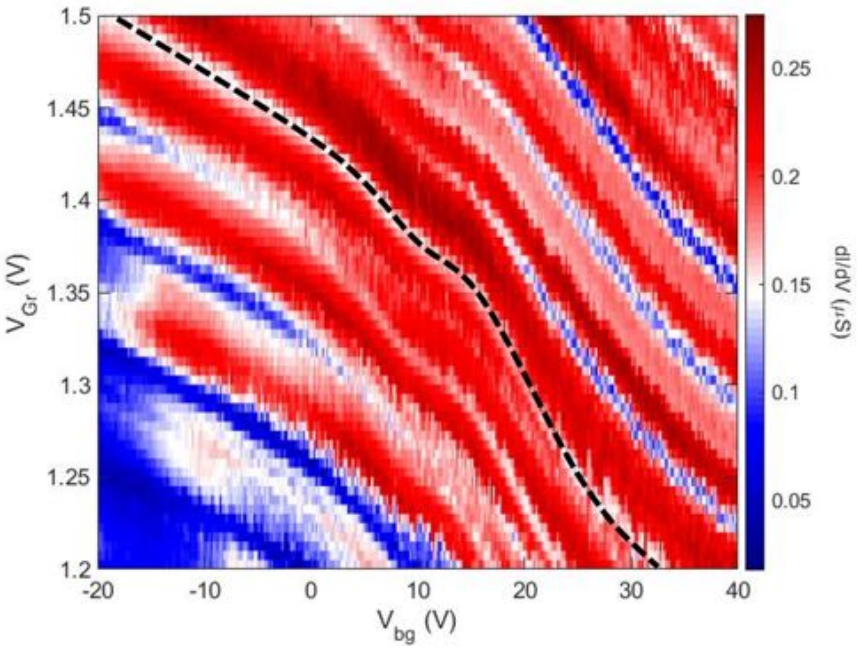


Figure 6.7: Local compressibility measurement with a changing 1D potential given by V_{Gr}. (a) the full map of the differentiated local compressibility measurement. The dotted black box shows the plot window of figure (b). (b) zooms in to a smaller region (black dotted box in figure (a)) to see the "kink" (one of which is marked by a dashed line) with clarity.

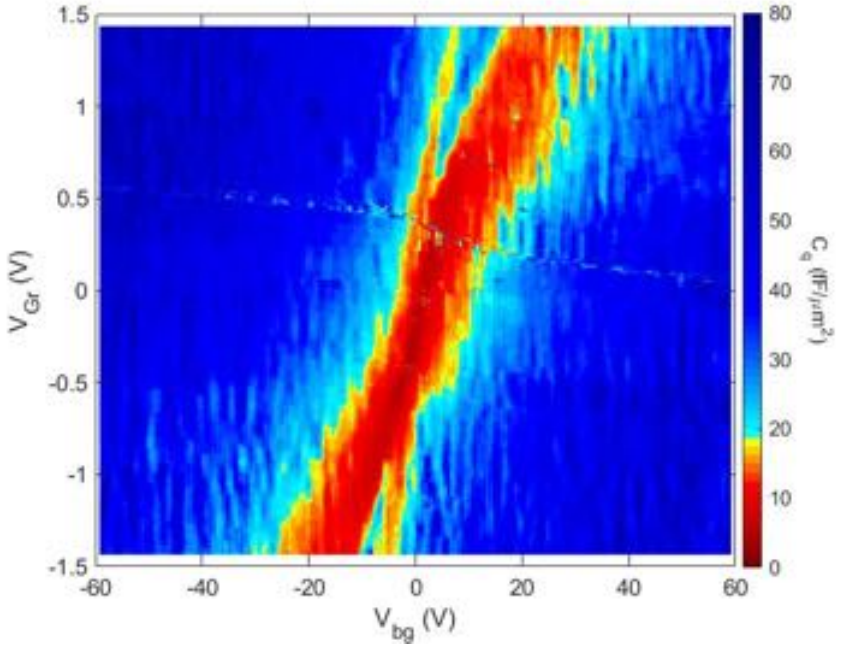
6.4 Local Compressibility Measurement of the Single Guided Mode

Figure 6.7a shows the local compressibility measurement of the hybrid CNT-Gr device. V_{bg} tunes the global Gr carrier density and V_{Gr} tunes the voltage difference between the CNT and Gr which consequently tunes the potential well depth. The general expected features at low potential energy are covered in chapter 2. We move on directly to examine the behavior at high potential energy. At around $V_{\text{Gr}} = 0.5$ V, if one looks closely to the trajectories of the CNT levels, one will see a new "kink" appearing at around $V_{\text{bg}} = 5$ V. The existence of a "kink" indicates a peak in DOS as explained in chapter 2. This "kink" becomes bigger as V_{Gr} increases. At the top limit of the map, the kink spans over 5 V wide in V_{bg}. This kink is a direct result of the VHS formed at the branch of a guided mode. Figure 6.7b shows a zoomed in version of the compressibility measurement. We zoom in to the range between $V_{\text{Gr}} = 1.2$ to 1.5 V and in this plot the kinks are obvious, with one of them outlined by a dashed black line.

Performing a quantum capacitance (C_q) extraction on this compressibility measurement plot (as detailed in chapter 2), we obtain figure 6.8a. In this figure, blue region indicates high C_q whereas red indicates low C_q. The red region in the figure roughly tracks the C_q minimum, which tells us the position of the Dirac peak. As we trace along the red region, from $V_{\text{Gr}} = 0$ to 1.5 V, we begin to see a new blue region emerge at around $V_{\text{Gr}} = 0.5$ V. As explained, this is due to the formation of the "kink" in the compressibility measurement. Reflecting what was observed in the measurement, the C_q of this newly formed resonance becomes higher and wider. Figure 6.8b is a tight binding numerical simulation where a Lorentzian potential was used (see appendix A.9). The formation and evolution of this resonance matches well with theory.

Figure 6.9a compares specific line cuts from the measurement and from theory and here we also

(a)



(b)

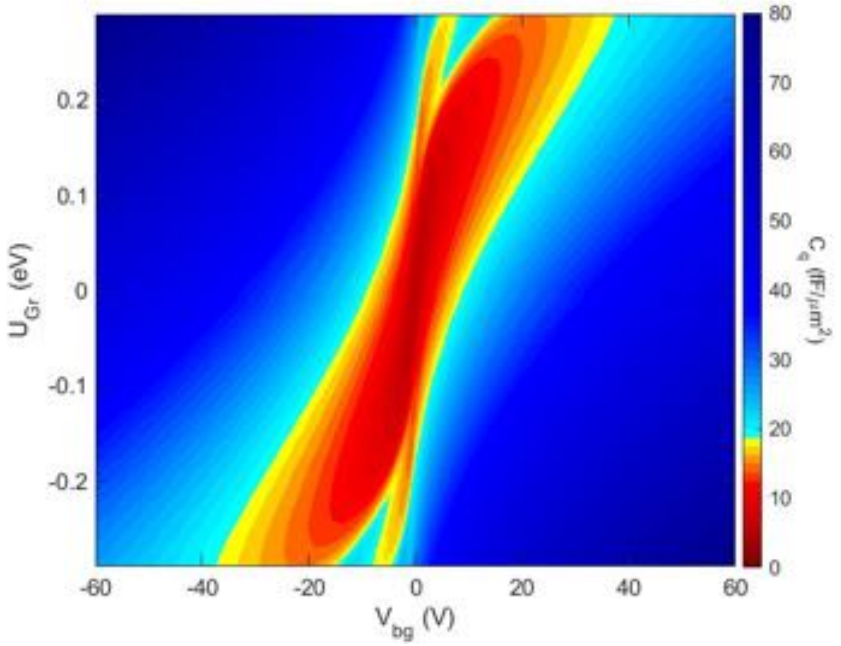


Figure 6.8: Quantum capacitance extraction and the simulation. (a) shows the experimental extraction and (b) is the numerical simulation. The red region roughly defines the Dirac peak or DOS minimum. In (a), at around $0.5V_{\mathrm{Gr}}$ and $-1V_{\mathrm{Gr}}$ a new blue region (peak in DOS) emerges and this corresponds to the single guided mode. (b) is plotted with respect to actual energy as opposed V_{Gr}.

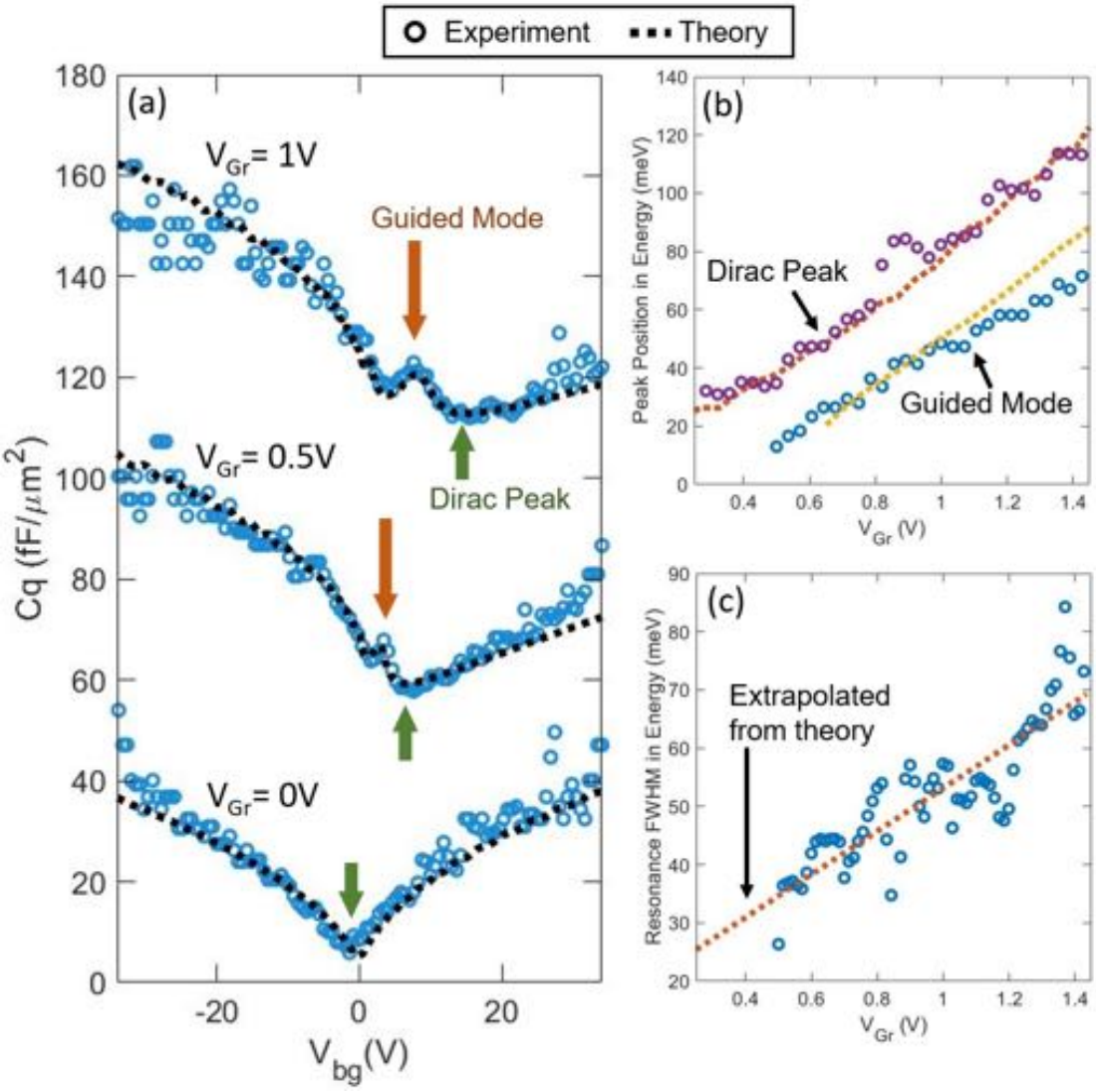


Figure 6.9: Comparison between experimental and theory regarding the quantum capacitance of a single guided mode. (a) line cuts at $V_{\text{Gr}} = 0, 0.5, 1$ V, which are offset by $0, 50, 100$ fF/μm^2 respectively. Blue circles correspond to experiment and black dashed line to theory. The orange arrow points to the resonance due to the guided mode and the green arrow points to the Dirac peak. (b) the tracking of the Dirac peak (or DOS minimum) and the guided mode resonance. Dashed lines are extracted from the simulation and hollow circles are from the experiment. The yellow dashed line and blue circles corresponding to the guided mode terminate because there are no detectable resonances. (c) the full width half maximum (FWHM) of the single guided mode resonance. The orange line is the linearly extrapolated fit of the FWHM of the resonances from the simulation.

see that they are very well matched, matching in both position and magnitude. Figure 6.9b shows the extracted position of the Dirac peak (minimum C_q) and the position of the resonance in both experiment and simulation. They all exhibit a linear behavior although the slope corresponding to the guided mode seems to deviate slightly as the CNT potential increases. The lines for the guided mode terminates suddenly for both experiment and simulation and this is because the resonance peak has become too small to be detected. Although according to the mode creation criterion (equation 6.5) there is a threshold in potential for the first mode to appear and this threshold is around 100 meV for our width potential ($\sim$ 4 nm) which is quite consistent with the loss of resonance detection. The failed detection, however, does not preclude the existence of the single guided mode simply because of the experimental and digital resolution. Figure 6.9c shows the extracted full width half maximum (FWHM) of the resonance. The blue circles are from experiment and the orange dashed line is the slope of the fit of the resonance's FWHM in the simulation. The slope from experiment matches well with the slope from theory. The linear behavior of the FWHM in the experimental range is completely reproducible in all samples with hBN that is 6 nm or thinner.

6.5 Connection with Atomic Collapse

Here we make a very brief connection with atomic collapse. As shown in chapter 4, the hybrid CNT-Gr device can be viewed as a 1D analogue for the atomic collapse problem where Gr provides the relativistic electrons and the CNT acts as a tunable nucleus. It is referred to as the 1D analogue because one can imagine viewing the system from the edge of the Gr while looking down along the CNT and the system reduces to simple 1D problem (i.e. to assume translational invariance along the CNT). By increasing V_{Gr}, we are analogously charging the nucleus. Then in the paradigm of atomic collapse, when the charge of the nucleus reaches supercriticality, relativistic electrons would begin to dive from

the positive continuum (conduction band) to the negative continuum (valence band). In the case of our device, the appearance of the resonance can be the cue for of the beginning of this dive. We see that it roughly takes $V_{\mathrm{Gr}} \approx 0.6$ V from the CNP (situated at $V_{\mathrm{Gr}} \approx -0.2$ V) for the resonance to appear. From the charge stability diagram measurement (see figure G.4), we see that on average an additional charge enters the CNT every $\Delta V_{\mathrm{Gr}} \approx 0.05$ V, meaning that to reach supercriticality it takes roughly ten charges for the 1D atomic collapse to happen. This here is assuming that the appearance of the resonance is the correct cue. It has been theorized that the supercriticality point for 1D atomic collapse is immediate, i.e., as soon as the nucleus is charged, states will begin to dive.[110] This would be in clear contradiction with some theories regarding the appearance threshold of a single guided mode in Gr, however.[86,111] Nonetheless, as soon as the resonance develops, the manner in which it dives down in energy is linear as the nucleus charge increases (compare with equation 4.61 and as shown in figure 6.9b). This is in qualitative agreement with the atomic collapse prediction. However, once the electron states have dove down into the negative continuum they begin to obtain an energy width that is increasing somewhat linearly as shown in figure 6.9c, in contradiction to equation 4.62. One can conceivably fit a parabolic curve to the peak width behavior given the large noise, however. All in all, the hybrid CNT-Gr device shows qualitatively a behavior that maybe consistent with atomic collapse. Keep in mind, strictly speaking the atomic collapse problem is rooted in the 3D Dirac equation with a 3D Coulomb potential. What we have here is simply a 1D analogue toy model that possesses a mathematical similarity with the equations for atomic collapse but not a mathematical equivalence. The reader is left to decide for him or herself whether or not this toy model can be taken seriously as an atomic collapse simulator.

6.6 Multimodes: Absence of Resonances in Wider Potentials

Figure 6.10 shows the case for a wider potential. The shown data come from a device with 30nm thick hBN separating the CNT and Gr. Qualitatively similar behavior is seen in devices with hBN that is 10nm or thicker. Additional data for these other devices are shown in the appendix G. The simulation is for a potential width that is 40nm wide. Figure 6.10a shows the dispersion relation of Gr underneath the CNT when the potential is at $U_{\mathrm{Gr}} = 0.45$ eV and here we see that the branches are essentially indistinguishable from the continuum. The modes formed all have merged into one another. So unlike the single mode which is very well isolated in momentum and energy space, these multi-modes can couple with one another as well as with the continuum easily, thereby making the modes not as well guided. Figure 6.10b shows the numerically calculated C_{q} and here we do not see any resonances developing and this is because there are no VHS coming from individual guided modes. Figure 6.10c shows a local compressibility measurement for the device around a high gate potential ($V_{\mathrm{Gr}} = 7.25$ V). The trajectories of the conductance peaks form smooth S curves which represent the Dirac point and there are no obvious "kinks" seen elsewhere. This signifies the absence of any resonance. However, as traced by the dark green and orange dashed lines, which are extrapolated from the sloped conductance lines far from the Dirac point, the slopes of the lines are mismatched. In this case, the line slopes to the left of the Dirac point ($V_{\mathrm{bg}} \approx 17$ V) are much flatter than the slopes to the right. Because the line slopes are directly correlated with the C_{q} value, C_{q} with respect to the V_{bg} will be asymmetric. This can be seen in figure 6.10d. Figure 6.10e show line cuts from the simulation at different potential well depths. As seen in the experiment, the C_{q} curves become more asymmetric.

The reasoning behind this electron-hole asymmetry can be understood from a simple electrostatic argument. For positive V_{Gr}, the CNT is negatively charged in comparison which will attract positively

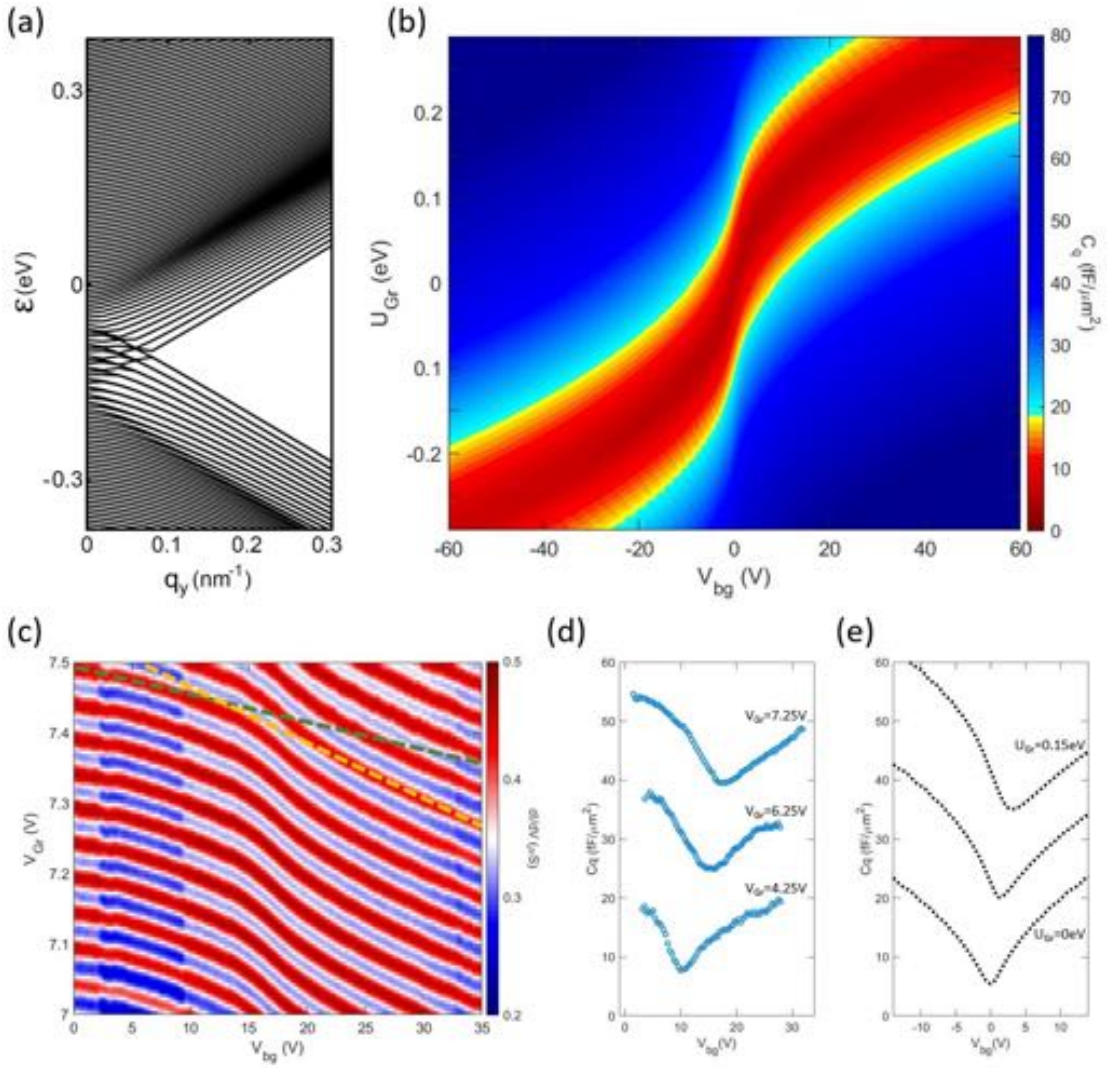


Figure 6.10: In the case of a wide potential, no resonances due to guided modes are formed. (a), (b), (c) are from simulations for a potential well width that is 40nm wide. (c) and (d) are experimental results from a device with 30nm thick hBN in between the CNT and Gr. (a) shows the dispersion relation of the Gr underneath the CNT. (b) plots the simulated quantum capacitance as a function of the potential well depth (U_{Gr}) and backgate potential (V_{bg}). No resonance can be seen. (c) is a local compressibility measurement done around $V_{\mathrm{Gr}} = 7.25$ V. There are no observable "kinks". What can be seen is the electron-hole asymmetry shown by the asymmetry in the slopes of the conductance lines to the left and right of the Dirac point ($V_{\mathrm{bg}} \approx 17$ V. A green dashed line is extrapolated based on the slope to the left of the Dirac peak and an orange dashed line for the slope on the right. The slopes about the Dirac point is clearly different. This is reflected in an asymmetry in C_{q} shown in (d). In (d), extracted C_{q} at $V_{\mathrm{Gr}} = 4.25, 6.25, 7.25$ V are plotted with an offset. C_{q} becomes more skewed with the left side rising up higher as V_{Gr} increases. (e) are linecuts from the simulation plot in (b). Here we also see the C_{q} curve becoming more asymmetric as the potential well depth increases.

charged holes and repel negatively charged electrons. This will increase the LDOS in the hole side while decreasing the LDOS at the electron side.[112] Otherwise, by looking at the formation of modes from theory, we see that for a potential well (not barrier or hill) that is created by a negatively charged CNT (positive V_{Gr}), the new modes develop and pile on the electron side (conduction band). This is best seen in figure 6.6. These additional states contribute to an increase in DOS and hence an increase in C_{q}.

The fact that we do not see resonances for wider potentials is in agreement with theory and this is a clear indication of the formation of only multimodes, much like the previous attempted works on guiding modes.[101,102] This provides further evidence that significantly sharp potential wells are required for the realization of a single bound state.

6.7 Briefly on Magnetic Field Dependence

Figure 6.11 shows the data from the 4 nm thick hBN device under a magnetic field of 1 T. In this magnetic field, we still see the resonance and the general behavior seems to not have altered. We expect the guided modes to be insensitive to the external field up until Gr enters the quantum hall regime. This is because, as shown in the result from Landau level collapse in chapter 5, Gr electrons have two distinct regimes under the influence of crossed electric and magnetic fields. There exists a critical magnetic field below which the Gr electrons will behave as though there was no magnetic field. Following this shown observation, we can conclude that the guided modes should be robust if the applied magnetic field is below the critical value.

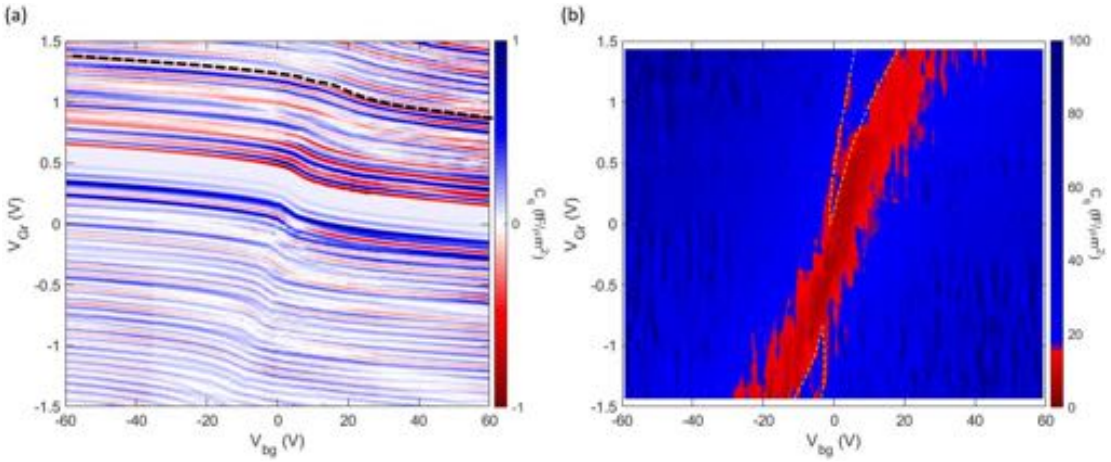


Figure 6.11: Compressibility measurement and extracted quantum capacitance at 1 T. (a) compressibility measurement and (b) extracted quantum capacitance both at 1 T. There is no signs of quantum hall oscillations and the guided mode resonance is still present (marked by a faint yellow line), demonstrating magnetic field robustness up to 1 T.

6.8 Conclusion

Our hybrid CNT-Gr devices with hBN as the dielectric spacer have let us fully examine the guided mode's dependence on potential well sharpness. In our local compressibility measurements for devices with 4 to 6 nm thick hBN we were able to observe a single resonance which originates from the existence of a single guided mode. In our devices that uses hBN that is 10 nm or thicker we were unable to see this resonance given the experimental limitations posed by the dielectric breakdown voltage of hBN. From our observations we conclude that to create a single guided mode in Gr, a sharp potential is required and this speaks to the importance of using the CNT as the potential modulator for future applications in the field of electron optics.

Spukhafte Fernwirkung [Spooky action at a distance].

Albert Einstein

7

Towards Probing Hydrodynamics via Coulomb Drag between a Carbon Nanotube and Graphene

Hydrodynamics is a universal description of a strongly correlated and rapidly thermalizing physical system. As the name implies, the governing equations are readily applicable to liquids which contain high numbers of constituents that are constantly colliding with one another, allowing the fast estab-

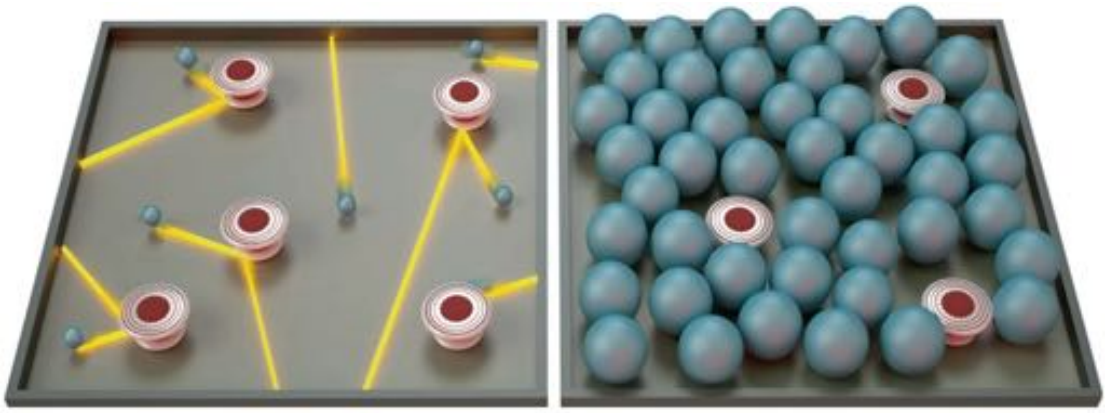


Figure 7.1: Non-hydrodynamic vs hydrodynamic flow regimes in graphene. On the left, electrons flow ballistically and occasionally scatter off impurities. This is the Knudsen regime. On the right, electrons are constantly colliding with one another and hence equilibrating in a macroscopic fluid described by the theory of hydrodynamics. This illustration is taken from the work by Zaanen.[113]

lishment of local equilibrium and hence satisfying the requirement of a hydrodynamic fluid. Electrons in matter in the right conditions can also behave exactly as a hydrodynamic fluid. The governing equations are blind to the type of constituents. So long as the particles, or molecules, or charges are highly interacting and are able to conserve their total energy and momentum, the many-body system can be accurately described as a single entity with macroscopic properties.

7.1 The Quest to Probe Electrohydrodynamics

The application of the theory of hydrodynamics to condensed matter can be traced back to 1968 when Radii Gurzhi predicted that the resistivity of a conductor would decrease with increasing electron-electron (ee) scattering rate due to the Poiseuille flow of electrons (as opposed to the Knudsen flow which describes electrons as ballistic particles in a disordered medium). This was called the Gurzhi effect.[115] Experimental verification of this prediction eventually came albeit slowly due to the limited systems that allowed hydrodynamics as well as the difficult conditions for it to be manifested.[116,117,8,118]

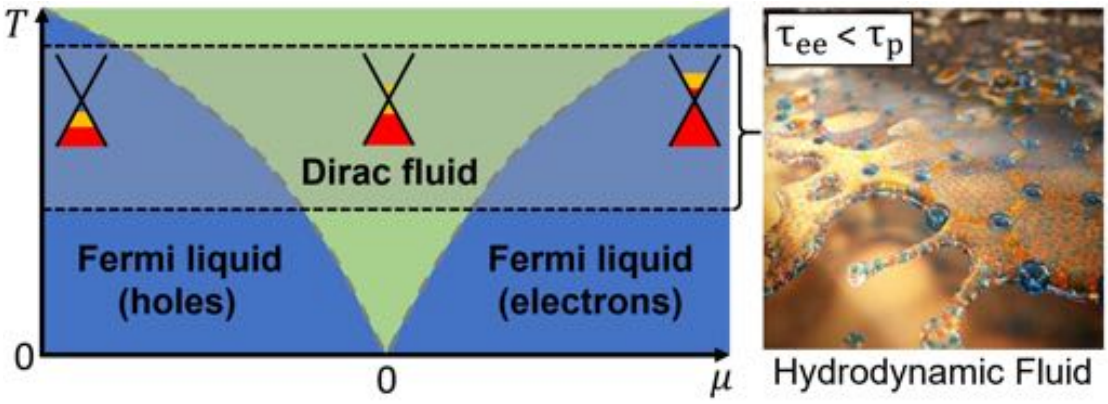


Figure 7.2: Electronic phase diagram of graphene. The yellow area corresponds to the $k_B T$ window (thermal smearing). In the Fermi liquid regime, graphene only consists of one type of carriers while in the Dirac fluid regime, both electrons and holes coexist. There exists a regime where the electron-electron scattering length is the dominant length scale. In this regime, the electron liquid in Gr behaves as a hydrodynamic fluid. Left picture is adapted from the work by Lucas et al.[114]

The advent of Gr resurrected the pursuit of hydrodynamics. In fact, Gr turned out to be an even more exotic hydrodynamic system due to its Dirac cone bandstructure. Thus far, the main experiments unveiled Gr electrons in the Fermi liquid regime as an extremely viscous liquid.[5,119] One of the key signatures of hydrodynamics in Gr is the negative potential that develops right next to the injection site as a result of viscous entrainment of the electrons in the fluid. This is referred to as negative vicinity resistance. In a setting with closed boundary conditions, the viscous electronic fluid would form whirlpools (see figure 7.3).[120,121,122,123] However, since these proposed experiments all rely on performing conventional transport measurements whereby inevitable stray currents are injected into the electronic system, measuring a negative vicinity resistance or spotting whirlpools can be difficult. Furthermore, electrons in the ballistic regime are also known to cause negative vicinity resistance and distinguishing the hydrodynamic regime from the ballistic regime is not straightforward.[124,125,126,127] Finally and perhaps most critically, one of the most intriguing properties of Gr is that at its Dirac point, Gr simultaneously hosts both electrons and holes which act as a relativistic plasma. This is

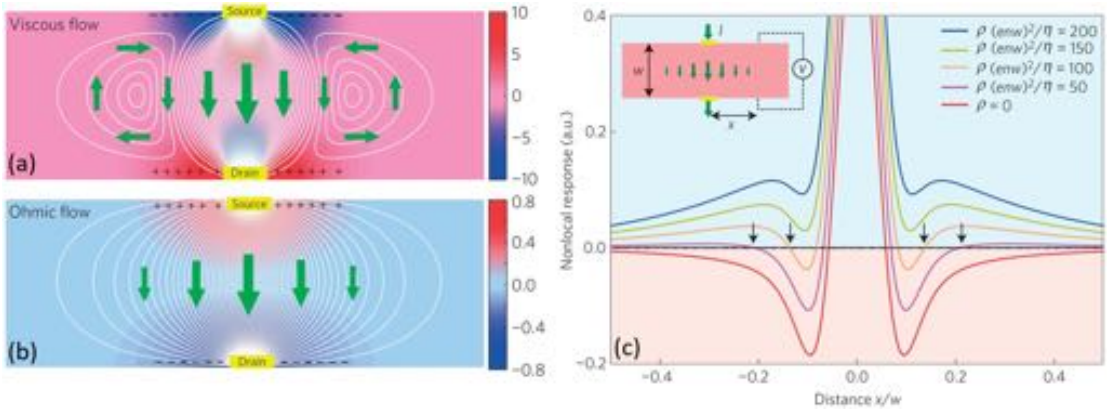


Figure 7.3: The two scenarios when injecting current directly into graphene. (a) the hydrodynamic signature of a viscous fluid is negative vicinity resistance. (b) when the fluid is purely ohmic, there are no vortices. (c) the longitudinal potential difference for different viscosities. The presence of a viscous fluid will cause a sign change along the transverse direction. The illustrations are adapted from the work by Levitov et al. [122]

known as the Dirac fluid and there are many exciting predictions regarding this fluid including the existence of the perfect fluid.[9,128,129] However, to date no conventional transport measurements have been able to directly probe this regime due to many associated experimental complications such as Joule heating.

7.1.1 Carbon Nanotube as the Perturbation Tool to Create a 1D Moving Boundary Condition

It has long been theorized that Coulomb drag in Gr can bring out hydrodynamics in Gr and there are unexplained observations already in Coulomb drag measurements that may be tied with hydrodynamics.[130,114,131,132] We seek to study hydrodynamics in a cleaner way and in a system more analogous to those in fluid mechanics. One hydrodynamic scenario that we can duplicate is the 2D lid-driven cavity. In this scenario, Gr acts as the 2D cavity hosting the electron liquid and the CNT becomes the 1D moving lid that serves to drive the liquid in Gr (see figure 7.4). If the CNT is placed close

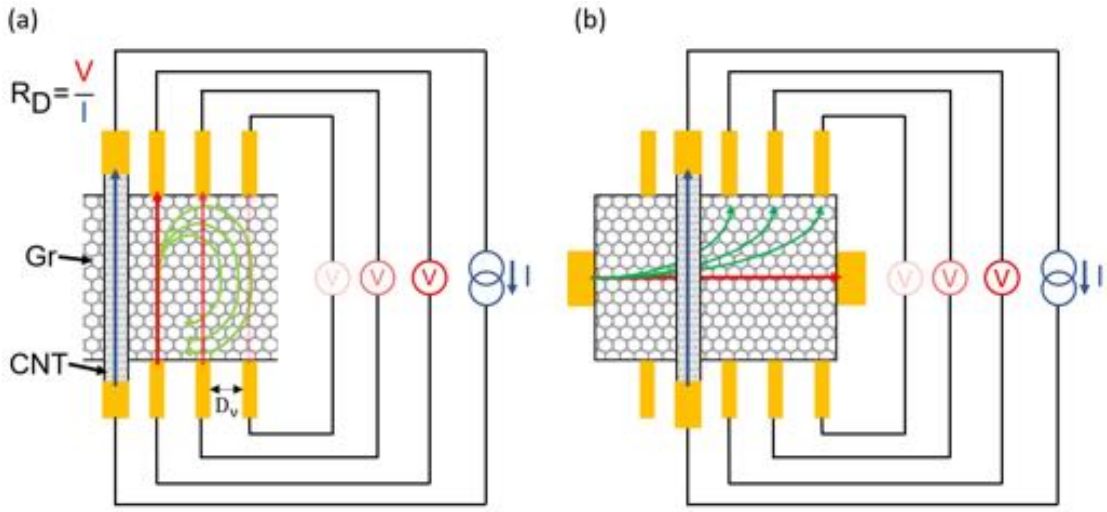


Figure 7.4: Carbon Nanotube for the creation of a 1D moving boundary condition. (a) and (b) show two different configurations of the Coulomb drag measurement between a carbon nanotube and graphene. With closely spaced leads (D_v is on the order of a few hundred nanometers which is the characteristic length scale of whirlpools of a hydrodynamic liquid in graphene, the device can possibly detect the spatial behavior of the electronic fluid.

enough to the Gr, the moving electrons in the CNT can drag electrons in Gr in the same direction via Coulomb drag. In doing this, we ensure that the driven electrons in Gr only travel in one direction and this eliminates the parasitic effects coming from stray currents that is otherwise introduced in conventional transport measurements.

But Coulomb drag turns out to be an extremely complex and convoluted phenomenon on its own right. The study of hydrodynamics has thus been put on hold and the theoretical background and simulations has been relegated to appendix E. The rest of this chapter will focus on the complexities regarding Coulomb drag.

7.2 Brief Introduction to Coulomb Drag

When two spatially separated conductors are placed in close proximity to one another, charge carriers in one can interact with charge carriers in the other via Coulomb interaction. By driving an electric current (I_{drive}) in the drive (active) layer, if the interlayer Coulomb interaction is the dominating force, charges in the drag (passive) can get dragged. Consequently, if the drag layer is left in an open circuit, the dragged charges will pile up on one end of the layer and an ensuing potential will develop across this drag layer. This potential is called the drag voltage (V_{drag}). The drag resistance is defined to be:

$$R_D = \frac{V_{\text{drag}}}{I_{\text{drive}}}. \tag{7.1}$$

We choose the sign of drag resistance to be negative when the two layers have the same carrier types and positive when they have opposite carrier type. This quantity quantifies the interlayer resistance of the double layer system, the double layer system being the two conductors. It also tells us how interactive the particles are in the system. Particularly, any true drag signal due to Coulomb drag is a direct consequence of interlayer electron-electron (e-e) interactions. The use of Coulomb drag in a double layer system can then be used to either probe or drive e-e interactions specifically.

The main challenge in performing Coulomb drag measurements is to decipher whether the measured drag signal is purely due to inter-layer Coulomb interactions. Because individual layers in a double layer system can talk to each other via other means such as capacitively, the measured drag signal can often be mistaken as other parasitic effects. In the following sections, I will lay down the basic principles of Coulomb drag and show some of its main signatures. I will then discuss possible parasitic signals. Finally, I will show the ongoing progress of our work regarding the Coulomb drag between

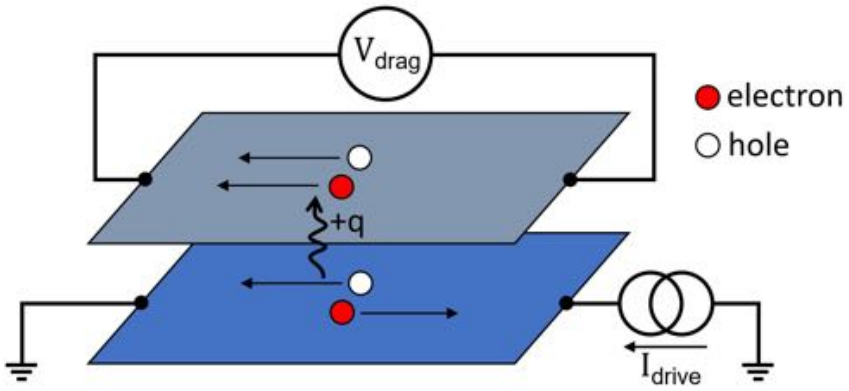


Figure 7.5: Illustration of momentum drag. A current is driven into the drive layer. The charge carriers in the drive layer, through electron-electron interactions, impart momentum to the carriers in the drag layer. As a result the carriers get dragged and a voltage develops across the drag layer. Note that electrons and holes both receive the same momentum so a detectable signal necessarily requires an asymmetry in the electron's and hole's transport properties.

graphene (Gr) and the carbon nanotube (CNT).

7.3 Coulomb Drag Mechanisms

Coulomb drag has been studied extensively since the early 1990's.[133,134,135,136,137] The experimental platform had been GaAs quantum wells and all of the experimental findings had been explained very successfully by momentum drag where e-e scattering mediated by Coulomb interaction transfers momentum from the drive layer to drag layer. The same type of drag measurements were performed after the advent of Gr. The double Gr layer system proved to be a much more versatile experimental platform because the individual layer carrier concentration can be tuned separately in addition to the application of magnetic field and interlayer displacement field. Most importantly though is that the separation between the two Gr layers can be much smaller than the interlayer spacing in traditional 2DEG double layer systems. This allowed the probing of the uncharted strong-coupling regime where

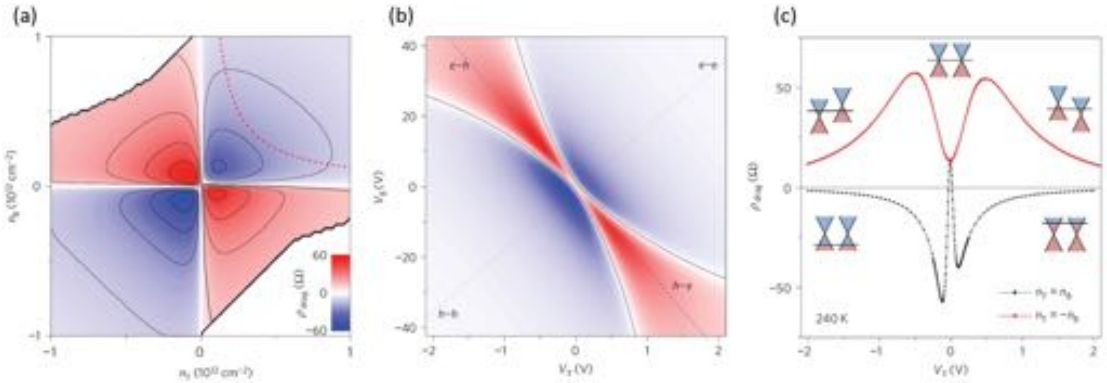


Figure 7.6: Momentum drag performed in a double graphene layer system by Gorbacehv et al.[138] (a) shows the drag signal with respect to drive and drag layer carrier densities. (b) shows the same drag signal except in terms of measurement parameters: the top and bottom gate voltages. In both (a) and (b), we see the characteristic negative drag for same doping type in the two layers and positive for opposite. The drag signal diminishes as either carrier density goes to zero. (c) shows the line cuts for the same and opposite layer doping. In (c) we see the anomalous positive drag signal at double charge neutral point. Momentum drag predicts the value to vanish.

the Fermi wavelength is the same order of magnitude as the separation distance. This means that the interlayer and intralyer interactions are roughly equally strong.

Recent measurements on Coulomb drag in the double Gr system yielded unexpected results however. At high carrier concentrations, Gr Coulomb drag could be explained well by momentum drag, but near the double charge neutrality point (DCNP), the drag signal was anomalously high and negative. See figure 7.6 for the experimental results.[138] This was just one of the anomalies. In the low temperature regime of double monolayer drag system as well as the double bilayer Gr system, there were results that do not follow the prediction by momentum drag.[139,140,141,132,142] Many of these mysterious findings bewildered the drag community. People have attempted to explain these new behaviors but at the time of writing, there is no definitive explanation for the anomalous behavior. Below, I will report the two popular mechanisms that together seem to sufficiently explain Gr Coulomb drag.

7.3.1 Momentum Drag

Momentum drag is considered to be the conventional drag mechanism. The gist of it is as mentioned before: e-e scattering mediated by Coulomb interaction transfers momentum from the drive layer to drag layer. Let us address this phenomenon more quantitatively.

In the semiclassical Drude treatment, we can break down the transport of a bilayer system into two coupled equations of motion:

$$\frac{d\mathbf{p_1}}{dt} = q_1(\mathbf{E_1} + \mathbf{v_1} \times \mathbf{B}) - \frac{\mathbf{p_1}}{\tau_1} - \frac{(\mathbf{p_1} - \mathbf{p_2})}{\tau_D}, \tag{7.2}$$

$$\frac{d\mathbf{p_2}}{dt} = q_2(\mathbf{E_2} + \mathbf{v_2} \times \mathbf{B}) - \frac{\mathbf{p_2}}{\tau_2} - \frac{(\mathbf{p_2} - \mathbf{p_1})}{\tau_D}, \tag{7.3}$$

where subscripts $\{1, 2\}$ denote the drive and drag layer respectively, $q_i = +|n_i e|, -|n_i e|$ for holes and electrons respectively, $\mathbf{p_i} = m_i \mathbf{v_i}$ and $\mathbf{B}$ is assumed to be homogeneous across the two layers. $\tau_{1,2}$ are the intralayer scattering times and τ_D the interlayer scattering time. Noting that in a drag measurement the drag layer is in an open circuit with no net current, so $v_2 = 0$, then the longitudinal drag resistivity in steady state is:

$$\rho_D = \frac{E_2}{e n_1 v_1} = -\frac{m_1}{n_1 n_2 e^2 \tau_D}, \tag{7.4}$$

under the premise that the bilayer system is composed of two 2D layers where $n = [\mathrm{m}^{-2}]$ and $\rho_D \cdot (L/W) = R_D$. The longitudinal drag resistivity is independent of magnetic field and has a sign that depends on the doping of the two layers. In the work regarding Coulomb drag, we will take the sign of drag resistance to be negative if layers have the same carrier type and positive if opposite. In this treatment, all the microscopic details are encoded in the phenomenological parameter τ_D.

To calculate the rate of change of momentum due to e-e scattering between drive (layer 1) and drag (layer 2) layers, we start from the collision integral integrated over transferred momentum:

$$\begin{aligned}\frac{d\mathbf{p_2}}{dt} &= \int \frac{d\mathbf{q}}{(2\pi)^2}\left(\hbar\mathbf{q}\left.\frac{df(\mathbf{q})}{dt}\right|_{\text{col}}\right) \\ &= \int \frac{d\mathbf{q}}{(2\pi)^2}\hbar\mathbf{q}\int \frac{d\mathbf{k_1}}{(2\pi)^2}\int \frac{d\mathbf{k_2}}{(2\pi)^2}\Big\{W_{\mathbf{k_1},\mathbf{k_1}+\mathbf{q},\mathbf{k_2},\mathbf{k_2}-\mathbf{q}}\Big\}\cdot \\ &\quad \Big\{f(\mathbf{k_1})\big[1-f(\mathbf{k_1}+\mathbf{q})\big]f_0(\mathbf{k_2})\big[1-f_0(\mathbf{k_2}-\mathbf{q})\big]- \\ &\quad f(\mathbf{k_1}+\mathbf{q})\big[1-f(\mathbf{k_1})\big]f_0(\mathbf{k_2}-\mathbf{q})\big[1-f_0(\mathbf{k_2})\big]\Big\}.\end{aligned} \tag{7.5}$$

Here, f is the non-equilibrium distribution function and f_0 is the Fermi distribution function. $\mathbf{q}$ is the interlayer momentum transfer and subscripts $1, 2$ denote drive and drag layer respectively. The transition probability is given by Fermi's Golden Rule:

$$W_{\mathbf{k_1},\mathbf{k_1}+\mathbf{q},\mathbf{k_2},\mathbf{k_2}-\mathbf{q}} = \frac{2\pi}{\hbar}|\mathbf{V}(\mathbf{q})|^2\delta(\varepsilon_{k_1}+\varepsilon_{k_2}-\varepsilon_{k_1+q}-\varepsilon_{k_2-q}). \tag{7.6}$$

For Gr there is a phase associated to the transition probability due to the non-trivial Berry phase as well an overall multiplier due to degeneracy (see appendix A.6). Hereinafter we will ignore these for simplicity. The final form for rate of momentum change is unaltered regardless. The Fourier transform of the unscreened interaction potential (the Coulomb potential) for carriers confined in a 2D plane is given by:

$$\mathbf{V}(\mathbf{q}) = \begin{bmatrix} \frac{2\pi e^2}{q} & \frac{2\pi e^2}{q}e^{-qd} \\ \frac{2\pi e^2}{q}e^{-qd} & \frac{2\pi e^2}{q} \end{bmatrix}. \tag{7.7}$$

The diagonal terms are the intralayer bare Coulomb potential and the off-diagonal terms are the interlayer bare Coulomb potential. e is the elementary charge and d is the interlayer separation. Here

we report the unscreened potential for simplicity. To be more accurate, the screened potential should be used and for that the reader is referred to the cited references.[143,131,144] The main point here is the exponential decay with respect to the interlayer momentum transfer and interlayer spacing. Even with the screening correction, the exponential decay is unchanged and this dependence is what limits the range of the interlayer momentum transfer. The maximum momentum that can be transferred is:

$$q \ll \frac{1}{d}. \tag{7.8}$$

Fortunately, for van-der Waals materials, the interlayer spacing can be as thin as a couple nanometers without the adverse effect of current leakage or tunneling. This is due to the confinement of the electron wavefunction in van-der Waals materials to a single or just a few atomic layers. The older generation of 2DEGS (2D electron gas), such as GaAs heterostructures, confined electrons in an engineered quantum well via modulation doping. These electrons have wavefunctions that extended on the order of high tens of nanometers away from the 2DEG plane. This severely limited the thinness of the spacing between layers. With hBN, a typical thickness of 2 nm is used in which case the transferable momentum can be about $0.5\ \mathrm{nm}^{-1}$. This is sufficient for any electron to reach any part of the Dirac cone at all experimentally reachable gate values. For reference, the cone diameter is about $0.5\ \mathrm{nm}^{-1}$ long at 0.3 eV and the two valleys in Gr are separated by about $10\ \mathrm{nm}^{-1}$.

Looking back at equation 7.5, it tells us that rate of momentum change comes from all possible phase space transitions given by Fermi's Golden Rule and weighted by their probabilities. The probabilities are given by the distribution functions; the delta function ensures energy conservation; the Coulomb potential determines the momentum transfer range. To evaluate the equation, a linearization of the non-equilibrium distribution function is performed while assuming $\mathbf{q} = m\mathbf{v}/\hbar$. After

the linearization, we get a very transparent form of equation for the rate of momentum transfer:

$$\frac{d\mathbf{p_2}}{dt} = \frac{\mathbf{v_1}}{2}\frac{1}{k_B T}\int \frac{d\mathbf{q}}{(2\pi)^2}(\hbar q)^2|V(q)|^2 \int \frac{d\mathbf{k_1}}{(2\pi)^2}\int \frac{d\mathbf{k_2}}{(2\pi)^2}\Big\{f_0(\mathbf{k_1})\big[1-f_0(\mathbf{k_1}+\mathbf{q})\big]\Big\} \\ \Big\{f_0(\mathbf{k_2})\big[1-f_0(\mathbf{k_2}-\mathbf{q})\big]\Big\}\delta(\varepsilon_{k_1}+\varepsilon_{k_2}-\varepsilon_{k_1+q}-\varepsilon_{k_2-q}). \tag{7.9}$$

With some additional mathematical tricks, we arrive at the generalized drag formula:[143]

$$\frac{d\mathbf{p_2}}{dt} = \frac{\mathbf{v_1}}{2}\frac{1}{k_B T}\int \frac{d\mathbf{q}}{(2\pi)^2}(\hbar q)^2|V(q)|^2 \int d\omega \frac{\mathrm{Im}[\chi_1(\mathbf{q},\omega)]\mathrm{Im}[\chi_2(\mathbf{q},\omega)]}{\sinh^2(\hbar\omega/k_B T)}. \tag{7.10}$$

We have introduced the Lindhard function (or polarization function) which represents the susceptibility of the electron gas (or liquid). It is given by:

$$\chi_i(\mathbf{q},\omega) = \int \frac{d\mathbf{k_i}}{(2\pi)^2}\frac{f_0(\mathbf{k_i})-f_0(\mathbf{k_i}+\mathbf{q})}{\hbar\omega-\varepsilon(\mathbf{k_i})+\varepsilon(\mathbf{k_i}+\mathbf{q})-i\eta}. \tag{7.11}$$

The imaginary part of the Lindhard function is associated with the dissipation or absorption of incoming energy by the electron gas via electron-hole pair excitations. It is a measure of how interactive the electron gas is and is given by:

$$\mathrm{Im}[\chi_i(\mathbf{q},\omega)] = \int \frac{d\mathbf{k_i}}{(2\pi)^2}\big[f_0(\mathbf{k_i})-f_0(\mathbf{k_i}+\mathbf{q})\big]\delta(\hbar\omega-\varepsilon(\mathbf{k_i})+\varepsilon(\mathbf{k_i}+\mathbf{q})). \tag{7.12}$$

To reach equation 7.10, a very important requirement must be fulfilled and it is that the electron and hole distribution functions must be asymmetric. Without this, the parts with distribution functions would result in zero.[143] This touches on one of the most important point regarding Coulomb drag and it is: Coulomb drag is highly sensitive to the electron-hole asymmetry in the system. This will be

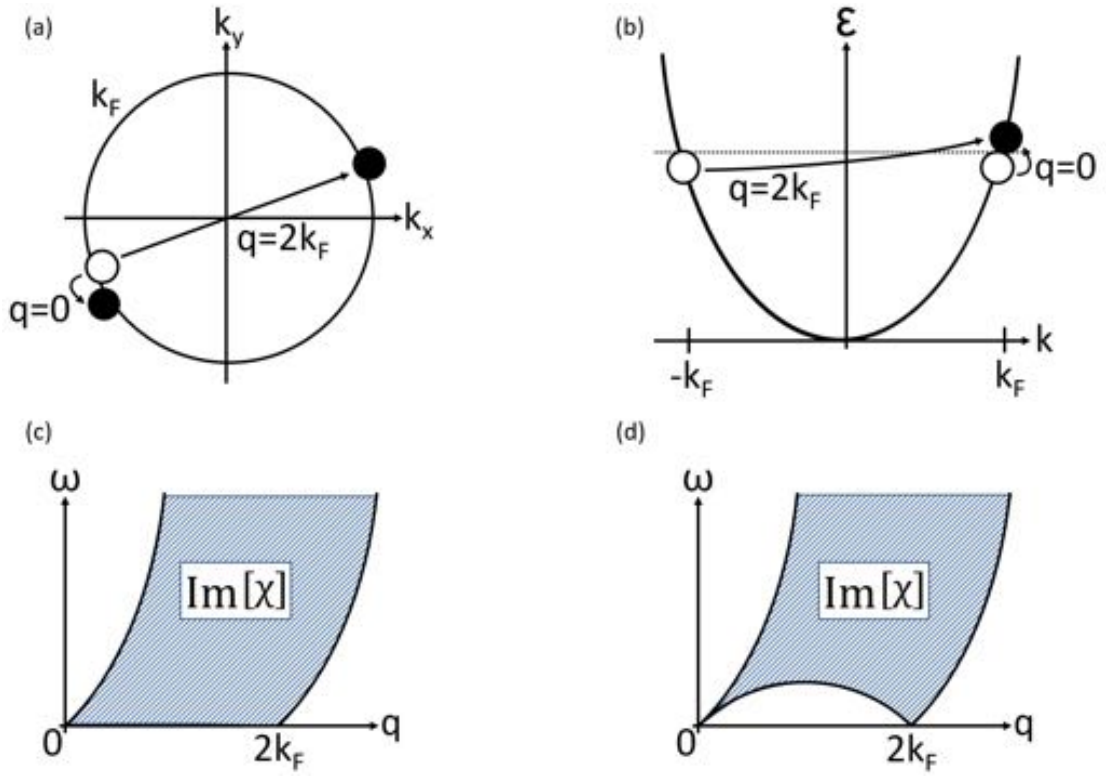


Figure 7.7: Electron-hole excitations for 1D and 2D systems. (a) and (b) are the 2D and 1D energy vs momentum space respectively with the momenta (q) associated with electron-hole excitations labeled. (c) and (d) are the 2D and 1D electron-hole excitation spectra respectively.

addressed again later. And as a last side note, the Berry phase and degeneracy information associated with Gr, if taken into account, would be embedded in the polarization function.

To compute the drag resistance, we use the fact that any measurable change in momentum in the drag layer will generate an electric field:

$$\frac{d\mathbf{p_2}}{dt} = n_2 e \mathbf{E_2}. \tag{7.13}$$

Together with equation 7.4, we get the generalized formula for drag resistivity due to momentum

drag:[143,131,145]

$$\rho_D = \frac{1}{n_1 n_2 e^2} \frac{1}{2k_B T} \int \frac{d\mathbf{q}}{(2\pi)^2} (\hbar q)^2 |V(q)|^2 \int d\omega \frac{\mathrm{Im}[\chi_1(\mathbf{q},\omega)]\mathrm{Im}[\chi_2(\mathbf{q},\omega)]}{\sinh^2(\hbar\omega/k_B T)}. \tag{7.14}$$

In the low temperature regime ($k_B T \ll \hbar\omega$) and only taking into account forward (small angle) scattering, so ignoring backscattering (large angle), which is arguably a valid assumption for most times due to the exponential decay dependence, the drag resistivity simplifies to:[143,131,145]

$$\rho_D = \frac{\hbar}{e^2} \frac{\pi^2 \zeta(3)}{16} \frac{(k_B T)^2}{\varepsilon_{F1}\varepsilon_{F2}} \frac{1}{q_{\mathrm{TF1}} q_{\mathrm{TF2}} k_{F1} k_{F2} d^4}, \tag{7.15}$$

where $\zeta(3) \approx 1.202$, q_{TFi} is the Thomas-Fermi screening wavevector, ε_{Fi} is the Fermi energy and k_{Fi} is the Fermi wavevector. Equation 7.14 and 7.15 together gives us much of the information regarding Coulomb drag. The T^2 dependence is if we ignored backscattering specifically. With backscattering in 2D, $T^2 \rightarrow T^2 \ln(\varepsilon_F / k_B T)$ but this logarithmic temperature correction is not easily noticeable experimentally.[144,146] Furthermore, backscattering in Gr is suppressed due to its Berry phase and large momentum transfer is less probable so overall we expect the temperature dependence to follow more closely with T^2. This overall T^2 dependence has the same root as the Fermi liquid e-e scattering temperature dependence. Interlayer Coulomb drag is just another manifestation of e-e scattering but instead of occurring in one layer, it is occurring in two layers. With this in mind, we expect the temperature dependence to be the first signature to go if one or both layers no longer behave as a Fermi liquid. Coulomb drag can thus be used as a probe to detect unconventional electronic regimes. For instance, for 1D-1D drag the temperature dependence is expected to be non-monotonic,[147,148,149] while for highly correlated systems such as exciton condensates the temperature dependence diverges as temperature decreases.[150,151,152]

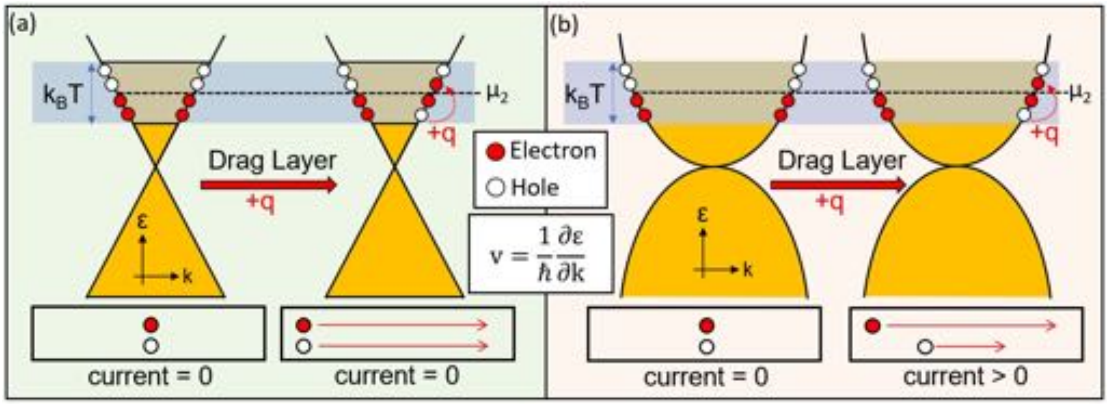


Figure 7.8: Comparison between electron-hole symmetric and electron-hole asymmetric systems. Shown are pictures depicting two different types of drag layers given momentum q. (a) shows 1D system with linear energy bands which necessarily means electron-hole symmetry around the Fermi energy for all energy values. (b) shows a 1D system with parabolic energy bands. Anywhere away from the charge neutrality point (where the bands touch), there will be electron-hole asymmetry around the Fermi energy. The curvature of the energy bands is what dictates the symmetry. A drag signal depends on the asymmetry of the system as it is required to generate an imbalance between electron and hole quasiparticles.

The drag resistance due to momentum transfer can be summarized to depend on the following factors:

1. phase space availability for scattering limited by k_BT,
2. the interaction term $V(q)$ which dictates q_{max} given d,
3. electron-hole asymmetry expressed through the asymmetric energy dependence of DOS and group velocity around the Fermi energy,
4. screening due to each layer's carrier concentrations as reflected by the $1/n$ dependence,
5. in-plane density fluctuations given by $\text{Im}[\chi(\mathbf{q}, \omega)]$.

The importance of electron and hole asymmetry was briefly touched upon previously but looking at the final form for momentum drag, the existence seems to be lost in writing. This goes back to the evaluation of the imbalance in the electron distribution functions. In moving forward from equation 7.5, there is the underlying assumption that the terms with distribution functions yield non-zero val-

ues. This can only happen if the electron and hole energy spectrum are asymmetric. This asymmetry requirement can partially be captured by figure 7.8. Specifically shown in the figure is the response of the drag layer. If electrons and holes were symmetric, the excited pairs of electrons and holes after a momentum gain would travel at the same velocity, resulting in a zero overall current. For the asymmetric case, electrons and holes would travel differently which results in an overall current. This concept is further expounded upon in the cited references.[153,154,155,156,136,143,131]

Under a magnetic field, it was shown that equation 7.14 can be rewritten in the following tensor form:[157]

$$\boldsymbol{\rho_D} \sim \boldsymbol{\rho_2} \frac{d\boldsymbol{\sigma_1}}{d(en_1)} \frac{d\boldsymbol{\sigma_2}}{d(en_2)} \boldsymbol{\rho_1}, \tag{7.16}$$

where the electron-hole symmetry can seen through the derivative of the intralayer conductivity ($\boldsymbol{\sigma_i}$) with respect to its layer carrier density (n_i). $\boldsymbol{\rho_i}$ is the resistivity tensor of layer i. This equation was also used as an accurate explanation for magneto-drag under strong magnetic fields,[158] substantiating the universality of equation 7.14 and the momentum drag description. For weak classical magnetic fields, one can obtain the magnetic field dependence through the expansion of the polarization function as well as momentum which has a B dependence. It is shown that the leading term for weak classical magnetic fields goes as B^2.[159] Indeed experiments showed that at high doping the drag signal does follow this B^2 dependence, but at the DCNP the drag signal showed an unexpected sign reversal. Instead of a positive increase in drag signal, it becomes negative very quickly.[158,138] Up to this point, momentum drag was able to explain very well all the physics seen at high doping but has come short when describing the physics at the DCNP. And this brings us to our next topic: energy drag.

7.3.2 Energy Drag

Gr Coulomb drag exhibits an anomalous positive signal at DCNP that has a non-monotonic temperature dependence and turns negative upon application of external magnetic field. This is not captured by momentum drag which predicts the signal to vanish at zero magnetic field at DCNP and to become positive under a magnetic field. The energy drag mechanism was proposed to explain the anomaly. The underlying principle behind energy drag is interlayer energy transfer. This notion is captured well by the following coupled heat transport equation:

$$-\boldsymbol{\nabla}\cdot(\kappa_1\boldsymbol{\nabla}\delta T_1)+a(\delta T_1-\delta T_2)+\lambda\delta T_1=-\boldsymbol{\nabla}\cdot\mathbf{j}_{\mathbf{1},\mathbf{q}}, \tag{7.17}$$

$$-\boldsymbol{\nabla}\cdot(\kappa_2\boldsymbol{\nabla}\delta T_2)+a(\delta T_2-\delta T_1)+\lambda\delta T_2=0. \tag{7.18}$$

Here, $\delta T(\mathbf{r})=T(\mathbf{r})-T_0$ where T_0 is the lattice temperature, which is the same as the global equilibrium temperature, and $T_i(\mathbf{r})$ is the absolute electron temperature in layer i. $\mathbf{j}_{\mathbf{1},\mathbf{q}}$ is the heat current in layer 1, the drive layer. κ_i is the thermal conductivity of layer i and is assumed to be a scalar for simplicity, λ is the electron-phonon scattering rate (or electron-lattice cooling rate), a is the interlayer energy transfer rate and can be acquired through the collision integral method as in equation 7.9, except now we integrate over energy instead. The energy relaxation rate is:

$$\begin{aligned}\frac{d\varepsilon}{dt}=\frac{2\pi}{\hbar}\int\frac{d\mathbf{q}}{(2\pi)^2}(\varepsilon(\mathbf{k}+\mathbf{q})-\varepsilon(\mathbf{k}))|V(q)|^2\\ \int\frac{d\mathbf{k_1}}{(2\pi)^2}\int\frac{d\mathbf{k_2}}{(2\pi)^2}\Big\{f_0(\mathbf{k_1})\big[1-f_0(\mathbf{k_1}+\mathbf{q})\big]\Big\}\\ \Big\{f_0(\mathbf{k_2})\big[1-f_0(\mathbf{k_2}-\mathbf{q})\big]\Big\}\delta(\varepsilon_{k_1}+\varepsilon_{k_2}-\varepsilon_{k_1+q}-\varepsilon_{k_2-q}).\end{aligned} \tag{7.19}$$

Like previously, with similar mathematical tricks, we get a form similar to the generalized momentum change equation (equation 7.10) which is:

$$\frac{d\varepsilon}{dt} = \frac{\pi}{\hbar}\int d\omega(N_2(\omega) - N_1(\omega))(\hbar\omega)\int \frac{d\mathbf{q}}{(2\pi)^2}|V(q)|^2 \mathrm{Im}[\chi_1(\mathbf{q},\omega)]\mathrm{Im}[\chi_2(\mathbf{q},\omega)], \tag{7.20}$$

where

$$N_i(\omega) = \frac{1}{e^{\hbar\omega/k_B T_i} - 1}, \tag{7.21}$$

is the Bose function at temperature of layer i. a, the interlayer energy transfer rate, is then equal to $d^2\varepsilon/dtdT$.[160] The main purpose of writing out the energy transfer equation is to show that energy and momentum drag bear a lot of similarities, the most important one being the Coulomb interaction mediated interlayer scattering. This is to distinguish energy drag from the pure thermoelectric phenomenon. Energy drag hinges on the presence of Coulomb interaction which is what provides a pathway for interlayer scattering. But as we shall see, the thermoelectric effect is also central to the energy drag description.

We begin with the coupled charge and energy transport equations describing a single isolated layer:[161,162]

$$\begin{bmatrix} -\mathbf{j} \\ \mathbf{j_q} \end{bmatrix} = \begin{bmatrix} e\mathbf{L_{11}}/T & e\mathbf{L_{12}} \\ \mathbf{L_{21}}/T & \mathbf{L_{22}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\nabla}\mu \\ \boldsymbol{\nabla}\frac{1}{T} \end{bmatrix}, \tag{7.22}$$

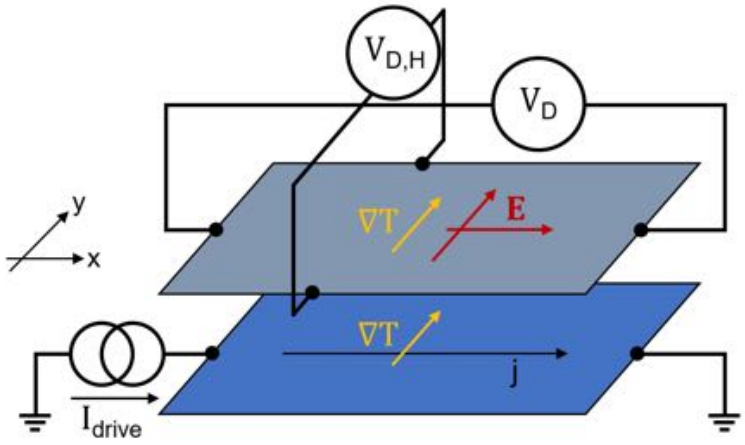


Figure 7.9: Illustration of energy drag. An electric current is driven into the drive layer. This electric current is accompanied with a heat current because of heat generation. The heat current causes a temperature gradient which, through Coulomb interaction, is duplicated in the drag layer. The temperature gradient in the drag layer in turn creates an electric field and hence a drag voltage. The drag voltage can develop in directions irrespective of the drive current or heat current.

In this notation, we have the following quantities for the kinetic coefficients, a.k.a., $\mathbf{L}$ matrix:

$$\mathbf{L_{11}} = \frac{T}{e^2}\boldsymbol{\sigma}, \tag{7.23}$$

$$\mathbf{L_{12}} = e\mathbf{Q}\mathbf{L_{11}}, \tag{7.24}$$

$$\mathbf{L_{22}} \approx T^2\boldsymbol{\kappa}, \tag{7.25}$$

where each $\mathbf{L}$ element is a tensor and from these elements we can construct the following observables:

$$\mathbf{S} = \frac{1}{eT}\mathbf{L_{12}}\mathbf{L_{11}^{-1}}, \tag{7.26}$$

$$\mathbf{Q} = \frac{1}{e}\mathbf{L_{12}}\mathbf{L_{11}^{-1}} = \frac{\mathbf{s}T}{e} = \mathbf{S}T, \tag{7.27}$$

$$\mathbf{j_q} = \mathbf{Q}\mathbf{j}. \tag{7.28}$$

Note that here $\mathbf{Q}$ is the heat per charge which we will refer to as heat hereinafter and $\mathbf{S}$ is the Seebeck coefficient (or thermopower), not entropy $\mathbf{s}$. The Seebeck coefficient ($\mathbf{S}$), heat ($\mathbf{Q}$), entropy ($\mathbf{s}$), electrical conductivity ($\boldsymbol{\sigma}$) and thermal conducitivity ($\boldsymbol{\kappa}$) are rank 2 tensors associated with the x, y directions in a single layer. $\mathbf{j_q}$ is the heat flux or heat current and $\mathbf{j}$ is the electric current, both of which are vectors.

Evoking Onsager reciprocity as well as assuming an isotropic layer, the off-diagonal kinetic coefficient has to obey:

$$\mathbf{L_{12}}(B) = \mathbf{L_{21}}^{\mathsf{T}}(B). \tag{7.29}$$

Then, because the drag layer is measured in an open circuit where $\mathbf{j} = 0$, from equation 7.22 we can solve for the resultant electric field:

$$\begin{aligned} \mathbf{E} &= \frac{\boldsymbol{\nabla}\mu}{e} = \frac{T}{e}\mathbf{L_{11}^{-1}L_{12}}\boldsymbol{\nabla}\frac{1}{T} \\ &= \frac{1}{T}\mathbf{L_{11}^{-1}QL_{11}}\boldsymbol{\nabla}T \\ &= \mathbf{Q}\frac{\boldsymbol{\nabla}T}{T}, \end{aligned} \tag{7.30}$$

where in the last step $[\mathbf{Q}, \mathbf{L}] = 0$, because of the layer isotropy assumption.[162] What we have arrived at up to this point is that for a single isolated layer, if there exists a temperature gradient within the layer, an electric field will develop, and this field is not restricted to just the direction of the temperature gradient which is why a longitudinal and transverse (Hall) drag signal can be measured. Now the question is, how can a temperature gradient form in the drag system? A somewhat phenomenological scenario can be used and that is through balancing the heat equation while assuming the double layer

system has strong interlayer coupling such that individual thermal conductivities add:

$$(\kappa_1 + \kappa_2)\boldsymbol{\nabla}\delta T = \mathbf{D}\mathbf{Q_1}\mathbf{j_1}. \tag{7.31}$$

$\mathbf{D}$ is like a sample specific form factor that determines the heat flow pattern for the double layer system, hence characterizing both layers at the same time. For isotropic heat flow, $\mathbf{D} = \mathbf{I_2}$, where $\mathbf{I_2}$ is the 2×2 identity matrix.[162,159] Solving for the above will net the temperature gradient of the whole double layer system. A more precise temperature gradient can be solved for from equations 7.17 and from this one can acquire a relationship specifically between $\boldsymbol{\nabla}\delta T_2$ and $\mathbf{j_{1,q}}$. The expression is boundary condition specific and the procedure is detailed in the cited reference.[162] What remains now is to calculate the drag resistance. Putting equations 7.28, 7.30, 7.31 all together and evoking definition of drag resistivity, one will find for energy drag:

$$\boldsymbol{\rho_D} = \frac{\mathbf{Q_2}\mathbf{D}\mathbf{Q_1}}{T(\kappa_1 + \kappa_2)}. \tag{7.32}$$

The drag resistivity depends largely on the heat tensor $\mathbf{Q}$. As a reminder, $\mathbf{j_q} = \mathbf{Q}\mathbf{j}$, so $\mathbf{Q}$ dictates the angle relation as well as magnitude relation between the electric and heat current. Figure 7.10 shows the behavior of Q for Gr. In the most ideal case where the heat flow is isotropic the drag resistance simplifies to:[162]

$$R_{D,H} \sim \frac{1}{T(\kappa_1 + \kappa_2)} Q_{1,xy}Q_{2,xx}, \quad R_D \sim \frac{L}{WT(\kappa_1 + \kappa_2)} Q_{1,xy}Q_{2,xy}, \tag{7.33}$$

where L and W are the length and width of the system. This equation is shown to help the reader digest roughly what to expect for the energy drag signal.

One striking behavior from the energy drag resistivity is this: for isotropic heat flow $\mathbf{D} = \mathbf{I_2}$, in

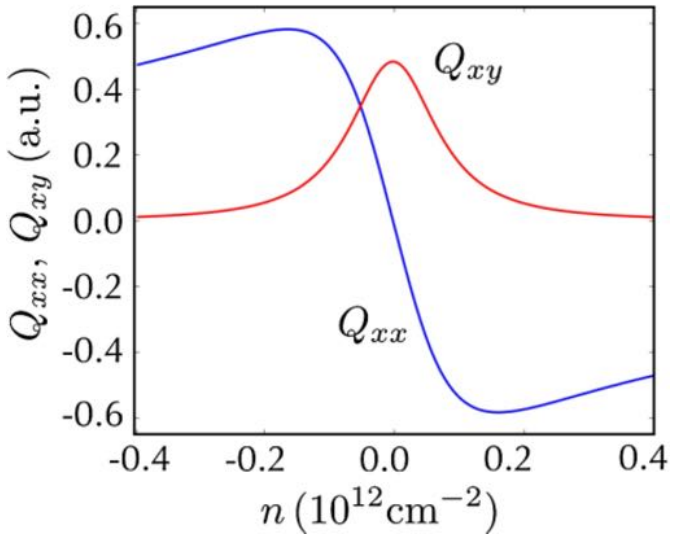


Figure 7.10: The heat characteristic of graphene. The plot is taken from the work by Song et al.[162]

which case, $[\mathbf{Q_1}, \mathbf{Q_2}] = 0$, and so $\boldsymbol{\rho_D}$ becomes layer symmetric, i.e., swapping drive (1) and drag (2) layers results in the same behavior; however, for anisotropic heat flow $\mathbf{D} \neq \mathbf{I_2}$, $\boldsymbol{\rho_D}$ becomes layer asymmetric, i.e., $[\mathbf{Q_1}, \mathbf{DQ_2}] \neq 0$ This is an important result with regards to drag. It hast been widely assumed that Onsager reciprocity implies layer symmetry. This is unfounded and completely false. In the next section, we will discuss about Onsager's reciprocal relations and see how it extends to the drag measurement.

Regarding the anomaly seen at the DCNP, let us first stress that energy drag as described so far does not enforce any non-zero signal at the DCNP. In fact, based on what is described so far we would expect the energy drag signal to also vanish near DCNP at zero magnetic field (mainly due to κ. See figure 7.11). In order to explain the anomaly both energy drag and disorder must be taken into account. When a charge current is applied to the drive layer, the intrinsic density inhomogeneity due to disorder will cause a spatially varying heating or cooling pattern. This is because the current going through

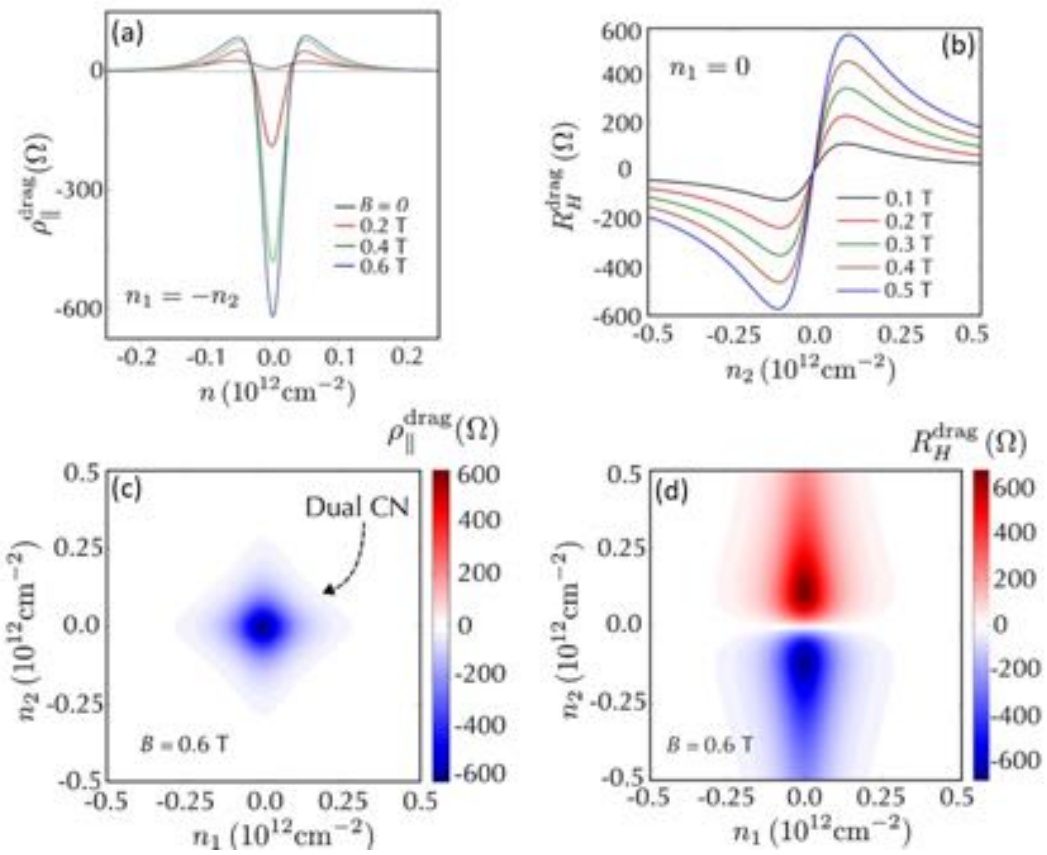


Figure 7.11: Energy drag transport characteristics. (a) is the experimental measurement of longitudinal drag in a Gr-Gr drag system at different magnetic fields, which can be reproduced well by theory of energy drag (see cited reference for comparison). Here we can see the drag signal becoming more negative as a function of increasing magnetic field, which is a key signature of energy drag. (b) is the theoretical simulation of Hall (transverse) drag in a Gr-Gr drag system at different magnetic fields. The overall magnitude increases with magnetic field. (c) and (d) are the respective simulated color plots, plotted with respect to the layer carrier densities and at 0.6 T. All plots are either taken or simulated at 150 K. All plots are adapted from the work by Song et al.[162]

the disordered layer will not be homogeneous. Instead, the microscopic direction of the current is dictated by the spatially varying electronic landscape. This non-uniform current is responsible for creating a spatially varying heating or cooling pattern, resulting in a temperature pattern $\delta T(\mathbf{r})$. If the thermal coupling between the drive and drag layer is strong, meaning that energy transfer rate between layers is fast, the drag layer will then efficiently duplicate the temperature pattern in the drive

layer: $\delta T_{\text{drag}}(\mathbf{r}) \approx \delta T_{\text{drive}}(\mathbf{r})$, after which, the thermoelectric effects will carry through and generate the anomalous signal. Through the same energy drag formalism detailed above, it is shown that the energy drag resistivity is given by:

$$\rho_D = \frac{1}{2T(\kappa_1 + \kappa_2)} \frac{\partial Q_1}{\partial \mu_1} \frac{\partial Q_2}{\partial \mu_2} \sum_q \frac{\langle \delta\mu_2(-\mathbf{q}) \delta\mu_1(\mathbf{q}) \rangle}{1 + lq^2}, \tag{7.34}$$

where l denotes the interlayer cooling length which can be given by the e-e scattering length and q is the momentum. Looking at figure 7.10 it is clear that the signal peaks at zero-doping since the derivative of Q is at a maximum. The most important quantity here is the correlator $\langle \delta\mu_2(-\mathbf{q}) \delta\mu_1(\mathbf{q}) \rangle$. This quantity describes how aligned the density inhomogeneity is between the drive and drag layers. The sign of the drag signal will depend entirely on whether the density inhomogeneity is co-aligned or anti-aligned. For a positive correlation, the energy drag signal is due to a disorder potential dominated by charge impurities. For a negative correlation, the energy drag signal is due to strain induced charge puddles. The anomalous drag signal at DCNP has been termed remnant drag. Based on experiments thus far, the anomalous drag signal is likely due to charge impurities.[160]

The magnetic field dependence of the energy drag resistivity is encoded in the heat tensor. To see this simply, in the diffusive case according to Mott formula and equation 7.27, we have:

$$\mathbf{Q} = \frac{\pi^2}{3e} (k_B T)^2 \boldsymbol{\sigma}^{-1} \frac{\partial \boldsymbol{\sigma}}{\partial \mu}. \tag{7.35}$$

And so the energy drag acquires its magnetic field dependence through the conductivity tensor ($\boldsymbol{\sigma}$). Figure 7.10 shows the behavior of $\mathbf{Q}$. Energy drag is predicted to have a $-B^2$ dependence. Compared to the momentum drag magnetic field dependence, they both scale quadratically but energy drag is expected to have a much larger prefactor in addition to it being negative. This is the reason for the

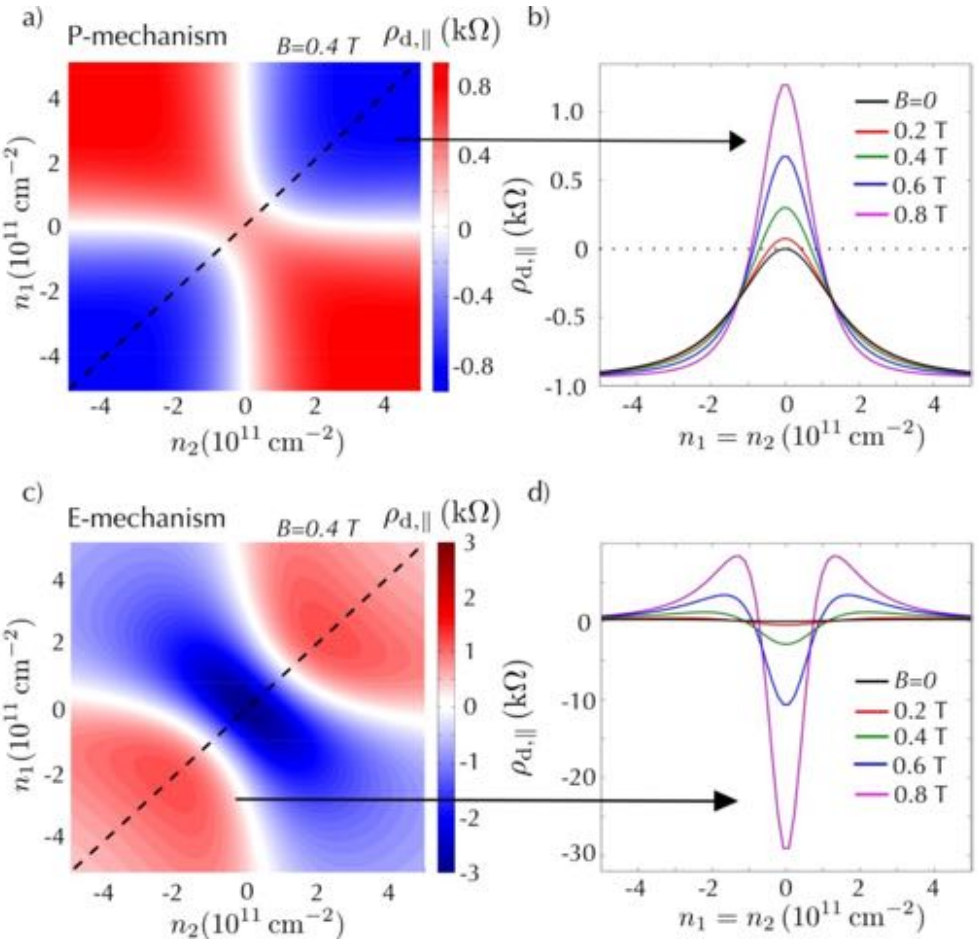


Figure 7.12: Comparison between momentum and energy drag at non-zero magnetic field. All plots are taken from the work by Song et al.[159]

observed anomaly in magnetic field dependence.[159] Figure 7.12 compares and contrasts the magnetic field dependence of momentum and energy drag. This will be the key identifier in determining the drag mechanism moving forward. Figure 7.13 shows the temperature dependence of energy drag at the DCNP. It shows a peculiar non-monotonic behavior. This is due to the competing terms among all the factors within the equation and results in a T^2 at low temperature and T^{-4} at high temperature with the transition point Δ given approximately by the Dirac peak disorder broadening energy.[160]

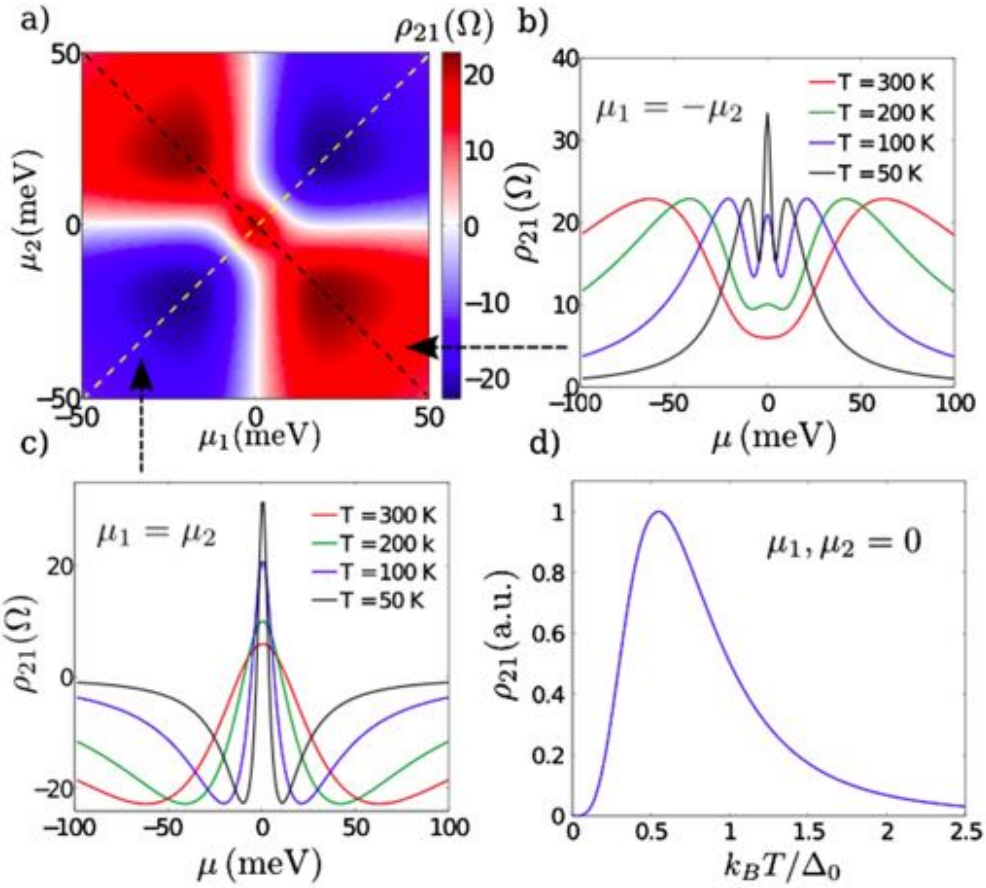


Figure 7.13: Temperature dependence of energy drag. (a) is the total drag signal, i.e. momentum and energy drag combined. (b) and (c) show the line cuts at same and opposite layer doping. For both of them we see the anomalous positive signal becoming more prominent as a function of decreasing temperature. (d) is the temperature dependence of this anomalous drag peak. All the plots are taken from the work by Song et al.[160]

7.4 Onsager Reciprocal Relations

In Onsager's seminal work on the reciprocity of different thermodynamic quantities, he proved a general and powerful set of reciprocal relations treating the symmetry in the mutual interference of two or more irreversible processes occurring simultaneously in a system. Onsager reciprocal relations are satisfied only when the flows (or fluxes) are expressed as a function of their respective conjugate ther-

modynamic forces. The relations are based on the fundamental principle of time-reversal symmetry, which is to say, "if velocities of particles are reversed simultaneously, the particles will retrace their former paths, reversing the entire succession of configurations."[163]

If an irreversible process expressed as rates of change of a set of measurable parameters α is able to be expressed in a linear form:

$$\frac{\delta \alpha_i}{\delta t} = \sum_j L_{ij} \gamma_j, \tag{7.36}$$

where

$$\frac{\delta s}{\delta t} = \sum_j \gamma_j \frac{\delta \alpha_j}{\delta t}, \tag{7.37}$$

then in equilibrium, Onsager's theorem states:

$$L_{ij}(B) = L_{ji}(-B). \tag{7.38}$$

γ causes the deviation from equilibrium state and is called the generalized force. In equilibrium the entropy s is at a extremum and hence γ is about zero. The reason the process is irreversible is $\delta s / \delta t \geq 0$. L are the kinetic coefficients which define the state of the system and L_{ij} is the measure of the interference of the i and j process. To summarize in point format, Onsager's reciprocity only works under the following conditions:

1) the process must be expressible in linear form and the state must be in equilibrium,

2) only the forces γ_j conjugate to α_j can be used and must satisfy relation 7.37.

The above underlines the ultimate conditions for Onsager's reciprocal relations to work. Many scenarios described by different pairs of conjugate variables (α, γ) have been individually investigated over the many years.[163,164,161,165,166] In the system regarding energy drag, the interlayer mutual interference

involves the charge and heat currents. Onsager's reciprocity would hold if we expressed the state of the system as the coupled heat and particle transport equation:

$$\begin{bmatrix} -\mathbf{j} \\ \mathbf{j_q} \end{bmatrix} = \begin{bmatrix} L_{11}/T & L_{12} \\ L_{21}/T & L_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{\nabla}\mu \\ \boldsymbol{\nabla}\frac{1}{T} \end{bmatrix}, \tag{7.39}$$

Here, the fluxes and forces are all adequately related through equation 7.37 and Onsager's theorem requires $L_{12}(B) = L_{21}(-B)$.[161] In another well investigated scenario, regarding the conduction of electricity, if the system can be described by:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}, \tag{7.40}$$

then $R_{12}(B) = R_{21}(-B)$.[166] In this system, the indices refer to different probes regardless of the geometry and separation. We can draw a connection between this and momentum drag, whereby the indices can refer to the drive and drag layers and Onsager's reciprocity would work as expected– we would expect layer symmetry in momentum drag measurements.

One must be wary when rendering the energy drag system in terms of the drive and drag layers. Attempting to describe the process like momentum drag would simply coarse grain the intricate details of energy drag. To see an example, let us work with a process that includes both an electrostatic potential and temperature gradient as the generalized forces occuring in multiple layers. Written in terms of interlayer conductivity where now the indices denote different layers, we have:

$$\begin{bmatrix} j_1 \\ j_2 \end{bmatrix} = - \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \begin{bmatrix} (\boldsymbol{\nabla}V + S\boldsymbol{\nabla}T)_1 \\ (\boldsymbol{\nabla}V + S\boldsymbol{\nabla}T)_2 \end{bmatrix}, \tag{7.41}$$

where S is the Seebeck coefficient assumed to be layer independent for simplicity. In this case, Onsager reciprocity is not clear. Despite the equation being in linear form, the generalized force is now composed of electrostatic potential and temperature gradients. The conjugate variable cannot be generalized into a single electric current. This is to say, in the above, equation 7.37 is not satisfied. In this case, we do not expect layer symmetry to hold (i.e. $\sigma_{12} \neq \sigma_{21}$) unless the temperature gradient vanishes. This idea extends to the description of energy drag where both temperature and electrostatics play a central role. A more careful approach must be taken to describe energy drag as there are many more forces at play and one must not mistake the lack of layer symmetry as being a violation of Onsager's reciprocal relations. This notion is emphasized in the work by Song et. al.[162]

As the concluding remark: Onsager reciprocity predicates on being able to write the process strictly in linear form where α and γ are conjugate variables related through equation 7.37. Regarding energy drag, Onsager reciprocity holds, but it is not manifested in terms of layer symmetry in contrast to momentum drag.

7.5 Parasitic Effects and the Drag Measurement Scheme

From the very beginning when Coulomb drag measurements were first performed, it was well-known that parasitic effects can easily mask the true drag signal if the measurements were not done right.[133,135] One of the main parasitic effects come from capacitive coupling between the drive and drag layers. An excitation bias in the drive layer will cause the drive layer to acquire a non-zero potential due to the contact resistances and its intrinsic resistance. For DC measurements, this potential will electrically gate the drag layer, creating a non-uniform electronic landscape if the layers are asymmetric like the CNT-Gr drag devices. For AC measurements, the problem is exacerbated. An oscillating potential in the drive layer will send an oscillating current into the drag layer through capacitive response. This

current will cause a measurable unwanted potential drop across the drag layer due to the layer and contact resistances. See figure 7.5 for the circuit explanation. This interlayer bias problem is often referred to as the common mode problem.

It is undeniable that low AC techniques offer huge benefits in terms of low noise and fast measurement rates. To take advantage of low AC techniques while circumventing the parasitic effects due to capacitive coupling, people have performed the measurements using resistance bridges to manually tune the potential of the drive and drag layer to be about the same as seen in figure 7.5. The problem with this is as the carrier densities in the layers change, the layer resistance as well as contact resistance change. This means that the interlayer potential difference will also change as the gates change. At the time of this writing, potential tuning is done by manually by hand. It is unrealistic to tune continuously as the measurement is running. One may conceivably develop a feedback system that can constantly fine tune the interlayer potential difference to be zero.

An alternative method is simply to perform slow DC measurements and this will solve the capacitive coupling problem. Performing DC measurements, however, creates problems of its own. Slow measurements are subject to thermoelectric signals flowing through the measurement system and the DC measurements are prone to slow thermal drifts that affect instruments references. And of course, DC measurements are not immune to any noise at any frequency that AC measurements can readily get rid of and it is particularly difficult to measure low amplitude signals.

For our drag measurements, we always start with an AC measurement but we almost always are hindered by a large capacitive phase in our drag signal regardless of how we try to tune the interlayer bias. This is not a surprise because the contact resistance to the CNT is at best (hundreds of kOhms) still at least three orders of magnitude larger than Gr's. What Gr Coulomb drag can get away with is not the same for CNTs. This means that CNT measurements are extremely prone to interlayer bias

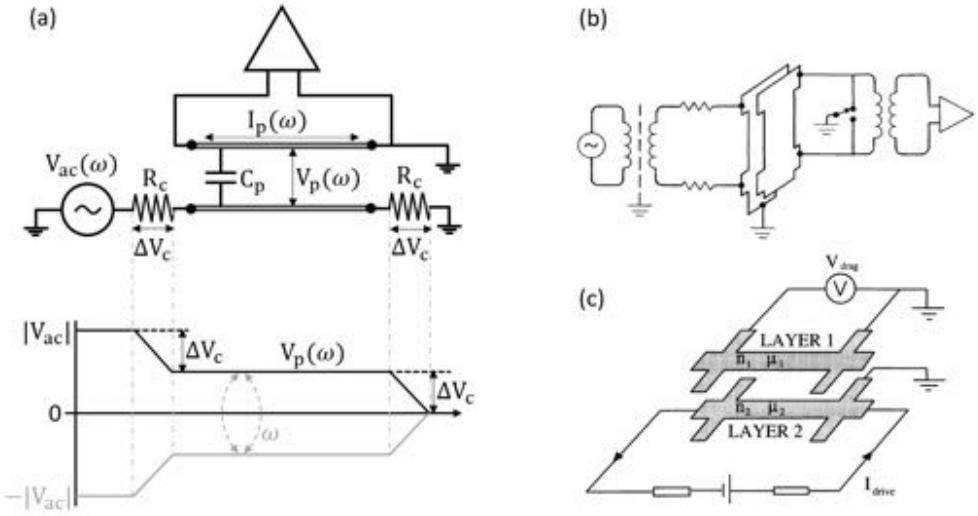


Figure 7.14: Parasitic signal is easily detectable in a drag measurement. Measurements must be done properly to avoid it. (a) shows the cause of a very common parasitic signal, commonly called the common-mode problem. It arises because of the interlayer bias generated by the high contact resistance at the contacts. As a result of the interlayer bias (V_p) and parasitic capacitance (C_p), the AC bias (V_{ac}) will transmit an AC current (I_p) into the drag layer which is then detected as a parasitic voltage because of the resistances in the drag layer. The picture below shows roughly the potential profile across the device in the drive layer. As one can see, the potential at the sample is not zero which is the cause of the interlayer bias. (b) shows the AC measurement scheme where the resistance bridge is used to tune the interlayer potential difference and the transformer at the bias is used to create a more symmetric type bias. Picture adapted from the work by Gramila et al.[133](c) shows the DC measurement scheme were the a floating DC current is driven into the drive layer. The reference for the drive layer was defined by a non-current carrying ohmic contact. DC measurements are performed to prevent any capacitive cross talk between layers. Picture adapated from the work by Hill et al.[135]

effects. We then resort to performing all our measurements with extremely slow symmetric DC bias (see figure 7.5). A typical fast sweep in bias takes 5 or more seconds. We also choose to perform DC measurements because we can carefully analyze the I-V characteristics of each data point, which is a detail lost in AC measurements and a mistake many people too willingly commit. By analyzing the I-V characteristic of each bias sweep, we can carefully see whether we are in the linear response regime, thus increasing the confidence in our measurements. To reduce noise, we do many iterations of the same sweep and then by performing an average. The methodology is: we always start by performing

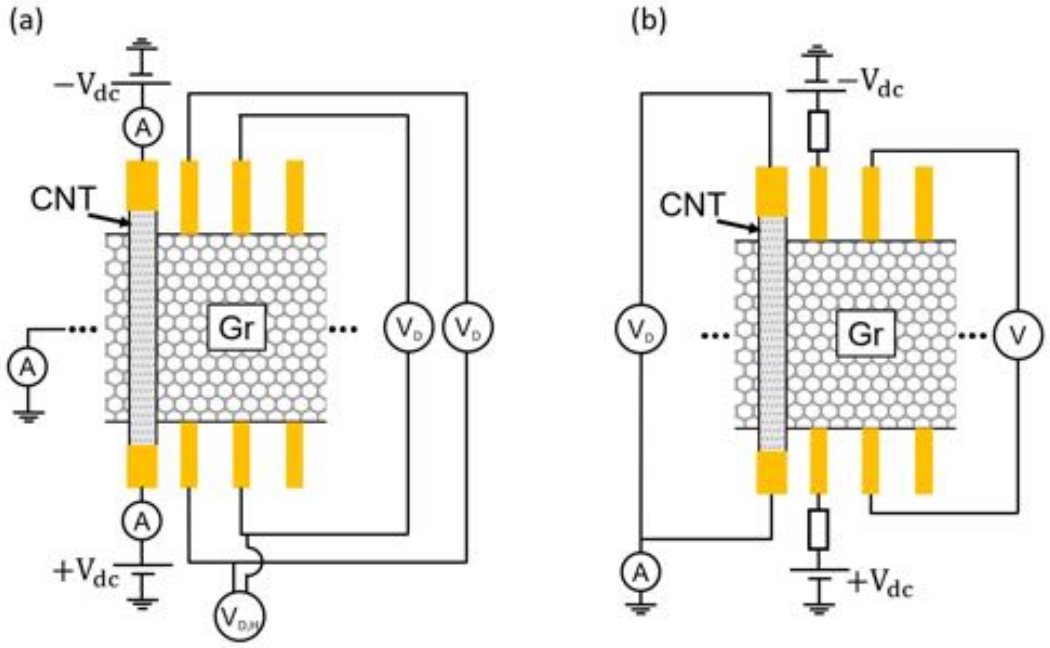


Figure 7.15: Drag device measurement scheme. (a) CNT is the drive layer and Gr is the drag layer. A DC voltage is applied symmetrically on the two ends of the CNT and the resulting current is monitored. Voltage bias is chosen because of the high contact resistance of the CNT. (b) Gr is the drive layer and CNT is the drag layer. A DC current is injected symmetrically on the two ends of the Gr. This is possible because of Gr's low resistance. The drag voltage is measured for both configurations. An ammeter to ground in the drag layer can be used to detect any parasitic tunneling current between the drive and drag layer.

an extremely slow and high resolution DC measurement with an extremely high number of iterations per data point. This is what we consider as our high-confidence measurement. A measurement that typically takes an AC measurement 10 seconds, our high confidence measurement will take over half a day. We then examine the effects of decreasing iterations and increasing bias sweep speed as well as amplitude to see if the high confidence measurement can be reproduced. Eventually we settle on a set of sweep parameters that optimizes speed, accuracy and resolution.

In our measurements, because of the generally small amplitude of our signal, it is tempting to drive large currents into the device for larger signal. This, however, must be done cautiously because large

currents can generate heating power that creates a parasitic temperature gradient in the drag layer that causes a thermoelectric voltage. This type of signal is difficult to distinguish from the energy drag signal which is a thermoelectric effect as well but has an origin in Coulomb interaction. As above, a precautionary measure is taken by examining the I-V characteristics and making sure joule heating (which shows up quadratically) does not come into play. However, even in the linear response regime the thermoelectric signal due to heating is difficult to identify and in the following section we will address this problem separately.

One final type of spurious signal is the mesoscopic fluctuations which seems to only occur at low temperature. The origin of these meoscopic fluctuations is shown to be due to quantum interference effects. It is a behavior seen more prominently in Gr Coulomb drag.[138,167,168,169] For Gr Coulomb drag, these fluctuations typically appear below $\sim$ 50 K. In our CNT-Gr drag measurement, since the main motivation is to investigate 1D-2D drag and with the hopes of being in the hydrodynamic regime, we will be focusing on higher temperatures ($>$ 100 K) to avoid the formation of CNT quantum dots so we are safe from these mesoscopic fluctuations.

7.6 Preliminary Results

As a preamble we remind the reader the that the device is composed of a 1D carbon nanotube in close proximity to a 2D graphene. This is a highly anisotropic device and if in the energy drag regime, we would expect a highly layer asymmetric behavior.

We first focus on the temperature dependence of the longitudinal Gr drag signal by driving the CNT as shown in figure 7.17. At high temperatures we see a very subtle bipolar behavior about the Dirac point reminiscent to the typical drag signal. However starting at about 230 K we begin to see a hump developing in the middle while the signal becomes larger in magnitude and more negative. Very

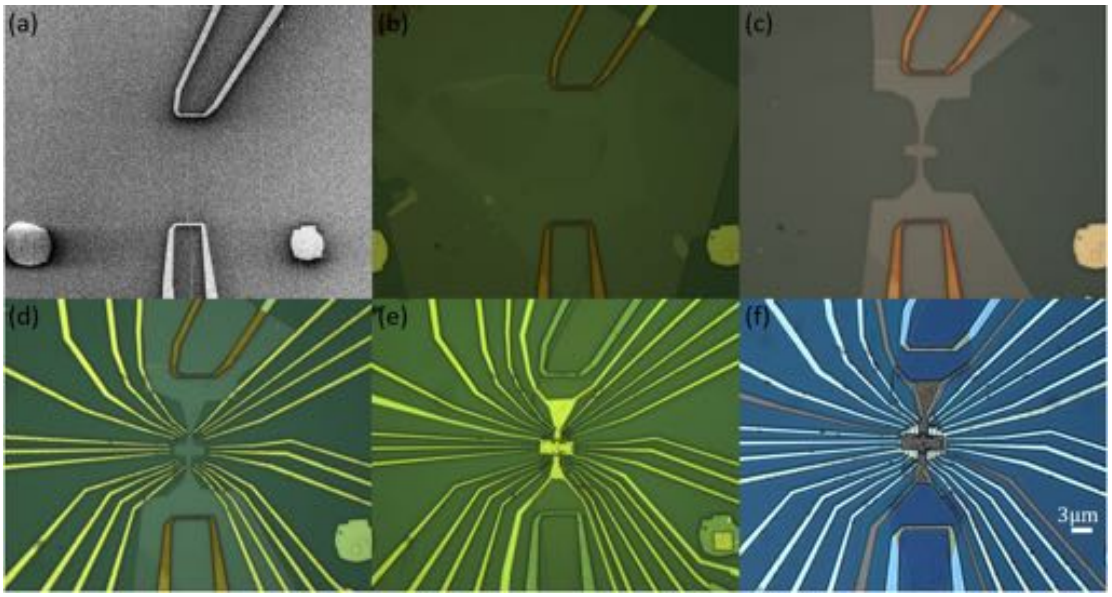


Figure 7.16: Device fabrication of the hybrid CNT-Gr drag device. (a) a Rayleigh characterized CNT is transferred onto a 1μm thick SiO_2 substrate and contacted using bridge leads. Bridge leads allow for the possibility of current annealing the contacts. (b) a hBN-Gr Stack is transferred onto the CNT with the thin hBN facing down. Typical hBN thicknesses range from 2 to 4 nm. (c) the device is etched to have Gr contacts sticking out. (d) metallic contacts are made onto the Gr contacts. (e) 80 nm thick Al_2O_3 is deposited via ALD and a top gate is deposited. (f) 40 nm thick Al_2O_3 is deposited again and the contact gates are deposited. Contact and ALD recipes are detailed in appendix F.

quickly, the hump becomes the main feature in an overall negative signal. Figure 7.18 shows a comparison between CNT drive and Gr drive signals. At 180K we see that layer reciprocity is grossly violated although it seems that the two signals are tending towards each other by 250K suggesting that at even higher temperatures there is a possibility of perfect layer reciprocity. The longitudinal drag results thus far suggests that there is a thermoelectric contribution that overtakes the device measurement due to the layer asymmetry.

The thermoelectric possibility is further substantiated if we look at the Hall drag measurement (figure 7.17b). As a reminder, here Hall drag is equivalent to a transverse drag measurement and does not require the application of a magnetic field (see figure 7.5) for the Hall drag measurement scheme).

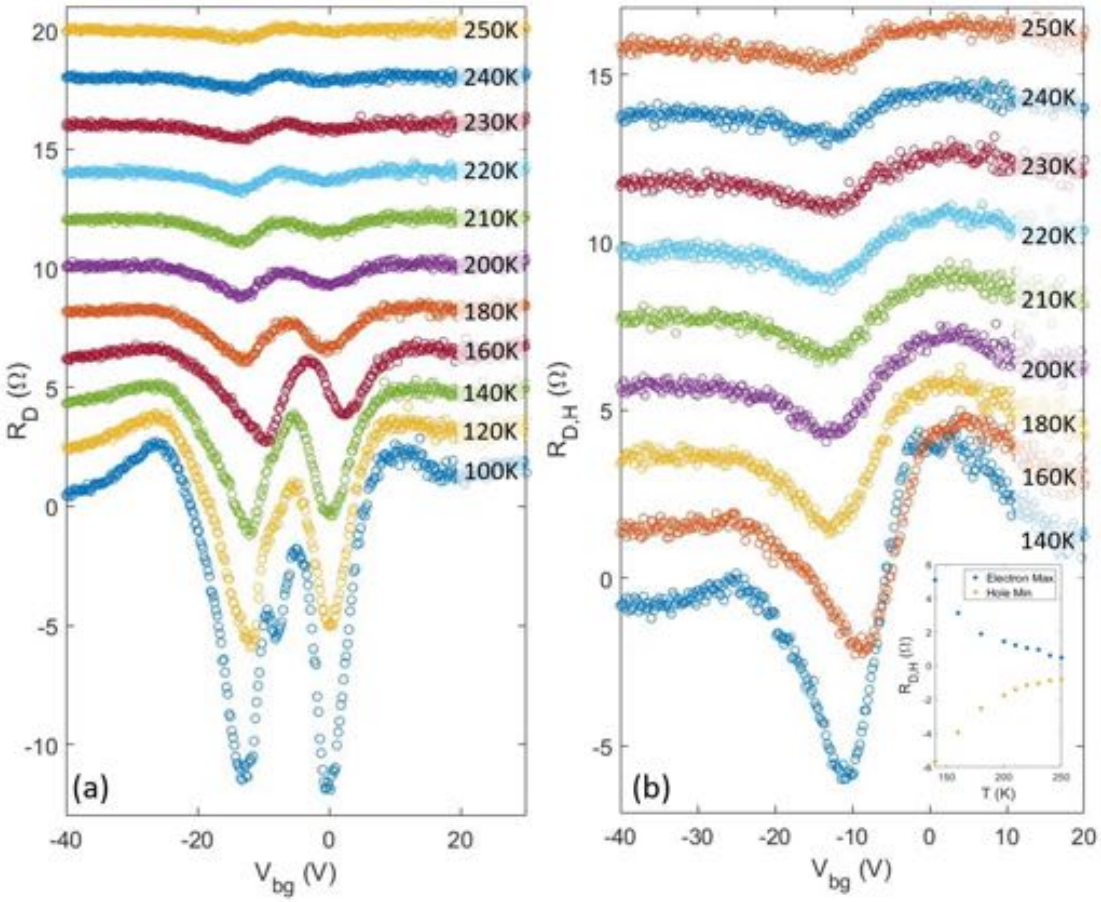


Figure 7.17: Coulomb drag in a larger hybrid CNT-Gr device. (a) is the longitudinal drag measurements while driving the CNT at different temperatures. (b) is the hall (transverse) measurment while driving the CNT at different temperatures. The inset shows the signal amplitude of the dips and peaks of the hall signal as a function of temperature. The charge neutrality point for the two plots are at $V_{\mathrm{bg}} = \sim 5$ V.

First, the existence of a hall (transverse) signal is most likely due to the voltage developed caused by a temperature gradient created perpendicular to the CNT. This signal increases in magnitude with decreasing temperature, reminiscent of the energy drag behavior.

The critical question that needs to be addressed here is whether this thermoelectric-like signal is a consequence of energy drag (where Coulomb interaction plays a central role) or simply just because of the temperature gradient caused by the heating power of the CNT. To identify this, a measurement

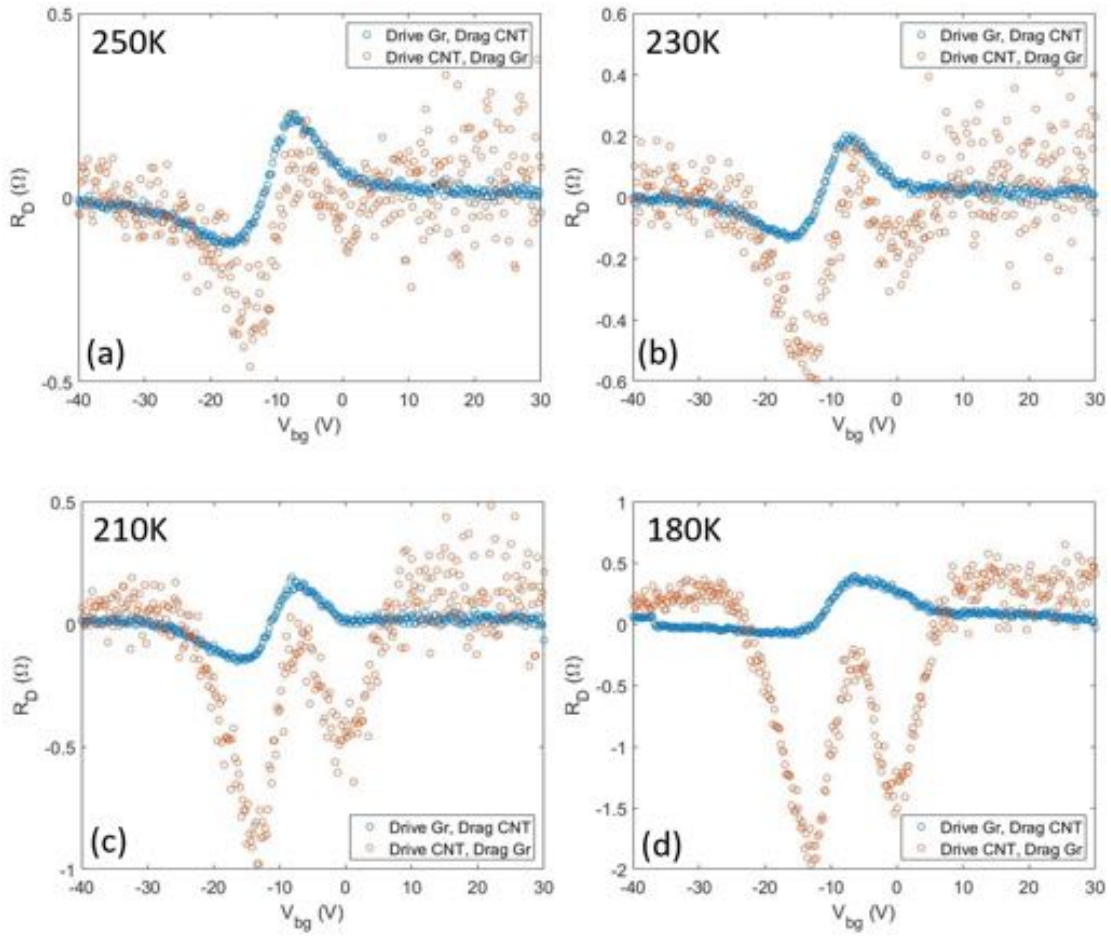


Figure 7.18: Measurements showing the violation of layer reciprocity when measuring drag in the hybrid device. The two measurements compared are the CNT drive and Gr drive drag signal. (a) through (d) show temperatures at 250, 230, 210 and 180 K respectively.

of magnetic field dependence may be enlightening.

A look at a particular sample which was fabricated to be much smaller and with leads that are denser and larger (figure 7.19e), shows that the onset of energy drag mechanism can be delayed. As seen in figure 7.19, the bipolar behavior exists from high temperatures to as low as 160 K below which the middle hump feature, as seen in figure 7.19, begins to intrude the signal. Figure 7.20 shows the comparison between the CNT and Gr drive signals and it shows that layer reciprocity is held well to

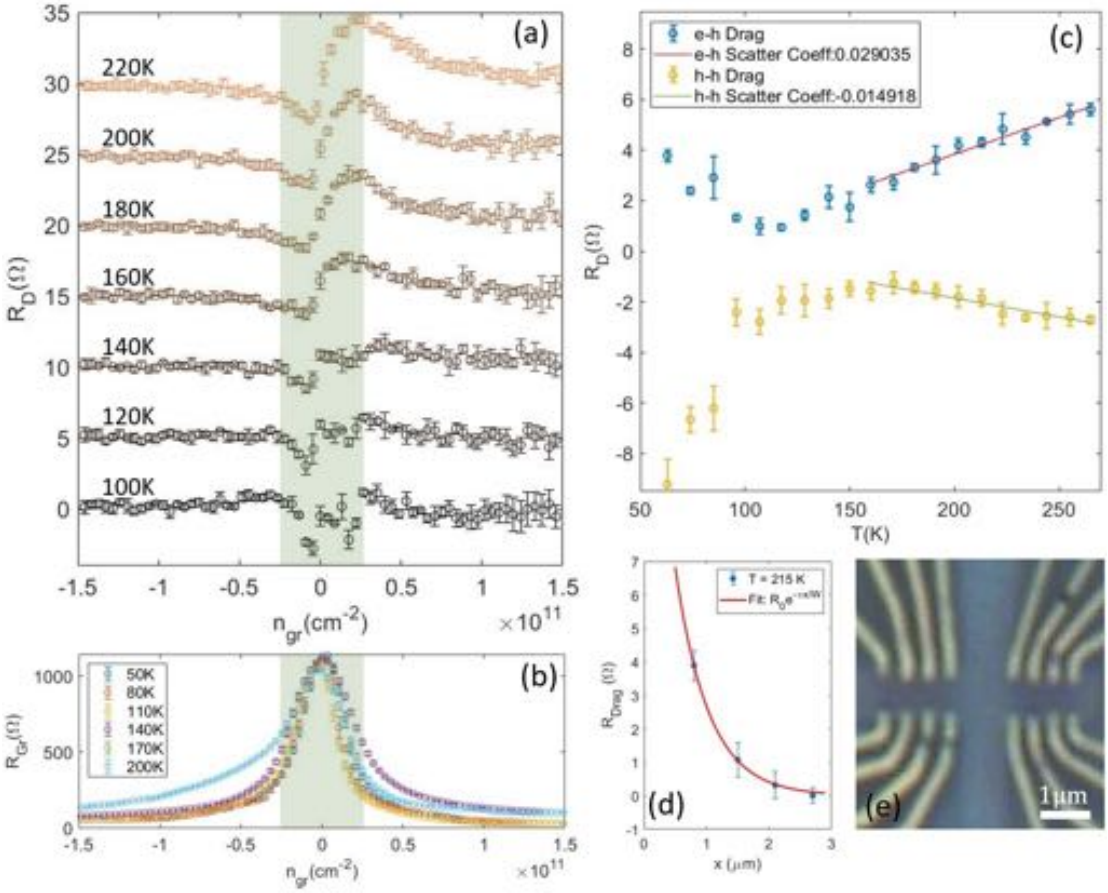


Figure 7.19: Coulomb drag in a smaller hybrid CNT-Gr device. (a) is the longitudinal drag measurements while driving the CNT at different temperatures. (b) is the Dirac peak of Gr. The shaded green region indicates the FWHM where electron-holes coexist. (c) is the signals' dips and peaks plotted at different temperatures. (d) is the spatial dependence of the drag signal fitted with a decaying exponential. (e) is the small device picture.

as low as 160 K. We attribute the layer asymmetry below this temperature to energy drag such as for the bigger device. A possibility for the delayed onset of energy drag is the reduced sample geometry and contacts configuration. Energy drag depends on the thermoelectric effect which is driven by a temperature gradient. If a temperature gradient is not sustained in a device, the energy drag signal will be small. In our small signal of size ($0.7 \times 2.5\,\mu\text{m}^2$), a large temperature gradient is harder to form because of its small size and because the large dense contacts serve to dissipate heat efficiently which

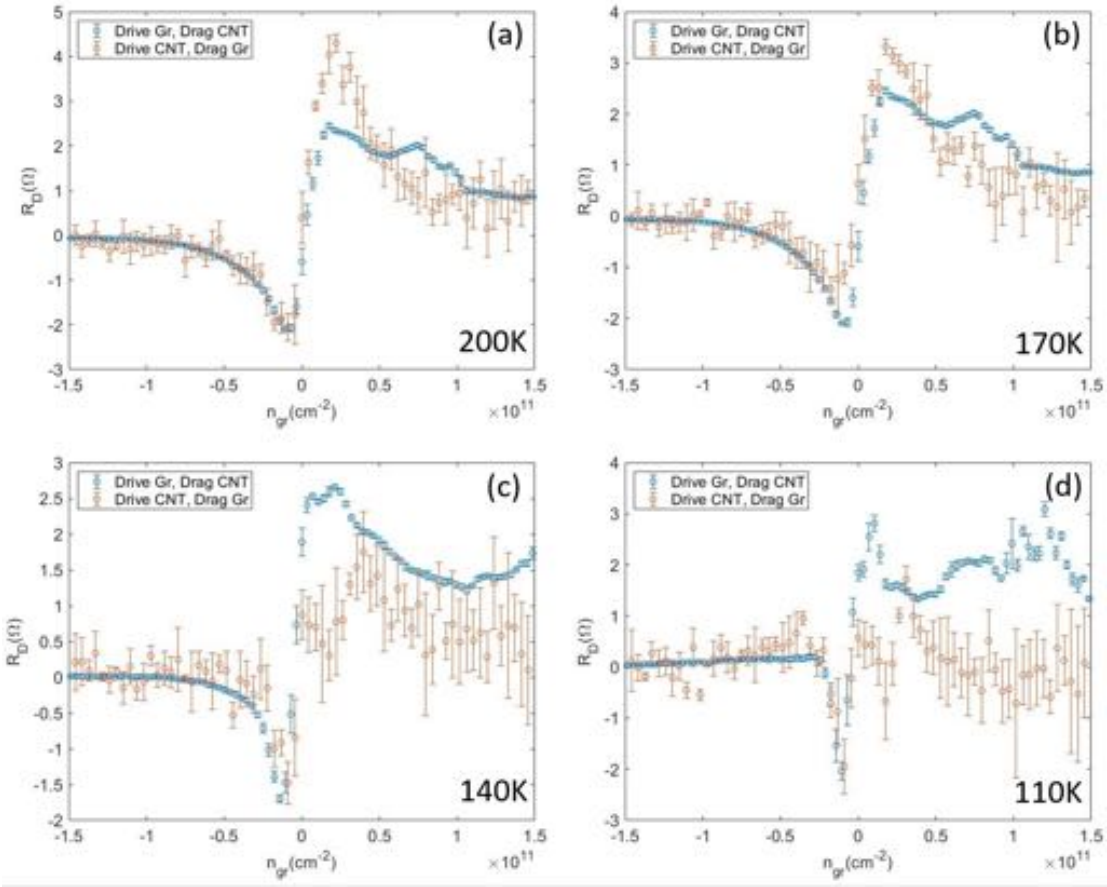


Figure 7.20: Measurements showing that at high temperature layer reciprocity can be followed. The two measurements compared are the CNT drive and Gr drive drag signal and this set of measurement is done in a much smaller device compared to the set of measurements shown in figure 7.18. (a) through (d) show temperatures at 200, 170, 140 and 110 K respectively.

consequently serves to cool the device uniformly. For the devices where energy drag seems to runaway right away even at high temperatures, the devices typically have Gr contacts (which is not as effective as thermal sink) and are roughly $2 \times 5\ \mu\text{m}^2$.

For the small device in the regime where layer reciprocity is held, we are inclined to say this is the regime of momentum drag. Looking closely at the data, the drag signal occurs within the full width half maximum (FWHM) of the Dirac peak as shown in figure 7.19b. This suggests that the signal is

due to the electron-hole plasma of Gr where there is a co-existence of electrons and holes. 7.19c shows the amplitude of the peak and dip of the drag signal at different temperatures. These amplitudes corresponds to the maximum scattering rates for hole and electron doped Gr. In this measurement, the CNT is hole-doped so the signal corresponds to either electron-hole or hole-hole drag. At high enough temperatures where layer reciprocity is held, the drag signal is linear with temperature. This is in contrast to the typical quadratic drag temperature dependence but this in fact qualitatively follows the expectation of electron-electron scattering at the Dirac point irrespective of carrier type.[170,171] Another explanation for the linear temperature dependence is phonon scattering which is expected to be linear as well after the Bloch-Grüneisen temperature. However, the coefficient of the linear temperature dependence coefficient for electron-phonon scattering is measured to be $\sim 0.14\Omega/K$, which is an order of magnitude larger than the measured coefficient in our drag device.[172] Figure 7.19d shows the spatial dependence of the measured longitudinal drag signal away from the CNT and an exponential decay function provides a good fit. Unfortunately, either due to poor spatial resolution, poor proximity to the CNT injection lead, or poor sample quality (although the sample mobility in this device is $\mu \sim 130,000\,\mathrm{cm}^2/V\cdot s$), we are unable to observe the sign reversal, which is one of the main signatures for hydrodynamics.

Drag attempts were also performed on bilayer Gr, which is expected to be in the hydrodyanmic regime more easily due to its parabolic dispersion. Figure 7.21 shows some representative plots for devices composed of a CNT and bilayer Gr that are well-reproduced across samples. Perhaps not to our surprise the behavior is different from the device with monolayer Gr, given that bilayer Gr has a record of exhibiting anomalous drag behaviors.[132,142,139,140] Like the monolayer devices, the bilayer Gr devices also have the Gr sitting on top of the the CNT. Consequently, the top gate (V_{tg}) controls purely the Gr because the CNT shows completely insensitivity to it due to its small size and Gr screening. The

backgate (V_{bg}) controls both layers' densities. Figure 7.21b shows the coarse map and there are two distinctive features. One is the bipolar behavior that follows the Dirac peak throughout (Dirac peak position is shown in figure 7.21a). The other feature happens around the double charge neutrality point (DCNP, where the CNT conductance dip is shown in figure 7.21c). The background bipolar trend is possibly indicative of the thermoelectric behavior because it seems to not depend at all on the drive layer carrier type. Not shown is that this bi-polar feature increases in amplitude with decreasing temperature. Figure 7.21d-e are longitudinal drag measurements and figure 7.21f is a hall measurement, both taken around the DCNP. Like any other behavior for the bilayer Gr device, we see that signal increases with decreasing temperature. More peculiarly is the development of six distinct regions in the drag measurements. The middle lobes develop at the region of Gr's Dirac point and switches sign upon crossing the charge neutrality point of the CNT. Outside of the middle lobes at the Dirac point the drag signal changes sign. A hall measurement at the same gate ranges reveal a bipolar behavior across the CNP of the CNT. The behaviors in general in the these drag devices are rich and complex and at the time of writing the understanding is still limited.

Thus far we have mainly focused on the CNT drive measurements. Here we speculate that this type of measurement is prone to be dominated by energy drag. When the CNT is driven, due to the substantial heat current generated through the high contact resistance (on the order of $0.1 - 1\ \mathrm{M}\Omega$), a subsequent temperature gradient according to equation 7.31 will develop in the drag layer. On the other hand, if Gr is driven, we gently argue that a temperature gradient is less likely to be created across the carbon nanotube. Ideally, the temperature gradient should develop perpendicular to the heat current which is parallel with the electric current. And since the current and CNT is facing the same direction (when we are driving Gr), the temperature gradient that should develop perpendicular to the current direction will not be established along the CNT. Of course the temperature gradient direction

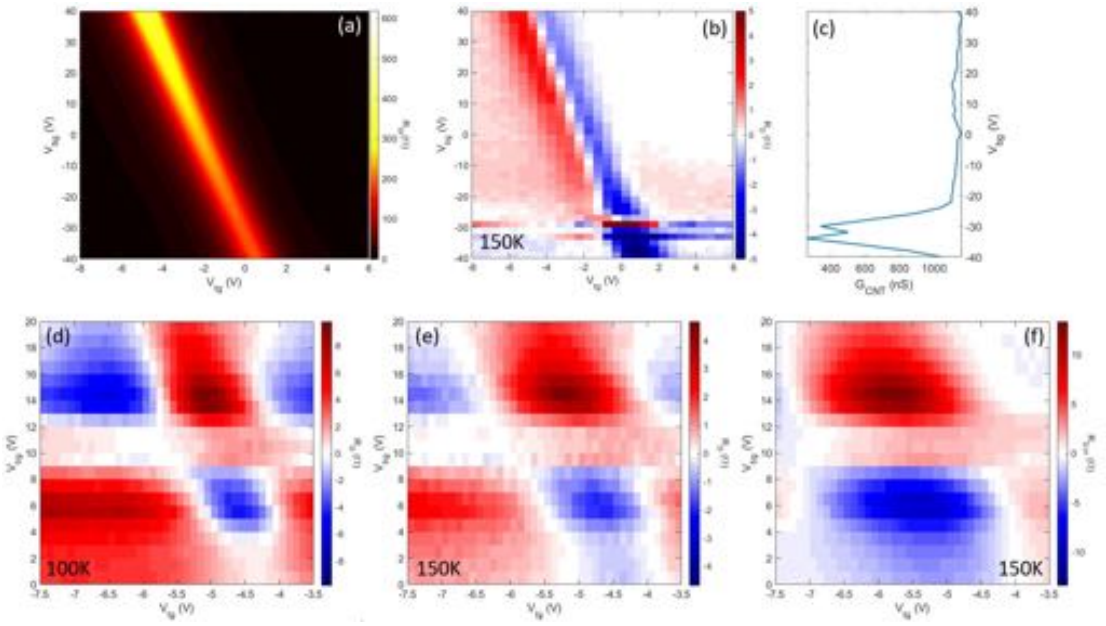


Figure 7.21: Drag measurements on hybrid CNT-bilayer Gr devices. (a) the Dirac peak of the Gr. (b) longitudinal drag measurement while driving the CNT and dragging Gr. (c) CNT conductance. (d) longitudinal drag at 100K and at (e) 150 K. (f) hall drag measurement at 150 K. The Dirac peak happens along the middle lobes in figures (d) to (f) going from $V_{\text{tg}} \approx 4.2$ to 5.3 V from bottom to top of the plot. For the same plots, the CNT conductance dip happens along the middle horizontal feature at $V_{\text{bg}} \approx 10$ V.

is complicated by the nature of $\mathbf{Q}$ and $\mathbf{D}$ but intuitively it seems unlikely that the heat current deviate much from the direction of the electric current. With this thought in mind, looking at the Gr drive signals shown in figure 7.18 and 7.20, they tend to hold the bipolar behavior longer, while at the same time, obeying the drag sign dictated by momentum drag as well as subtly decreasing in amplitude with decreasing temperature. These trends possibly suggest that Gr drive may be better in bringing out the momentum drag mechanism.

Finally, we entertain the idea that the components of the hybrid device do not fall under the category of a trivial Fermi liquid. For one, the 1D CNT can be considered as a Luttinger liquid. The drag resistance of Luttinger liquid has been investigated extensively and one of the key signatures is its non-monotonic temperature dependence and its diverging drag signal towards zero temperature.[147?]

Similarly, Gr can be in the hydrodynamics regime in which case it is also expected to show a non-monotonic behavior with a sharp cross-over at intermediate temperatures.[173] Considering these two highly relevant behaviors, it is not unreasonable to say the drag signal we observe could potentially be a manifestation of the two highly-interacting states of matter.

While it is possible that all the non-trivial drag behavior, especially at low CNT densities where screening is minimal, is due to the highly-interacting state of matter, a more concrete study of the drag behavior will need to be done regardless. Additional experiments on semiconducting CNTs will certainly be enlightening because its curvature in energy dispersion can give rise to a larger drag signal—a theoretical prediction that has yet to be shown. Furthermore, with a semiconducting CNT we can be more confident about the carrier type even at low doping thereby making the analysis of the drag signal more certain. A thorough study on the layer density dependence should be done to see more properly the roles of the drive and drag layers. Currently, we are hindered by the hysteretic effects of the CNT which cause confusion in the CNT carrier type as well as its large contact resistances which cause difficulty in performing measurements that are completely free of parasitic thermoelectric effects. Once with an improved contact recipe for the CNT and with definite access to the different carrier types in CNT, along with knobs such as the magnetic field, the problem of Coulomb drag between the CNT and Gr can certainly be tackled with clarity. And if one were to be able to isolate the directional momentum drag effect, perhaps the goal of observing hydrodynamics will come within reach.

8

Conclusion: Applications and Outlook

What we have developed in the pursuit of the hybrid carbon nanotube-graphene device is a system that can deterministically select the right kind of carbon nanotube followed by an immediate transfer of it to a target substrate. This serves to be a precious tool moving forward for the field of nanotechnology. Currently, nanofabrication has largely relied on the top-down approach in creating nanoscale structures. In this type of process, the introduction of fabrication residue is inevitable, not to mention the edge and surface roughness of deposited substances are impossible to smooth. The demonstration of the physical manipulation of individual carbon nanotubes is a big step toward building nanostructures deterministically from the bottom-up approach. Because we are utilizing crystalline products

that at the same time is one of the most microscopic objects to date, there is tremendous potential in building even more perfect nanostructures than ever before.

However, as demonstrated in the fabrication chapter, the process is far from perfect and there remains plenty to improve on. Regarding the carbon nanotube transfer, currently a polymer is still required to actively grab the carbon nanotube from the growth chip to the target substrate. To completely throw out the use of polymers, one can make growth chips that have pairs of electrodes on the two ends of the chips. Carbon nanotubes will grow over these electrodes (tested to be possible) and the method of transfer would simply be running a large current between the pair of electrodes at the two ends. The large current will cut the carbon nanotube and the carbon nanotube will fall onto the target substrate, thereby removing the need to use a polymer as a grabbing mechanism.

On the selection of carbon nanotubes, the on-chip Rayleigh setup has proven to be extremely useful. The idea behind on-chip Rayleigh is the crossed polarization method where the imaging of the carbon nanotube relies on the fact that the dipole moment of the carbon nanotube can help transmit light through a crossed polarizer.[174] Currently the Rayleigh setup is equipped to do it but the components will need to be improved in order for the setup to be more consistently useful. Being able to do so will facilitate the imaging of carbon nanotubes. SEM, AFM, EFM etc. all of which are extremely slow, inconvenient and potentially invasive, can be completely removed from the overall process.

Employing something as small as a the carbon nanotube will prove to be rewarding for many 2D systems. For the first time, a local gate that is on the order of 2D material's Fermi wavelength can be used. This undoubtedly will open up many new opportunities and possibly a new paradigm of physics. The reason for this is electronic channels have always had a hard time interacting with each other in a controllable and precise manner. Previously metal gates were used and these objects are not as narrow and not as smooth as a carbon nanotube. Metal gates also have a large density of states that

serve to screen out interactions. A carbon nanotube is possibly the least invasive 1D object and using it will finally allow electronic channels to feel each others presence more substantially.

Among the many types of narrow gating experiments, one clear direction is towards engineering the artificial Luttinger liquid. By using the carbon nanotube as a gate to deplete the region underneath it in a quantum hall system, one can create counter-propagating chiral edge states traveling along the sides of the carbon nanotube. Given that the edge states can possibly interact, this fulfills the recipe for a Luttinger liquid. If one were able to confidently engineer a Luttinger liquid, a huge rewarding venue opens up and this is because Luttinger liquid is one of the most highly correlated systems and also one that is studied least convincingly. Being able to engineer a Luttinger liquid will give one access to the rich physics of this highly correlated system with a high degree of control. Its coveted property of spin-charge separation may possibly be harnessed this way.

Recently, the flat-band physics have captivated the whole physics community.[175] The flat-band was first created in a double graphene layer system where the two layers are twisted to the so-called "magic angle". With it people saw a Mott insulator and superconductivity.[176,177] The carbon nanotube itself can intrinsically host flat bands as seen in the bandstructure of zig-zag carbon nanotubes (figure B.3). Coupled with the fact that the carbon nanotube is a more correlated system than graphene, one can expect something possibly even more exotic than what was observed. On the same topic of flatband, the 1D guided mode in graphene is shown to have a cyclic behavior. By changing the potential of the 1D gate, one will eventually create a 1D flatband in graphene at the so-called "magic" potential.[107] These 1D flatbands is a completely uncharted territory and can potentially yield extremely ground breaking results.

It has been mentioned, and discussed many times with many theoretical suggestions in literature all suggesting that graphene can be used as a platform to study general relativity. A big part of this

thesis along with many of other works in literature have demonstrated its striking connection with special relativity in flat spacetime. There is a call for efforts in trying to study general relativity in curved spacetime. With the nanotube, local perturbations can be created in the potential landscape in graphene electrostatically. The nanotube can also create large physical modulations of the graphene lattice by placing the graphene on top of a large diameter carbon nanotube. From the perspective of general relativity, these perturbations can be interpreted as the curving of a flat spacetime. Together with the local probing capability of the carbon nanotube, one can possibly study the dynamics of particles in these curved regions.

More directly related with the work demonstrated in this thesis, there are plenty of immediate continuation studies. Using the carbon nanotube as a local probe to measure local density of states will certainly unveil plenty of properties that conventional transport cannot. This will especially be true for materials that are sensitive to the destructive process of nanofabrication in which conventional transport cannot evade. So long as the materials are quantum mechanical in that the density of states are small and varying, the carbon nanotube will be able to pick out the slightest changes that would otherwise be hidden in the global transport measurement. Some rewarding choices of materials to perform local probing are the transition metal chalcogenides (TMCs) and few layer high temperature superconductors such as Bismuth strontium calcium copper oxide (BSCCO).

For the single guided mode, there are plenty of engineering questions that need to be answered. To list some of the essential ones: what is the actual lifetime or mean free path of the guided modes? Are they actually ballistic over hundreds of microns? How do they survive at kinks or sharp turns along the gated region, i.e. what is its angle of acceptance or critical angle? How can one inject and receive information out of this guided mode? And once with these foundational questions answered, one can actually try to integrate these guided modes into electron-optics. There is by now a very sufficient set

of electron-optics tools that one can use. Demonstrating the on-chip optical capability on graphene will be a huge step in the field of information processing. Simply demonstrating the transmission of information across a graphene sheet will be extremely impactful.

As for hydrodynamics in graphene, one would first have to find more conclusive and convincing signatures of its existence. Thus far, the hydrodynamic behavior in graphene has mostly befuddled the physics community. Most of the preliminary results shown across the community cannot be distinguished from ballistic electrons. The smoking gun signature of hydrodynamics will need to be established and the Coulomb drag in graphene may possibly be the platform for this search. But regardless of whether hydrodynamics is the main goal or not, drag between a 1D wire and a 2D conductor is still not at all understood. The complexity introduced by the 1D wire alone, not even considering the complications from a Coulomb drag measurement, is overwhelming to unravel. Likely the convincing way moving forward is to strip everything down to the bare minimum and study each piece individually first.

This brings us to the final point. The carbon nanotube is simply full of wonder and surprises still waiting to be discovered and understood. Any proclaimed correlated 1D physics is generally not easily accepted. Behaviors of any sort are generally seen with stubborn doubt and any explanation are followed by unyielding suspicion. This is because the technology is not there. The carbon nanotube is simply too microscopic and it really is like a blackbox. There will be need to be improvements on nanofabrication, mainly on establishing transparent electrical contacts, in order to move forward steadily with 1D carbon nanotube physics. As a quick digression, if superconducting contacts can be made with high transparency, the nanotube will be great a platform for the realization of Kitaev chains and superconducting qubits in general. With the contact hurdle tackled, carbon nanotube will undoubtedly be a revolutionizing material. Like graphene that has now widely become the universal

choice for 2DEG studies, the nanotube has plenty of reasons to be considered the universal 1D electronic platform. This will eventually come with urgency because the evolution of scientific studies shows that there is a growing need to understand materials with an increasingly microscopic perspective. This is especially true nowadays when we are struggling to keep up with Moore's law. The carbon nanotube will inevitably draw larger attention as it is one of the most robust and microscopic semiconductor that occurs naturally.

As a concluding remark for this thesis, what we have unraveled and seen so far is merely the surface and glimpse of what is really hidden within the bonds between the atoms. It is certainly a space not of vacuum but full of unworldly entities and events. Although the thesis marks the end of my time in graduate school, my curiosity and passion for the world of mesoscopic physics will stay with me forever. I am certainly a very proud participant in pushing the frontier of science, regardless of how insignificant it may seem in the large scheme of things. But as we all know by now, everything is built on small baby pieces of work. Even the likes of graphene or carbon nanotube, no matter how small they are in our eyes, are still made of uncountable number of unfathomably small building blocks. I cannot wait for what is in store for us to discover. Maybe after several more theses, things will come together nicely like puzzle pieces and maybe one day it will all come together and take the whole world aback, and then push humanity light years forward... maybe one day...

Epilogue: Prologue to the Appendix

I have come across many textbooks, review works and theses but only few of the multitude I felt was all-encompassing enough for my own taste and for the work that I do. Often times, I find myself referring to references after references within references. My mission here is simply to create as complete a thesis as I can, describing all the relevant basics in a way that any undergraduate physics major who has taken the introductory physics classes can understand, and of course, without needing to refer to another source. I will try my best to elucidate the physical meaning of the mathematics behind the derivations whenever possible in hopes of making this an enlightening read. I will definitely avoid the thought and phrase: "it is trivial to show".

The later few appendices are meant for researchers in the field looking to have a consolidated library of methods and recipes. Like what was said, throughout my graduate school career, I have found myself just salvaging bits and pieces of information scattered here and there. Hopefully this can save some of them some time. Finally, I am writing this for those who have always demonstrated interest in my work. Unfortunately, I can only do this to a certain level of simplicity otherwise this thesis will just be too lengthy and wordy than it already is. I hope I have done a reasonable job at this and that it maybe helpful for someone in the future.

For those who want a more rigorous and complete derivation of the electronic properties of graphene, I highly recommend the book *Graphene: Carbon in Two Dimensions* by M. Katsnelson[178] as well as

the review paper *Electronic Properties of Graphene* by A. H. C. Neto et al.[179] As for carbon nanotubes, I highly recommend the book *Carbon Nanotube and Graphene Device Physics* by H.-S. P. Wong et al.[180] and the review paper *Quantum Transport in Carbon Nanotubes* by E. A. Laird et al.[15] Much of the background knowledge I have regarding either material are acquired from the sources mentioned above. As for the fundamentals required to understand the advanced topics, they are all from the book *Solid State Physics* by N.W. Ashcroft and N. D. Mermin.[181] This book, I have read religiously, over and over again. It was the book that first introduced me to solid state physics as an undergraduate, the book that conditioned me as a first year graduate student, the book that prepared me for all my qualification exams, the book that I constantly fell back to throughout graduate research and still do as I am writing this thesis. Without further ado, let's move onto the appendices.

A

Overview of the Electronic Properties of Graphene

In this appendix, I will give a brief overview of the electronic properties of graphene. I will mostly follow the tight binding model in which some naive but arguably insignificant approximations are made. Consequently, the electronic description provided here is not complete. The reader will be notified of the approximations and hence flaws of this model. This overview is meant to only cover the basics essential to the understanding of this thesis and I am writing this assuming readers have at

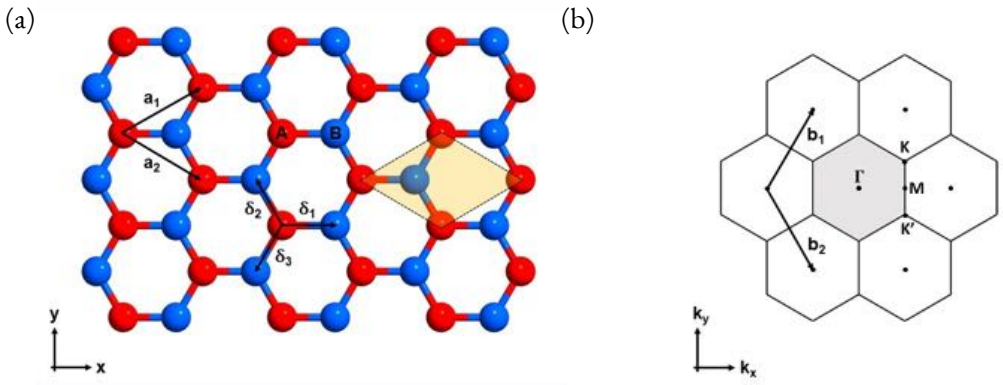


Figure A.1: Illustration of the direct and reciprocal lattices of graphene, both of which have a hexagonal Bravais lattice. (a) The direct lattice in real space. The coloring scheme shows how it can be broken down into two sublattices. Sublattice A is colored red and sublattice B is colored blue. The yellow shaded parallelogram is the primitive unit cell defined by the real space Bravais vectors $\mathbf{a_1}$ and $\mathbf{a_2}$. The nearest neighbour vectors are given by $\boldsymbol{\delta_1}$, $\boldsymbol{\delta_2}$ and $\boldsymbol{\delta_3}$. (b) The reciprocal lattice in momentum space. The gray shaded hexagon in the middle is the first Brillouin Zone. $\mathbf{b_1}$ and $\mathbf{b_2}$ are the two reciprocal lattice vectors. The high symmetry points are labeled as $\boldsymbol{\Gamma}$, $\mathbf{K}$, $\mathbf{K}'$ and $\mathbf{M}$ where $\mathbf{K}$ and $\mathbf{K}'$ are two inequivalent positions that have Dirac cones with opposite helicity.

least taken an undergraduate solid state physics course. I will, however, try to elucidate the physical meaning of the mathematics behind the derivations whenever possible in hopes of making this an enlightening read.

A.1 Graphene Lattice

Graphene is a single layer of carbon atoms arranged in a honeycomb lattice as shown in figure A.1a. This honeycomb lattice is systematically characterized by a hexagonal Bravais lattice with a basis of two atoms which are labeled A and B (color coded red and blue respectively). This means that atoms A and B separately make up a hexagonal sublattice. The Bravais lattice vectors are:

$$\mathbf{a_1} = \left(\frac{\sqrt{3}a}{2}, \frac{a}{2}\right), \quad \mathbf{a_2} = \left(\frac{\sqrt{3}a}{2}, -\frac{a}{2}\right), \tag{A.1}$$

where $a = |\mathbf{a_1}| = |\mathbf{a_2}|$ is the Bravais lattice constant of graphene and has a value of about 2.46 Å. These Bravais lattice vectors define the primitive unit cell which is the parallelogram shaded in yellow. Each carbon has three nearest neighbors, all of which are occupants of the other sublattice. These nearest-neighbor vectors are:

$$\boldsymbol{\delta_1} = \left(\frac{a}{\sqrt{3}}, 0\right), \quad \boldsymbol{\delta_2} = \left(-\frac{a}{2\sqrt{3}}, -\frac{a}{2}\right), \quad \boldsymbol{\delta_3} = \left(-\frac{a}{2\sqrt{3}}, \frac{a}{2}\right), \tag{A.2}$$

where the carbon to carbon bond length or the nearest neighbour distance, $a_{cc} = |\boldsymbol{\delta_1}| = |\boldsymbol{\delta_2}| = |\boldsymbol{\delta_3}|$, is about 1.42 Å.

The reciprocal lattice is the Fourier transform of the direct lattice, taking the dimension reciprocal of real space which is spatial frequency, more commonly known as the wavevector $\mathbf{k}$— a quantity that is directly tied with crystal momentum, given by $\hbar\mathbf{k}$. This crystal momentum is a conserved quantity that characterizes the allowed states in an electronic system with a periodic potential. It gives information on how electrons propagate in a crystalline material as Bloch waves. Following the culture in the field of condensed matter physics, from hereinafter, I will refer to crystal momentum simply as momentum and will denote it as $\mathbf{k}$ unless noted otherwise. But behold, I would like to remind the reader that there is a distinct difference between crystal momentum and momentum. To be brief, momentum is generally applied to free electrons whereas crystal momentum is specific to electrons under the influence of a periodic potential. Since we are almost never dealing with free electrons in condensed matter, the usage of momentum will largely be absent and so when one speaks of momentum, it is almost always in reference to crystal momentum.[182]

There are only a finite number of unique momenta the electrons are allowed to have. These allowed momenta are all contained within the first Brillouin Zone (BZ). This BZ is a primitive unit cell in

momentum space, which is also called the Wigner-Seitz cell. This cell forms the basis of the recirpocal lattice that spans the entire momentum space. The reciprocal lattice vectors are:

$$\mathbf{b_1} = \left(\frac{2\pi}{\sqrt{3}a}, \frac{2\pi}{a}\right), \quad \mathbf{b_2} = \left(\frac{2\pi}{\sqrt{3}a}, -\frac{2\pi}{a}\right), \tag{A.3}$$

which are obtained via the reciprocity condition that fulfills the Fourier transform relation between the direct and reciprocal lattice.[183] Figure A.1b depicts the momentum space that electrons in graphene live in. The hexagon that is shaded in grey is the first BZ. At the center is the origin of momentum space where momentum is zero. This point is conventionally labelled $\mathbf{\Gamma}$. $\mathbf{M}$ denotes the BZ edge and is thus a high symmetry point. The most important points in the graphene BZ are the $\mathbf{K}$ and $\mathbf{K}'$ points. These two points are located at the two inequivalent corners of the BZ and are described by the following vector with respect to $\mathbf{\Gamma}$:

$$\mathbf{\Gamma K} = \left(\frac{2\pi}{\sqrt{3}a}, \frac{2\pi}{3a}\right), \quad \mathbf{\Gamma K'} = \left(\frac{2\pi}{\sqrt{3}a}, -\frac{2\pi}{3a}\right). \tag{A.4}$$

The lowest-lying energy states of graphene with respect to the Fermi level exist here. These states form a cone-like structure and is called the Dirac cone. All of what we study in graphene is associated with this cone-like dispersion relation. It is given the name Dirac cone because the nature of this cone-like disperion relation resembles that of massless relativistic fermions predicted by Paul Dirac.

There will be much to talk about regarding the Dirac cone but we will first have to see how it arises.

A.2 Graphene Bandstructure

Given the information of graphene's direct and reciprocal lattice, we can proceed to find the dispersion relation or bandstructure of graphene. The popular method of obtaining the dispersion relation is the tight-binding method. This method assumes that electrons do not interact with one another or any other objects (such as impurities, phonons etc.) except for the energy potential from the graphene lattice, forming the so-called Fermi gas. This method is based solely on the superposition of wavefunctions for atoms at each lattice site and requires that the overlap of these wavefunction be small or that the wavefunction is "tightly bound" to the atom. A common name for this method is: linear combination of atomic orbitals (LCAO).

This method poses the bare minimum treatment for describing electrons in a periodic potential. Despite this, the results here are able to describe electrons in graphene extremely well. This is largely due to the power of the Fermi liquid theory which states that at sufficiently low temperatures interacting fermions, regardless of how strong the interaction is, can behave as a Fermi gas with a one-to-one correspondence, i.e., all the individual properties of a Fermi gas can be mapped directly to individual properties of the Fermi liquid. For example, spin, charge and momentum of fermions in a Fermi gas remains unchanged as interaction is turned on, while dynamical properties such as effective mass, specific heat and spin susceptibility are renormalized to new values.[184]

The above is to point out that the tight-binding model can take one very far in terms of understanding graphene, but of course, there are its limits and there are plenty of phenomena that the tight-binding model fails to address, one of which is hydrodynamics, subject to intense research efforts at the time of writing and is further detailed in the main text of this thesis. So going back to the tight-binding model of graphene, we begin by stating the time-independent Schrödinger's equation in the

bra-ket notation:

$$H\left|\Psi(\mathbf{k},\mathbf{r})\right\rangle = \varepsilon(\mathbf{k})\left|\Psi(\mathbf{k},\mathbf{r})\right\rangle, \tag{A.5}$$

where Ψ is the wavefunction and H is the Hamiltonian operator given by:

$$H = \frac{\hbar^2}{2m}\nabla^2 + \sum_{n=1}^{N}[U(\mathbf{r}-\mathbf{R}_n) + U(\mathbf{r}-\mathbf{R}_n-\boldsymbol{\delta_1})]. \tag{A.6}$$

The first term represents the kinetic energy and the second term is the potential energy from the graphene lattice, where $\mathbf{R}_n = \sum_{\{n_1,n_2\}\in n} n_1\mathbf{a_1} + n_2\mathbf{a_2}$ is the generalized Bravais lattice vector, N is the number of primitive unit cells and n is a counter that counts all the primitive unit cells (which are given by the coordinates $\{n_1, n_2\}$). The two terms in the potential energy corresponds to the two basis atoms and $\boldsymbol{\delta_1}$ is the location of the sublattice B atom with respect to the sublattice A atom within the same primitive unit cell (see figure A.1a).

As a result of the translational symmetry of a Bravais lattice, the wavefunction has to satisfy Bloch's theorem as shown in equation 8.6 in Ashcroft and Mermin.[185] What this theorem basically says is that the wavefunction of an electron gains a phase factor associated with its momentum as it translates throughout the crystal and this simply reflects the periodicity or translational symmetry of the crystal lattice. Couple this idea with the fact that graphene has a basis of two atoms per lattice site (which form the basis of sublattice A and B), the wavefunction has the following form:

$$\left|\Psi(\mathbf{k},\mathbf{r})\right\rangle = C_A(\mathbf{k})\left|\Phi_A(\mathbf{k},\mathbf{r})\right\rangle + C_B(\mathbf{k})\left|\Phi_B(\mathbf{k},\mathbf{r})\right\rangle. \tag{A.7}$$

Here $\left|\Phi_A(\mathbf{k},\mathbf{r})\right\rangle$ and $\left|\Phi_B(\mathbf{k},\mathbf{r})\right\rangle$ each are the linear combinations of atomic orbitals of every carbon at sublattice A and B respectively. Note here the use of two separate weights: C_A and C_B, which are only

functions of $\mathbf{k}$ and not $\mathbf{r}$. This entails that the sublattices are separate entities with different energetics, but within the same sublattice all sites are energetically equivalent. Expanding the wavefunction fully to encompass the whole crystal lattice with N unit cells while respecting Bloch's theorem, they become:

$$|\Phi_A(\mathbf{k},\mathbf{r})\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{i\mathbf{k}\cdot\mathbf{R}_{A_n}} \left|\phi(\mathbf{r}-\mathbf{R}_{A_n})\right\rangle, \tag{A.8}$$

$$|\Phi_B(\mathbf{k},\mathbf{r})\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{i\mathbf{k}\cdot\mathbf{R}_{B_n}} \left|\phi(\mathbf{r}-\mathbf{R}_{B_n})\right\rangle, \tag{A.9}$$

where $\left|\phi(\mathbf{r}-\mathbf{R}_{A_n})\right\rangle$ and $\left|\phi(\mathbf{r}-\mathbf{R}_{B_n})\right\rangle$ are the localized atomic orbitals of each individual carbon situated at positions $\mathbf{R}_{A_n}$ and $\mathbf{R}_{B_n}$, the Bravais lattice vectors for sublattice A and B. Note that $\mathbf{R}_{B_n} = \mathbf{R}_{A_n} + \boldsymbol{\delta_1}$ (see figure A.1a).

The atomic orbitals of choice can vary but we know the following: graphene is composed of carbon with sp^2 hybridization where each carbon's $2s$ orbital mixes with its own $2p_x$ and $2p_y$ orbitals and overlaps with other sp^2 hybridized oribtals from other carbons via σ bonding, pairing up all of these orbitals' electrons in the process. These σ bonds form the backbone of graphene's honeycomb lattice and the orbitals responsible for these σ bonds are energetically stable. This leaves the $2p_z$ orbitals with unpaired valence electrons, which are then readily available to participate in the sea of electrons also known as the conduction band. Knowing this, we can safely designate the $2p_z$ orbital as the sole orbital contributing to graphene's wavefunction.

We now have all the information to derive the dispersion relation (or equivalently, the eigenenergies). Following the spirit of the tight binding model, we can isolate the eigenenergy, $\varepsilon(\mathbf{k})$, by operating on the equation with $\langle\Psi(\mathbf{k},\mathbf{r})|$ as such:

$$\langle\Psi(\mathbf{k},\mathbf{r})|\, H\, |\Psi(\mathbf{k},\mathbf{r})\rangle = \varepsilon(\mathbf{k})\, \langle\Psi(\mathbf{k},\mathbf{r})|\Psi(\mathbf{k},\mathbf{r})\rangle, \tag{A.10}$$

where $\varepsilon(\mathbf{k})$ is now in relation to scalars as opposed to vectors. We are now essentially computing the expected value for energy at each $\mathbf{k}$. To continue we just insert equations A.7 through A.9 into equation A.10 but as one can imagine the equation will get long and messy. We can circumvent this problem if we transcribed the equations into matrix form where we choose $\{\Phi_A(\mathbf{k},\mathbf{r}), \Phi_B(\mathbf{k},\mathbf{r})\}$ as our basis. In this case, the wavefunction (from equation A.7) becomes a column vector:

$$\mathbf{\Psi}(\mathbf{k}) = \begin{bmatrix} C_A(\mathbf{k}) \\ C_B(\mathbf{k}) \end{bmatrix}, \tag{A.11}$$

and the Hamiltonian in this basis becomes the Hamiltonian matrix (noting that the real space parameter is integrated out when calculating the matrix elements):

$$\begin{aligned} \mathbf{H}(\mathbf{k}) &= \begin{bmatrix} \langle \Phi_A(\mathbf{k},\mathbf{r})| \, H \, |\Phi_A(\mathbf{k},\mathbf{r})\rangle & \langle \Phi_A(\mathbf{k},\mathbf{r})| \, H \, |\Phi_B(\mathbf{k},\mathbf{r})\rangle \\ \langle \Phi_B(\mathbf{k},\mathbf{r})| \, H \, |\Phi_A(\mathbf{k},\mathbf{r})\rangle & \langle \Phi_B(\mathbf{k},\mathbf{r})| \, H \, |\Phi_B(\mathbf{k},\mathbf{r})\rangle \end{bmatrix} \\ &= \begin{bmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{bmatrix} = \begin{bmatrix} H_{AA} & H_{AB} \\ H_{AB}^* & H_{AA} \end{bmatrix} = \mathbf{H}^\dagger(\mathbf{k}), \end{aligned} \tag{A.12}$$

and lastly, we call the following the overlap matrix:

$$\begin{aligned} \mathbf{S}(\mathbf{k}) &= \begin{bmatrix} \langle \Phi_A(\mathbf{k},\mathbf{r})|\Phi_A(\mathbf{k},\mathbf{r})\rangle & \langle \Phi_A(\mathbf{k},\mathbf{r})|\Phi_B(\mathbf{k},\mathbf{r})\rangle \\ \langle \Phi_B(\mathbf{k},\mathbf{r})|\Phi_A(\mathbf{k},\mathbf{r})\rangle & \langle \Phi_B(\mathbf{k},\mathbf{r})|\Phi_B(\mathbf{k},\mathbf{r})\rangle \end{bmatrix} \\ &= \begin{bmatrix} S_{AA} & S_{AB} \\ S_{BA} & S_{BB} \end{bmatrix} = \begin{bmatrix} S_{AA} & S_{AB} \\ S_{AB}^* & S_{AA} \end{bmatrix} = \mathbf{S}^\dagger(\mathbf{k}). \end{aligned} \tag{A.13}$$

We first note that both matrices are hermitian (because they are observables) and that the overlap be-

tween atoms in sublattice A is identical to overlap between atoms in sublattice B, and hence the above equalities.

The overlap matrix is a measure of how tightly bound the atomic orbitals are to their respective atoms. If we assume that atomic orbitals are tightly bound enough that neighbouring atomic orbitals do not overlap, then:

$$S_{lm} = \delta_{lm}. \tag{A.14}$$

But this is a first order approximation and in reality there is some amount of overlap. This overlap contributes to graphene's slight electron-hole asymmetry (as we shall later see) and this order of correction is similar to that coming from the next nearest neighbour (nnn) hopping term (which we will cover shortly). Such corrections, however, are quite small relative to our measurement sensitivity and when the corrections do become non-negligible, that only happens in an energy range that we cannot typically access. So bottom line is, it is generally safe to assume that equation A.14 is true and that electron-hole symmetry is preserved in graphene. The reader shall be notified in cases where this assumption is not valid.

Continuing on, with the matrix reformulation, we have the following eigenvalue problem to solve:

$$\mathbf{H}(\mathbf{k})\mathbf{\Psi}(\mathbf{k}) = \varepsilon(\mathbf{k})\mathbf{S}(\mathbf{k})\mathbf{\Psi}(\mathbf{k}), \tag{A.15}$$

and so the solution is obtained from the secular equation:

$$\det[\mathbf{H}(\mathbf{k}) - \varepsilon(\mathbf{k})\mathbf{S}(\mathbf{k})] = 0, \tag{A.16}$$

which is a condition to satisfy for non-trivial solutions to exist (wavefunction not being zero). Writing

the secular equation out explicitly and solving for the eigenenergy specifically, we have:

$$\varepsilon(\mathbf{k}) = \frac{\varepsilon_0(\mathbf{k}) \pm \sqrt{\varepsilon_0(\mathbf{k})^2 - 4(S_{AA}(\mathbf{k})^2 - |S_{AB}(\mathbf{k})|^2)(H_{AA}(\mathbf{k})^2 - |H_{AB}(\mathbf{k})|^2)}}{2(S_{AA}(\mathbf{k})^2 - |S_{AB}(\mathbf{k})|^2)}, \quad \text{(A.17)}$$

with

$$\varepsilon_0(\mathbf{k}) = (2H_{AA}(\mathbf{k})S_{AA}(\mathbf{k}) - S_{AB}(\mathbf{k})H^*_{AB}(k) - H_{AB}(\mathbf{k})S^*_{AB}(\mathbf{k}). \quad \text{(A.18)}$$

Above is the dispersion relation for graphene written out in the most generic way according to the tight binding model. Previously we mentioned that the overlap matrix contributes to electron-hole asymmetry and if we look at the role of S_{AB} then we will see why. The quantity S_{AB} here serves to shift a perfectly symmetric square root function with respect to energy causing the reflection symmetry along the energy axis to vanish. But again, this correction is very small and for the sake of elegance and simplicity (as opposed to an otherwise messy and tedious derivation), from here onwards, we will assume equation A.14 to be true. Looking back at the dispersion relation, we see that there are a total of four quantities we need to determine but we just took care of two, namely $S_{AA} = 1$ and $S_{AB} = 0$. So now let's move on to the remaining two: the Hamiltonian matrix elements. For the diagonal terms we have:

$$H_{AA} = \frac{1}{N}\sum_{l=1}^{N}\sum_{m=1}^{N} e^{i\mathbf{k}\cdot(\mathbf{R}_{A_l} - \mathbf{R}_{A_m})} \left\langle \phi(\mathbf{r} - \mathbf{R}_{A_m}) \right| H \left| \phi(\mathbf{r} - \mathbf{R}_{A_l}) \right\rangle, \quad \text{(A.19)}$$

$$= \frac{1}{N}\sum_{l=1}^{N}\sum_{m=1}^{N} e^{i\mathbf{k}\cdot(\mathbf{R}_{A_l} - \mathbf{R}_{A_m})} \varepsilon_{2\mathrm{p_z}}\delta_{lm} = \varepsilon_{2\mathrm{p_z}}, \quad \text{(A.20)}$$

here we have utilized the fact that $H\left|\phi(\mathbf{r} - \mathbf{R}_{A_l})\right\rangle = \varepsilon_{2\mathrm{p_z}}\left|\phi(\mathbf{r} - \mathbf{R}_{A_l})\right\rangle$ where we remind the reader that here $\left|\phi(\mathbf{r} - \mathbf{R}_{A_l})\right\rangle$ is the $2p_z$ atomic orbital at position $\mathbf{R}_{A_l}$ and $\varepsilon_{2\mathrm{p_z}}$ is its energy, which

in this context is called the onsite energy. As for the off-diagonal terms:

$$H_{AB} = \frac{1}{N}\sum_{l=1}^{N}\sum_{m=1}^{N} e^{i\mathbf{k}\cdot(\mathbf{R}_{B_l}-\mathbf{R}_{A_m})} \left\langle \phi(\mathbf{r}-\mathbf{R}_{A_m})\right| H \left|\phi(\mathbf{r}-\mathbf{R}_{B_l})\right\rangle , \tag{A.21}$$

$$\approx \frac{1}{N}\sum_{l=1}^{N}\sum_{n=1}^{3} e^{i\mathbf{k}\cdot\boldsymbol{\delta_n}} t_{\mathrm{nn}} = (e^{i\mathbf{k}\cdot\boldsymbol{\delta_1}} + e^{i\mathbf{k}\cdot\boldsymbol{\delta_2}} + e^{i\mathbf{k}\cdot\boldsymbol{\delta_3}})t_{\mathrm{nn}}. \tag{A.22}$$

In the above, we used the approximation that only nearest neighbors (nn) would contribute to a non-zero value for the hopping (or transfer or overlap) integral, $\left\langle \phi(\mathbf{r}-\mathbf{R}_{A_m})\right| H \left|\phi(\mathbf{r}-\mathbf{R}_{B_l})\right\rangle$, which means that $\boldsymbol{\delta}_{n\in 1,2,3}$ are the three vectors that connect a sublattice A atom to its three nearest neighboring sublattice B atoms (see figure 1.1a). The quantity t_{nn} is the hopping (or transfer or overlap) energy between nearest neighbors given by the corresponding hopping (or transfer or overlap) integral. Reported values for t_{nn} range from -2.7 eV to -3.3 eV.[180] Note that the energy here is negative which reflects the attractive nature of the crystal lattice periodic potential. Here we use t_{nn} to accentuate that the hopping energy is only for nearest neighbors. Another common label for the nearest neighbor hopping energy is given by γ_0.

Now we can finally plug everything we know into equation A.17 and obtain the dispersion relation of graphene:

$$\varepsilon(\mathbf{k}) = \pm t_{\mathrm{nn}}\sqrt{1 + 4\cos\frac{\sqrt{3}a}{2}k_x\cos\frac{\sqrt{a}}{2}k_y + 4\cos^2\frac{a}{2}k_y}. \tag{A.23}$$

In arriving at this formula, we have set $\varepsilon_{2\mathrm{p_z}} = 0$ and this can be done without loss of generality or accuracy. The reason is any energy reference point is arbitrary and in the convention of condensed matter physics, the Fermi level is designated as zero energy. For graphene, the Fermi level exists at the $\mathbf{K}$ and $\mathbf{K}'$ points in the BZ (see figure 1.1b). This is because undoped graphene has exactly one donor electron per atom which results in one electron per unit cell (a unit cell consisting of the two basis

atoms). Accounting for spin degeneracy this means graphene's energy band is intrinsically half filled, which causes the Fermi level to rest at exactly where the valence and conduction bands touch, making graphene a semi-metal. Figure A.2a shows the bandstructure of graphene plotted using equation A.23 and figure A.2c shows the energy contour lines of the bandstructure. Figure A.2d shows the line cut of the bandstructure along the high symmetry points and one can see that at the $\mathbf{K}$ point the dispersion looks linear and that the conduction and valence bands touch. Note that, without going through derivation, the valence (lower energy) band corresponds to the π (bonding) p_z orbitals and the conduction (higher energy) band corresponds to the π^* (antibonding) p_z orbitals. The $\mathbf{K}$ and $\mathbf{K}'$ points are also known as the Dirac points (for a reason that will soon be clear) and are referred to as the $\mathbf{K}$ and $\mathbf{K}'$ valleys. The remaining four corners of the BZ are simply part of either the $\mathbf{K}$ and $\mathbf{K}'$ points since they can all be joined with a reciprocal lattice vector.

A.3 Linear Dispersion of Graphene: Massless Dirac Fermions

Let us now examine the energies around the Fermi level, the Dirac points, more closely. To do so, first we recognize that the dispersion relation around the Dirac point is linear. Then going back to the Hamiltonian of graphene (equation A.12), while noting that the diagonal terms are zeros leaving only non-zero off-diagonal terms (given by equation A.21), we perform a Taylor expansion around the Dirac point. For the reader's convenience, the original Hamiltonian is rewritten down below:

$$\mathbf{H}(\mathbf{k}) = \begin{bmatrix} 0 & H_{AB} \\ H_{BA} & 0 \end{bmatrix} = \sum_{n=1}^{3} t_{\mathrm{nn}} \begin{bmatrix} 0 & e^{i\mathbf{k}\cdot\boldsymbol{\delta}_{\boldsymbol{n}}} \\ e^{-i\mathbf{k}\cdot\boldsymbol{\delta}_{\boldsymbol{n}}} & 0 \end{bmatrix}. \tag{A.24}$$

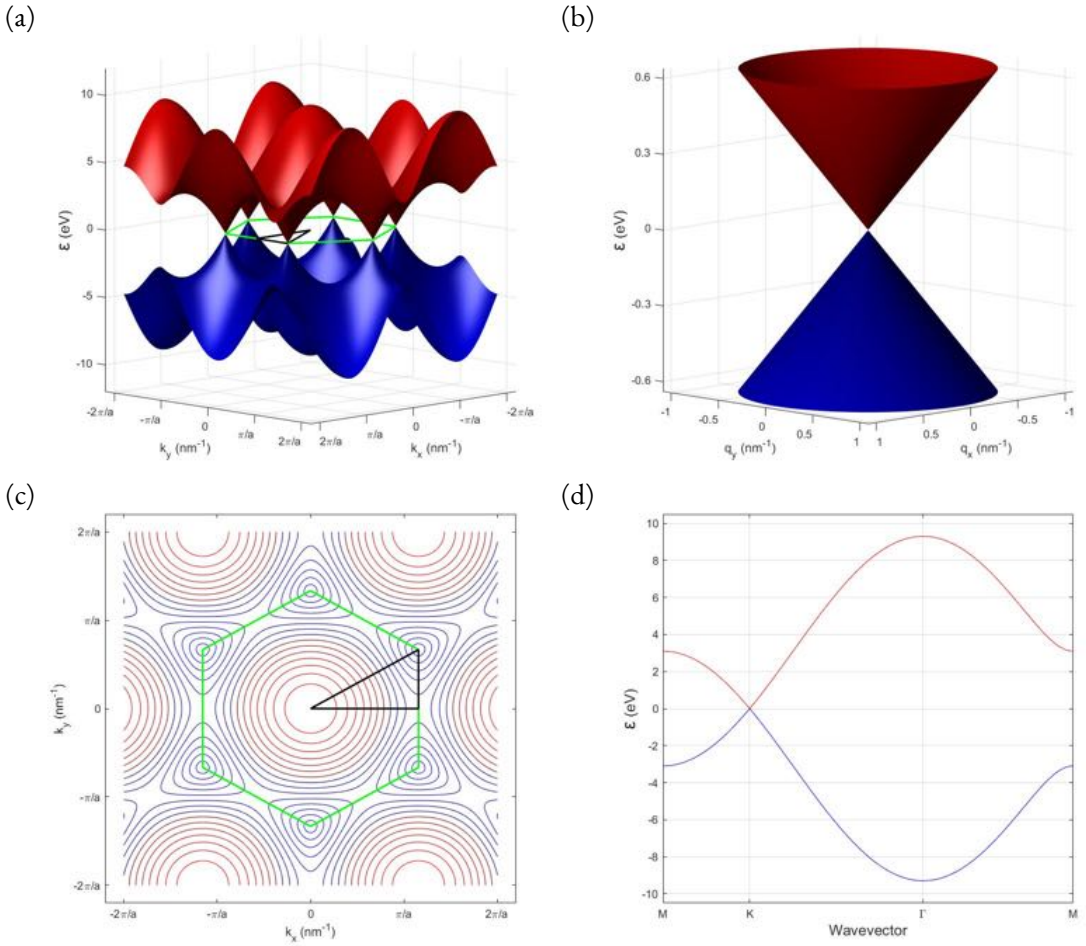


Figure A.2: Illustration of the dispersion relation (ε vs $\mathbf{k}$) of graphene. (a) The bandstructure of graphene in the first BZ. Red denotes the conduction band and blue denotes the valence band. (b) The Dirac cone, i.e., the cone structure at the $\mathbf{K}$ and $\mathbf{K}'$ points. Here the dispersion relation is linear and symmetric. (c) The projection of the graphene bandstructure into the 2D momentum plane which shows the 2D Fermi surface of graphene. The contour lines represent constant energies. The green line marks the first BZ and the black line traces through the high symmetry points in graphene. $\mathbf{\Gamma}$ defines the origin of the BZ, $\mathbf{K}$ and $\mathbf{K}'$ are any two neighboring corners of the BZ. $\mathbf{M}$ is the midpoint between $\mathbf{K}$ and $\mathbf{K}'$ along the BZ edge. (d) The dispersion relation along the black line as seen in figure (c) which traces through the high symmetry points $\mathbf{\Gamma}$, $\mathbf{K}$, $\mathbf{K}'$ and $\mathbf{M}$.

To ease the expansion, we define a wavevector with respect to the Dirac point as:

$$\mathbf{q} = \mathbf{k} - \mathbf{K}, \quad \mathbf{q}' = \mathbf{k} - \mathbf{K}', \tag{A.25}$$

then, focusing only on the $\mathbf{K}$ point and only on positive values, the expansion of the Hamiltonian becomes:

$$\begin{aligned} \mathbf{H}(\mathbf{q}) &\approx \frac{\sqrt{3}t_{\mathrm{nn}}a}{2} \begin{bmatrix} 0 & e^{-i5\pi/6}(q_x - iq_y) \\ e^{i5\pi/6}(q_x + iq_y) & 0 \end{bmatrix} (1 + \mathcal{O}(q/K)^2) \\ \rightarrow \mathbf{H}(\mathbf{q}) &= \frac{\sqrt{3}t_{\mathrm{nn}}a}{2} \begin{bmatrix} 0 & q_x - iq_y \\ q_x + iq_y & 0 \end{bmatrix}. \end{aligned} \tag{A.26}$$

In the above sequence, we have chosen to ignore the higher order terms as well as the phase factors and where we only address the positive solution. The higher order terms are ignored because the value is negligible compared to the first order term in the linear regime. Higher order terms, of course, become more significant as we go farther away the Dirac points. As for the phase factor, it is meaningless and does not contribute to any observable quantities. We can easily see this by trying to compute the eigenenergy of the above Hamiltonian where we will see that the prefactor will be absolute squared in which case the phase factor will vanish (the eigenenergy will be shown briefly). In performing the expansion around $\mathbf{K}$, we have effectively converted the Hamiltonian basis: $\{\Phi_A(\mathbf{k}, \mathbf{r}), \Phi_B(\mathbf{k}, \mathbf{r})\} \rightarrow \{\Phi_A(\mathbf{q}, \mathbf{r}), \Phi_B(\mathbf{q}, \mathbf{r})\} \equiv \{\Phi_{A,\mathbf{K}}, \Phi_{B,\mathbf{K}}\}$. What this entails is the wavefunctions now become slowly varying with spatial frequency $\mathbf{q}$ (see section A.6 for the wavefunction around the Dirac cones).

Rewriting the Hamiltonian conventionally for both $\mathbf{K}$ and $\mathbf{K}'$ points with bases $\{\Phi_{A,\mathbf{K}}, \Phi_{B,\mathbf{K}}\}$ and $\{\Phi_{A,\mathbf{K}'}, \Phi_{B,\mathbf{K}'}\}$ respectively while taking into account the positive and negative solutions, we

have:

$$\begin{aligned}\mathbf{H}(\mathbf{q}) &= \pm\hbar v_F \begin{bmatrix} 0 & q_x - iq_y \\ q_x + iq_y & 0 \end{bmatrix}, \\ \mathbf{H}(\mathbf{q}') &= \pm\hbar v_F \begin{bmatrix} 0 & q'_x + iq'_y \\ q'_x - iq'_y & 0 \end{bmatrix},\end{aligned} \tag{A.27}$$

where positive and negative values correspond to the conduction and valence bands respectively which means that the Fermi level crosses the Dirac point. The Fermi velocity is defined as:

$$v_F = \frac{\sqrt{3}t_{\text{nn}}a}{2\hbar} \approx c/300\,(\text{m/s}), \tag{A.28}$$

with c being the speed of light $\approx 3 \times 10^8$ m/s. Put more simply, the Hamiltonian for a single valley (in this case for $\mathbf{K}$ valley in particular as it is traditionally the one that is referenced) can be written as:

$$\mathbf{H}(\mathbf{q}) = \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q}, \tag{A.29}$$

where $\boldsymbol{\sigma}$ are the Pauli matrices given by:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \tag{A.30}$$

and where σ_z is intentionally left out to highlight the fact that the Hamiltonian does not (intrinsically) depend on it since $\mathbf{q}$ only exists in the 2D xy plane. The reader must be cautioned that the Hamiltonian can in fact acquire a σ_z term and more on this will be discussed in the next section. As

for the Hamiltonian of the $\mathbf{K}'$ valley, it is simply related to $\mathbf{K}$ valley through $\mathbf{H}(\mathbf{q}') = \mathbf{H}^*(\mathbf{q})$. This relation implies that the two valleys are linked via a time reversal operation (see section A.5). Consolidating the Hamiltonian for the $\mathbf{K}$, $\mathbf{K}'$ valleys into an overall low energy Hamiltonian, in the basis of $\{\Phi_{A,\mathbf{K}}, \Phi_{B,\mathbf{K}}, \Phi_{A,\mathbf{K}'}, \Phi_{B,\mathbf{K}'}\}$ we have:

$$\mathbf{H}(\mathbf{q}) = \hbar v_F \begin{bmatrix} \boldsymbol{\sigma} \cdot \mathbf{q} & 0 \\ 0 & \boldsymbol{\sigma}^* \cdot \mathbf{q} \end{bmatrix}, \tag{A.31}$$

and the eigenenergy of the Hamiltonian in the linear regime within each valley is:

$$\varepsilon(\mathbf{q}) = \pm \hbar v_F |\mathbf{q}|. \tag{A.32}$$

The eigenenergy as we can see is isotropic around the Dirac point, depending only on the magnitude of $\mathbf{q}$. The positive and negative values correspond to the conduction and valence bands respectively.

The graphene Hamiltonian in the linear regime (equation A.29) is a 2D analogue of the Dirac Hamiltonian for massless fermions except with the velocity of light c replaced by graphene's Fermi velocity v_F. The similarity between the disperison relation of ultra-relativistic particles ($pc \gg mc^2$) like light ($\varepsilon = \hbar\omega = \hbar ck = pc$) and that of the electrons in graphene make graphene a playground on which one could study various quantum relativistic effects. Because of this, graphene is often coined "CERN on one's desk" (CERN is in reference to the Large Hadron Collider in Switzerland).[11]

A.4 Corrections to the Linear Dispersion: Massive Dirac Fermions

Focusing only on the **K** point, the Hamiltonian can be more fully written as:

$$\begin{aligned}\mathbf{H}(\mathbf{q}) &= \hbar v_F \boldsymbol{\sigma} \cdot \mathbf{q} + m^* v_F^2 \sigma_z \\ &= \begin{bmatrix} m^* v_F^2 & \hbar v_F (q_x - i q_y) \\ \hbar v_F (q_x + i q_y) & -m^* v_F^2 \end{bmatrix},\end{aligned} \tag{A.33}$$

where essentially a so-called "mass term" is added to the original Hamiltonian and this new Hamiltonian now describes the "massive" Dirac Fermions of graphene. m^* is termed the global effective mass of graphene and what this serves to reflect is the different onsite energies experienced by the two sublattices. The eigenenergy of the "massive" Hamiltonian is:

$$\varepsilon(q) = \pm\sqrt{\left(\hbar v_F |\mathbf{q}|\right)^2 + \left(m^* v_F^2\right)^2}, \tag{A.34}$$

making this closer to Einstein's energy-momentum relation (the use of σ_z is required for this to work, otherwise it will be Klein tunneling problem where electrons will be immune to the energy barrier introduced. Klein tunneling is covered in the main chapters of this thesis). Here we can see the reason the additional term is referred to as the mass term: it precisely behaves as the rest mass of the system. And what this mass term does is open a bandgap at the Dirac point of graphene with a band gap energy of:

$$\Delta = 2 m^* v_F^2. \tag{A.35}$$

A.5 Additional Remarks on the Graphene Hamiltonian

A general remark on the Hamiltonian of graphene: the intrinsic full Hamiltonian is time-reversal and inversion symmetric. In any crystal lattice, the periodic potential always forces a bandgap opening at the high symmetry points in the BZ edges where bands cross[186]. There is an exception when said band crossings are protected by symmetries such as the time reversal and inversion symmetry of graphene. One popular saying goes as: symmetry breaking lifts degeneracy. And so, if symmetry is maintained, degeneracies such as the Dirac point can be protected.

The effect of time-reversal is to go back in time. A negation in time sensitive quantities such as momentum ($k \to -k$) and time evolution (which is always tied with an imginary unit: $it \to -it$) serves to backtrack in time. As for inversion, we are simply inverting the spatial coordinates ($x \to -x$, which implies also $k \to -k$ since real and reciprocal space directions are tied) and in graphene's case this amounts to simply switching sublattice A and B. These two together enforce the existence of a gapless mode (the Dirac point at the BZ corner) in graphene. The conditions forced onto the Gr Hamiltonian are:

$$\mathbf{H}(\mathbf{k}) = \mathbf{H}^*(-\mathbf{k}), \tag{A.36}$$

$$\mathbf{H}(\mathbf{k}) = \mathbf{H}(-\mathbf{k}). \tag{A.37}$$

Here we stress that the conditions refer to the full Hamiltonian (implied by the usage of $\mathbf{k}$ instead of $\mathbf{q}$) and not the low-energy-approximation Hamiltonian. All of this information is captured in the wealth of information provided throughout the bandstructure derivation. The moment a symmetry-breaking term is introduced into the Hamiltonian, the above conditions will no longer hold and the

band crossing at the Dirac cone will no longer be protected. A simple case of breaking time reversal symmetry is the application of a magnetic field and a possible case of inversion symmetry breaking is when we put Gr on a piece of hexagonal-Boron Nitride (hBN) with a small or no rotation angle with respect to one another. In this case, the periodic potential coming from the hBN lattice causes a difference in potential felt by sublattice A and B. The key here is to be able to induce a short range potential variation that has a length scale close to the carbon to carbon distance.

A.6 Graphene's Wavefunction around the Dirac Point: A Pseudo-Spin

The graphene Hamiltonian around the Dirac points (equation A.27) can be rewritten in polar coordinates as:

$$\begin{aligned}\mathbf{H}(\mathbf{q}) &= \hbar v_F|\mathbf{q}| \begin{bmatrix} 0 & e^{-i\theta_{\mathbf{q}}} \\ e^{i\theta_{\mathbf{q}}} & 0 \end{bmatrix}, \\ \mathbf{H}(\mathbf{q}') &= \hbar v_F|\mathbf{q}'| \begin{bmatrix} 0 & e^{i\theta_{\mathbf{q}'}} \\ e^{-i\theta_{\mathbf{q}'}} & 0 \end{bmatrix},\end{aligned} \tag{A.38}$$

where $|\mathbf{q}|^2 = q_x^2 + q_y^2$ and the angle is given by:

$$\theta_{\mathbf{q}} = \tan^{-1}\left(\frac{q_x}{q_y}\right), \tag{A.39}$$

in which case it is easy to see that the respective eigenfunctions (which are coefficients of the wavefunctions) of the Hamiltonians in the linear regime around the Dirac points in momentum space are:

$$\begin{aligned}\boldsymbol{\Psi}(\mathbf{q}) &= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-i\theta_{\mathbf{q}}/2} \\ \pm e^{i\theta_{\mathbf{q}}/2}\end{bmatrix}, \\ \boldsymbol{\Psi}(\mathbf{q}') &= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{i\theta_{\mathbf{q}'}/2} \\ \pm e^{-i\theta_{\mathbf{q}'}/2}\end{bmatrix},\end{aligned} \tag{A.40}$$

where the positive is for the electron states in the conduction band and negative is for the hole states in the valence band. The bases of the two wavefunctions (which are the spatial component of the wavefunctions) are: $\{\Phi_{A,\mathbf{K}}, \Phi_{B,\mathbf{K}}\}$ and $\{\Phi_{A,\mathbf{K}'}, \Phi_{B,\mathbf{K}'}\}$ (given by equation A.8 and A.9 with $\mathbf{k} \to \mathbf{q}$ and $\mathbf{q}'$ respectively for $\mathbf{K}$ and $\mathbf{K}'$ valleys). Note again that the two wavefunctions are linked via a time reversal operation just like the Hamiltonians around the Dirac points, i.e., $\boldsymbol{\Psi}(\mathbf{q}') = \boldsymbol{\Psi}^*(\mathbf{q})$. As for the wavefunction's coefficients in real space representation:

$$\begin{aligned}\boldsymbol{\Psi}(\mathbf{r}) &= \langle \mathbf{r}|\boldsymbol{\Psi}\rangle \\ &= \int_0^{\mathbf{k}_F} d\mathbf{k}\, \langle \mathbf{r}|\mathbf{k}\rangle \langle \mathbf{k}|\boldsymbol{\Psi}\rangle \\ &\approx \int_0^{\mathbf{q}_F} d\mathbf{q}\, \langle \mathbf{r}|\mathbf{q}\rangle \langle \mathbf{q}|\boldsymbol{\Psi}\rangle + \int_0^{\mathbf{q}'_F} d\mathbf{q}'\, \langle \mathbf{r}|\mathbf{q}'\rangle \langle \mathbf{q}'|\boldsymbol{\Psi}\rangle \\ &= \frac{1}{\sqrt{2}}\left(\int_0^{\mathbf{q}_F} d\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{r}} \begin{bmatrix} e^{-i\theta_{\mathbf{q}}/2} \\ \pm e^{i\theta_{\mathbf{q}}/2}\end{bmatrix} + \int_0^{\mathbf{q}'_F} d\mathbf{q}' e^{i\mathbf{q}'\cdot\mathbf{r}} \begin{bmatrix} e^{-i\theta_{\mathbf{q}'}/2} \\ \pm e^{i\theta_{\mathbf{q}'}/2}\end{bmatrix}\right).\end{aligned} \tag{A.41}$$

In the above steps, the states that contribute to the total wavefunction are occupied states with momentum up to $\mathbf{k}_F$ and we have assumed the Fermi level is low enough such that the Fermi surface can

be thought of as two separate circles. It is more useful to retain the momentum information, then the coefficients of the wavefunction with a specified momentum at the $\mathbf{K}$ and $\mathbf{K}'$ valleys respectively are:

$$\begin{aligned}\mathbf{\Psi}(\mathbf{q},\mathbf{r}) &= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{-i\theta_{\mathbf{q}}/2} \\ \pm e^{i\theta_{\mathbf{q}}/2}\end{bmatrix} e^{i\mathbf{q}\cdot\mathbf{r}}, \\ \mathbf{\Psi}(\mathbf{q}',\mathbf{r}) &= \frac{1}{\sqrt{2}}\begin{bmatrix} e^{i\theta_{\mathbf{q}'}/2} \\ \pm e^{-i\theta_{\mathbf{q}'}/2}\end{bmatrix} e^{i\mathbf{q}'\cdot\mathbf{r}}.\end{aligned} \tag{A.42}$$

Here we see that the wavefunctions' coefficients are merely plane waves (with the phase information regarding the sublattices retained), which is expected because the Gr Hamiltonian commutes with momentum.

The structure of the wavefunction is like that of a spinor desribing an electron spin where the two vector components are related and defined by a rotation angle which describes the spin direction. However, in graphene's case, the two components correspond the individual sublattices (top component refers to the A sublattice and bottom component refers to the B sublattice) and to distinguish graphene's wavefunction from the actual spinor, it is instead called the "psuedo-spin". The electrons are equally distributed between sublattice A and B except that there is a relative quantum mechanical phase difference between the sublattices. The existence of the sublattice phase difference may have implications when in the quantum limit during the hopping between sublattices, electrons' wavefunctions may constructively or destructively interfere. More importantly, however, is that there exists a non-trivial Berry's phase (or geometric phase) coming from the division of two in the phase. This means that a full rotation in momentum space about the Dirac point will have the wavefunction acquire a π phase as opposed to the expected 2π phase (and hence no phase). This Berry phase is re-

sponsible for many outstanding transport behaviors unique to graphene. One such is the suppressed backscattering. We can see this by using Fermi's Golden Rule (accurate only to first order) to calculate the rate of scattering between the initial state with momentum $\mathbf{q_i}$ to the final state with momentum $\mathbf{q_f}$ mediated by a long range scattering potential $V(\mathbf{r})$:

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar} \left| \left\langle \mathbf{\Psi}(\mathbf{q}_f, \mathbf{r}) \right| V(r) \left| \mathbf{\Psi}(\mathbf{q}_i, \mathbf{r}) \right\rangle \right|^2 \delta(\varepsilon_f - \varepsilon_i). \tag{A.43}$$

The delta function is a restatement of energy conservation. The transition matrix element using graphene's wavefunction takes the following form:

$$\left| \left\langle \mathbf{\Psi}(\mathbf{q}_f, \mathbf{r}) \right| V(r) \left| \mathbf{\Psi}(\mathbf{q}_i, \mathbf{r}) \right\rangle \right| = |V(\mathbf{q}_i - \mathbf{q}_f)| \cos(\theta_{\mathbf{q}_i, \mathbf{q}_f}/2), \tag{A.44}$$

where $\theta_{\mathbf{q}_i, \mathbf{q}_f}$ is the angle between the initial and final states in momentum space with the Dirac point as the origin. In the case of backscattering, $\theta_{\mathbf{q}_i, \mathbf{q}_f} = \pi$, the transition matrix element vanishes and the scattering rate (equation A.43) becomes zero— backscattering is suppressed. Note that the scattering potential must be long range for the above to be valid. Long range here means much larger than the lattice constant of graphene, or equivalently in momentum language, $\mathbf{q} \ll \mathbf{K}$ (momentum being inversely proportional to real space). This makes sense since if momentum scattering is large, intervalley ($\mathbf{K}$ and $\mathbf{K}'$) scattering becomes possible which renders equation A.43 invalid.[187,188,189]

More on the wavefunction is that the electron and holes are helical because they are eigenfunctions

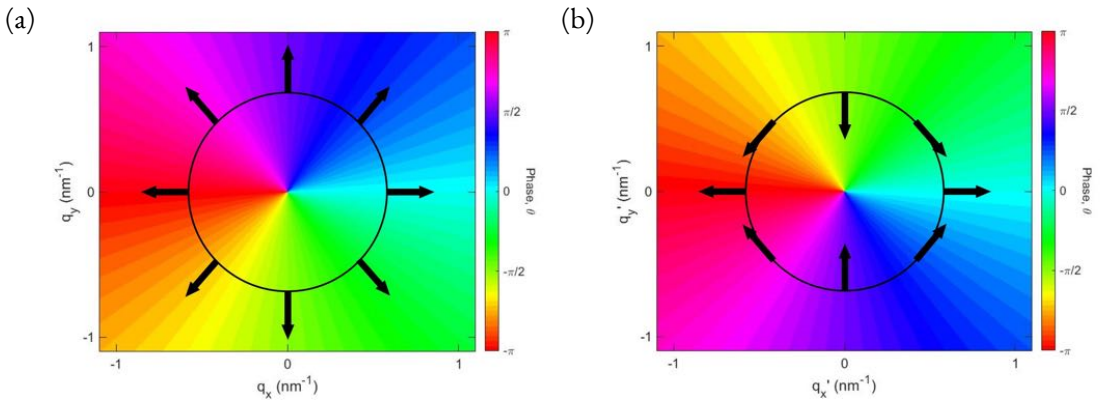


Figure A.3: Illustration of the evolution of graphene pseudo-spin (specifically of the conduction bands of the $\mathbf{K}$ and $\mathbf{K}'$ valleys. (a) The pseudo-spin of the conduction band of the $\mathbf{K}$ valley rotates counter-clockwise around the $\mathbf{K}$ point, corresponding to right-handed (positive) helicity. (b) The pseudo-spin of the conduction band of the $\mathbf{K}'$ valley rotates clockwise around the $\mathbf{K}'$ point, corresponding to left-handed (negative) helicity.

of the helicity operator:

$$\boldsymbol{\sigma} \cdot \frac{\mathbf{q}}{|\mathbf{q}|} \boldsymbol{\Psi}(\mathbf{q}) = \pm \boldsymbol{\Psi}(\mathbf{q}), \tag{A.45}$$

$$\boldsymbol{\sigma} \cdot \frac{\mathbf{q}'}{|\mathbf{q}'|} \boldsymbol{\Psi}(\mathbf{q}') = \mp \boldsymbol{\Psi}(\mathbf{q}'), \tag{A.46}$$

where the eigenvalue (or helicity) is positive for electrons and negative for holes. The helicity is also inverted in the opposite valley. This equation tells us that the pseudo-spin of graphene is locked with its momentum with respect to the Dirac point (see figure A.3). This helicity becomes crucial in instances such as Klein tunneling which will be touched upon in the main chapters of the thesis.

A.7 Low-Energy Density of States

The density of states in d dimensions is given by:

$$g(\varepsilon) = \frac{g}{L^d} \sum_{i=0}^{N} \delta(\varepsilon - \varepsilon(\mathbf{k}_i)) \tag{A.47}$$

where L is the length of the system and g is the degeneracy of the system (not to be confused with density of states which is energy dependent and will always be referred to as $g(\varepsilon)$ and not g). In the continuum limit, i.e., $\Delta k \to dk$ and for $d = 2$, we have:

$$g_{\text{2D}}(\varepsilon) = \frac{g}{4\pi^2} \int_0^\infty dk\, 2\pi k \cdot \delta(\varepsilon - \varepsilon(\mathbf{k})). \tag{A.48}$$

To find the low energy density of states of graphene, i.e. if we only care about states near the two Dirac cones, substitute $g = 2 \times 2 = 4$ for the spin and valley degeneracy, and use graphene's linear dispersion given by equation A.31 ($\varepsilon = \hbar v_F |\mathbf{q}|$), where we make the change of $\mathbf{k} \to \mathbf{q}$, we acquire this simple expression:

$$g_{\text{Gr}}(\varepsilon) = \frac{2\varepsilon}{\pi \hbar^2 v_F^2}, \tag{A.49}$$

in units of number per energy per area. The low energy density of states of graphene has a linear dependence on energy, which is strikingly different from the constant density of states ($g_{\text{2D,parabolic}}(\varepsilon) = m/(\pi\hbar^2)$) for systems with a parabolic dispersion.

A.8 Effective Mass

Because the electrons near the Dirac point follow a linear dispersion, the effective mass cannot be calculated by the traditional definition ($m^*_{\text{parabolic}} = (\partial^2\varepsilon/\partial k^2)^{-1}$, assuming isotropic parabolic dispersion relation, otherwise effective mass is given by a tensor) that is only appropriate for materials with parabolic dispersions. Instead, when people speak of the effective mass of graphene, they implicitly refer to the cyclotron effective mass which is the mass deduced when measuring the material's cyclotron resonance frequency:

$$\omega_c = \frac{eB}{m^*}, \tag{A.50}$$

where B is the applied magnetic field and m^* is the cyclotron effective mass. Because the cyclotron frequency is determined by the period of a closed orbit in momentum space:[190]

$$\tau_c = \frac{\hbar^2}{eB}\frac{\partial A(\varepsilon)}{\partial \varepsilon}, \tag{A.51}$$

where $A(\varepsilon)$ is the area in momentum space bound by the closed path formed by the electron in orbit with energy ε, the cyclotron effective mass can be written as:

$$m^* = \frac{\hbar^2}{2\pi}\frac{\partial A(k)}{\partial k}\frac{\partial k}{\partial \varepsilon}, \tag{A.52}$$

where we have taken $\omega_c = 2\pi/\tau_c$. Applying this to graphene's linear dispersion (equation A.31) and using $A(k) = \pi k^2$, graphene's effective mass near the Dirac point is simply:

$$m^* = \frac{\hbar k}{v_F}. \tag{A.53}$$

Note that this is can be easily acquired by treating graphene's dispersion relation the same as relativistic energy:

$$\varepsilon = pv_F = \hbar k v_F, \tag{A.54}$$

where $p = m_{\text{rel}} v_F$ and where v_F has replaced the role of speed of light c, in which case the relativistic mass (m_{rel}) is equivalent to the cyclotron effective mass.

A.9 Source Code for the Dispersion Relation of Graphene with 1D Potential and Magnetic Field

Python code developed by Jean Damien Pillet for the dispersion relation of graphene with a 1D potential and magnetic field. This simulation follows the tight-binding approach. Source codes regarding other simulations found in this thesis are all performed with Matlab and can be provided upon reasonable request.

```
from numpy import*
from scipy.linalg import*
from matplotlib.pyplot import*
import time
import scipy.sparse.linalg as slinalg
import scipy.sparse as sparse
import csv

no=400
kyo=arange(-51,51)*0.02
b=0
kappao=0.2*arange(0,51)*0.15*20/50
do=10
DOSlist=[]
LDOSlist=[]
EFlist=[]
MyCoeflist=[]
Energy=[]

start = time.time()

for kappa in kappao:

Mylist=[]
Mylist2=[]
MyCoef2=[]

  for ky in kyo:
      Coupling=-1j*(-eye(2*no,k=no+1)+(1+ky)*eye(2*no,k=no)-\
```

```
        (-eye(2*no,k=-(no+1))+(1+ky)*eye(2*no,k=-no)))
      Magnetic=-1j*b*concatenate((concatenate((diag(arange(no)*0),\
        diag(arange(no)+1-no/2)),axis=0),concatenate((-diag(arange(no)+1-no/2),\
        diag(arange(no)*0)),axis=0)),axis=1)
      Potential=kappa*diag(do**2/((concatenate((arange(no)+1,\
        arange(no)+0.5))-no/2)**2+do**2))
      temp=eigh(Coupling-Magnetic+Potential)
      Mylist.append(temp[0])
      Mylist2 = concatenate((Mylist2,temp[0]),axis=0)
      MyCoef2 = concatenate((MyCoef2,dot(abs(transpose(temp[1]))**2,\
        concatenate((1/((arange(no).transpose()-no/2)**2+do**2),\
        1/((arange(no).transpose()-no/2)**2+do**2)),axis=0))),axis=0)
      if ky*1-floor(ky*1) == 0:
        print(ky)

DOS=[]
Myw=0.01

Energy.append(Mylist2)
for En in arange(-400,400)*0.0025:
DOS.append(sum(1/((En-Mylist2)**2+Myw**2)))

DOSlist.append(DOS)

LDOS=[]
Myw=0.01

for En in arange(-400,400)*0.0025:
LDOS.append(sum(MyCoef2*1/((En-Mylist2)**2+Myw**2)))

LDOSlist.append(LDOS)

EF=[]
Myw=0.01

for En in arange(-400,400)*0.0025:
EF.append(sum(MyCoef2*1/pi*arctan((En-Mylist2)/Myw)))

EFlist.append(EF)

MyCoeflist.append(MyCoef2)

print(kappa)
plot(DOS/DOS[0])
plot(LDOS/LDOS[0])
show()
plot(EF)
show()
ylim(-1,1)
plot(sort(Mylist))
show()
```

B

Overview of the Electronic Properties of Carbon Nanotubes

The single-wall carbon nanotube (SWCNT) is a piece of graphene (Gr) rolled up into a hollow tube. In this appendix, the electronic properties of carbon nanotube (CNT) will be derived based off of this idea as is conventionally done. As such, only the single particle picture will be addressed. The novel many-body effects tied with CNTs are mentioned throughout the main text (especially the beginning of chapter 2). The reader is referred to their citations for more information.

B.1 Briefly about the Carbon Nanotube

There are a finite number of ways to roll up Gr and produce a unique SWCNT. Figure A.3 is the so-called periodic table of CNTs which shows the many ways of wrapping a piece of Gr to form CNTs of different chiralities. Chirality is a concept used to characterize and distinguish the numerous types of SWCNT. More information on this concept will be provided later. The periodic table shown is truncated. In reality it extends to chirality indices that are much higher and to values that are, theoretically speaking, uncapped. However, based on the Kataura plot, which is a plot that experimentally relates nanotube diameters and their bandgap energies, as well as being the most extensive record of SWCNT data, the highest reported observable chirality index so far is $(n, m) = (37, 1)$, while, the diameter of the largest reported SWCNT is about 2.9995 nm with a chirality index of $(33, 9)$.[191]

As suggested above, SWCNT has an extremely small diameter. Typically we expect SWCNT to be on the order of 1nm in radius. More remarkably, the length of a SWCNT can reach up to half a meter, making its aspect ratio of $132,000,000$ to 1 unequalled by any other material in the world.[192] This huge anisotropy certainly qualifies carbon nanotubes in general as quasi-one-dimensional objects. Even more importantly, its small radius confines electrons to such a degree that its energy subbands are separated on the order of one to three electron volts (eV) (as we will see later). This is what truly sets the carbon nanotube apart from other supposedly quasi-one-dimensional objects where typically many energy subbands naturally contribute to electron transport. On top of this, carbon nanotube has the inherent properties of its mother material, graphene. Among the properties, for metallic SWCNTs specifically (we will see later why this does not apply to semiconducting CNTs), their suppressed backscattering due to its non-trivial Berry's phase make them exceptional candidates for high mobility electronics.

B.2 Carbon Nanotube Lattice

The carbon nanotube (CNT) retains the graphene (Gr) lattice constant and primitive lattice vectors $\mathbf{a_1}$ and $\mathbf{a_2}$. (Note that in Figure B.1 the lattice shown is rotated 30° away from the lattice shown in figure A.1a for the sake of presentation convenience. The primitive basis vectors stay unchanged.) The additions to the geometrical parameters for CNTs due to its more complex structure are: the chiral vector $\mathbf{C}$, the translation vector $\mathbf{T}$, and the chiral angle θ. The chiral vector describes how the Gr is wrapped to become a CNT. It points from one lattice site to another one that it will connect to upon wrapping the graphene. The vector is:

$$\mathbf{C} = n\mathbf{a_1} + m\mathbf{a_2}, \tag{B.1}$$

where n and m take on positive integer values with $0 \leq m \leq n$, and together they form the chirality index (n, m), which uniquely identifies each CNT. The circumference of a CNT is simply given by $|\mathbf{C}|$. The translation vector (or unit lattice vector) defines the periodicity of the carbon nanotube along its axis. It has the following form:

$$\mathbf{T} = t_1\mathbf{a_1} + t_2\mathbf{a_2}, \tag{B.2}$$

for integer values of t_1 and t_2. For a given CNT chirality their allowed values are given by the condition:

$$\mathbf{C} \cdot \mathbf{T} = 0. \tag{B.3}$$

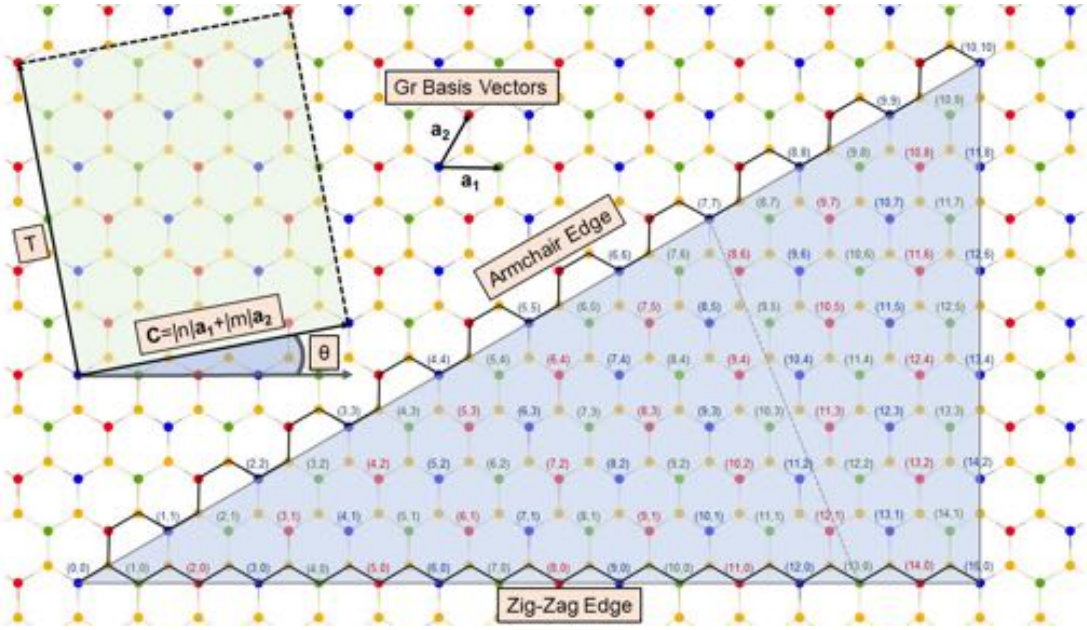


Figure B.1: The periodic table of carbon nanotubes. Shown is the graphene direct lattice, which is the same as figure A.1a except it is rotated by 30°. At the top shows the graphene basis vectors which is also used as the basis for the carbon nanotube's chiral vector. A carbon nanotube is formed by picking an origin (in the case of this schematic it is the blue atom in the bottom left labeled $(0,0)$) and folding it to another atom in the same sublattice (colored blue, green or red), hence creating a hollow cylindrical structure: the carbon nanotube. The atom picked for the origin to fold to has coordinates (n,m), which defines the chirality index. $\mathbf{C}$ is the chiral vector, $\mathbf{T}$ is the translation vector (or unit lattice vector) and θ is the chiral angle. The bottom edge in bold black is the zig-zag edge and the other bold black line defines the armchair edge. The chiral angle is defined with respect to the zig-zag edge, meaning that carbon nanotubes with zig-zag edges necessarily have $m = 0$ and $\theta = 0$. Following the same line of reasoning, carbon nanotubes with armchair edges necessarily have $n = m$ and $\theta = \pi/6$. The grey dashed line demarcates nanotube diameters of about 1nm. Finally, metallic nanotubes are formed whenever blue atoms are chosen for folding and semiconductors are formed otherwise.

Lastly, the chiral angle defines the angle between $\mathbf{C}$ and $\mathbf{a_1}$ which is just:

$$\cos(\theta) = \frac{\mathbf{C}\cdot\mathbf{a_1}}{|\mathbf{C}||\mathbf{a_1}|} = \frac{2n+m}{2\sqrt{n^2+nm+m^2}}. \tag{B.4}$$

The wavevectors defining the Brillouin zone (BZ) of the CNT which contains all the unique crystal momenta are given by the reciprocity condition just like how it was done for Gr. $\mathbf{K}$ is conventionally

used to denote the general reciprocal lattice vector but to avoid confusion with the K and K' valley points of Gr, $\mathbf{G}$ (which is also commonly used) will be used instead to denote the general reciprocal lattice vector. The reciprocal lattice vectors which are algebraically derived from the reciprocity relation are:

$$\begin{aligned}\mathbf{G}_{\parallel} &= \frac{1}{N_H}(m\mathbf{b_1} - n\mathbf{b_2}),\\ \mathbf{G}_{\perp} &= \frac{1}{N_H}(-t_2\mathbf{b_1} + t_1\mathbf{b_2}),\end{aligned} \tag{B.5}$$

where $\mathbf{G}_{\parallel}$ and $\mathbf{G}_{\perp}$ correspond to the wavevectors pointing along and around the CNT respectively and N_H is the number of hexagons (the Gr Wigner-Seitz cell) in a CNT unit cell:

$$N_H = \frac{|\mathbf{C} \times \mathbf{T}|}{|\mathbf{a_1} \times \mathbf{a_2}|}. \tag{B.6}$$

A more detailed description of the CNT lattice is given by chapter 4 of *Carbon Nanotube and Graphene Device Physics* by H.-S. P. Wong et al.[193]

B.3 Dispersion Relation of Carbon Nanotubes: Quantized Graphene

With all the CNT lattice information down, acquiring the dispersion relation is relatively straightforward. Once again, a carbon nanotube is a Gr wrapped into a cylinder. This means that what was originally a 2D momentum space effectively becomes 1D because of the imposed periodic boundary condition that exists in the newly formed azimuthal angle (in terms of cylindrical coordinates). The newly imposed periodic boundary condition is (as a requirement for the Bloch wave traversing around the CNT cylinder to be continuous):

$$e^{i\mathbf{k}_{\perp}\cdot\mathbf{C}} = 1, \tag{B.7}$$

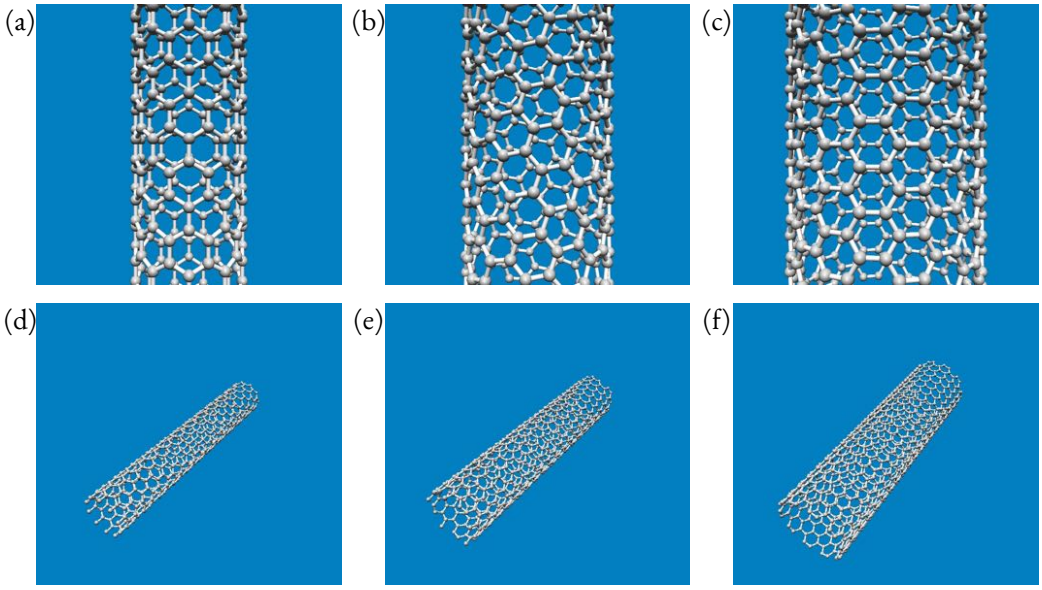


Figure B.2: The three different structural types of CNTs. Illustrations on the left column (a), (d) show a zig-zag CNT with chiral index $(10, 0)$, the middle column (b), (e) show a chiral CNT with chiral index $(10, 5)$, the right column (c), (f) show an armchair CNT with with chiral index $(10, 10)$. (a)-(c) show the top view of the CNTs. (d)-(f) show the slanted side view of the CNTs. Take note that the ends of the CNT indicate their respective structural type. Zig-zag and armchair are terms given to describe how the CNTs terminate. Chiral CNTs are ones that are neither zig-zag nor armchair types.)

where $\mathbf{k}_\perp$ is the wavevector perpendicular to the cylindrical surface of the CNT which describes the circumferential momentum and is forced to take on discrete values of:

$$k_\perp = \frac{2\pi j}{|\mathbf{C}|}, \quad j = 0, 1, ..., N_H - 1, \tag{B.8}$$

where j is an integer restricted to values that make the wavevector unique in reciprocal space. As for the wavevector along the axis of the CNT, it is given by the Born-von Karman (or macroscopic periodic) boundary condition on the Bloch wavefunctions that propagate along the axis of the CNT (of length

$\mathbf{L} = N_{uc}\mathbf{T}$ where N_{uc} is the number of unit cells along the CNT):

$$e^{i\mathbf{k}_{\parallel}\cdot N_{uc}\mathbf{T}} = 1, \tag{B.9}$$

which results in the set of wavevectors:

$$k_{\parallel} = \frac{2\pi l}{N_{uc}|\mathbf{T}|}, \quad l = 0, 1, ..., N_{uc} - 1, \tag{B.10}$$

where l is an integer restricted to values that make the wavevector to be in the first Brillouin zone (BZ). Due to $N_{uc} \gg 1$, the wavevector can be considered as a continuous variable within the CNT BZ that is conventionally centered symmetrically around zero:

$$k_{\parallel} = \left(-\frac{\pi}{|\mathbf{T}|}, \frac{\pi}{|\mathbf{T}|} \right). \tag{B.11}$$

Combining all the information above, the general reciprocal space vector (or wavevector) for a CNT is:

$$\begin{aligned} \mathbf{k} &= k_{\parallel}\frac{\mathbf{G}_{\parallel}}{|\mathbf{G}_{\parallel}|} + k_{\perp}\frac{\mathbf{G}_{\perp}}{|\mathbf{G}_{\perp}|} \\ &= k_{\parallel}\frac{\mathbf{G}_{\parallel}}{|\mathbf{G}_{\parallel}|} + \frac{2\pi j}{|\mathbf{C}|}\frac{\mathbf{G}_{\perp}}{|\mathbf{G}_{\perp}|}, \end{aligned} \tag{B.12}$$

for $-\pi/|\mathbf{T}| < k_{\parallel} < \pi/|\mathbf{T}|$ and $j = 0, 1, ..., N_H - 1$, where $\mathbf{G}_{\parallel}/|\mathbf{G}_{\parallel}|$ is the axial unit wavevector and $\mathbf{G}_{\perp}/|\mathbf{G}_{\perp}|$ is the circumferential unit wavevector. This expression tells us the allowed momenta of an arbitrary CNT with $k_{\parallel}$ being the linear momentum and j being the discretized angular momentum quantum number. Each value of j has a line cut that is $2\pi/|\mathbf{T}|$ long and exists somewhere in the

reciprocal space of graphene. Each of these line cuts corresponds to an individual 1D subband of the CNT. Figure B.3 shows examples of the line cuts in the reciprocal space of graphene and their corresponding energy subbands. To get these energy subbands, one simply imposes the wavevector restriction (equation B.12) onto the dispersion relation of graphene (equation A.23). One will find that the chirality index (n, m) defines a unique set of subbands. It is interesting to note that for the zig-zag CNTs shown in figure B.3, the bandstructures host flat energy bands. Recently, it was found that two layers of graphene twisted at a "magic" angle relative to one another modifies the double layer graphene bandstructure to host flat energy bands. As a result of these flat energy bands, graphene can demonstrate highly correlated electronic phases such as becoming a Mott insulator or a superconductor. In light of this, zig-zag CNTs can possibly open up opportunities for exploring 1D flat energy band physics.

B.4 Metallic Versus Semiconducting Carbon Nanotubes

A CNT with no bandgap such as the one shown in figure B.3b are metallic CNTs. Ones with a bandgap like the one shown in figure B.3a are semiconducting CNTs. As alluded to from the previous section, whether the bandstructure of a CNT has a bandgap or not depends purely on the line cuts in the reciprocal lattice of Gr— if a line cut crosses the Dirac point of Gr, the CNT is metallic, otherwise it is semiconducting. The condition for this is straightforward:

$$k_{\perp} = |\mathbf{\Gamma K}|\cos(\theta). \tag{B.13}$$

(See equation A.4, B.4 and B.8 for variable definitions) This condition ensures that at least a part of the line cut will overlap the Dirac point. Writing the equation out one will find that the equation can

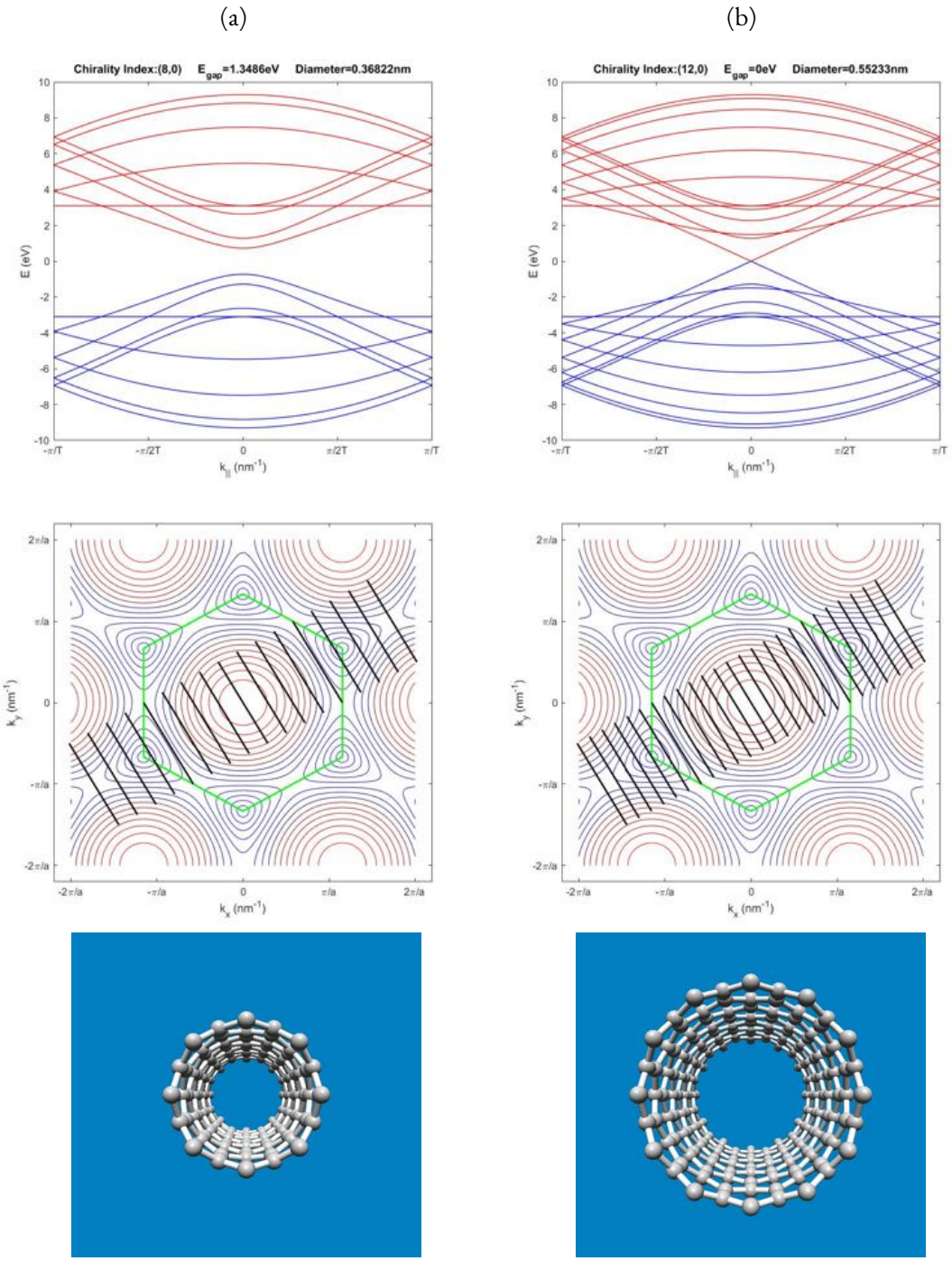


Figure B.3: Examples of zig-zag CNTs. Top: Dispersion relation. Middle: Fermi lines. Bottom: CNT end view. Left (a) and right (b) columns correspond to $(n, m) = (8, 0)$ and $(12, 0)$ respectively.

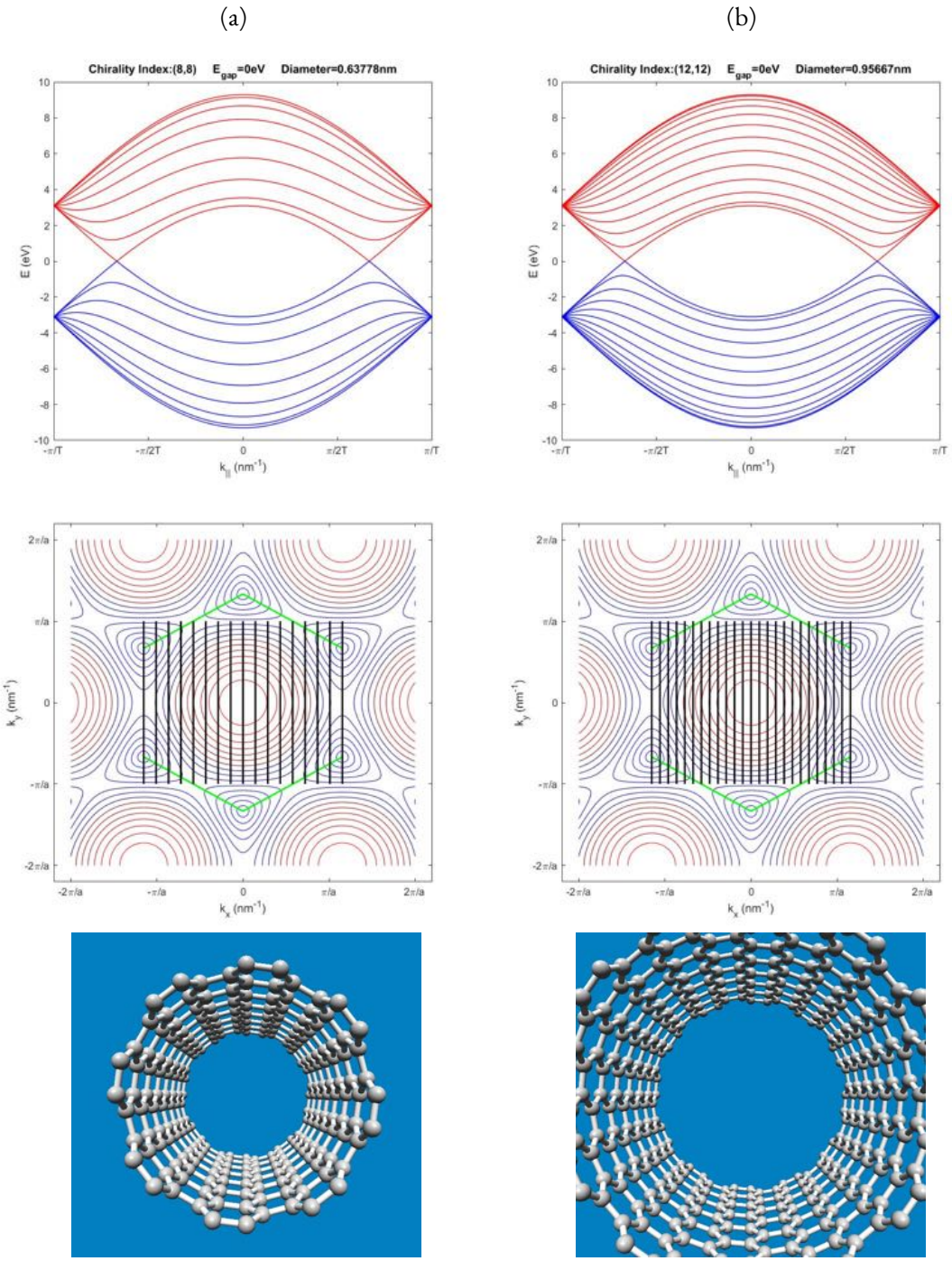


Figure B.4: Examples of armchair CNTs. Top: Dispersion relation. Middle: Fermi lines. Bottom: CNT end view. Left (a) and right (b) columns correspond to $(n, m) = (8, 8)$ and $(12, 12)$ respectively.

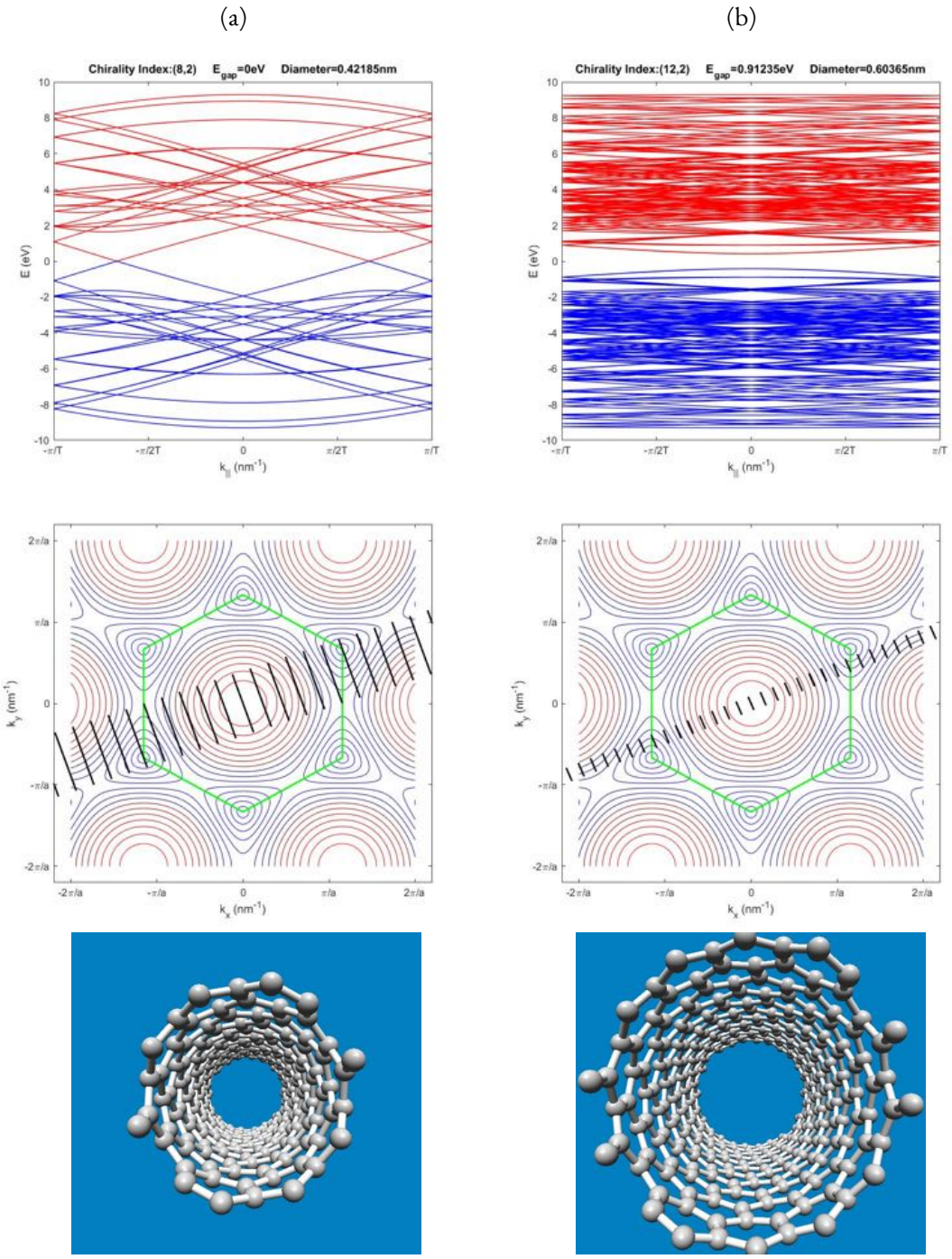


Figure B.5: Examples of chiral CNTs. Top: Dispersion relation. Middle: Fermi lines. Bottom: CNT end view. Left (a) and right (b) columns correspond to $(n, m) = (8, 2)$ and $(12, 2)$ respectively.

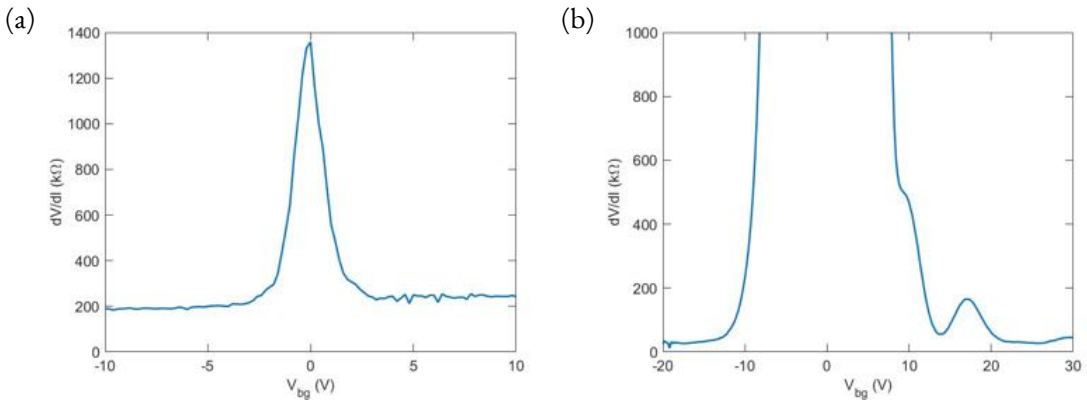


Figure B.6: The transport properties of a metallic and semiconducting CNT. (a) the resistance of a metallic CNT. It shows a peak resembling the Dirac peak of Gr. (b) the resistance of a semiconducting CNT. It has a bandgap where the CNT becomes completely insulating.

only be satisfied if and only if the following is true:

$$n - m = 3l, \quad l = 0, 1, 2, ... \tag{B.14}$$

for l being any integer including 0. This is the celebrated condition for when a CNT is metallic. The CNT is semiconducting whenever this condition is not satisfied. Looking more into this condition, one can deduce that if there is an equal probability of obtaining each chirality in a CNT growth, then a CNT has a $1/3$ chance of being metallic and $2/3$ chance of being semiconducting. The semiconducting CNT can be further classified into two types: type 1 and 2. Type 1 semiconducting CNTs are those with a chirality that satisfies $\text{mod}(n - m, 3) = 1$ and type 2 semiconducting CNTs are when $\text{mod}(n - m, 3) = 2$ (one can see that metallic CNTs are then just type 0 CNTs). There is a need for this type of classification because it turns out there is an underlying difference in physical behavior between the two. This will be further discussed in section B.6.

B.5 The Density of States and Other Important Quantities of Carbon Nanotubes

Starting from the general definition of density of states (equation A.45), one can show that in 1D, the density of states (DOS) can be rewritten in the following way:

$$g_{1\mathrm{D}}(\varepsilon) = \frac{g}{\pi}\left|\frac{\partial k}{\partial \varepsilon}\right|_{\varepsilon}, \tag{B.15}$$

where g is the degeneracy (again, this is not to be confused with density of states which is energy dependent and will always be referred to as $g(\varepsilon)$ and not g) and the expression is quite self-explanatory. Acquiring the DOS of different CNTs is just a matter of doing the calculation, and unfortunately, there is no way to obtain a universal analytical expression for all CNTs. There are, however, some approximated forms that are accepted as universal forms of DOS for the lowest energy subband(s) of both metallic and semiconducting CNTs. Here we just cite the results for reference.[194] Metallic CNTs' first energy subband has a universal DOS that goes as:

$$g_{\mathrm{CNT,M}}(\varepsilon_F) = g_{\mathrm{CNT}} = \frac{8}{\sqrt{3}a\pi\gamma_0} = \frac{8}{hv_F} \approx 2\,\mathrm{nm}^{-1}\mathrm{eV}^{-1}, \tag{B.16}$$

which is also a material constant and hence given the label g_{CNT}. $\gamma_0 \approx 3.1$ eV is the nearest neighbor hopping energy, $a \approx 2.46$ Å is the Gr Bravais lattice constant and v_F is the Fermi velocity given by equation A.28. The approximated universal form for the lowest subbands of semiconducting CNTs are derived from zig-zag CNTs. It turns out that zig-zag CNTs which has a chiral index of $(n, 0)$ are decent representatives of CNTs with chiral index (n, m). And so the semiconducting CNTs all have

a DOS that roughly follows:

$$g_{\text{CNT,SC}}(\varepsilon, j) \approx \frac{g_{\text{CNT}}\gamma_0}{\varepsilon_{\text{vh2}}} \frac{|\varepsilon|}{\sqrt{\varepsilon^2 - \varepsilon_{\text{vh1}}^2}}, \tag{B.17}$$

where the energies at the van Hove singularities (ε_{vh}) are given by:

$$\varepsilon_{\text{vh1}}(j) = \pm\gamma_0 \left| 1 + 2\cos\left(\frac{j\pi}{n}\right) \right| \xrightarrow{\text{zig-zag},(m=0)} \pm\gamma_0 \left| 1 + 2\cos\left(\frac{ja}{d}\right) \right|, \tag{B.18}$$

$$\varepsilon_{\text{vh2}}(j) = \pm\gamma_0 \left| 1 - 2\cos\left(\frac{j\pi}{n}\right) \right| \xrightarrow{\text{zig-zag},(m=0)} \pm\gamma_0 \left| 1 - 2\cos\left(\frac{ja}{d}\right) \right|, \tag{B.19}$$

with j being the subband index, a being the Gr Bravais lattice constant, and d being the CNT diameter given by:

$$d = \frac{|\mathbf{C}|}{\pi} = \frac{a\sqrt{n^2 + nm + m^2}}{\pi} \xrightarrow{\text{zig-zag},(m=0)} \frac{an}{\pi}. \tag{B.20}$$

Note that j the subband index (or angular momentum quantum number) starts at 0 which corresponds to the first linecut that crosses the Γ point at the center of the Gr BZ. The subband neighboring the bandgap (or the subband closest to the BZ edge of Gr, which is the lowest energy subband) always has the subband index of $j = \text{round}(2n/3)$.[195] As for why there are two van Hove singularities, each linecut or subband of the CNT roughly follows a sinusoidal form with a minimum and maximum (see any CNT bandstructure, e.g. in figure B.7 which is shown in the folded-zone scheme, one can imagine in the repeated-zone scheme each subband will have a minimum and maximum). $\varepsilon_{\text{vh1}}(j)$ refers to the the extremum that is closer to the Fermi level ($\varepsilon = 0$). This is the more relevant energy parameter as this gives us information regarding the bandgap. Equations B.16, B.17 and B.18 together constitute the universal DOS for CNTs. Regarding the semiconducting CNT DOS given by equation B.17, it is accurate up to the bottom of the third subband.

Knowing the DOS allows one to calculate the carrier density:

$$n_{\text{SC,M}} = \int_{-\infty}^{\infty} d\varepsilon f(\varepsilon) g_{\text{CNT,(SC,M)}}(\varepsilon, j), \tag{B.21}$$

where $f(\varepsilon)$ is the Fermi distribution. For a metallic CNT it is straightforward. Using the sommerfeld expansion, for a metallic CNT in the first subband we get:

$$n_{M,e(h)} = \pm g_{\text{CNT}} |(\varepsilon - \varepsilon_F)|, \tag{B.22}$$

but for the semiconducting CNT, the computation becomes grueling and must be done numerically. Below we cite the semi-empirical analytical formula for carrier density that takes into account the first two subbands. It has been shown that typically only up to the first two subbands are significant for the charge transport of CNTs, and so the equations below are mostly sufficient:[28]

$$n_e \approx \sum_{q=1}^{2} n_e(q) = \sum_{j=1}^{2} \frac{2N_0 e^{x_{\text{n}}}}{1 + Ae^{\alpha x_{\text{n}} + \beta x_{\text{n}}^2 + \gamma x_{\text{n}}^3}}, \quad x_{\text{n}} = \frac{2\varepsilon_{\text{F}} - \varepsilon_{\text{g}}}{2k_B T}, \tag{B.23}$$

$$n_h \approx \sum_{q=1}^{2} n_h(q) = \sum_{j=1}^{2} \frac{2N_0 e^{x_{\text{p}}}}{1 + Ae^{\alpha x_{\text{p}} + \beta x_{\text{p}}^2 + \gamma x_{\text{p}}^3}}, \quad x_{\text{p}} = \frac{-2\varepsilon_{\text{F}} - \varepsilon_{\text{g}}}{2k_B T}. \tag{B.24}$$

Above are the two-subband carrier densities for electrons and holes respectively. α, β, γ and A are fitting parameters. The fitting parameters that yield a high accuracy are:$\alpha = 0.88, \beta = 3.57 \times 10^{-3}, \gamma = -4.23 \times 10^{-5}, A = 0.66$. These parameters will be the ones used for the simulations in the chapter two of the main thesis. Continuing, q is the subband number (which merely serves as a counter for the number of subbands away from the Fermi level and is not to be confused with the actual subband index j. The subband index closest to the Fermi level is given by $j = \text{round}(2n/3)$[195]),

ε_g is the bandgap of the CNT and N_0 is the effective density of states given by:

$$N_0 = g_{\text{CNT}}\gamma_0 \frac{(2\varepsilon_{\text{vh1}} + k_B T)}{4\varepsilon_{\text{vh2}}} \sqrt{\frac{2\pi k_B T}{\varepsilon_{\text{vh1}}}} \approx \frac{g_{\text{CNT}}}{4} \sqrt{\pi k_B T \varepsilon_{\text{g}}}, \tag{B.25}$$

and where the simplification is made through the substitution: $\varepsilon_{\text{vh1}} + k_B T/2 \approx \varepsilon_{\text{g}}/2$ and $\varepsilon_{\text{vh2}} \approx 2\gamma_0$.[196] It follows that the approximated algebraic equation in the non-degenerate limit ($\mu < k_B T$) for a single subband is given by:[28]

$$n_{e(h)} = n_i e^{\pm \mu / k_B T}, \tag{B.26}$$

which we can see closely resembles the carrier density for 3D bulk semiconductors. Here, n_i is the intrinsic carrier density, ε_g is the bandgap of the CNT. The remaining important parameters of the CNT are as follows (where we have included the all-important effective mass, m):[195]

$$n_i = \frac{2 g_{\text{CNT}} \gamma_0 (\varepsilon_g + k_B T)}{\varepsilon_g + 4\gamma_0} \sqrt{\frac{\pi k_B T}{\varepsilon_g}} e^{-\varepsilon_g / 2 k_B T}, \tag{B.27}$$

$$\begin{aligned} \varepsilon_g &= 2\varepsilon_{\text{vh1}}(j) = 2\gamma_0 \left| 1 + 2\cos\left(\frac{\pi j}{n}\right) \right| \\ &\approx \frac{2 a_{cc} \gamma_0}{d}, \end{aligned} \tag{B.28}$$

$$m = \frac{4\hbar^2}{3\gamma_0 a^2} \frac{\varepsilon_{\text{g}}}{2\gamma_0 + \varepsilon_g}. \tag{B.29}$$

We caution the reader to be careful with a_{cc} and a. The former is the carbon-carbon bond length ≈ 1.42 Å and latter is the Gr Bravais lattice constant ≈ 2.46 Å. And yet again, we remind the reader that the subband index closest to the Fermi level is given by $j = \text{round}(2n/3)$[195]. Figure B.7 shows some examples of the DOS of different CNTs (computed numerically though equation B.15). The main feature one should know is that at the bottom of the subbands where the bands flatten out the

DOS show sudden peaks. These are the van Hove singularities (VHS). They are important as they determine the characteristic optical absorption and emission spectra that distinguish different CNTs from one another. Figure B.8 shows the zoomed-in DOS of a metallic and semi-conducting CNT. In these plots, the first three optical transitions are marked and denoted by E_{11}, E_{22}, E_{33} respectively. The nature of the optical transitions are captured by Rayleigh spectroscopy which is essential in our characterization of CNT and is further addressed in chapter 3 of the thesis. Note that the metallic CNT has a constant non-zero DOS within the E_{11} optical transition in contrast to the vanishing DOS within the bandgap of a semiconducting tube.

B.6 A Little Further into the Specifics of Carbon Nanotubes

If we look at the Kataura plot (see figure B.8), we can see that two branches are formed for each optical transition trend. The branches become more widely split as we go into higher optical transitions. The two branches which belong to type 1 and 2 CNTs also alternate in their relative positions within an optical transition trend. For example, $E_{11}^2 > E_{11}^1$ but $E_{22}^1 > E_{22}^2$. It turns out that type (or mod) 1 vs 2 semiconducting CNTs are expected to have drastically different optical behavior. This stems from the fact that the transition energies exist at opposing sides of the $\mathbf{K}$ point in the BZ for the two types of CNT. And according to both theoretical calculation and experimental observation, the optical absorption is inhomogeneous around $\mathbf{K}$.[197,198] This line of thought brings us to another important observation and that is the experimental optical transitions deviate from the theoretical one predicted by tight-binding theory. The bare-minimum theoretical prediction from the tight-binding model is:

$$E_{ii} \approx \frac{2ja_{cc}\gamma_0}{d}, \tag{B.30}$$

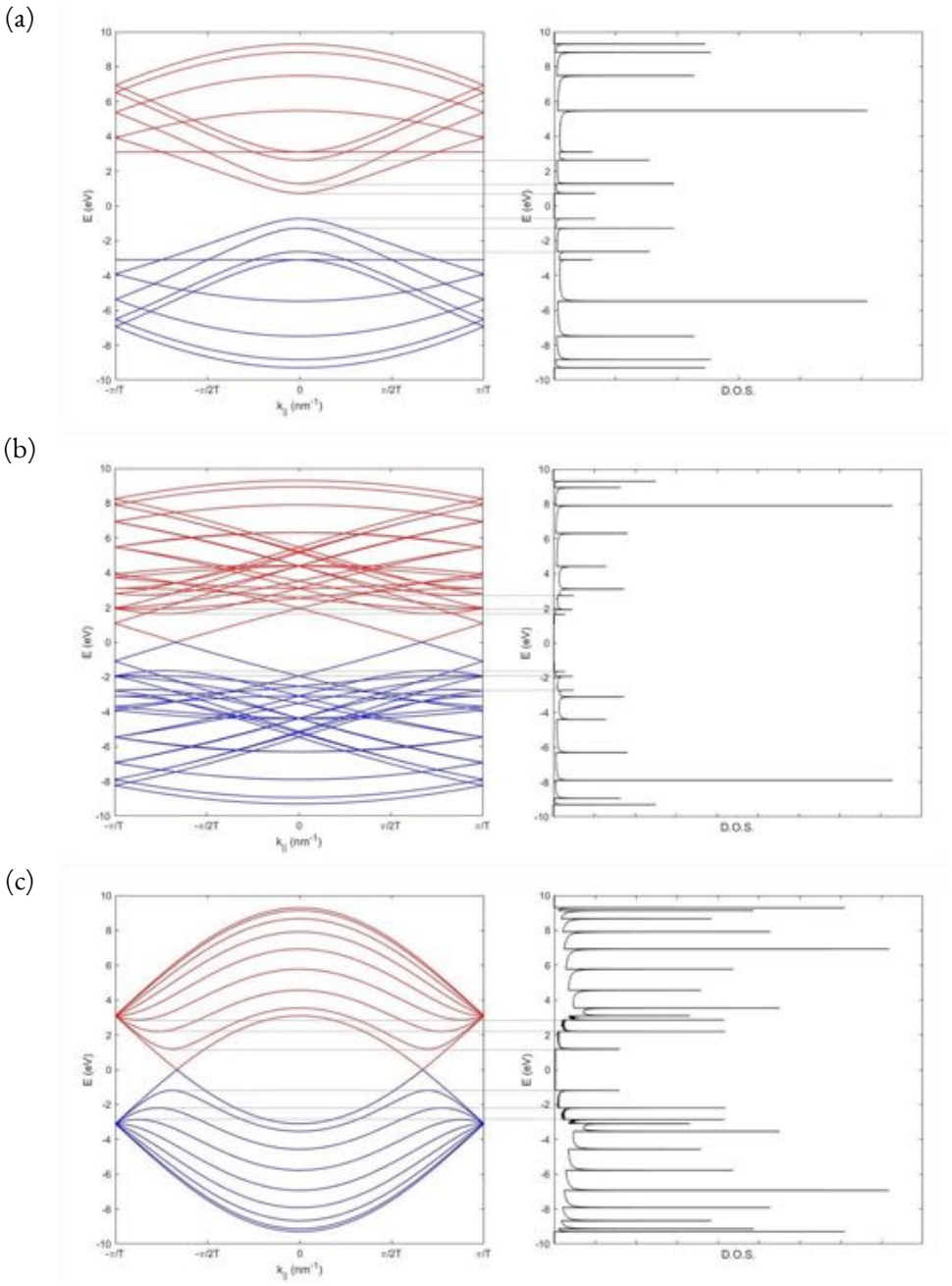


Figure B.7: The bandstructures of carbon nanotubes and their corresponding density of states (DOS). The first three Van Hove singularities of each band and their DOS are indicated by a gray dashed line. (a) A zig-zag $(8, 0)$ CNT. (b) A chiral metallic $(8, 2)$ CNT. (c) An armchair $(8, 8)$ CNT.

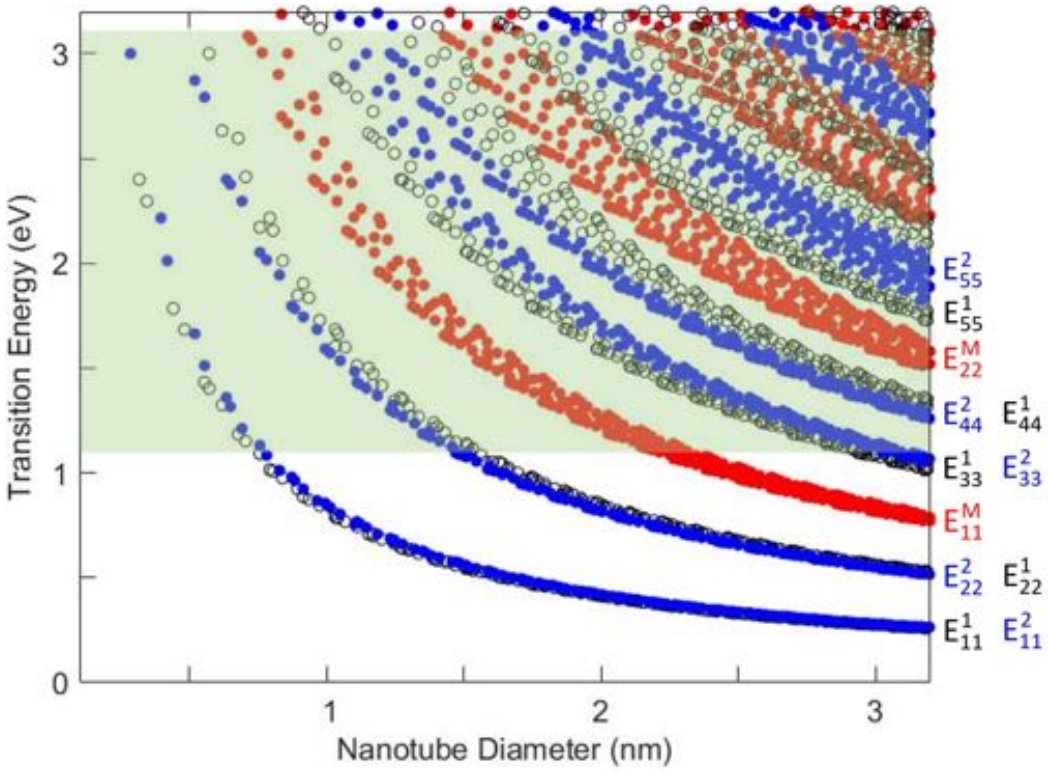


Figure B.8: The Kataura plot and depiction of the CNT transition energies. The Kataura plot is generated from the calculation of tight-binding theory. The parameters assumed are: $t_{\text{nn}} = 2.9\,\text{eV}$ (the nearest neighbor hopping energy) and $a_{\text{cc}} = 0.144\,\text{nm}$ (the carbon to carbon bond length)[59]. The black circles correspond to type 1 semiconducting CNTs, the blue points correspond to type 2 semiconducting CNTs and the red points correspond to metallic CNTs. These points are the transition energies of taking an electron from the valence band to the conduction band as shown in figure 3.7. The first five optical transitions are labeled. The shaded green area corresponds to the CNTs that our Rayleigh setup can see.

where j is an integer, which is tied with, i, the optical transition number, γ_0 is the nearest neighbor hopping parameter, and d is the CNT diameter. When $j = 1, 2, 4, 5...$ the corresponding transition energies are $E_{11}, E_{22}, E_{33}, E_{44}...$ which are the transition energies for semiconducting CNTs. When $j = 3, 6...$, the corresponding transition energies are $E_{11}^{M}, E_{22}^{M}...$ which are the transition energies for metallic CNTs. Up to the time of this writing, there are still ongoing efforts in describing the optical transitions more accurately. The possible causes for deviation are electron interactions and

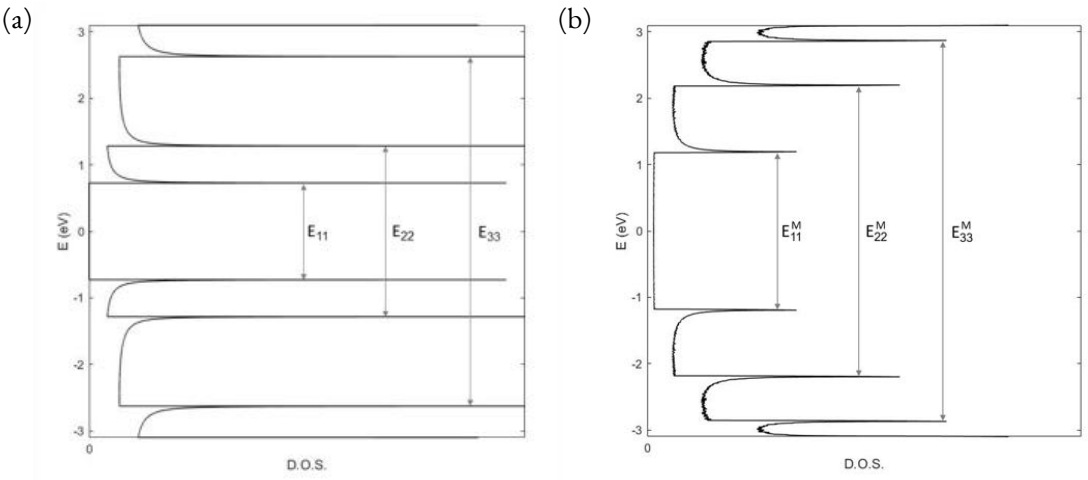


Figure B.9: Zoomed-in density of states plots of (a) a $(8, 0)$ semiconducting zig-zag carbon nanotube and of (b) a $(8, 8)$ metallic armchair carbon nanotube. E_{11}, E_{22}, E_{33} are the first three allowed optical transitions. These are the energies that are required to excite an electron from the valence band to the conduction band while obeying the selection rule. Note that the metallic carbon nanotube has a constant non-zero density of states around the charge neutrality point as opposed to the vanishing density of states in the bandgap of a semiconducting nanotube.

spin-orbit coupling due to curvature effects. So far, revised descriptions are all empirical models and here we stay away from going into the explicit equations. The interested reader is referred to the cited references.[60,61,62]

It is important to note that a metallic CNT shares the same electronic properties as Gr such as suppressed backscattering due to its non-trivial berry phase. When a particle in a right moving state scatters to a left moving state, the wavefunction of the left moving particle acquires a $\pi/2$ phase which destructively interferes with the incident right moving particle. This again is only because the right and left moving branches are on the two sides of the Dirac point. This, however, does not extend to semiconducting CNTs simply because the line cut does not run over the Dirac point, which means that the right and left moving branches are not directly on opposite sides of the Dirac point. This results in a phase less than $\pi/2$ which then leads to a non-zero backscattering probability. See figure B.6 for

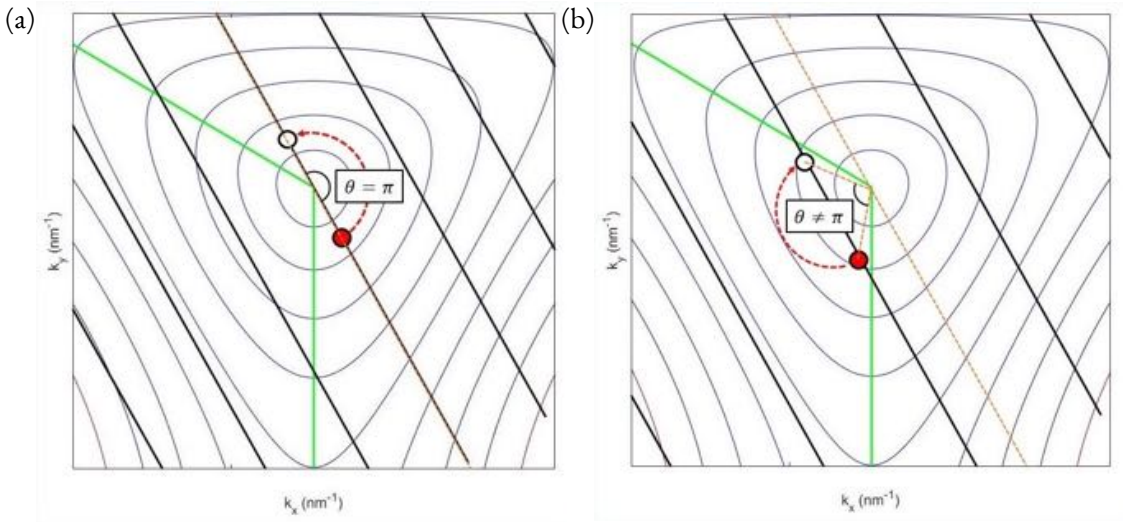


Figure B.10: Line cuts (straight black lines) of metallic and semiconducting carbon nanotubes in the reciprocal lattice space of graphene and around the Dirac point (zoomed in images of the contour plots from figure B.3). The green line marks the BZ boundary and the corner is the Dirac point. (a) Line cuts of a metallic CNT where we see the line cuts across the Dirac cone. A backscattering event will take a particle from the red colored state to the empty state and the relative angle between the states is π (with the Dirac point as the reference point), which when inserted into equation A.42 yields a vanishing scattering rate, meaning that backscattering is not possible. (b) Line cuts of a semiconducting CNT where no lines cut across the Dirac cone. A backscattering event is characterized by a non-π angle. The beige dotted line is drawn to help distinguish a π and non-π angle relative to the Dirac point. In this case, equation A.42 will yield non-zero which means there can be a backscattering rate.

a more pictorial explanation. Due to this discrepancy, metallic CNTs demonstrate ballistic electron transport which results in perfect transmission across the CNT, in contrast to semiconducting CNTs that have much higher resistance in general.

C

Overview of the Electronic Transport of Carbon Nanotube Quantum Dots

When electrons are confined into an area that is around or smaller than its Fermi wavelength (the de-Broglie wavelength at the Fermi level, $\lambda_F = 2\pi/k_f$), the electrons lose their wave-like property and exhibit quantization effects. This makes sense because roughly speaking the electrons are unable to complete a full period of an oscillation (strictly speaking this is not necessarily the case but what follows still holds). Quantities like wavelength and phase, which are essential to bring out wave-interference

effects, are no longer fit descriptions. The quantum dot is a case where electrons are confined in all dimensions, effectively rendering it an artifical atom. These electrons form bound (or resonant, localized, stationary) states and exhibit discrete energy levels, like an atom. Such is called the quantum confinement effect. More to this is that as a conducting object size shrinks, its capacitance with its surroundings becomes smaller. As this happens, the electrons in the object become less susceptible to the ongoings of the outside world and instead feel the presence of its neighbouring electrons much more strongly. This results in an increased Coulomb effect which causes electrons to become much more correlated with one another. Consequently, the electronic transport of the system is expected to be drastically unconventional. In this appendix, we will cover the key properties as well as the signature behaviors of a quantum dot.

C.1 Confining the Electrons in a Carbon Nanotube Even More

The carbon nanotube (CNT) already confines electrons to one dimension (1D) and as a result the electronic energies are a set of 1D subbands as we saw in appedix B.3. The remaining translationally-invariant axis, the longitudinal direction of the carbon nanotube, can also be confined in which case the 1D energy subbands will also become discretized. The method of confinement is naturally done through laying electrical contacts down onto the CNT. In the case where these contacts are not transparent, it becomes difficult for electrons to hop in and out of the CNT. By increasing this difficulty, the electrons eventually behave as though they are confined within the space between the contacts, i.e. they tend to live in the CNT longer than the rate at which they enter and leave the CNT, and this inevitably means the electrons within the CNT will feel each other much more (a more quantitative description will follow). Because the longitudinal unit wavevector follows the usual wavevector

condition for a particle-in-a-box problem which is

$$\Delta k_{||} = \frac{\pi}{L}, \tag{C.1}$$

where L is the confinement length of the CNT, we see that as L shrinks, the discrete nature of the wavevector becomes more apparent. Consequently, if we look at a CNT with a linear dispersion (any metallic CNT or a degenerately doped CNT where even a semiconducting CNT recovers its linear dispersion), the energy spectrum also becomes discretized into units of energy:

$$\Delta\varepsilon_{\text{conf}} = \hbar v_F \Delta k_{||} = \frac{\hbar v_F \pi}{L}. \tag{C.2}$$

This is called the confinement energy which is a direct result of the longitudinal confinement of the CNT. There are points to make here. First, this is only an approximation of the confinement energy. In reality the confinement energy depends a lot on the type of confinement. For example, there can be hard and soft confinement and each yields different confinement energies. The above expression is only used to roughly find out the length of the CNT quantum dot and is best suited for hard confinement (ensuring the accuracy of the momentum discretization condition) and when the CNT had a linear dispersion prior to discretization as suggested above. But even so, there are early experimental reports that suggest under hard confinement and for linear dispersion the confinement energy goes as $\Delta\varepsilon_{\text{conf}} \approx \hbar v_F/2L$ instead.[29,30,31] Second, the natural thought of a quantum dot can mislead one to think that it has to be quasi-zero dimensional, i.e., confined and equally small in all dimensions. This is vague. Regarding confinement there is a rough condition that certainly needs to be fulfilled (addressed in the next section). As for the smallness— this is where the ambiguity comes in as sizes are all relative. In literature, one would often see the term 1D quantum dot which is sometimes labeled

semi-inappropriately as a quantum wire. What a 1D quantum dot means is that electrons are confined in all directions but the aspect ratio of the confined space is as though it were 1D. A CNT quantum dot is almost always used as a 1D quantum dot. This is because the longitudinal confinement length is hardly ever as small as its typical diameter of 1nm. And this brings us to the third point. The CNT devices in this thesis are all longer than $\sim$ 200nm. This gives us $\Delta\varepsilon_{\text{conf}} \lesssim$ 1meV. As we will see very soon, the characteristic energy in a quantum dot system is the charging energy (ε_C) and it is typically tens of meV which is much greater than the confinement energy. As a result, the contribution from $\Delta\varepsilon_{\text{conf}}$ is relatively insignificant in the work presented in this thesis.

Following up briefly the second point, there is a subtle distinction between a 1D quantum dot and a quantum wire. First, a quantum dot means that the particle's motion is quantized in all three dimensions such that $dk_x dk_y dk_z \rightarrow \Delta k_x \Delta k_y \Delta k_z$ whereas a quantum wire means that it's quantized only in two dimensions such that $dk_x dk_y dk_z \rightarrow \Delta k_x \Delta k_y dk_z$. The nuance of a 1D quantum dot is that it is inherently a quantum wire but the electrons in the unquantized direction is confined to the extent where the discrete nature of the particles come out (we shall see how this arises in the next section). The 1D quantum dot is simply used to highlight the anisotropy of the system yet still accounting for the fact that the system still displays Coulomb blockade.

C.2 Conditions to Form and Detect a Quantum Dot

The discrete nature of the electrons (quantized packets of charge) is not so easy to observe. To see this discrete nature is tantamount to realizing a quantum dot where charges are only allowed to move in and out one by one. As mentioned earlier, the condition to fulfill is $L \lesssim \lambda_F$, where L denotes sample size. But this in reality is quite tricky to achieve. How do we maintain this condition while electrically reading out the state of the system? Electrically accessing a quantum dot inevitably means attaching

(a) (b)

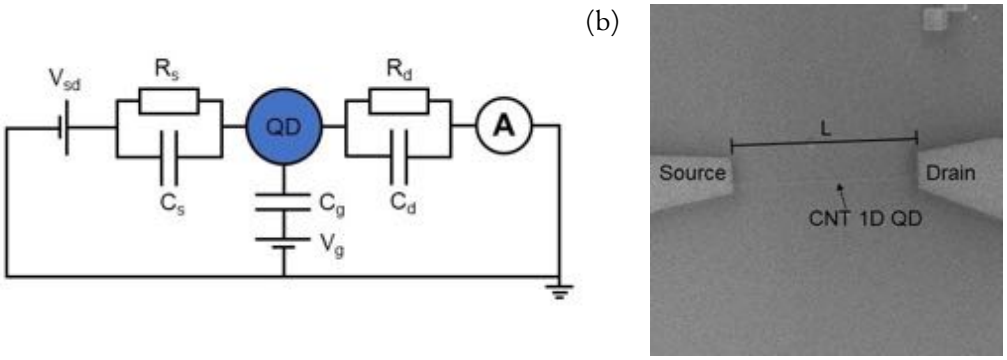


Figure C.1: Schematic of a quantum dot device (single electron transistor). (a) The circuit of a quantum dot device. Here, s, d and g correspond to source, drain and gate respectively. QD (blue island) is the quantum dot. For mathematical convenience, indices are assigned as $\{\text{QD, s, d, g}\} \rightarrow \{0, 1, 2, 3\}$. The resistor and capacitor in parallel is a lumped element that represents the electrical contacts of the source and drain with the quantum dot. (b) SEM image of a real carbon nanotube quantum dot device. It is simply a CNT with electrical leads placed on top. The device length is L. The source and drain contacts have characteristic resistances and capacitances. The gate for this device is a global backgate that comes from the doped Si underneath the SiO_2 substrate which the device rests on.

macroscopic sized electrical contacts to the system which will then destroy the condition one set out to satisfy. This, as it turns out, is not necessarily the case as long as

$$h\Gamma < \varepsilon_C, \tag{C.3}$$

where h is Planck's constant, Γ is the electron tunneling rate between the electrical contact and the quantum dot and ε_C is the charging energy of the quantum dot. The idea is that if the electrons have a hard time leaving the quantum dot, hence having an extremely low rate of tunneling, they are effectively considered trapped in the quantum dot. But what is considered low enough? As suggested by equation C.3, this threshold is determined by ε_C, the charging energy.

C.2.1 Charging Energy

We know that the number of charges in a highly metallic object (i.e. with large density of states hence allowing one to safely neglect the effects from quantum capacitance) is determined by its geometrical capacitance with its surroundings through:

$$Q = Nq = CV, \quad q = \begin{cases} e & V \geq 0 \\ & \\ -e & V < 0 \end{cases}, \tag{C.4}$$

where Q is the total charge, N is the total number of charges, q is the elementary charge, $e = |e| = +1.602176634 \times 10^{-19}$ C is the physical constant for an elementary charge (written in a way to stress that it represents a positive charge value), C is the total geometrical (or classical) capacitance between the object of interest and its surroundings, and V is the electric potential difference or voltage between the object and its surroundings. Now let us consider a small conducting island that is capacitively coupled to the outside world. Assuming that the outside world is the ground reference, the total electrostatic energy (a type of potential energy) of the object relative to ground is simply the sum of the energy of all the charges held at a specific potential:

$$\begin{aligned} U_E &= \int_0^Q dq' V(q') \\ &= \frac{Q^2}{2C}, \end{aligned} \tag{C.5}$$

where we have made use of equation C.4. Here, $dq' = |e| dN'$ and N' is a dummy variable for number of charges. The charging energy is conventionally defined as the energy required to charge an object

with just one elementary charge and it is:

$$\varepsilon_C \equiv \left.\frac{dU_E}{dN}\right|_{N=1} = \left. N\frac{e^2}{C}\right|_{N=1} = \frac{e^2}{C}. \tag{C.6}$$

The charging energy is an energy scale that roughly tells us how much energy is needed to overcome the electrostatic barrier (the Coulomb repulsion from all the other existing electrons is one big factor) in order for the new electron to become accepted in the the small conducting island. If the capacitance between the island and its surrounding is very small, meaning that the island is in a sense very distant from the outside world, it becomes difficult to transfer an electron in.

C.2.2 The Time Scale of Electron Movement

In a small conducting island that is coupled to the outside world as shown in figure C.1, there are the source and drain contacts which have defining resistances and capacitances. On top of those, there is also a backgate that is only capacitively coupled to the quantum dot. Imagine in this case if one were to charge the quantum dot by sending one extra electron into the island, normally the time it would take for an object to be charged continuously to a certain voltage is given by a characteristic charging time, the $\tau = RC$ time constant (the same time applies for discharging), where R and C are the lumped (the total representative) resistance and capacitance of the system respectively. However, here we are dealing with a discrete charge where either the charge is in the quantum dot or not— there is no in between. So more accurately, this time constant τ represents the time frame between events which we relabel as Δt. Another way of saying this in the context of a quantum dot is: Δt represents the average time an electron spends idle either inside or outside the quantum dot before making a statistical action to either switch states or stay put. Such is the constitution of quantum mechanics where microscopic

events are probabilistic. An electron (if undisturbed) cannot deterministically be at a certain state but is only expected to be in a certain state with some degree of uncertainty. This uncertainty is given by the energy-time uncertainty relation:

$$\Delta\varepsilon\Delta t \geq \frac{\hbar}{2}. \tag{C.7}$$

This relation states that there exists a minimum energy uncertainty $\Delta\varepsilon$ for a given time frame Δt. The longer the time frame, the lower the energy uncertainty. In the case of a quantum dot, this time frame is given by the RC time constant and so substituting $\Delta t = RC$, we get for the minimum in energy uncertainty:

$$\Delta\varepsilon_{\text{min}} = \frac{\hbar}{2RC}. \tag{C.8}$$

This is the energy window the electron can fall into around its expected energy value. The longer Δt is, the more defined and accurate one can know about the electron's energy. We can put this into a more physical (classical) context: the electron spends RC amount of time in a given state, after which, it faces a certain probability of either moving away or staying put. If RC is large, the electron spends a significant amount of time localized in a state, which then enables an observer to more easily spot said electron, hence increasing the certainty. On the other hand, if RC is small, the electron spends essentially no time in any given state and this translates to electrons flowing freely in and out of the quantum dot. The electrons are never, not even for a brief moment, trapped inside it. As the observer, we can never tell exactly if the electron is ever there. This is a reflection of high uncertainty.

C.2.3 The Resistance Requirement to Form a Quantum Dot

Let us now put what we have learnt all together. To see the quantization of electron behavior coming from the charging energy (which is the energy needed to overcome the electrostatic barrier to add

an individual electron), the electron energy uncertainty cannot mask the charging energy (otherwise electrons will just move randomly in and out as though there were no barrier); hence, one condition that must be satisfied to realize a quantum dot is:

$$\varepsilon_C = \frac{e^2}{C} \gg \Delta\varepsilon_{\text{min}} = \frac{\hbar}{2RC}. \tag{C.9}$$

This condition can be approximately rewritten in terms of a familiar physical constant:

$$R \gg \frac{h}{e^2} \equiv R_Q. \tag{C.10}$$

R_Q is the quantum of resistance which has the value of $R_Q = 25.812807557(18)\,\text{k}\Omega$. We remind the reader that this condition is qualitative (we have indiscriminately taken away a factor of π to arrive at the form above but this is inconsequential due to the qualitative nature of the condition). We can now go back to equation C.3, where we can infer $\Gamma \sim 1/\tau$. The question of how low of a tunneling rate is required to observe a quantum dot is now qualitatively answered (in terms of resistance): the contact resistance to the quantum dot must be much larger than the intrinsic quantum of resistance of a single charge carrier channel.

C.2.4 The Temperature Requirement to Form a Quantum Dot

If being able to discern and resolve the charging energy is what ultimately dictates whether we see single electron behavior, then another another energy that can mask charging energy is thermal smearing. Temperature must be low enough such that electrons cannot be thermally activated and be aided to

overcome the charging energy. So another condition that must be fulfilled to have a quantum dot is:

$$k_B T < \varepsilon_C. \tag{C.11}$$

C.3 Coulomb Blockade

The signature feature of a quantum dot is the Coulomb blockade which refers to individual charges becoming blocked from going in or out of the small metallic island. The simplest model that describes this phenomenon is the constant interaction model which assumes the total energy of the system is purely of composed of single particle energies (ε_i) and the electrostatic energy of the dot (U_E) and nothing more (so we are neglecting the energy arising from other electron correlations besides Coulomb repulsion). Written out the total energy as a function of number of particles (N) in the dot is given by:

$$\varepsilon(N) = \sum_{i=1}^{N} \varepsilon_i + U_E(N). \tag{C.12}$$

The single particle energies are given by the quantized dispersion relation due to confinement effects and so simplistically the difference between the single particle energies is equivalent to the confinement energy ($\varepsilon_i - \varepsilon_{i-1} = \Delta\varepsilon_{\text{conf}}$). For the CNT, it means that the dispersion relation (detailed in appendix B.3) is imposed with the momentum restriction such as equation C.2. Turning to the electrostatic energy, it will be more useful to describe the energy in terms of the experimental knobs in our device. For an actual electronic device like the one shown in figure C.1, the source, drain and gate will each contribute to the electrostatic potential of the dot via their respective capacitance. Looking at the charges first:

$$Q_i = \sum_j C_{ij} V_j. \tag{C.13}$$

Here we have decomposed the contributions from each electrical object labeled by subscripts $\{i, j\} = \{0, \text{s}, \text{d}, \text{g}\} \Leftrightarrow \{0, 1, 2, 3\}$ which denote dot, source, drain and gate respectively. Q_i is then equal to the charge in object i. Because we are interested in the dot itself, the total charge in the dot is then:

$$Q_0 = C_{00}V_0 + \sum_{j=1}^{3} C_{0j}V_j, \tag{C.14}$$

written in such a way to see the following step easier. To ultimately find the electrostatic potential of the dot, we isolate V_0:

$$V_0 = \frac{Q_0 - \sum_{j=1}^{3} C_{0j}V_j}{C_{00}}. \tag{C.15}$$

C_{00} is called the self capacitance and because of conservation of charge, $C_{00} = -\sum_{j=1}^{3} C_{0j}$. Now following equation C.5 we can compute the electrostatic energy:

$$\begin{aligned} U_E &= \int_0^Q dQ_0 V_0(Q_0) \\ &= \frac{Q^2}{2C_{00}} - Q\frac{\sum_{j=1}^{3} C_{0j}V_j}{C_{00}} \end{aligned} \tag{C.16}$$

The total energy written in terms of charge number (substituting $Q = Ne$) is then:

$$\varepsilon(N) = \sum_{i=1}^{N} \varepsilon_i + \frac{e^2N^2}{2C_{00}} - eN\sum_{j=1}^{3} \frac{C_{0j}}{C_{00}}V_j \tag{C.17}$$

The next quantity of interest is the electrochemical potential, which is the change in energy of the system when adding an electron (taking into account both single particle and electrostatic energies). Because of the discrete nature of the energy levels, this quantity is calculated through a subtraction

between the system's energy before and after the particle addition:

$$\begin{aligned}\mu_E(N) &= \varepsilon(N) - \varepsilon(N-1) \\ &= \varepsilon_N + \frac{e^2}{C_{00}}\left(N - \frac{1}{2}\right) + e\sum_{j=1}^{3}\frac{C_{0j}}{C_{00}}V_j \\ &= \varepsilon_N + \frac{e^2}{C_{00}}\left(N - \frac{1}{2}\right) - e\sum_{j=1}^{3}\alpha_j V_j.\end{aligned} \tag{C.18}$$

μ_E denotes electrochemical potential (μ denotes chemical potential). Note that this isn't the same as the chemical potential of a system which is typically given by the single particle energy of the highest occupied state. The subtlety in their difference lies in the fact that the electrochemical potential accounts for the electrostatic component which for a quantum dot is at the core of its description and can not be at all neglected. If one were to only talk about the chemical potential, then the chemical potential of the quantum dot would indeed simply be the single particle energy of the electron being added into the system, i.e. $\mu = \varepsilon_N$. But again, this quantity is practically useless when describing a quantum dot where the electrostatic energy plays a huge role in the energetics of the system.

Above we have introduced the following quantity:

$$\alpha_i = -\frac{C_{0i}}{C_{00}}, \tag{C.19}$$

which is always a positive quantity because the self-capactiance (C_{00}) is always negative. This is the lever arm of object (typically gate) i. The role of the lever arm is to give the effects from the different objects a weight. The most affecting objects are the gates built in specifically as a knob to control the number of charges in the quantum dot. Gates that are more capacitively coupled to the dot will have a greater effect in altering the dot's electrochemical potential. Objects like the source and drain,

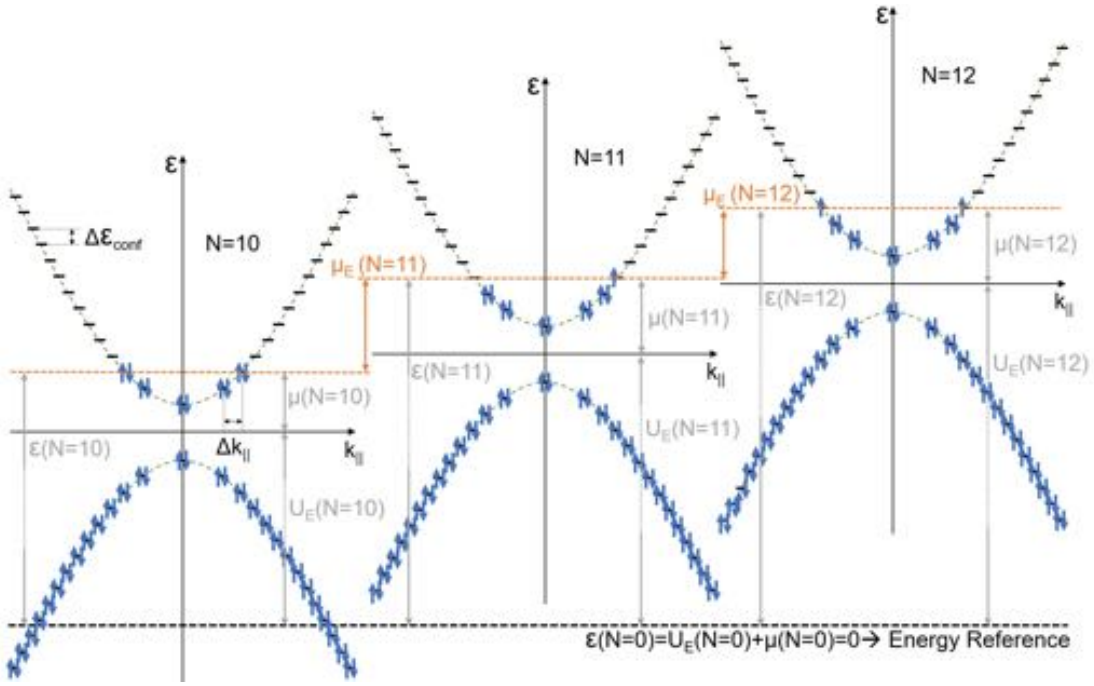


Figure C.2: The energy definitions for a quantum dot are indicated. Shown here going from left to right is the movement of the band and change in energies as the number of electrons in the quantum dot is increased from 10 to 12. The dotted line at the bottom of the graph is at an arbitrary position used only to indicate the energy reference of the system conventionally defined when $N = 0$. Going from $N = 10$ to $N = 11$, there is both an electrostatic energy and confinement energy contribution in shifting the electrochemical potential. Going from $N = 11$ to $N = 12$, however, the confinement energy contribution is absent because electron of the same single-particle energy is added. There is still an overall increase in the energy of the system (i.e. the entire energy band is shifted up) because of the electrostatic potential energy gained.

however, are not completely negligible and can also have a gating effect when they are biased to a non-zero voltage and they must also be accounted for. They are typically what cause the asymmetry to the Coulomb diamonds as we will see later.

Knowing the electrochemical potential of the dot, we can now understand its transport properties. The electrochemical potential is the energy change in the system as an electron is added which directly translates to the amount of energy an electron needs in order for it to enter the dot. Because the electrochemical potential is discretized due to the countable number of charge, we can view the

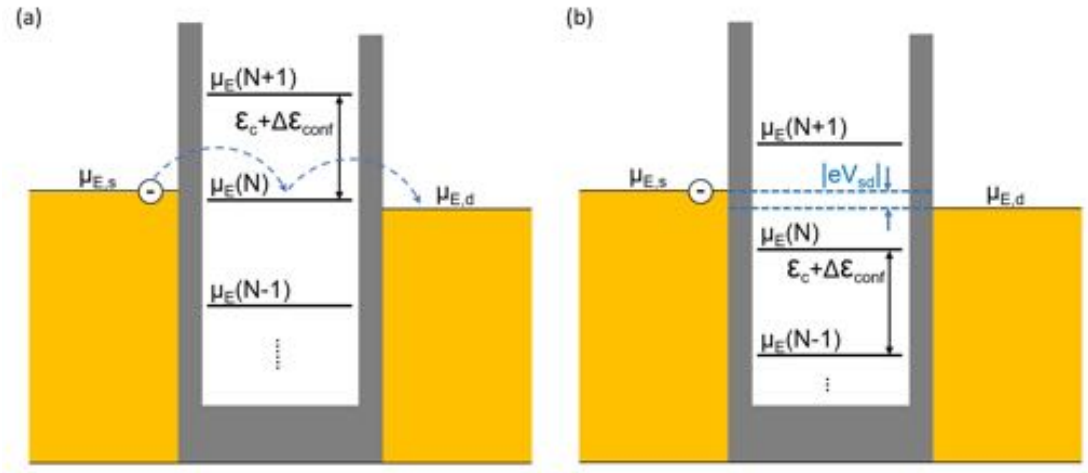


Figure C.3: The transport mechanism of a quantum dot. (a) shows when conduction is allowed through the quantum dot. In this case, the electrochemical potential of the quantum dot is aligned with energy window between the source and drain electrochemical potentials. This is called the degenerate state and electrons can flow freely across. By tuning the gate voltage, the electrochemical potential of the quantum dot will move. (b) shows when the electrochemical potential falls outside the energy window defined by the source and drain. The energy levels here are completely misaligned and electrons are blocked from entering the quantum dot. All the energy levels below the source drain electrochemical potentials are filled, thereby Pauli blocking electrons, while the next energy level higher up is beyond reach.

electrochemical potential as discrete energy levels that mimic those of an atom. See figure C.2 for pictorial definitions of the energies related to a quantum dot. In this figure, it shows that in each single particle state, despite them being two fold degenerate (due to spin), to add an electron, an electrostatic barrier needs to be overcome no matter what. The entry price here is the charging energy.

Let us first consider the case when the source drain bias is small such that $eV_{\text{sd}} \ll e^2/C_{00} \equiv \varepsilon_C$, still assuming that temperature is also small where $kT \ll e^2/C_{00} \equiv \varepsilon_C$. This is so that we allow ourselves to only consider one energy level at a time. The difference between the electrochemcial potentials of the source and drain defines an energy window: $|eV_{\text{sd}}| = |\mu_{E,\text{s}} - \mu_{E,\text{d}}|$. This energy window is created if we directly bias either the source and drain with V_{sd} (see figure C.1). When the electrochemical potential of the quantum dot falls within this energy window, i.e. when $\mu_{E,\text{s}} \leq \mu_E(N) \leq \mu_{E,\text{d}}$

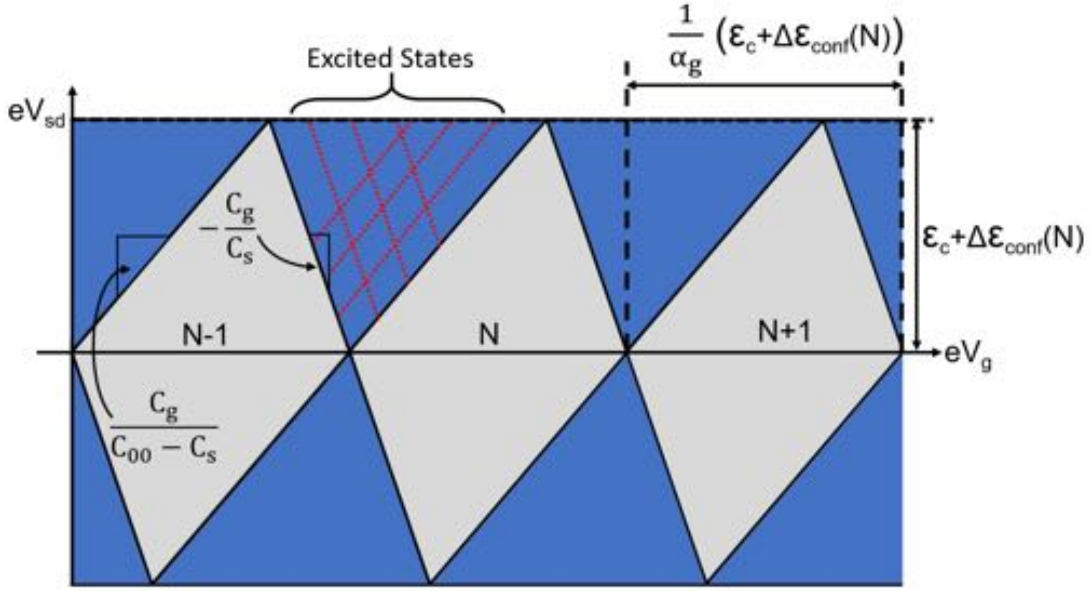
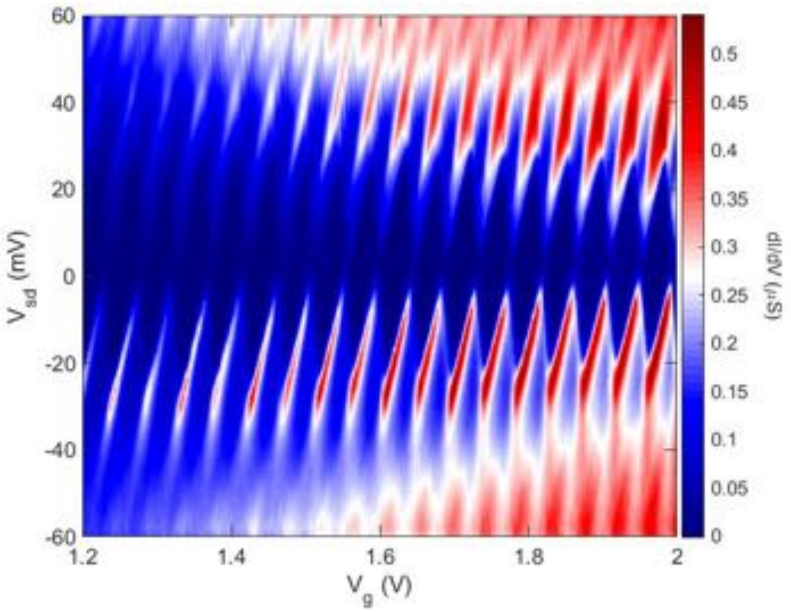


Figure C.4: Shown are three consecutive Coulomb diamonds corresponding to $N-1$ to $N+1$ number of charges. The lines demarcating the diamonds provide critical information on the quantum dot under study. The information that can be deduced include: self-capacitance (C_{00}), source (and/or drain) capacitance ($C_{s(d)}$), gate capacitance (C_g), the lever arm (α_g), charging energy (ε_c) and confinement energy ($\Delta\varepsilon_{conf}$).

(for $V_{sd} > 0$) or $\mu_{E,s} \geq \mu_E(N) \geq \mu_{E,d}$ (for $V_{sd} < 0$), the energy levels are considered aligned. Electrons from either the source and drain now have sufficient energy to enter and escape the dot and electrons can now flow freely through the dot. This is the degenerate state of the system. However, as soon as the electrochemical potentials of the dot and of the source drain are misaligned, done via tuning the gate voltage, electrons from the source or drain will not be able to go into the dot. This is because despite the electron energy in the source or drain being possibly higher, there simply is no room in the dot to accommodate an extra electron due to Pauli's exclusion principle. The only way is if the energy alignment occurs again done via tuning the gate voltage. Figure C.3 shows the mechanism regarding transport through a dot.

From the energy alignment argument, we thus expect a peak in conductance whenever the electro-

(a)



(b)

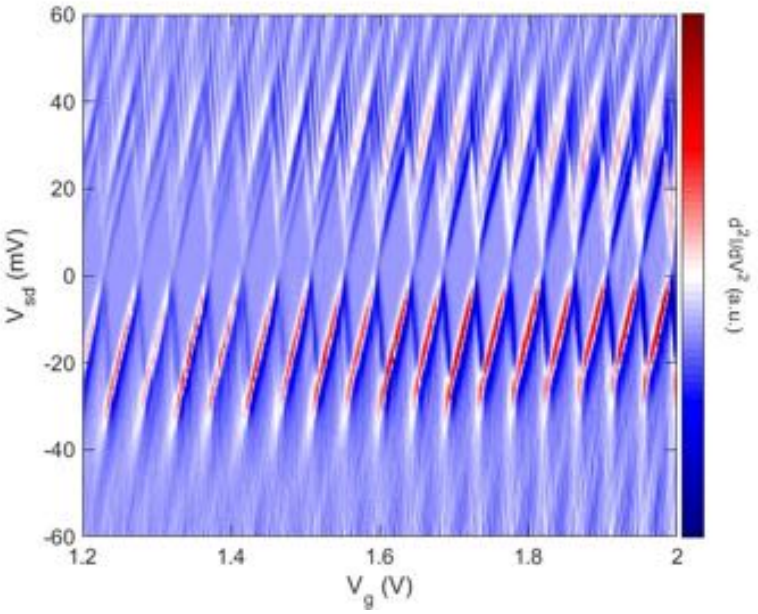


Figure C.5: A charge stability diagram showing CNT Coulomb diamonds. (a) The differential conductance of the CNT plotted with respect to gate voltage and source-drain bias. (b) The second order differential conductance of the CNT. This is done to bring out the more subtle features of plot (a). Here the Coulomb diamonds exhibit high periodicity indicating little contribution from confinement energy. The area inside a diamond correspond to Coulomb blockade where electrons are blocked. The charging energy is about 28mV. The array of lines outside of and above the Coulomb blockade region are excitation lines. See main text for more information.

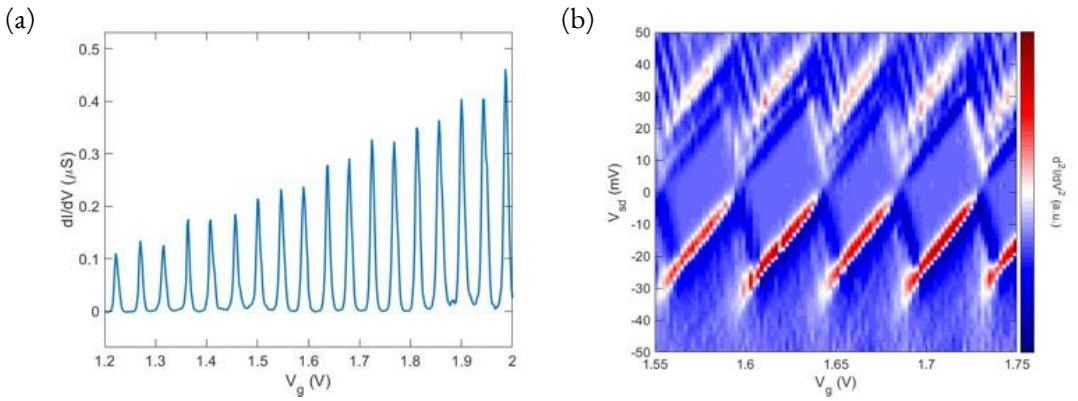


Figure C.6: A charge stability diagram showing CNT Coulomb diamonds. (a) The differential conductance of the CNT plotted with respect to gate voltage and source-drain bias. (b) The second order differential conductance of the CNT. This is done to bring out the more subtle features of plot (a). Here the Coulomb diamonds exhibit high periodicity indicating little contribution from confinement energy. The area inside a diamond correspond to Coulomb blockade where electrons are blocked. The charging energy is about 28mV. The array of lines outside of the Coulomb blockade region are excitation lines. See main text for information.

chemical potential is aligned with the energy window defined by the source and drain. In the small bias regime ($V_{\text{sd}} \ll \varepsilon_C$), for $V_{\text{sd}} > 0$, the condition as mentioned before is: $\mu_{E,\text{s}} \leq \mu_E(N) \leq \mu_{E,\text{d}}$. Because in our measurement setting we typically put the drain to be at the ground potential, we can assume that $\mu_{E,\text{d}} = 0$ and $-eV_{\text{sd}} = \mu_{E,\text{s}}$. The condition then becomes: $-eV_{\text{sd}} \leq \mu_E(N) \leq 0$. Using the expression of the electrochemical potential (equation C.18), the condition can be broken down into:

$$\begin{aligned} V_{\text{sd}} &\leq \frac{\alpha_{\text{g}}}{1-\alpha_{\text{s}}} V_{\text{g}} + \frac{1}{e(1-\alpha_{\text{s}})}\left[\varepsilon_N + \frac{e^2}{C_{00}}\left(N+\frac{1}{2}\right)\right], \\ V_{\text{sd}} &\geq -\frac{\alpha_{\text{g}}}{\alpha_{\text{s}}} V_g + \frac{1}{e\alpha_{\text{s}}}\left[\varepsilon_N + \frac{e^2}{C_{00}}\left(N+\frac{1}{2}\right)\right]. \end{aligned} \tag{C.20}$$

We remind the reader that the subscript of object i has $\{i\} = \{0, \text{s}, \text{d}, \text{g}\} \Leftrightarrow \{0, 1, 2, 3\}$. These set of equations (together with the ones for $V_{\text{sd}} < 0$) define the conduction region in the charge stability diagram. The charge stability diagram is the electronic transport characteristic of a quantum dot.

The reason it is called so is charges can sit stably within the quantum dot in the right conditions and may also become unstable (which is to say conduction happens) in other conditions. The regions in which charges sit stably are called Coulomb diamonds (see figure C.5). Inside the diamonds are regions where Coulomb blockade is happening. The equality of these equations define the borderlines of the Coulomb diamonds as seen in figure C.4. Differentiation of V_{sd} with respect to the gate parameter V_{g} then yields the slope:

$$\begin{aligned} \frac{dV_{\mathrm{sd}}}{dV_{\mathrm{bg}}}^{+} &= \frac{\alpha_g}{1-\alpha_s} = \frac{C_g}{C_{00}-C_s}, \\ \frac{dV_{\mathrm{sd}}}{dV_{\mathrm{bg}}}^{-} &= -\frac{\alpha_g}{\alpha_s} = -\frac{C_g}{C_s}, \end{aligned} \tag{C.21}$$

where the positive defines the positive slopes and negative for the negative slopes. The tip or height of the Coulomb diamond is where the two borderline equations meet. It is also the difference between two sequential electrochemical potentials:

$$\begin{aligned} \mu_E(N) - \mu_E(N-1) &= \underbrace{\varepsilon_N - \varepsilon_{N-1}}_{\Delta\varepsilon_{\mathrm{conf}}(N)} + \frac{e^2}{C_{00}} \\ &= \Delta\varepsilon_{\mathrm{conf}}(N) + \varepsilon_C. \end{aligned} \tag{C.22}$$

The gate value at which the conduction occurs at $V_{\mathrm{sd}} = 0$ can be derived from equation C.18. Setting $V_{\mathrm{s}} = V_{\mathrm{d}} = 0$ and $\mu_E(N) = 0$ (because now everything is aligned with ground) and solving for V_{g}, one finds the gate value at which conduction occurs for the N electron case is:

$$V_{\mathrm{g}}(N) = \frac{1}{e\alpha_{\mathrm{g}}}\left[\varepsilon_N + \frac{e^2}{C_{00}}\left(N - \frac{1}{2}\right)\right]. \tag{C.23}$$

The length of the Coulomb diamond along the gate axis is then given by:

$$V_g(N) - V_g(N-1) = \frac{1}{e\alpha_g}(\Delta\varepsilon_{conf}(N) + \varepsilon_C) = \frac{\Delta\varepsilon_{conf}(N)}{e\alpha_g} + \frac{e}{C_{00}}. \qquad \text{(C.24)}$$

Note that this is written in terms of voltage. Figure C.4 shows all the relevant quantities in terms of the energy scale for the charge stability diagram. Without going into too much detail, the lines extending out of the Coulomb diamonds are the excitation lines. They correspond to the excited states accessed within the dot. In the constant interaction picture considering only single particle energies, these excited states come from the next single particle energy states that are now within reach given that V_{sd} is big enough. One can think that these lines are arising from another set of shifted Coulomb diamonds— shifted by the confinement energy equivalent along the gate axis. Remember, the excitation lines of the Nth Coulomb diamond are describing the excited states of the same N charges. The charges are only re-configuring and no extra charges are coming in and hence no extra charging energy needs to be paid. However, these excited states are never simply described by the single particle picture. In reality, these states are formed by the collective behavior of many electrons and so the origins of the excited states are much more complex in nature. Figures C.5 and C.6 show examples of Coulomb diamonds in a charge stability diagram.

C.4 Density of States of a Carbon Nanotube Quantum Dot

In zero dimension (0D), the density of states (DOS) is equal to the number of states (N) per energy level. The DOS of a carbon nanotube quantum dot (CNT QD) according to the general definition

of DOS (equation A.47) is simply:

$$g_{0\mathrm{D}}(\varepsilon) = g \sum_{N=0}^{\infty} \delta(\varepsilon - \varepsilon_N), \tag{C.25}$$

where g is the degeneracy (not to be confused by the DOS, $g(\varepsilon)$) and ε_N is the discretized energy spectrum due to quantum confinement. Assuming a linear energy spectrum discretized by the hard confinement energy (as in equation C.2) then:

$$\varepsilon_N = \frac{\hbar v_F \pi N}{L}. \tag{C.26}$$

However, as mentioned earlier, the CNT devices used in this thesis has a very small $\Delta\varepsilon_{\text{conf}}$ where $\Delta\varepsilon_{\text{conf}} \sim k_B T \ll \varepsilon_C$ (for $T = 1$ K), in which case, the Dirac delta peaks of the 0D DOS will begin to merge and we recover the 1D CNT DOS. In the linear regime which is an approximation we will make for the CNT QD, it becomes the same as the metallic CNT 1D DOS:

$$g_{\mathrm{CNTQD}}(\varepsilon) = \frac{gL}{\hbar v_F \pi}, \tag{C.27}$$

which is an intuitive expression, saying there is one state (multiplied by a degeneracy) for every unit of confinement energy. This raises an important point which is that the CNT QD used in this thesis all fall under the classical Coulomb blockade regime (as opposed to quantum Coulomb blockade regime) where the confinement energy due to size quantization is insignificant compared to the charging energy. And so the quantization effects resulting in quantized energy levels are merely coming from the Coulomb repulsion of the discrete charges.

C.5 Conductance of a Carbon Nanotube Quantum Dot

The electrons in an isolated quantum dot have wavefunctions that are completely localized within the quantum dot. But when a quantum dot is in contact with metallic reservoirs, i.e., the source and drain electrodes, the wavefunctions of the electrons in the quantum dot will overlap with the wavefunction of the electrons in the source and drain. The overlap is what dictates the transparency, or conduction, across the barrier between the quantum dot and metallic reservoir. This overlap is quantified by the energy coupling to the source and drain which given by $\hbar\Gamma_{\mathrm{s},i}$ and $\hbar\Gamma_{\mathrm{d},i}$ respectively. i denotes the coupling to the energy level i of the quantum dot. Another term for overlap is hybridization or mixing of wavefunctions. This overlap effectively causes the energy levels to spread to a width of:

$$\hbar\Gamma_i = \hbar\Gamma_{\mathrm{s},i} + \hbar\Gamma_{\mathrm{d},i} = \frac{\hbar}{\tau}, \tag{C.28}$$

and here τ is the lifetime of the electrons in the quantum dot (as in equation C.3). Depending on the lifetime of the electrons, the conduction across the quantum dot can be different.

In principle the conductance of a QD is given by the Landauer Buttiküer equation:

$$G = \frac{e^2}{h} MT. \tag{C.29}$$

The conductance depends on the number of available channels or modes (M) where each channel contributes a quantum conductance weighted by a transmission coefficient (T). In a QD the available channels are the discrete energy levels within the source drain bias energy window. The transmission coefficient is what contains the dynamical details. Thermal broadening and tunnel rates at each barrier must be taken in account and the calculation of the transmission is non-trivial. Below we report the

two limiting regimes.

C.5.1 Weak-Coupling Regime

The weak-coupling regime is when:

$$\hbar\Gamma \ll k_B T \ll \varepsilon_C. \tag{C.30}$$

Assuming that only a single energy level contributes to conduction, then the conductance for energy level i is given by:[199]

$$G_i = \frac{ge^2}{4k_B T}\left(\frac{1}{\Gamma_{\mathrm{s},i}} + \frac{1}{\Gamma_{\mathrm{d},i}}\right)^{-1} \cosh^{-2}\left(\frac{\alpha_{\mathrm{g}}(V_{\mathrm{g}} - V_{\mathrm{g},i})}{2k_B T}\right). \tag{C.31}$$

The distinctive behavior of the conductance in this regime is its inverse dependence on temperature and the broadening of the conductance peak as a function of increasing temperature. In this regime, the transport can be describe by single electrons moving in and out sequentially of the dot.

C.5.2 Strong-Coupling Regime

In the strong-coupling regime,

$$k_B T \lesssim \hbar\Gamma \ll \varepsilon_C, \tag{C.32}$$

in which case, the conductance is shown to follow the famous Breit-Wigner formula:[200]

$$G = \frac{ge^2}{h}\left(\frac{1}{\Gamma_{\mathrm{s}}} + \frac{1}{\Gamma_{\mathrm{d}}}\right)^{-1} \frac{h^2\Gamma}{\alpha_{\mathrm{g}}^2(V_{\mathrm{g},i} - V_{\mathrm{g}})^2 + (h\Gamma/2)^2}. \tag{C.33}$$

The electrons in this regime is treated identically as the resonant tunneling across a double barrier and in this regime the tunneling is no longer just a single particle event. The transport of electrons across a level is given by the simultaneous tunneling processes of other electrons through other levels. These levels can be virtual intermediate states. A localized charge is no longer a valid concept in this regime. The conductance is thus irrespective of the transparency of individual energy levels (hence the absence of subscript i) but is a result of the inclusion of many levels. This is why g, the degeneracy of each state, is necessarily included in the equation. The conductance is independent of temperature and the different conductance peaks are identical as long as the strong coupling condition holds.

C.6 Other Quantum Dot Behaviors: Gate-Defined and Disorder-Defined Quantum Dots as well as Fabry-Perot Interferences

One of the reasons the CNT QD is unique is it can become a QD naturally simply because of the tunnel barriers at the contact interface. There are cases where the contact is sufficiently good that the contacts become perfectly transparent. In this case, the CNT QD will lose its coulomb blockade behavior and instead behave as a 1D conducting wire. A CNT QD in this case can only be made by electrostatically gating the CNT in between the contacts. For instance, if the CNT and the contacts are both p-doped, then the CNT must be n-doped to create pn junctions at the contacts. These pn junctions act as a tunnel barriers which then helps fulfill the QD formation requirement. The height of the tunnel barrier in this case roughly depends on the bandgap and the strength of the gating. Figure C.7 illustrates the behavior of such a QD. The bipolar side is where the QD exists. The unipolar side is where no tunneling barrier exists and so the dot becomes a simple conducting wire.

The formation of a CNT QD can also happen because of disorder such as impurities on or near the CNT as well as physical kinks and defects on the CNT. Because CNT is a 1D object, electrons

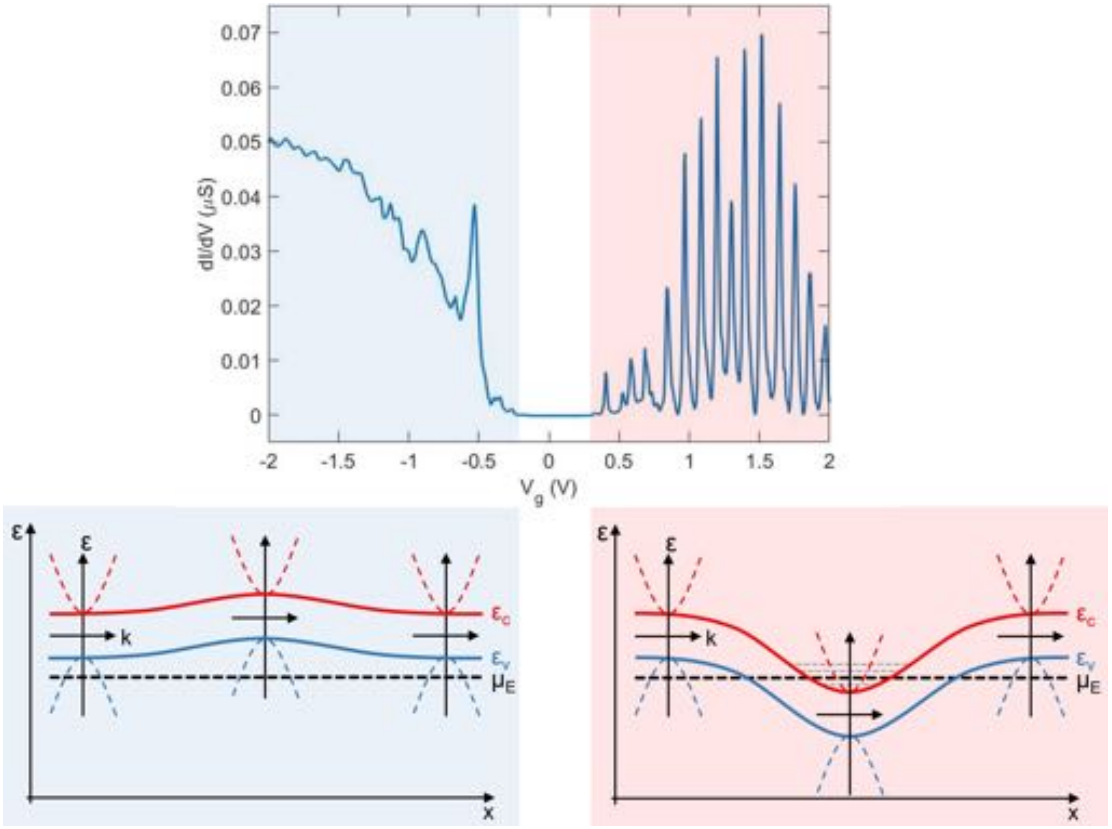


Figure C.7: The top graph shows the conductance of a gated semiconducting CNT. The blue shaded area corresponds to the CNT being unipolar (ppp) while the red shaded area corresponds to the CNT being bipolar (pnp). In the unipolar regime, electrons can flow freely across the CNT without any barriers, while in the bipolar regime, due to the pn junction the electrons have to go through two barriers. The pn junction interface significantly increases the resistance and hence causes confinement of electrons in the quantum well bounded by the pn junctions. The bottom left and right pictures show the energy diagrams of the unipolar and bipolar regimes respectively.

are highly sensitive to anything that sits in their path. The imperfections will cause backscattering or behave as tunnel barriers and hence form a QD. These QDs typically do not demonstrate regular Coulomb blockade peaks because disorders are likely to cause the formation of multiple QDs in series. These QDs will all behave differently and their transport properties get stacked electrically in series. Figure C.8a shows an example of a disordered QD.

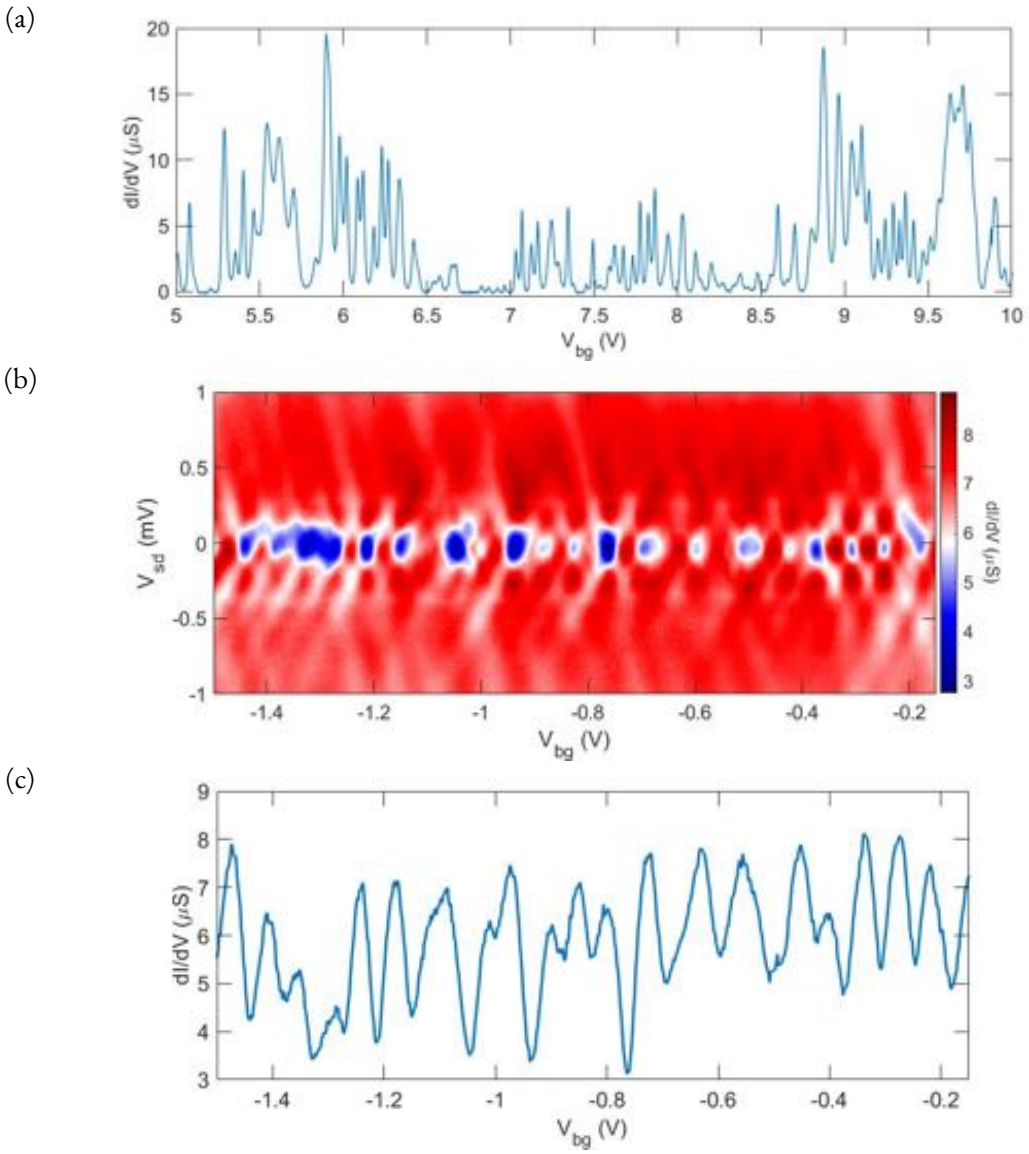


Figure C.8: Other quantum dot behaviors. (a) a disordered QD which is identified by irregular Coulomb blockade peaks. (b) Fabry-Perot regime of a QD where the transport across the dot is governed by interference effects. This happens when the transmission at the interface is at least close to transparent. This regime is identified by the checker-board like features as well as high conductance that never reaches zero. (c) line cuts of (b) at zero DC bias. The conductance is always non-zero and the conductance exhibits a wave-like oscillation rather than sharp Coulomb blockade peaks.

There are plenty of other QD behaviors that are not addressed and the reader is encouraged to dig into literature. Here I will end with a brief description of a QD in the Fabry-Perot regime. This regime is when the contact resistance is $R < R_Q$ (in contrast to equation C.10), or equivalently $h\Gamma \gtrsim \varepsilon_C$. In this regime, the rate at which electrons pass through the contacts are fast. Electrons are no longer a localized object but instead become extended and will show wave-like properties such as constructive and destructive interference.[201] Since electrons cannot be pinpointed at a particular location, the description of Coulomb repulsion coming from an island is no longer valid. This is evident when we see that the QD never becomes blocked indicated by the absence of vanishing conductance as seen in figure C.8c. Figure C.8b shows a measurement of a CNT QD in the Fabry-Perot regime.

D

The Integer Quantum Hall Effect of Graphene

The quantum hall effect is a phenomenon that exists specifically in two dimensions (2D). It is an important phenomenon because of a few reasons. One, it is a macroscopic manifestation of quantum mechanics. Systems that show this are rare and graphene (Gr) being one of the few that do. Two, the electron edge channels are robust and topologically protected one dimensional (1D) states with different characters. These can be used as building blocks to engineer more exotic structures or simply

can be used as a 1D platform for 1D studies. Three, in the quantum hall regime, due to the localizing effect of the required applied magnetic field, the kinetic energy of the charge carriers are quenched. This means that the electrons are now no longer running wild but they stay put and are sensitive to the presence of one another. Consequently, the dynamics of this body of electrons are revamped and their behaviors become much more delicate. This ultimately gives rise to much more exotic electronic phases governed by many-body physics.

D.1 Classical Regime: Weak Magnetic Fields

The force on an electron under the influence of an electric and magnetic field is given by:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{D.1}$$

In a 2D planar geometry with electrons under a longitudinal electric field bias, the application of a weak magnetic field perpendicular to the plane will bend the electrons trajectories. In steady state, charges will build up along the transverse sides of the device and this gives rise to a Hall voltage that develops in the transverse direction. If the (positive) current (attributed to holes) is going in the +x direction (I_x), with the magnetic field pointing in the +z direction (B_z), and if the device has length L that charges move longitudinally along and width W which is in the transverse direction, the Hall voltage is then:

$$V_H \equiv V_{xy} = v_x B_z W. \tag{D.2}$$

In the spirit of Ohm's law ($V = IR$), the Hall resistance is given by:

$$R_{xy} = \frac{V_H}{I_x} = \frac{v_x B_z W}{-nev_x W} = -\frac{B_z}{ne} \tag{D.3}$$

It is interesting to note that Hall resistivity is the same as Hall resistance in 2D:

$$R_{xy} = \frac{V_H}{I_x} = \frac{E_y W}{j_x W} = \frac{E_y}{j_x} = \rho_H \equiv \rho_{xy}. \tag{D.4}$$

This is not the case for longitudinal resistance where $R_{xx} = \rho_{xx} \cdot (L/W)$. The Hall coefficient is defined as:

$$R_H = \frac{E_y}{j_x B_z} = -\frac{1}{ne}, \tag{D.5}$$

where $j_x = I_x/W = -nev_x$ in 2D, $e = |e|$ is the elementary charge and n is the carrier density. Here we follow the convention for the labeling of the Hall coefficient. Note that R_H is not resistance but has units of m^2/C. In fact, it is a quantity that is independent of any scattering times and hence insensitive to friction. This coefficient is useful when determining the carrier type and density in an electronic device. A positive R_H means hole doping and negative means electron doping.

D.2 Quantum Regime: Strong Magnetic Fields

The behavior of the device in strong magnetic fields will exhibit quantization effects in part because electron orbits are closed, becoming atomic-like, and the energies discretize accordingly. More specifically, the Hall resistance will develop plateaus at integer (or fractional for fractional quantumm hall) multiples of the quantum of resistance, h/e^2. This is the main signature of the quantum hall effect. In order to be in such a regime, the criterion of strong magnetic field must be met and this happens when the electrons are allowed to complete its cyclotron orbit without getting interrupted. This will be explained more quantitatively in a later section. Let us first briefly draw the physical picture of what's happening.

D.2.1 Brief Semiclassical Description

When the applied magnetic field becomes strong enough, the trajectories of the electrons will close on itself and form closed orbits. These are called cyclotron orbits. The condition for closed orbit is given by the Bohr-Sommerfeld quantization condition:

$$\oint_C d\mathbf{r} \cdot \mathbf{p}(\mathbf{r}) = h(N + \gamma), \tag{D.6}$$

which comes from the requirement of the quantization of angular momentum. h is Planck's constant, the quantized unit of angular momentum; N is an integer; A is the circumference of the orbit; the γ factor is the Maslov index and has a value of $1/2$ to account for the zero-point motion.[202] The center of these orbits are guiding centers and they either drift along equipotential lines or are pinned down by disorder in the bulk of the sample. These equipotential lines are all created by the impurities littered across the bulk of the sample. The drift velocity is given by:

$$\mathbf{v_d} = \frac{1}{q}\frac{\mathbf{F} \times \mathbf{B}}{B^2} = \frac{\mathbf{E} \times \mathbf{B}}{B^2}, \tag{D.7}$$

which is equal for both positive and negative charges. Equipotential lines that are closed on its own, disconnected from others, form the electron puddles. As the magnetic field is swept or when the electron density is changed, the sizes of the cyclotron orbits and the number of participating electrons will change and this directly affects the coverage of electron puddles (localized pockets of electrons) throughout the sample. This ultimately causes the bulk of the sample to become conducting or insulating. When it is insulating, electrons are trapped inside disconnected electron puddles, unable to leak out and jump from one place to another. When it is conducting the electrons are able to percolate

through the sample because the electron puddles are big enough to merge, paving a way to go from one end of the sample to the other. In the midst of the bulk transitions, at the edges of the sample the cyclotron orbits are unable to close on itself but will instead bounce along the edge, hence creating a skipping orbit. These skipping orbits form the 1D chiral edge states that are always carrying a current flow in one direction. Such is the picture of what's happening in the quantum hall regime, but one can only go so far when attempting to model it semiclassicaly. The description becomes especially difficult if one has to account for the quantum mechanical details of Gr (such as valley physics) or electron-electron interactions. Landau quantization, on the other hand, can beautifully capture the phenomenon without needing to delve into the messy microscopic details. We begin by looking at the Gr Hamiltonian with the addition of a magnetic field.

D.2.2 Single Particle Graphene Landau Levels

To understand the nature of the Landau levels (LL) of Gr, it is important to include both $\mathbf{K}$ and $\mathbf{K}'$ valleys. And so we start from equation A.31 and introduce the magnetic field via substituting the kinetic momentum ($\hbar\mathbf{q}$) with the more appropriate canonical momentum ($\mathbf{\Pi}$) as warranted by Hamiltonian mechanics: $\hbar\mathbf{q} \to \mathbf{\Pi} \equiv \hbar\mathbf{q} + e\mathbf{A}$. We then acquire:

$$\mathbf{H}(\mathbf{q}) = v_F \begin{bmatrix} 0 & \Pi_x - i\Pi_y & 0 & 0 \\ \Pi_x + i\Pi_y & 0 & 0 & 0 \\ 0 & 0 & 0 & \Pi_x + i\Pi_y \\ 0 & 0 & \Pi_x - i\Pi_y & 0 \end{bmatrix}, \tag{D.8}$$

with the basis of $\{\Phi_{A,\mathbf{K}}, \Phi_{B,\mathbf{K}}, \Phi_{A,\mathbf{K}'}, \Phi_{B,\mathbf{K}'}\}$ and the corresponding eigenfunction has the form:

$$\mathbf{\Psi}(\mathbf{q}) = \begin{bmatrix} \psi_{A,\mathbf{K}}(\mathbf{q}) \\ \psi_{B,\mathbf{K}}(\mathbf{q}) \\ \psi_{A,\mathbf{K}'}(\mathbf{q}) \\ \psi_{B,\mathbf{K}'}(\mathbf{q}) \end{bmatrix}, \tag{D.9}$$

where $\psi_{(A,B),(\mathbf{K},\mathbf{K}')}(\mathbf{q})$ are the coefficients of the Gr wavefunction. Focusing on the $\mathbf{K}$ valley, writing out the system of eigenvalue equations (while dropping out the dependencies of the eigenfunctions) we have:

$$v_F(\Pi_x + i\Pi_y)\psi_{A,\mathbf{K}} = \varepsilon\psi_{B,\mathbf{K}}, \tag{D.10}$$

$$v_F(\Pi_x - i\Pi_y)\psi_{B,\mathbf{K}} = \varepsilon\psi_{A,\mathbf{K}}. \tag{D.11}$$

Combining the two nets us two decoupled eigenvalue equations:

$$v_F^2(\Pi_x - i\Pi_y)(\Pi_x + i\Pi_y)\psi_{A,\mathbf{K}} = \varepsilon^2\psi_{A,\mathbf{K}}, \tag{D.12}$$

$$v_F^2(\Pi_x + i\Pi_y)(\Pi_x - i\Pi_y)\psi_{B,\mathbf{K}} = \varepsilon^2\psi_{B,\mathbf{K}}. \tag{D.13}$$

Note that the order of the operators must be obeyed because they do not commute. Looking specifically at the equation for $\psi_{B,\mathbf{K}}$, with some rearrangement and writing out the canonical momentum

explicitly, in the Landau gauge where we choose $\mathbf{A} = (-B_z y, 0, 0)$, and for $p_i = \hbar q_i$, we get:

$$\begin{aligned}
\varepsilon^2 \psi_{B,\mathbf{K}} &= v_F^2\big((p_x - eB_z y + ip_y)(p_x - eB_z y - ip_y)\big)\psi_{B,\mathbf{K}} \\
&= v_F^2\big((p_x - eB_z y)^2 + i[p_y, p_x - eB_z y] + p_y^2\big)\psi_{B,\mathbf{K}} \\
&= v_F^2\big((p_x - eB_z y)^2 - e\hbar B_z + p_y^2\big)\psi_{B,\mathbf{K}} \\
\frac{1}{2m}\left(\left(\frac{\varepsilon}{v_F}\right)^2 + e\hbar B_z\right)\psi_{B,\mathbf{K}} &= \left(\frac{p_y^2}{2m} + \frac{(eB_z)^2}{2m}\left(y - \frac{p_x}{eB_z}\right)^2\right)\psi_{B,\mathbf{K}}
\end{aligned} \tag{D.14}$$

where we used the commutation relations: $[p_x, p_y] = 0$ and $[p_y, y] = -i\hbar$, and at the end we deliberately put the right hand side of the equation into a quantum harmonic oscillator form. The strategy for solving for the Landau level (LL) energies is to realize that the Hamiltonian with magnetic field can be cast into the same structure as a harmonic oscillator and since the solution of the harmonic oscillator is known, one can come up with an analogous solution for the Hamiltonian with magnetic field. Because the quantum harmonic oscillator oscillating around y_0 has the following eigenvalue equation:

$$\left(\frac{p_y^2}{2m} + \frac{1}{2}m\omega^2(y - y_0)^2\right)\psi = \hbar\omega\left(N + \frac{1}{2}\right)\psi, \tag{D.15}$$

where N is the quantum number which can take on integer values: $0, 1, 2\ldots$. This allows us to convert equation D.14 into:

$$\frac{1}{2m}\left(\left(\frac{\varepsilon_{\mathrm{LL},\mathbf{K},B}}{v_F}\right)^2 + e\hbar B_z\right)\psi_{B,\mathbf{K}} = \hbar\omega_c\left(N + \frac{1}{2}\right)\psi_{B,\mathbf{K}}, \tag{D.16}$$

where we have relabeled $\varepsilon \to \varepsilon_{\mathrm{LL},\mathbf{K},B}$ for better representation. Note here that B_z is the perpendicular magnetic field with respect to the 2D plane and B is referring to the sublattice. A further note here from a mathematical perspective is that we have decoupled the eigenvalue equation and now by

solving for the independent eigenvectors, we can acquire the full set of eigenvalues. What remains now is to solve for $\varepsilon_{\mathrm{LL},\mathbf{K},B}$. Below are the expressions for the key parameters encountered:

$$\omega_c = \frac{eB_z}{m}, \tag{D.17}$$

$$y_0 = \frac{\hbar q_x}{eB_z} = l_B^2 q_x, \tag{D.18}$$

$$\begin{aligned}\varepsilon_{\mathrm{LL},\mathbf{K},B} &= \mathrm{sgn}(N) v_F \sqrt{2\hbar e B_z |N|} \\ &= \mathrm{sgn}(N) \frac{\hbar v_F}{l_B} \sqrt{2|N|} \qquad \text{(D.19)} \\ &= \mathrm{sgn}(N) \hbar \omega_D \sqrt{|N|},\end{aligned}$$

$$l_B = \sqrt{\frac{\hbar}{eB_z}} = \frac{26\mathrm{nm}}{\sqrt{B_z[T]}}, \tag{D.20}$$

$$\omega_D = v_F \sqrt{\frac{2eB_z}{\hbar}}. \tag{D.21}$$

ω_c is the standard 2DEG (two dimensional electron gas) cyclotron frequency (for parabolic dispersion relations), y_0 defines the characteristic distance between the centers of cyclotron orbits and also tells us the relation between momentum space and real space, $\varepsilon_{\mathrm{LL},\mathbf{K},B}$ is the Landau level energy for the B sublattice of the $\mathbf{K}$ valley with N now being the Landau level index, and ω_D is the Dirac cyclotron frequency. Above we have introduced a new characteristic length scale, l_B, which is the magnetic length. More will be said about this later. Unlike before, N is now a quantum number that can take on both positive and negative integer values: $\ldots -1, 0, 1 \ldots$. This is a consequence of the 2×2 Hamiltonian (for a single valley) admitting two eigenvalues (one positive and one negative) for a given N. The positive and negative solution is now dictated by the sign of N which represents the energies of electrons and holes respectively. The interesting case here is $N = 0$— the zeroth Landau level (zLL) where $\varepsilon_{\mathrm{LL}} = 0$. At this level, the LL is shared equally by both electrons and holes. This is to say: when

the zLL is half filled, there are only holes occupying the LL; when the zLL is completely filled, it has both electrons and holes; and when the zLL is empty, the zLL has neither holes or electrons. This is in complete contrast to conventional band physics where filled bands are occupied by electrons, empty bands are filled by holes and partially filled bands have both electrons and holes. The significance of this is it gives rise to the anomalous nature of the Gr's quantum hall effect which will be covered more in-depth shortly.

As hinted, the zLL is more fascinating and complex compared to the higher order LLs. To see this, let us repeat the same procedure above but for $\psi_{A,\mathbf{K}}$. Here we get:

$$\begin{aligned}\varepsilon_{\mathrm{LL},\mathbf{K},A} &= \mathrm{sgn}(N) v_F \sqrt{2\hbar e B_z (|N|+1)} \\ &= \mathrm{sgn}(N)\hbar\omega_D \sqrt{|N|+1},\end{aligned} \tag{D.22}$$

for $N = \ldots -1, 0, 1 \ldots$. As one can see, the zLL ($\varepsilon_{\mathrm{LL}} = 0$) is absent in this case. So far what is shown is that the Hamiltonian can be decoupled into two independent eigenvalue equations, hence admitting two independent eigenvectors. The two sets of eigenvectors share the same eigenvalues except for the zLL. This then means that besides the eigenvector for the zLL, all other eigenvectors of the Hamiltonian can be the superposition of the decoupled eigenvectors with the same eigenvalue. For the zLL, the only allowed eigenvector is a non-zero $\psi_{B,\mathbf{K}}$ and vanishing $\psi_{A,\mathbf{K}}$ (for the K valley) with eigenvalue $\varepsilon_{\mathrm{LL}} = 0$. From this, we can deduce that the nonzero LLs should be twice more degenerate than the zLL.

Similar solutions can be derived for the $\mathbf{K}'$ except the roles of sublattice A and B are switched. Below we summarize the solutions regarding the Gr Hamiltonian in magnetic field. The Gr wave-

functions are taken from the citation.[203] For the $\mathbf{K}$ valley:

$$\mathbf{\Psi}_{N,\mathbf{K}}(q_x) = \frac{C_N}{\sqrt{W}} e^{iq_x x} \begin{bmatrix} \text{sgn}(N)(-i)\phi_{|N|-1}(q_x) \\ \phi_{|N|}(q_x) \\ 0 \\ 0 \end{bmatrix}. \tag{D.23}$$

For the $\mathbf{K}'$ valley:

$$\mathbf{\Psi}_{N,\mathbf{K}'}(q_x) = \frac{C_N}{\sqrt{W}} e^{iq_x x} \begin{bmatrix} 0 \\ 0 \\ \phi_{|N|}(q_x) \\ \text{sgn}(N)(-i)\phi_{|N|-1}(q_x) \end{bmatrix}, \tag{D.24}$$

with the parameters defined as follows:

$$C_N = \begin{cases} 1 & \text{for } N = 0 \\ \frac{1}{\sqrt{2}} & \text{for } N \neq 0 \end{cases}, \quad \text{sgn}(N) = \begin{cases} 0 & \text{for } N = 0 \\ \frac{N}{|N|} & \text{for } N \neq 0 \end{cases}, \tag{D.25}$$

$$\phi_{|N|}(y, q_x) = \frac{1}{\sqrt{2^N N!}} \left(\frac{1}{\pi l_B^2}\right)^{\frac{1}{4}} \exp\left(-\frac{1}{2}\frac{(y - q_x l_B^2)^2}{l_B^2}\right) H_{|N|}\left[\frac{(y - q_x l_B^2)}{l_B}\right].$$

$H_{|N|}$ are the Hermite polynomials and $\phi_{|N|}$ are the eigenfunctions of the harmonic oscillator. The behavior of the electrons are just harmonic oscillators centered around y_0 and their wavefunctions are like strips localized around y_0 with plane waves extending towards x. Because $0 < y < L_y$, the allowed q_x values are: $0 \leq q_x \leq L_y/l_B^2$ (from equation D.18).

Taking it all together, the wavefunction at $N \neq 0$ has non-zero amplitudes for both A and B

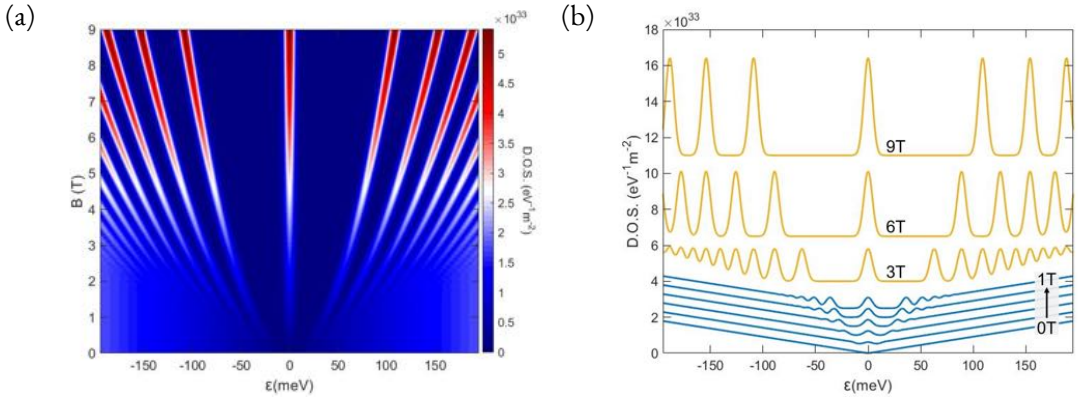


Figure D.1: Landau levels of monolayer graphene as a function of magnetic field. (a) shows the continuous evolution of the LLs in the integer quantum hall regime, assuming that the LLs retain their four-fold degeneracy, i.e., unbroken degeneracy. The LLs show its non-linear evolution along the energy axis and this along with a $N = 0$ LL are trademarks of monolayer graphene's anomalous quantum hall effect. (b) shows the line cuts of the LLs at different magnetic fields.

sublattice, meaning that particles are equally populated in both sublattices. However, for the $N = 0$ case, $\mathbf{K}$ valley particles only occupy the B sublattice whereas $\mathbf{K}'$ valley particles only occupy the A sublattice. This means that the valley spin (isospin) becomes equivalent as the sublattice (pseudospin). This is a form of highly exotic electron correlation effect and as a result the zLL has triggered many extensive studies.

D.3 Landau Level Degeneracy

Thus far the system has been treated as a $L_x = L$ by $L_y = W$ box. Operating under the Landau gauge still, all the states are piled up along the y direction as described by the wavefunction in the previous section. States are localized along y and are extended along x. Due to the boundary condition, momentum along the x is quantized into units of $\Delta q_x = 2\pi/L_x$. Because $0 < y < L_y$, in the Landau gauge, the range for q_x is $0 \leq q_x \leq L_y/l_B^2$ (from equation D.18). To enumerate the total

number of states, it comes down to seeing how many states fit given the physical size of our system:

$$N_{\mathrm{LL}} = g\frac{L_x}{2\pi}\int_0^{L_y/l_B^2} dq_x = g\frac{L_x L_y}{2\pi l_B^2} = g\frac{B_z A}{\Phi_0}. \tag{D.26}$$

Here g is the degeneracy of each state. For graphene $g = 4$ due to spin and valley degeneracies. Φ_0 is the flux quantum, and l_B is the magnetic length. The equations simply reads: the number of states in one Landau level is equal to the number of flux quanta threading through the sample surface A (multiplied by additional degeneracies g). If there are N filled LLs, then there will be N times more states. Below we summarize the key parameters:

$$\Phi_0 = \frac{h}{e} = 2\pi l_B^2 B_z, \tag{D.27}$$

$$l_B = \sqrt{\frac{\hbar}{eB_z}} \approx \frac{26\mathrm{nm}}{\sqrt{B_z}}, \tag{D.28}$$

$$n_{\mathrm{LL}} = \frac{N_{\mathrm{LL}}}{A} = g\frac{B_z}{\Phi_0}, \tag{D.29}$$

$$g_{\mathrm{LL}} = \frac{g}{2\pi l B^2}, \tag{D.30}$$

$$\nu = \frac{n_e}{n_{\mathrm{LL}}}. \tag{D.31}$$

Φ_0 is the quantum of flux. $2\pi l_B^2$ is interpreted as the area of a cyclotron orbit. So the magnetic length l_B is roughly the size of the orbit. Each cyclotron orbit allows one quantum of flux Φ_0. n_{LL} is the carrier density (or equivalently the total degeneracy) of each LL, taking into account the material's degeneracy g. g_{LL} is the DOS of each LL. ν is the filling factor, which tells us how many LLs are completely filled, counting degenerate LLs as separate ones. This can be translated to mean the number of electrons per quantum of flux. n_e is the total carrier density in the system.

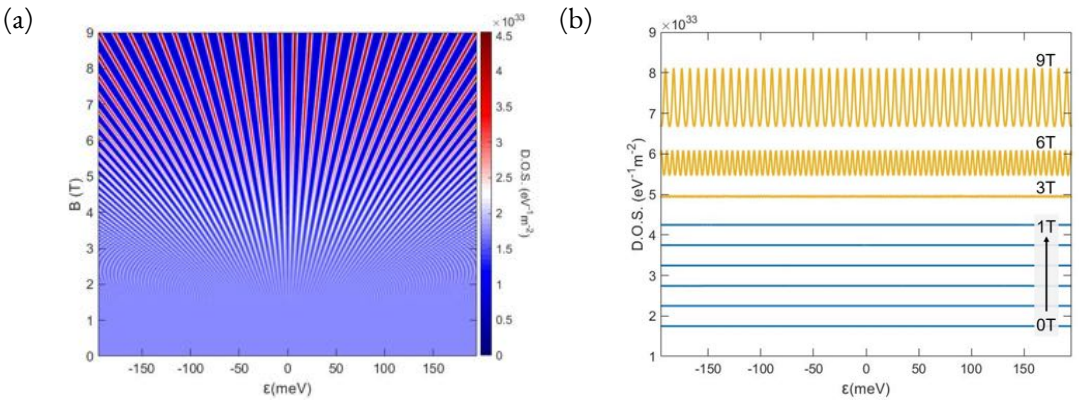


Figure D.2: Landau levels of a massive system with parabolic dispersion relation. (a) shows the continuous evolution of the LLs in the integer quantum hall regime, assuming that the LLs have a two-fold degeneracy. In contrast to monolayer Gr, the LLs show a linear evolution along the energy axis and has no Landau levels at zero energy. (b) shows the line cuts of the LLs at different magnetic fields. The density of the LL is much higher due to the cyclotron frequency. This is the prohibiting factor for observing quantum hall at high temperatures. The parameters used for this fan diagram are based on the electron doped GaAs properties.

D.4 Relativistic vs Non-Relativistic Landau Levels Comparison

Note that the cyclotron frequencies gives us a couple pieces of information. It tells us the rate of orbital motion and this is the critical parameter that defines whether we are in the quantum/strong field regime. Quantum hall effect can only be observed if the cyclotron orbit can complete and so the following conditions must be satisfied:

$$\omega_{C,D}\tau \gg 1, \tag{D.32}$$

$$\hbar\omega_{C,D} \gg k_B T, \tag{D.33}$$

where τ stands for any scattering times and $k_B T$ is the thermal smearing that can mask quantization effects. $\hbar\omega_{C,D}$ is the characteristic energy, defining the spacing between the quantized Landau levels. It is worth noting that $\hbar\omega_C \approx 1$ K, whereas $\hbar\omega_D \approx 1500$ K. This means that quantum hall is

not only more easily observable in Gr (or relativistic systems), it can also be observed in room temperature.[204] Figures D.1 and D.2 show the LLs of a relativistic system (monolayer Gr) and a non-relativistic system (massive particle system with parabolic dispersion relation) respectively. The LLs for the non-relativistic system assumes the electron-doped GaAs properties.

D.5 Quantum Hall Edge States and Transport Properties

In the incompressible state, the bulk of the quantum hall system is insulating and its transport is dictated by the edge channels. First for convenience, we introduce the following ladder operators:

$$\begin{aligned} \mathbf{a} &= \frac{l_B}{\sqrt{2}\hbar}(\Pi_x - i\Pi_y), \\ \mathbf{a}^\dagger &= \frac{l_B}{\sqrt{2}\hbar}(\Pi_x + i\Pi_y), \end{aligned} \tag{D.34}$$

At the boundary of the device a confining potential $V_{\text{conf}}(y)$ must be imposed to the Gr Hamiltonian. We choose the confinement direction to be along y to respect the translation invariance along x assumed in our choice of Landau gauge. Focusing on a single valley we have:

$$\begin{aligned} \mathbf{H}(\mathbf{q}) &= v_F \boldsymbol{\sigma} \cdot \boldsymbol{\Pi} + V(y)\sigma_z \\ &= \xi\sqrt{2}\frac{\hbar v_F}{l_B} \begin{bmatrix} 0 & \mathbf{a} \\ \mathbf{a}^\dagger & 0 \end{bmatrix} + \begin{bmatrix} V(y) & 0 \\ 0 & -V(y) \end{bmatrix}, \end{aligned} \tag{D.35}$$

where $\boldsymbol{\sigma} = (\sigma_x, \sigma_y), \xi = +1, -1$ for $\mathbf{K}, \mathbf{K}'$ valleys respectively. It is important to note that with this shorthand notation for addressing the different valleys, to keep the equations consistent, the bases for the $\mathbf{K}, \mathbf{K}'$ valleys become $\{\Phi_{A,\mathbf{K}}, \Phi_{B,\mathbf{K}}\}$ and $\{\Phi_{B,\mathbf{K}'}, \Phi_{A,\mathbf{K}'}\}$ respectively. The confinement comes in the form of a mass term in order for it to be consistent with the bandgap opening for relativistic

particles (otherwise it will become a Klein tunneling problem where electrons are immune to the potential. Same reasoning as in section A.4). We can arrive at an analytical solution if we expand the confinement potential around y_0:

$$V(y) \approx V(y_0) + \left.\frac{\partial V_{\text{conf}}}{\partial y}\right|_{y_0} + \ldots, \tag{D.36}$$

and keeping only the zeroth order term. This zeroth order term corresponds to the average of the potential that each harmonic oscillator (or extended strip) feels. Remember $y_0 = l_B^2 q_x$. The eigenvalues of the Hamiltonian with $V(y) = V(y_0)$ are:[205]

$$\varepsilon_{\text{LL},N\neq 0} = \pm\sqrt{V(y_0)^2 + 2\left(\frac{\hbar v_F}{l_B}\right)^2 N}, \tag{D.37}$$

$$\varepsilon_{\text{LL},N=0} = \xi V(y_0). \tag{D.38}$$

With the potential profile at the edges, we can now compute the velocity of the charged particles for point y_0 in the x direction (the direction along the extended states) which is given by:

$$v_x(q_x) = \frac{1}{\hbar}\frac{\partial \varepsilon_{\text{LL}}}{\partial q_x} = \frac{l_B^2}{\hbar}\frac{\partial V(y_0)}{\partial y_0} = \frac{1}{eB_z}\frac{\partial V(y_0)}{\partial y_0}, \tag{D.39}$$

and so the total current for a single Landau level is:

$$\begin{aligned} I_x &= \int dq_x \lambda v_x(q_x) \\ &= \frac{-e}{L_x}\frac{L_x}{2\pi}\int_0^{L_y/l_B^2} dq_x v_x(q_x) \\ &= \frac{-e}{2\pi l_B^2}\int_0^{L_y} dy_0 \frac{1}{eB_z}\frac{\partial V(y_0)}{\partial y_0} \\ &= \frac{-e}{2\pi\hbar}\big(V(L_y) - V(0)\big) \\ &= -\frac{e}{h}\Delta\mu. \end{aligned} \tag{D.40}$$

In the above we assume an electron current where $\lambda = -|e|/L_x$ is the line charge density. For positive current flow, it is the reverse. The Hall voltage and Hall resistance respectively becomes:

$$V_H = \frac{\Delta\mu}{-e}, \tag{D.41}$$

$$R_{xy} = \frac{V_H}{I_x} = \frac{h}{e^2} = R_Q. \tag{D.42}$$

These are the values for a single LL (with a degeneracy of one). R_Q is the quantum of resistance with a value of $R_Q = 25.812807557(18)\text{k}\Omega$. The description extends to many LLs in which case each LL contributes an additional channel (multiplied by the appropriate degeneracies). Below we write down useful quantities mostly pertaining to Gr:

$$\sigma_{xy} = 4\left(N + \frac{1}{2}\right)\frac{e^2}{h}, \tag{D.43}$$

$$G_Q = \frac{e^2}{h}, \tag{D.44}$$

$$\nu = 4\left(N + \frac{1}{2}\right) = \pm 2, \pm 6, \pm 10.... \tag{D.45}$$

Here the factor of four in front is due to the spin and valley degeneracy: $g = g_s g_v = 2 \times 2 = 4$. Compared to the conventional Hall conductance there is a $1/2$ shift in conductivity and this is due to the Berry phase of Gr.

D.6 Berry Phase and Graphene's Unconventional Filling Factors

Berry phase is the additional phase that the wavefunction acquires while traversing a full orbit in momentum space. It is a consequence of the wave property of quantum mechanical objects and it cannot be captured by the semiclassical method. Gr is peculiar partly for its non-trivial Berry phase. Typical wavefunctions would acquire a full 2π phase when making a full orbit but not Gr. Detailed in A.6, Gr acquires instead a π phase and one can trace that this to Gr's lattice. We won't go into the details about Berry phase and here we will simply acknowledge its value for Gr. For the experimental and theoretical details on Berry phase, the reader is referred to the cited references.[206,187]

To see how this factors into Gr's anomalous quantum hall, we start with Bohr-Sommerfeld's quantization condition (equation D.6) which was written in terms of angular momentum, and write it in terms of phase by dividing through with $\hbar$. Along with Peierl's substitution to account for the magnetic field ($\mathbf{p} \to \mathbf{p} + e\mathbf{A}$) we get:

$$\oint_C d\mathbf{r} \cdot \mathbf{k} + \frac{e}{\hbar} \oint_C d\mathbf{r} \cdot \mathbf{A} + \Gamma_B = 2\pi(N + \gamma). \tag{D.46}$$

An extra Γ_B phase was added to account for the Berry phase acquired when integrating a full circle. $\Gamma_B = \pi$ for Gr. The first term is the Bloch phase (or de-Broglie phase). The second term is the Aharonov-Bohm phase. All these phases must add up to be equal to the phase required for construc-

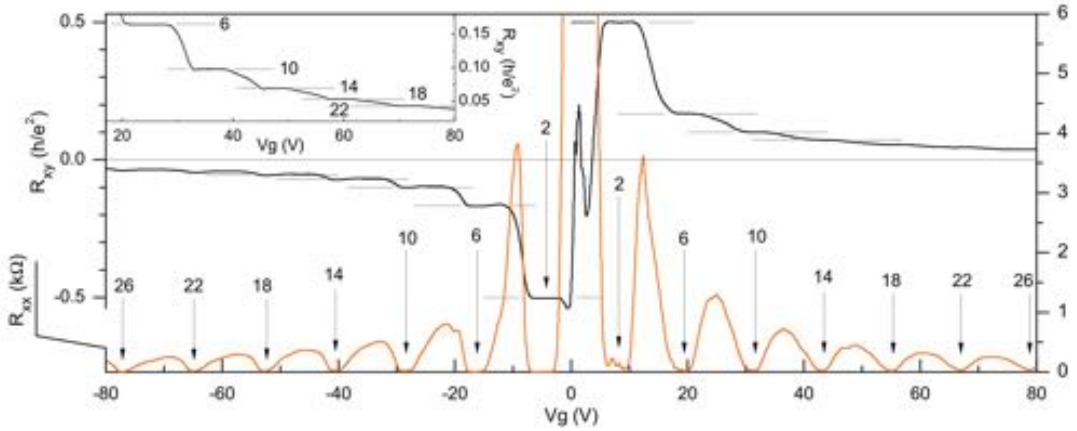


Figure D.3: Measurement of Quantum Hall Effect in graphene. Plot is adapted from the work by Zhang et al.[206] The four terminal longitudinal resistance measurement (R_{xx} colored orange) goes to zero when Landau Levels are filled and becomes non-zero once a Landau level becomes partially filled. It becomes zero because when Landau levels are filled, edge states carry the conducting charges across the sample and these states are ballistic and do not drop in potential until equilibrated with the source or drain. The zeros in R_{xx} thus indicate the full filling of Landau levels. The sequence is called the filling factor where each factor refers to the number of Landau levels filled. At the zeros in R_{xx}, R_{xy} will be a plateau. This is because when all Landau levels are filled, the bulk is insulating, i.e. charges are localized. Consequently, there cannot be any charge movement transversely and hence there cannot be any change in potential drop.

tive interference such that an orbit can exist. From semiclassical quantum mechanics we know:

$$\hbar\dot{\mathbf{k}} = -e(\mathbf{v} \times \mathbf{B}), \tag{D.47}$$

$$\hbar\mathbf{k} = -e(\mathbf{r} \times \mathbf{B}) + \text{const.}. \tag{D.48}$$

Plugging equation D.48 into D.46 and making use of Stoke's theorem where $\nabla \times \mathbf{A} = \mathbf{B}$, we get:

$$\begin{aligned} \mathbf{B} \cdot \oint_C d\mathbf{r} \cdot \mathbf{r} - \int_S d\mathbf{S} \cdot \mathbf{B} &= \frac{2\pi\hbar}{e}(N + \gamma - \gamma_B) \\ 2\Phi - \Phi &= \Phi_0(N + \gamma - \gamma_B) \\ \Phi &= \Phi_0(N + \gamma - \gamma_B) \end{aligned} \tag{D.49}$$

where S is the orbit area and $\Phi_0 = h/e$ is the quantum of flux. $\gamma_B = \Gamma_B/2\pi$ is the contribution from Berry phase. For Gr, $\gamma_B = 1/2$. Dividing throughout by B, we would get the equation in terms of area in real space. Converting that into momentum space by using equation D.48, we get:

$$\pi k_N^2 = (N + \gamma - \gamma_B)\frac{2\pi eB}{\hbar} = \begin{cases} N\frac{2\pi eB}{\hbar} & \text{for Gr} \\ (N + \frac{1}{2})\frac{2\pi eB}{\hbar} & \text{for parabolic bands} \end{cases}, \tag{D.50}$$

Remember, $\gamma = 1/2$ is due to the zero point motion. $\gamma_B = 1/2$ for Gr and $\gamma_B = 0$ for parabolic bands that has no Berry phase. From here we can see how this affects the filling factor. The conditions above dictate the positions of LLs. For Gr there exists one at the origin of the Dirac cone ($k = 0$) whereas for materials with zero Berry phase it starts away from the origin ($k \neq 0$). The quantity N is the Landau level index as seen throughout the appendix. As a result of the Berry phase, it becomes $1/2$ shifted down and coincidentally to the Dirac point. This ultimately prevents the condition of full filling of LLs at zero energy and hence the filling factor cannot be zero for Gr.

D.7 Briefly on Highly Correlated Phases

The quantum hall effect is a very fascinating phenomenon. It is mainly so because in the quantum hall regime, the quantizing magnetic field serves to quench the kinetic energy of the electrons. This can be seen by the fact that LLs are flat (kinetic energy is a direct consequence of dispersing bands because they are responsible for giving particles a velocity). The quenching of kinetic energy means particles' behavior are completely dominated by its potential energy which come from electric and magnetic fields. The integer quantum hall effect described in this section is a single particle picture of the quantum hall effect. What we see is that the magnetic field is localizing the electrons, but not

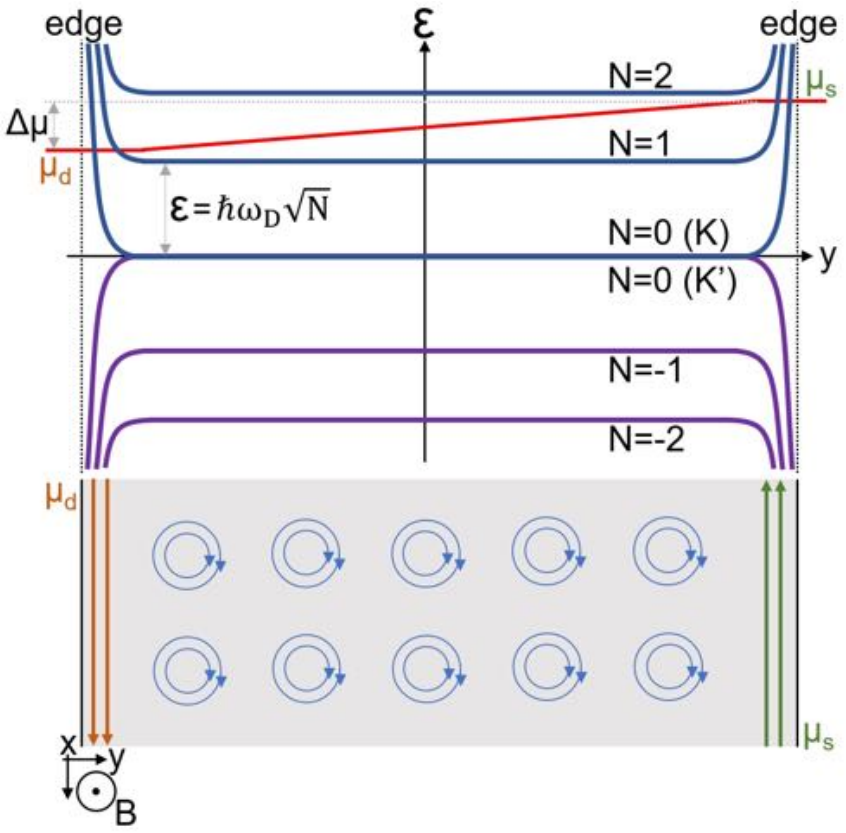


Figure D.4: This is a simplistic semiclassical picture of the quantum hall effect. Here the cyclotron orbits make up the insulating bulk. They are drawn to be pinned down in a regular pattern but this is not necessarily the case in reality. Not drawn are impurities and their associated equipotentials throughout the bulk. The cyclotron orbit motion is superimposed on a drift trajectory determined by the drift velocity. This drifting trajectory is along equipotential lines. Also not drawn is the tilt of the LLs with respect to real space. The existence of a Hall voltage necessarily means the existence of an electric field which serves to shift energy levels with respect to real space. Also not indicated precisely is the chemical potential line. The chemical potential can only change in value at the insulating (or equivalently, gapped or incompressible) regions of the sample. This is because conducting channels (also known as compressible stripes) behave like a perfect metal. Perfect metals can only have a constant potential hence the chemical potential must be constant within the conducting channels. Finally, this picture is a semiclassical interpretation. Electrons do not have traceable trajectories but have a probability cloud. This picture refers to when $\nu = 2$, which means two orbiting electrons per flux quantum. The probability of the electrons positions can be plotted using the wavefunctions from equation D.23. As described in the main text, the zLL has contributions from the two valleys. The behavior of the LLs for each valley are opposite and hence the different bending direction at the edges. The bending is due to the confinement potential. Not shown are the degeneracies of the LLs.

completely. If magnetic fields go even higher, the kinetic energy can be completely quenched. The chaotic and unregulated motions of the electrons will be gone and electrons will establish ordered phases due to electron-electron interactions. In this regime, the physics becomes purely a many-body problem and many exotic phenomena can appear such as the fractional quantum hall effect. Much of the work in this thesis falls onto the single-particle regime. It will definitely be interesting to bring the experiments described here to this many-body regime.

E

The Hydrodynamic Description of Electrons in Graphene

Hydrodynamics, as the name suggests, is a theory that models the dynamics of water. But, as a matter of fact, water is only one of many types of liquid that the theory can aptly describe. It turns out that liquids in general are all very similar to one another. For instance, water and honey are both liquids in that they deform continuously. So long as we ignore the constituents of said liquids, they are both just something "flowy" except one flows more quickly than the other. Hydrodynamics is blind to the

atoms or molecules that make up these entities. Instead, it coarse grains all the microscopic mechanisms into several macroscopic properties described by a universal set of macroscopic laws fitted with phenomenological parameters. The only underlying assumption of this theory is that the object of interest is continuous in real space. This then begs the question: what does it mean to be continuous? In this chapter we shall see what the requirements are.

E.1 Introduction to Hydrodynamics in Graphene

Let us first lay out explicitly the definitions of variables and nomenclatures.

$$
\begin{aligned}
n &= N/\text{Volume} = \text{number density } [\text{m}^{-3}] \\
\rho &= mn = \text{mass density } [\text{kg} \cdot \text{m}^{-3}] \\
\mathbf{p} &= mn\mathbf{v} = \rho\mathbf{v} = \text{momentum density or mass flux } [\text{kg} \cdot \text{m}^{-2} \cdot \text{s}^{-1}] \\
\mu &= \eta = \text{dynamic viscosity } [\text{kg} \cdot \text{m}^{-1} \cdot \text{s}^{-1}] \\
\nu &= \mu/\rho = \text{kinematic viscosity } [\text{m}^{2} \cdot \text{s}^{-1}] \\
\phi &= -\boldsymbol{\nabla}\mathbf{E} = \text{electric potential } [\text{V}] \\
P &= -ne\phi = \text{pressure } [J \cdot \text{m}^{-3}] \\
\mathbf{J} &= n\mathbf{v} = \text{particle number current density } [\text{s}^{-1} \cdot \text{m}^{-2}] \\
\text{Re} &= |\mathbf{v}|L/\nu = \rho|\mathbf{v}|L/\mu = \text{Reynold's number} \\
\gamma_p &= 1/\tau_p = \text{momentum relaxation rate } [\text{s}^{-1}] \\
D_\nu &= \sqrt{\nu\tau_p} = \text{vorticity diffusion length}
\end{aligned}
\tag{E.1}
$$

where N is the number of particles, $e = 1.602... \times 10^{-19}$(C) is the elementary charge, m is mass (kg), $\mathbf{E}$ is the electric field (V/m) and $\mathbf{v}$ is velocity (m/s). Note that even though we are working in a so-called 2D system, the natural units of the observable quantities are measured and reported in 3D.

The foundation of fluid dynamics* is based on two governing equations and they are: the continuity and momentum equation (better known as the Navier-Stokes equation). These two equations form the basis of fluid dynamics because they embody completely the fundamental physical principles of fluid flow problems. The following are the two respective equations:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = 0, \tag{E.2}$$

$$\frac{\partial \mathbf{v}}{\partial t} = -(\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{v} - \frac{1}{\rho}\boldsymbol{\nabla} P + \nu \nabla^2 \mathbf{v} + \mathbf{f}, \tag{E.3}$$

where we suppressed the explicit variable dependencies to avoid obscuring the equations. All kinematic variables (P, $\mathbf{v}$) as well as mass density (ρ) are functions of temperature (T), density (n), position ($\mathbf{r}$) and time (t). The continuity equation (equation E.2) is a restatement of conservation of mass globally while the Navier Stokes equation (equation E.3) is the application of Newton's second law on a fluid element. It states that the velocity of a fluid element can change in response to convection ($-(\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{v}$), spatial variation in pressure ($-\boldsymbol{\nabla} P/\rho$), viscous drag force ($\nu \nabla^2 \mathbf{v}$), and other external forces ($\mathbf{f} = \mathbf{F}/m$, force per unit mass, which is the acceleration of the fluid element).

*Hydrodynamics is a subset of fluid dynamics which is a branch of fluid mechanics that falls under the big umbrella of continuum physics. Hydrodynamics deals specifically with liquids in motion, in contrast to aerodyanmics, which deals with gases in motion, but both are part of fluid dynamics. Dynamics, the study of the causation of motion, is one of three branches of mechanics, the other two being statics and kinematics. Statics is concerned with an object's stationary properties and kinematics is about the properties of a moving object irrespective of why the object was set in motion (this is where the subtle distinction between dynamics and kinematics lies. Dynamics focuses on the interactions, such as force, that can cause things to move). And finally, continuum physics encompasses the study of anything that is continuous, ignoring intentionally the discrete nature of reality.

To describe fluid flow in graphene (Gr), we first note its Reynold's number:

$$\frac{(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v}}{\nu\nabla^2\mathbf{v}} \approx \mathrm{Re} \equiv \frac{|\mathbf{v}|L}{\nu} \approx \frac{(10^6\,\mathrm{m}\cdot\mathrm{s}^{-1})\cdot(10^{-6}\,\mathrm{m})}{(0.1\,\mathrm{m}^2\cdot\mathrm{s}^{-1})} = 10 \ll 10^3, \tag{E.4}$$

where we used Gr's Fermi velocity (v_F) for its characteristic velocity and assumed the characteristic device length to be on the order of a micron. The value for kinematic viscosity was taken from the cited references where the value was both calculated theoretically and measured experimentally[207,5]. The Reynold's number is a characterization of the relative strength between the inertial force and the viscous force. If the viscous force dominates such as in the case for Gr, the fluid flow would be laminar. The threshold of 10^3 is approximately the order of magnitude at which turbulence begins to set in. The low Re for Gr means that we can safely ignore the convective term, hence allowing us to linearize the Naiver-Stokes equation (by setting $(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v} = 0$). Additionally, we make two extra assumptions and they are: sample is isothermal and the fluid is incompressible. The two assumptions are reasonable— the small size of the device is unlikely to be subject to large temperature variation and the equilibrium electron density is dictated electrostatically by the global gates. These two assumptions allow us to say that the kinematic variables are not a function of density or temperature (i.e., $P = P(\mathbf{r})$ and $\mathbf{v} = \mathbf{v}(\mathbf{r})$) and that mass density is constant (ρ =constant). Finally, to account for the possible presence of momentum relaxing scattering events, we have $\mathbf{f} = -\gamma_p\mathbf{v}$. And so putting everything together, in steady-state transport, the governing equations can be simplified to:

$$\boldsymbol{\nabla}\cdot\mathbf{v}(\mathbf{r}) = 0, \tag{E.5}$$

$$\frac{1}{\rho}\boldsymbol{\nabla}P(\mathbf{r}) = \nu\nabla^2\mathbf{v}(\mathbf{r}) - \gamma_p\mathbf{v}(\mathbf{r}). \tag{E.6}$$

By multiplying both equations by the number density (n) and rearranging terms, we can rewrite the equations in terms of quantities that are more experimentally relevant:

$$\boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{r}) = 0, \tag{E.7}$$

$$\frac{\sigma}{e}\boldsymbol{\nabla}\phi(\mathbf{r}) + D_\nu^2 \nabla^2 \mathbf{J}(\mathbf{r}) = \mathbf{J}(\mathbf{r}). \tag{E.8}$$

where $\mathbf{J}(\mathbf{r})$ is the particle number current density ($\mathrm{s}^{-1}\cdot\mathrm{m}^{-2}$) and not the charge current density ($\mathrm{C}\cdot\mathrm{s}^{-1}\cdot\ \mathrm{m}^{-2}$) and where $\sigma = ne^2\tau_p/m$ is the Drude conductivity. Note that in the absence of viscosity, equation E.6 reduces to the local version of Ohm's Law:

$$e\mathbf{J}(\mathbf{r}) = \sigma\mathbf{E}(\mathbf{r}). \tag{E.9}$$

In the case of 2D flow in the xy plane, we can follow a very common methodology in the field of hydrodynamics and introduce the following transformation:

$$v_x = \frac{\partial\psi}{\partial y}, \quad v_y = -\frac{\partial\psi}{\partial x}, \tag{E.10}$$

$$\zeta = \frac{\partial v_x}{\partial y} - \frac{\partial v_y}{\partial x}, \tag{E.11}$$

where $\psi = \psi(x, y)$ is the stream function and $\zeta = \zeta(x, y)$ is the vorticity. The new set of functions is merely another way to capture the velocity profile in 2D while maintaining the condition imposed by the continuity equation. Expressing the equations in terms of the stream function and vorticity will significantly simplify and decouple the system of equations. To see this, let us return to using the

kinematic variables and rewrite equations E.5 and E.6 in terms of their x, y components:

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0, \tag{E.12}$$

$$\frac{1}{\rho}\frac{\partial P}{\partial x} = \nu\left(\frac{\partial^2 v_x}{\partial x^2} + \frac{\partial^2 v_x}{\partial y^2}\right) - \gamma_p v_x, \tag{E.13}$$

$$\frac{1}{\rho}\frac{\partial P}{\partial y} = \nu\left(\frac{\partial^2 v_y}{\partial x^2} + \frac{\partial^2 v_y}{\partial y^2}\right) - \gamma_p v_y. \tag{E.14}$$

We differentiate equation E.13 with respect to y and differentiate equation E.14 with respect to x. Subtracting them and writing the equation in terms of the stream function and vorticity will yield equation E.15. Likewise, we differentiate equation E.13 with respect to x and differentiate equation E.14 with respect to y. Adding them and writing the equation in terms of the stream function and vorticity will yield equation E.16.

$$\nabla^2 \zeta = \frac{\gamma_p}{\nu}\zeta, \tag{E.15}$$

$$\nabla^2 P = 0. \tag{E.16}$$

The final equation that completes the full set of equations for the three variables (ψ, ζ, P) comes from combining equation E.10 and E.11 to get:

$$\nabla^2 \psi = \zeta. \tag{E.17}$$

Equations E.15, E.16, E.17 encompasses all the information regarding the dynamics of the system. As expected, γ_p, the momentum relaxing contribution, is in direct competition with ν, the kinematic viscosity.

To the next most important detail— boundary conditions. The tangential component of the

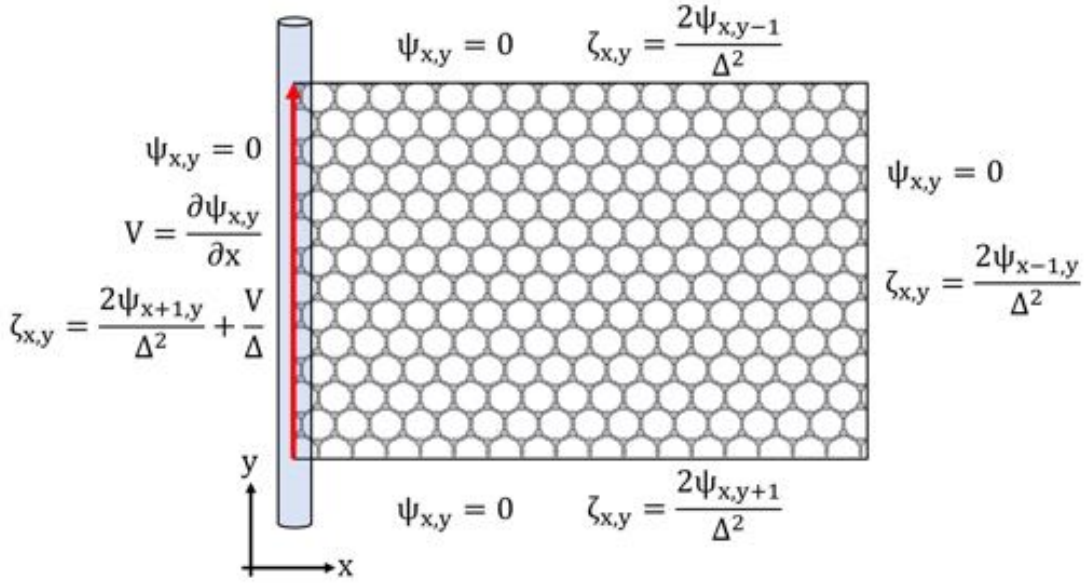


Figure E.1: Boundary conditions for a system with a moving edge (indicated by red arrow). Δ is the step size in the simulation. The shown equations correspond to the conditions for their associated edges.

boundary condition can by summarized generically by:

$$v^{||} - v_b^{||} = l_b \frac{\partial \mathbf{v}}{\partial \mathbf{n}}, \tag{E.18}$$

where $\mathbf{n}$ is the unit vector normal to the boundary, $v^{||}$ is the fluid velocity tangential to the boundary that we are trying to determine, $v_b^{||}$ is the fluid velocity along and at the boundary and l_b is the characteristic slip length. If we assume the no-slip boundary condition, $l_b = 0$, we get the obvious result: $v^{||} = v_b^{||}$, and this will be the ongoing assumption. The perpendicular component of the boundary condition at a wall boundary will be $v_b^{\perp} = 0$. This can safely be assumed if we expect all the fluid elements to be contained within the hard boundaries.

Armed with all the necessary equations and information regarding boundary conditions, we can simulate grapehen electrons being dragged by moving electrons in a carbon nanotube (as shown in

figure 7.4a). This is completely analogous to the driven-lid cavity problem except here we have a momentum relaxation term. Below we summarize the boundary conditions pertaining to the system shown in figure E.1

$$\begin{aligned} &\text{Static Boundary (no slip and no perpendicular flow)} : v_b^{||} = 0, v_b^{\perp} = 0 \\ &\rightarrow \psi|_b = 0, \left.\frac{\partial \psi}{\partial x}\right|_b = v_y|_b = 0, \left.\frac{\partial \psi}{\partial y}\right|_b = v_x|_b = 0. \end{aligned} \tag{E.19}$$

The exact boundary conditions applied to the simulation are acquired by the Taylor expansion of the streamline function. Focusing on the bottom edge of the system shown in figure E.1 and using the information we already have, we get:

$$\begin{aligned} \psi_{x,y+1} &= \psi_{x,y} + \Delta \left.\frac{\partial \psi}{\partial y}\right|_{x,y} + \frac{\Delta^2}{2} \left.\frac{\partial^2 \psi}{\partial y^2}\right|_{x,y} + ... \\ &\approx 0 + 0 + \frac{\Delta^2}{2} \left.\frac{\partial^2 \psi}{\partial y^2}\right|_{x,y} \\ &= \frac{\Delta^2}{2} \zeta. \end{aligned} \tag{E.20}$$

For the moving edge, the first derivative of the streamline function is retained and given by a fixed velocity. The method of finite difference was performed to solve equations E.15 and E.17 while accounting for the aforementioned boundary conditions. The simulations are shown in figure E.2.

According to the simulations of a graphene with a 1D moving boundary condition, unlike the current injection in graphene, vortices will always be formed whether the electron liquid is hydrodynamic or not. The vortex size, however, does depend on the viscosity strength. As implied by equation E.15, kinematic viscosity (ν) and the momentum relaxing scattering time ($\tau = 1/\gamma_p$) is in direct competition. The vortex diffusion length ($D_\nu = \sqrt{\nu\tau}$) is a good quantity that characterizes the prominence

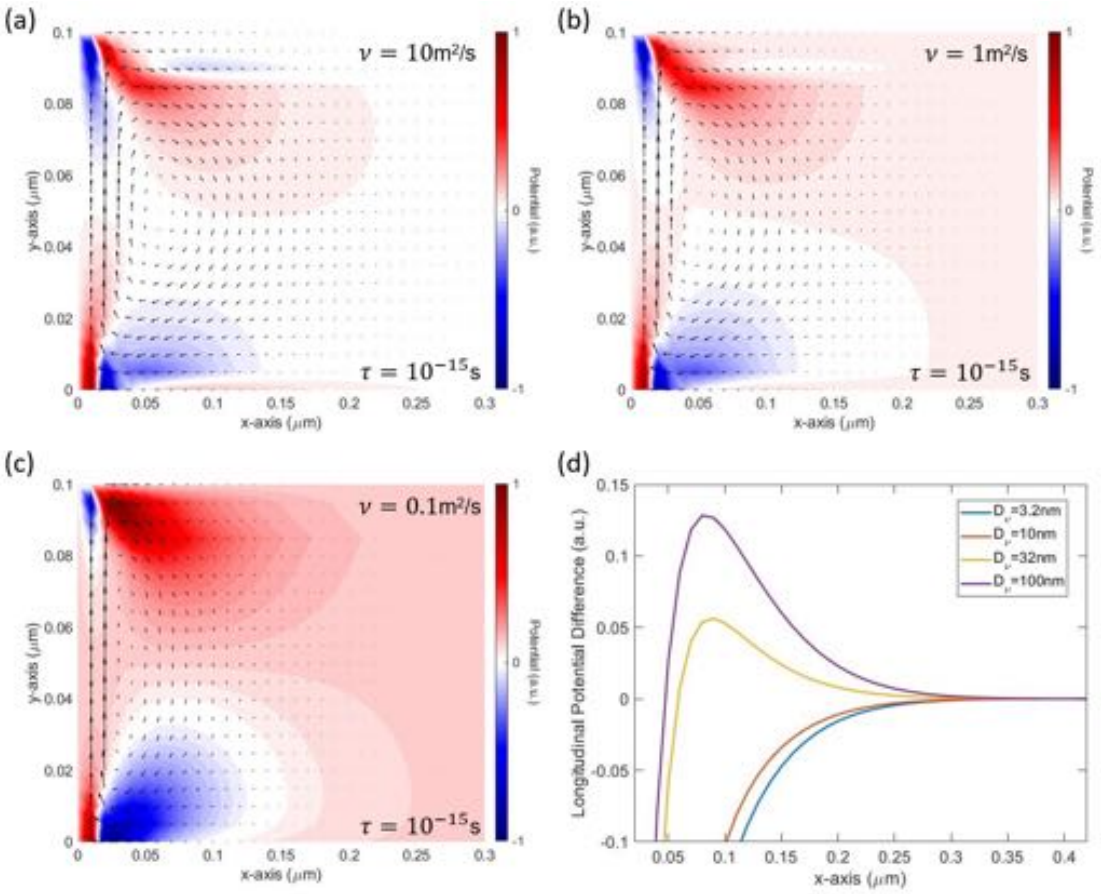


Figure E.2: The hydrodynamic simulation of the electronic liquid in graphene with a moving boundary condition. (a) to (c) show the different simulations based on the equations E.15 and E.17. ν is the kinematic viscosity and τ is the momentum-relaxing scattering time (giving rise to the Ohmic contribution). Vortices are formed regardless of the strength of viscosity and Ohmic contributions. The size of the vortex does change, however. The larger the viscosity, the larger the vortex. ν and τ are directly competing against each other. $D_\nu = \sqrt{\nu\tau}$ is the vorticity diffusion length and characterizes the prominence of hydrodynamics.

of hydrodynamics. According to the simulations, when hydrodynamics become more prominent, i.e. when $D_\nu \gtrsim 30$ nm, we would expect a sign change along the transverse spatial direction of the drag device, much like the signature from the direct current injection method. We can take this to be an indication of the existence of a hydrodynamic liquid.

E.2 The Hydrodynamic Window in Monolayer Graphene

The one criterion for any system to behave as a hydrodynamic fluid is that constituents must collide amongst themselves fast enough such that a thermal equilibrium is established locally in a space that is much larger than the characteristic size of the individual constituents as well as the sizes of their events (such as the mean free path, i.e. distance between a collision). This renders the motion of the constituents redundant in describing the global system's behavior. Instead, the system can be described well by a macroscopic thermally averaged quantity known as thermodynamic variables. Such quantities relevant to our study are: density, temperature and velocity. The three quantities form the complete set of variables in our equation of state.

Said in a different way, to create a hydrodynamic fluid, electrons cannot lose significant amount of energy or momentum to foreign entities such as phonons and impurities. Compactly the condition for hydrodynamics is:

$$\tau_{ee} \ll \tau_p \tag{E.21}$$

where τ_{ee} is the electron-electron scattering time and τ_p is the scattering time associated with any momentum-relaxing events. The two main types of collision that contribute to τ_p are: τ_{e-ph} (electron-phonon scattering, albeit being possibly quasi-elastic but still inelastic in nature), $\tau_{e-\mathrm{imp}}$ (electron-impurity scattering, albeit if impurity is large and heavy enough the scattering can be elastic and momentum conserving). In Gr, there exists a temperature window where the hydrodynamic condition is held. In the following, we will summarize the information regarding the different scattering rates.

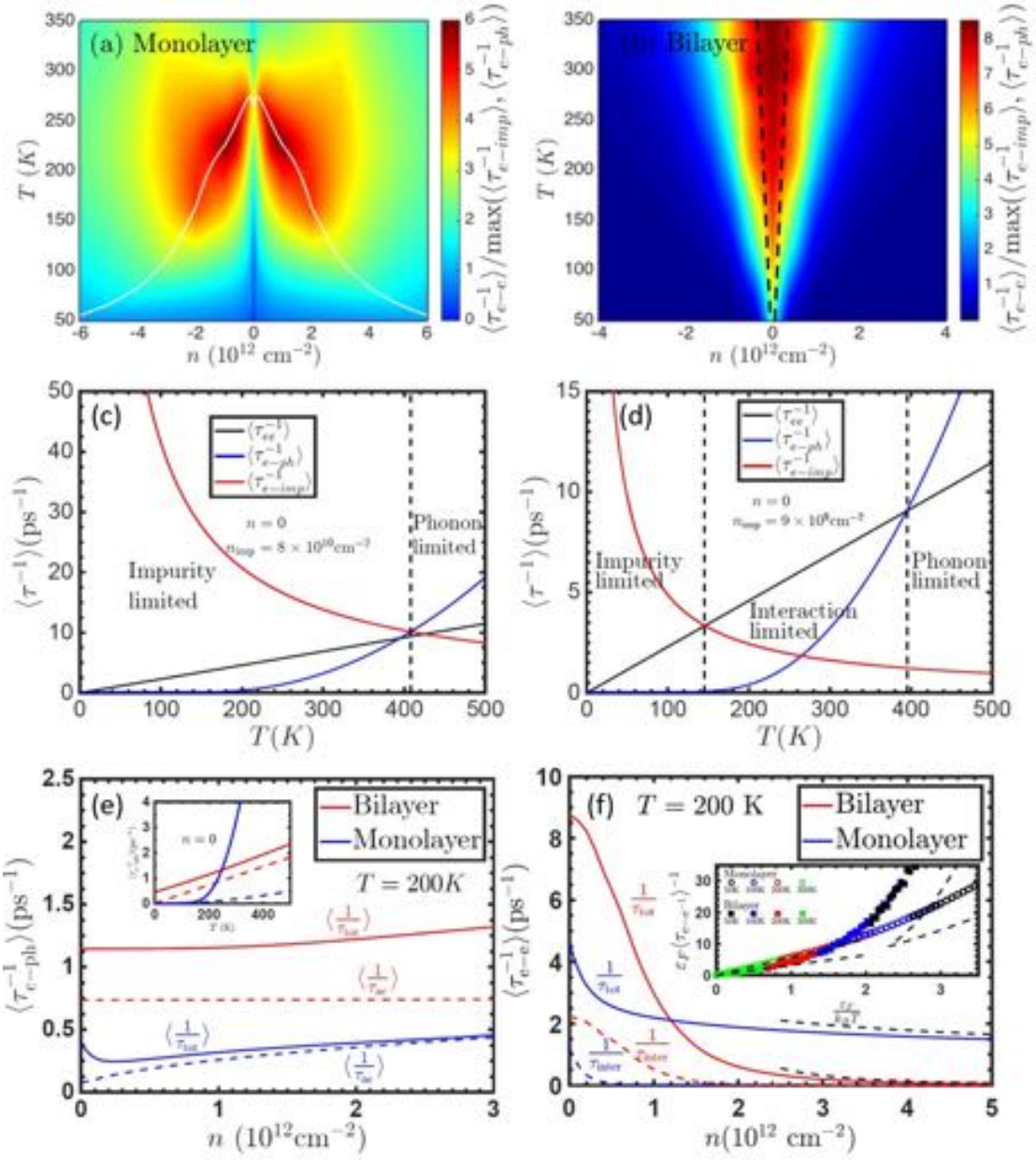


Figure E.3: Summary of scattering rates in graphene extracted from the works by Ho et al.[208,209] . (a) and (b) show the window (red) where e-e scattering dominates for mono and bilayer Gr respectively. (c) and (d) contrasts the hydrodynamic window opening for disordered (on SiO_2) and clean (on hBN) charge neutral Gr respectively. (e) and (f) show the $e-ph$ and $e-e$ scattering rates respectively (total is given by solid lines where blue and red are for mono and bilayer Gr respectively) at different carrier densities. From these plots we see that hydrodynamics is expected above 150K in clean Gr and is more prominent closer to CNP.

E.2.1 Scattering Rates Comparison

The momentum relaxing scattering time associated with e-imp scattering is a value typically acquired experimentally through the assumption of the Drude model:

$$\tau_{e-\text{imp}} = \frac{\hbar}{ev_F}\mu\sqrt{\pi n}, \tag{E.22}$$

where μ is the mobility of the sample. The associated length is taken to to be the same as the mean free path. The e-e scattering time varies depending whether Gr is near or away from the Dirac point. Its behavior is given by:[114]

$$\tau_{ee} = 0.1\,\text{ps}\frac{100\,\text{K}}{T}\cdot\left(1, \frac{\varepsilon_F}{k_B T}\right), \tag{E.23}$$

where the higher value inside the parenthesis is taken. This is to say that e-e scattering scales as T^{-1} when in the Dirac fluid regime as opposed to the T^{-2} in the Fermi liquid regime (see figure 7.2). The e-ph scattering time scales as:[172]

$$\tau_{e-ph} \propto \begin{cases} T^{-4} & \text{for } T \lesssim 0.1T_{BG} \\ T^{-1} & \text{for } T \gtrsim 0.1T_{BG} \end{cases}, \tag{E.24}$$

where $T_{BG} \approx 1000$ K is the Bloch-Grüneisen temperature. The behavior of the phonon scattering tells us that it only becomes appreciable at high enough temperature. The rough temperature where phonons become experimentally appreciable for a hBN encapuslated Gr device is usually assumed to be about $0.1T_{\text{BG}} \approx 100$ K.[172,210] Below we very roughly summarize the order of magnitudes of what to expect for the different scattering mechanisms:

	$\mu = 10$k	$\mu = 50$k	$\mu = 100$k
$\tau_{e-\text{imp}}$	0.1 ps	0.5 ps	1 ps
$l_{e-\text{imp}}$	100 nm	500 nm	1 μm
	$T = 10$ K	$T = 100$ K	$T = 200$ K
τ_{ee}	1 ps	0.1 ps	0.05 ps
l_{ee}	1 μm	100 nm	50 nm
τ_{e-ph}	1 ns	10 ps	1 ps
l_{e-ph}	10 mm	10 μm	1 μm

This table is only meant as a qualitative guide. Figure E.3 shows the theoretical calculation of the different scattering rates. Looking at the different scattering rates, the hydrodynamic window opens up at around 150 K.

E.3 From Semiclassical Electron Transport to Hydrodynamics

We start from the celebrated and intuitive Boltzmann equation which comes from the time derivative of the first order expansion of the non-equilibrium distribution function $f(\mathbf{r}, \mathbf{k}, t)$:

$$\frac{\partial f}{\partial t} = -\mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} - \mathbf{F} \cdot \frac{1}{\hbar} \frac{\partial f}{\partial \mathbf{k}} + \left. \frac{\partial f}{\partial t} \right|_{\text{col}}, \tag{E.25}$$

where $f = f(\mathbf{r}, \mathbf{k}, t)$, $\mathbf{F} = \mathbf{F}(\mathbf{r})$ is the external force, $\mathbf{v} = \partial \varepsilon(\mathbf{k}) / \partial(\hbar \mathbf{k})$ is mean velocity (group velocity in the semiclassical sense) and where we have added the collision term at the end. Under the relaxation time approximation:

$$\left. \frac{\partial f(\mathbf{r}, \mathbf{k}, t)}{\partial t} \right|_{\text{col}} = \frac{f_0(\mathbf{r}, \mathbf{k}) - f(\mathbf{r}, \mathbf{k}, t)}{\tau(\mathbf{k})}, \tag{E.26}$$

where f_0 is the Fermi function which represents the distribution at thermal equilibrium:

$$f_0(\mathbf{r},\mathbf{k}) = \frac{1}{e^{(\varepsilon(\mathbf{k})-\mu(\mathbf{r}))/k_B T(\mathbf{r})}+1}. \tag{E.27}$$

The approximation states that the particles emerging from a collision will tend toward the thermal equilibrium distribution setting. Under this assumption $\tau(\mathbf{r},\mathbf{k})$ characterizes the time in which the system returns to its equilibrium configuration. We can easily see this when no external force is applied in which case the solution to the Boltzmann equation is simply (for $\mathbf{F}=0$):

$$f(\mathbf{r},\mathbf{k},t) = f(\mathbf{r},\mathbf{k},t=0)e^{-t/\tau(\mathbf{r},\mathbf{k})} + f_0(\mathbf{r},\mathbf{k}). \tag{E.28}$$

Going from the Boltzmann equation to the Navier Stokes equation, a couple underlying notions are assumed. The first is that the system is no longer quantum mechanical, meaning that we would have to forgo the Fermi statistics and assume the classical Maxwell-Boltzmann statistics instead. Distribution functions are now given by the Maxwell-Boltzmann distribution:

$$f = \left(\frac{m}{2\pi k_B T}\right)^{3/2} 4\pi v^2 e^{-\frac{mv^2}{2k_B T}}. \tag{E.29}$$

The second is along the same lines as the first and that is the nature of the system is no longer discrete but considered a continuum. This means the discrete quantum number that represents momentum, $\mathbf{k}$, is no longer fit. Instead, a more macroscopic and continuous quantity is assumed. Generally $\mathbf{p} = m\mathbf{v} = \hbar\mathbf{k}$ where $\mathbf{p}$ is continuous.

The connection between the Boltzmann equation and hydrodynamics is straightforward in concept. If we take the Boltzmann equation, multiply it by some generic quantity that depends on the

same phase space, which we call $A(\mathbf{r}, \mathbf{k})$, and then integrate over all momentum ($\mathbf{p} = \hbar\mathbf{k}$) we would get:

$$\int d^3(\hbar k)\, A(\mathbf{r}, \mathbf{k}) \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \frac{\partial}{\partial \mathbf{r}} + \mathbf{F} \cdot \frac{1}{\hbar} \frac{\partial}{\partial \mathbf{k}} \right) f = \int d^3(\hbar k)\, A(\mathbf{r}, \mathbf{k}) \left. \frac{\partial f}{\partial t} \right|_{\text{col}}, \tag{E.30}$$

where functions introduced earlier have the same dependencies as before. The motivation behind this is to create an equation that describes how this observable quantity A flows or changes in time and space while at the same time eliminate the momentum dependence. Because the number density of particles is given by:

$$n(\mathbf{r}, t) = \int d^3(\hbar k)\, f(\mathbf{r}, \mathbf{k}, t), \tag{E.31}$$

and the momentum-averaged quantity is defined to be:

$$\langle A(\mathbf{r}, t) \rangle = \frac{1}{n(\mathbf{r}, t)} \int d^3(\hbar k)\, A(\mathbf{r}, \mathbf{k}) f(\mathbf{r}, \mathbf{k}, t), \tag{E.32}$$

equation E.30 can be rewritten into (with the use of integration by parts):

$$\frac{\partial}{\partial t} \langle nA \rangle + \frac{\partial}{\partial \mathbf{r}} \cdot \langle n\mathbf{v}A \rangle - n \left\langle \mathbf{v} \cdot \frac{\partial A}{\partial \mathbf{r}} \right\rangle - n \left\langle \mathbf{F} \cdot \frac{1}{\hbar} \frac{\partial A}{\partial \mathbf{k}} \right\rangle = \int d^3(\hbar k)\, A \left. \frac{\partial f}{\partial t} \right|_{\text{col}}, \tag{E.33}$$

where $A = A(\mathbf{r}, t)$ now. The above equation is the so-called master equation which is the key to obtaining the governing equations of hydrodynamics. If we consider the collisionless case ($\partial f / \partial t|_{\text{col}} = 0$), setting $A = 1$ will return the continuity equation (equation E.2). In the same collisionless scenario,

if we set $A = \hbar\mathbf{k} = m\mathbf{v}$, we get the momentum-averaged Boltzmann equation which is:

$$\frac{\partial}{\partial t}(mn\langle\mathbf{v}\rangle) + \frac{\partial}{\partial \mathbf{r}} \cdot mn\langle\mathbf{v}\,\mathbf{v}\rangle - \langle n\mathbf{F}\rangle = 0, \quad \text{for } \frac{\partial f}{\partial t}\bigg|_{\text{col}} = 0, \tag{E.34}$$

where $\mathbf{v}\mathbf{v} \equiv \mathbf{v} \otimes \mathbf{v} \equiv \mathbf{v}\mathbf{v}^{\mathsf{T}}$ is the dyadic (or outer) product that results in a second rank tensor. We expand the second term using this relation:

$$\begin{aligned}
\langle \boldsymbol{\delta}\mathbf{v}\boldsymbol{\delta}\mathbf{v}\rangle &= \langle(\mathbf{v} - \langle\mathbf{v}\rangle)(\mathbf{v} - \langle\mathbf{v}\rangle)\rangle \\
&= \langle\mathbf{v}\,\mathbf{v}\rangle - \langle\mathbf{v}\langle\mathbf{v}\rangle\rangle - \langle\langle\mathbf{v}\rangle\mathbf{v}\rangle + \langle\langle\mathbf{v}\rangle\langle\mathbf{v}\rangle\rangle \\
&= \langle\mathbf{v}\,\mathbf{v}\rangle - \langle\mathbf{v}\rangle\langle\mathbf{v}\rangle - \langle\mathbf{v}\rangle\langle\mathbf{v}\rangle + \langle\mathbf{v}\rangle\langle\mathbf{v}\rangle \\
&= \langle\mathbf{v}\,\mathbf{v}\rangle - \langle\mathbf{v}\rangle\langle\mathbf{v}\rangle,
\end{aligned} \tag{E.35}$$

hence turning the collisionless momentum-averaged Boltzmann equation into:

$$\frac{\partial}{\partial t}(\rho\langle\mathbf{v}\rangle) + \rho\frac{\partial}{\partial \mathbf{r}} \cdot (\langle\mathbf{v}\rangle\langle\mathbf{v}\rangle + \langle\boldsymbol{\delta}\mathbf{v}\boldsymbol{\delta}\mathbf{v}\rangle) - \langle n\mathbf{F}\rangle = 0, \quad \text{for } \frac{\partial f}{\partial t}\bigg|_{\text{col}} = 0. \tag{E.36}$$

Let us introduce conventional definitions associated with hydrodynamics. The stress tensor is defined

as:

$$
\begin{aligned}
\boldsymbol{\sigma} &= -\rho\langle \boldsymbol{\delta}\mathbf{v}\boldsymbol{\delta}\mathbf{v}\rangle \\
&= \begin{bmatrix} \sigma_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma zz \end{bmatrix} \\
&= \underbrace{-\begin{bmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{bmatrix}}_{\text{Hydrostatic}} + \underbrace{\begin{bmatrix} \sigma_{xx}+P & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_{yy}+P & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_{zz}+P \end{bmatrix}}_{\text{Deviatoric}} \\
&= -P\mathbf{I}_3 + \boldsymbol{\tau}
\end{aligned}
\tag{E.37}
$$

written in the $\{\mathbf{r}\} = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} = \{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\}$ basis where σ_{ii} is the normal stress component and τ_{ij} is the shear stress component. The first index of a component refers to the normal of the surface of the fluid element the stress is applied onto and the second index refers to the direction of the stress. Pressure (or hydrostatic stress) is defined to be the mean isotropic component of the stress that is normal and pointing inwards into the fluid element:

$$
P = -\frac{1}{3}(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}), \tag{E.38}
$$

where the diagonal stress component is given by:

$$
\sigma_{ii} = \frac{\partial\langle v_i\rangle}{\partial x_i} \tag{E.39}
$$

and the shear stresses are:

$$\tau_{ij} = \mu\left(\frac{\partial\langle v_i\rangle}{\partial x_j} + \frac{\partial\langle v_j\rangle}{\partial x_i}\right) + \delta_{ij}\lambda\boldsymbol{\nabla}\cdot\langle\mathbf{v}\rangle \tag{E.40}$$

Thus far we have retained $\langle\ldots\rangle$ to show that the quantity being dealt with is the momentum-averaged quantity (given by equation E.32). This momentum-average will now be assumed for the kinematic variables when dealing with hydrodynamics and for convenience, the $\langle\ldots\rangle$ will be dropped. λ is the second viscosity or bulk viscosity. It is a viscosity associated with the change in volume of the liquid. Substituting the definition of stress tensor in terms of pressure and shear stress, we get:

$$\frac{\partial}{\partial t}(\rho\mathbf{v}) + \rho(\boldsymbol{\nabla}\cdot\mathbf{v})\mathbf{v} + \rho(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v} + \nabla P - \boldsymbol{\nabla}\cdot\boldsymbol{\tau} - n\mathbf{F} = 0. \tag{E.41}$$

If we assume that we are dealing with an incompressible fluid, which means that $\lambda = 0$ and $\boldsymbol{\nabla}\cdot\mathbf{v} = 0$, then $\boldsymbol{\nabla}\cdot\boldsymbol{\tau} = \mu\nabla^2\mathbf{v}$, and the equation becomes:

$$\frac{\partial\mathbf{v}}{\partial t} = -(\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v} - \frac{1}{\rho}\boldsymbol{\nabla}P + \nu\nabla^2\mathbf{v} + \mathbf{f}, \quad \text{for } \left.\frac{\partial f}{\partial t}\right|_{\text{col}} = 0, \tag{E.42}$$

where $\nu = \mu/\rho$ and $\mathbf{F}/m$. This is the incompressible Navier-Stokes equation in the absence of momentum non-conserving collisions (same as equation E.3). The point of this section is to show that Navier Stokes is nothing but a special case of the Boltzmann equation. Specifically, it is the Boltzmann equation with the additional condition that momentum is conserved— electrons obeying semiclassical transport can be aptly described by hydrodynamics.

F

Fabrication and Synthesis Recipes

A big part of this work is device fabrication and material synthesis. Throughout my journey in graduate school, many different recipes have been concocted and optimized. Below, I will write all the recipes that have been essential to this thesis in hopes of creating a consolidated library for students' and researchers' reference.

Figure F.1: The pick up steps for assembling an hBN encapsulated graphene Gr stack. (a) shows a hBN picked up by PPC. The pick up was done at 60°C (b) shows a Gr picked up subsequently by the hBN on PPC. Note the amount of bubbles spread throughout the Gr. The pick up was done at 60°C (c) shows the hBN-Gr stack put on top of another hBN on a substrate. Note now the bubbles in the Gr region are all gone. This is done by pushing the bubbles out and is automatically done by putting down the stack slowly at $90 - 110^{\circ}$C.

F.1 Polymers

The handling of quantum materials (also called Van der Waals materials) have been done by polymers. Polymers can be transparent, soft, adhesive and almost residue-free. They can also create adiabatic, smooth and controllable movements upon heating and cooling. These are important properties that allow us to deal with quantum materials.

F.1.1 Polypropylene Carbonate and the Flake Transfer/Pick Up Recipe

Polypropylene carbonate (PPC) is the go-to polymer for picking up anything of interest. It has a glass transition temperature when between 25 to 45 $^{\circ}$C. Above 45°C, PPC begins to melt and becomes more liquid-like. The transition from glass to liquid is smooth. We do our pick-ups at 60°C in the Rayleigh transfer stage. At this temperature PPC is soft/liquid-like enough to not cause any mechanical stress on the flakes and at the same time adhesive enough to pick up the flakes. When putting down flakes or when bubbles need to be pushed out in a stack, flakes are approached onto the target area at 90°C to 110°C (see figure F.1). This method has reliably given us large area bubble free stacks.

When putting materials down by melting the PPC and delaminating the PPC completely from the transfer slide, we heat up the PPC to at least 120°C (many times up to 140°C). PPC is prepared with the following recipe:

— 15 g of PPC

— 100 ml of anisol

— dissolve PPC into anisol by mixing them in a closed bottle with a magnetic stirrer

— heat to 165°C on the hot plate and stir at rpm overnight

Note that anisol evaporates at above 154°C. Heating the hot plate to 165°C means that the anisol will be close to evaporating as the anisol does not truly reach the hot plate temperature. Some anisol will inevitably evaporate in which case the closed bottle will prevent the anisol from escaping. After the heating, the dissolved PPC is filtered with a $0.2\ \mu$m PTFE filter. When spin coating PPC onto a substrate, the substrate is oxygen (O_2) plasma cleaned to ensure that PPC can delaminate cleanly and smoothly. The spin coat rate varies from 1000 to 6000 rpm depending on application (with 2min post-bake). 2000 rpm is the most frequently used and yields roughly about a micron thick PPC. The thickness of PPC will change depending on the concentration of PPC in the anisol (which will change over time) during the time of spinning.

F.1.2 Inverted Stacking

Thin hBN ($\sim$ 10 nm or thinner) is very difficult to be picked up by PPC. To make devices with thin hBN on top, we first pick up the thick bottom hBN and pick up the thin hBN last. Thin hBN can be picked up easily by other flakes on PPC. The procedure is to then remove the entire PPC film from the PDMS transfer slide (see section below). This is done simply by peeling off the 3M scotch tape the PPC was sticking onto initially (it is wise to create a non-sticky handle in the 3M scotch tape prior).

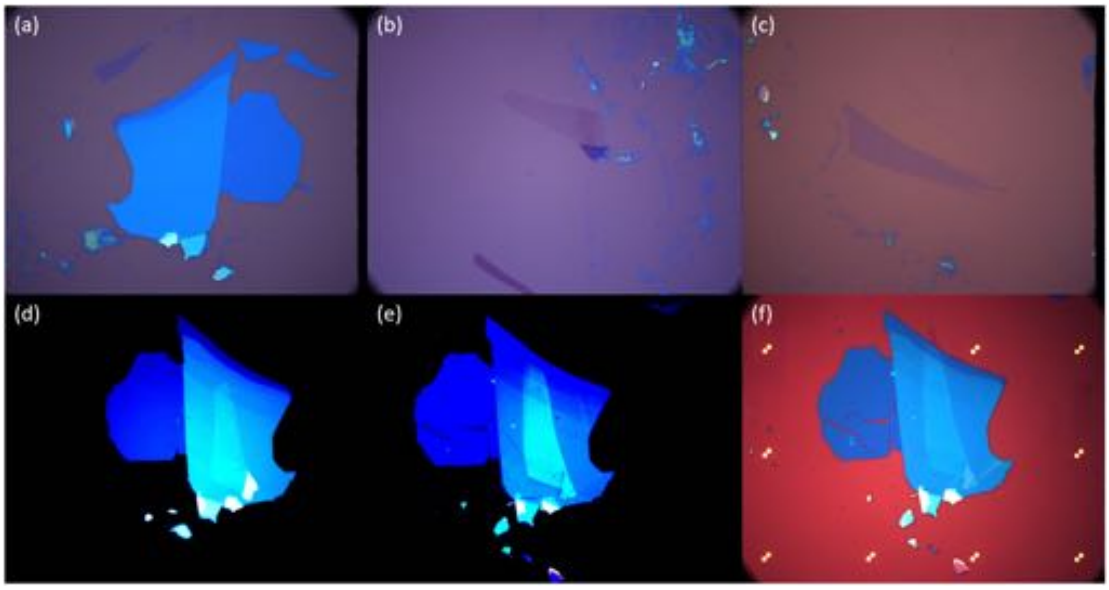


Figure F.2: The inverted stack method. (a), (b), (c) are the bottom hBN, Gr, top thin BN respectively. (d) shows the bottom hBN and Gr being picked up first. Note the bubbles in the Gr region. (e) shows the top hBN being picked up last. In this step, bubble pushing was performed, hence the disappearance of the bubbles. (f) shows the stack that is put on a substrate and annealed. Note how the stack is now inverted with respect to the flakes' original orientation.

The PPC film can be kept intact with extremely high success rate during the peeling process. The key is to make sure the PPC film is warm and not cold. Immediately after any transfer where the PPC film is typically heated up to 90°C is enough. Otherwise, if the PPC has gone cold, it can be warmed to 60°C. Once it is peeled, the PPC is transferred directly onto the target substrate. The substrate is then carefully heated up to 120°C where the PPC will then settle smoothly on its own. Figure F.3 compares the put down procedure of the regular stacking and inverted stacking methods.

F.1.3 Polydimethylsiloxane and the Transfer Slide

Polydimethylsiloxane (PDMS) is a soft and transparent polymer used as a moldable gel to support either PPC films or to hold on to chips while transferring. The recipe for it is as follows:

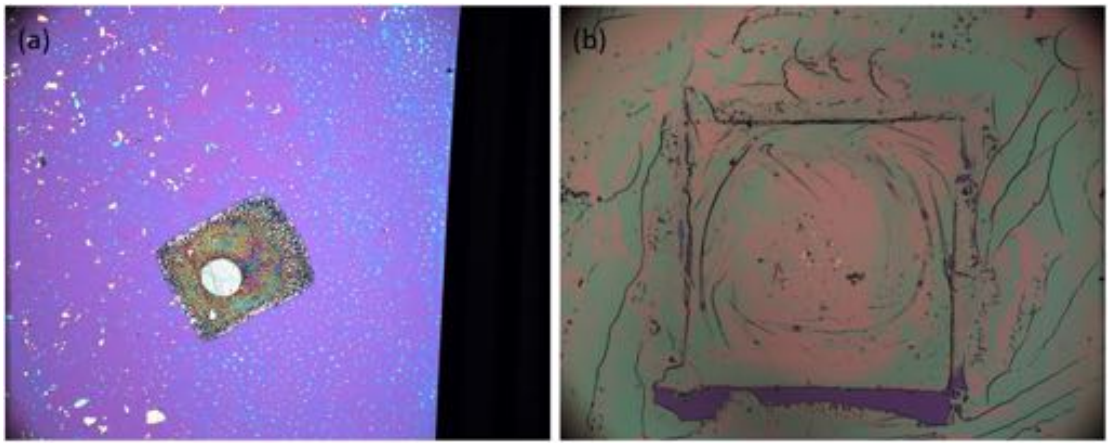


Figure F.3: Regular vs inverted stacking put down comparison. (a) shows what the put down looks like for regular stacking where the stack is put right-side up and the PPC is de-laminated onto the substrate. (b) shows what the put down looks like for the inverted stacking method. In this method, the entire PPC with the flakes is peeled off completely from the PDMS transfer slide. It is then directly transferred onto the substrate and heated to 120°C slowly until it settles naturally. Here you can see the entire PPC film on the substrate.

—1 g Sylgard 184 silicone elastomer base

—10 g Sylgard 184 silicone curing agent

—stir the above two until extremely well-mixed (else the PDMS will be non-uniform in adhesiveness)

—pour the mixed gel into a glass petri dish and place in a vacuum dessicator for 30mins to get rid of the bubbles (doing it longer will cause the gel to slowly solidify into PDMS)

—heat the gel in a glass petri dish in an oven at 160°C for two hours

The amount of PDMS mentioned above is typically used to fill three Petri dishes that are 47mm wide in diameter. The transfer slide consists of a glass slide on top of which is the PDMS and then a transparent scotch tape taped over it to anchor the PDMS down. The transparent scotch tape also serves to delaminate PPC more controllably (PPC does not delaminate from PDMS well). The PPC that is spun on a chip, a 3M scotch tape with a window is put and stuck on top of it. If the chip with PPC was originally clean (or was treated with oxygen plasma), by peeling the 3M scotch tape off, the PPC

will come off with it. There will be now a free-standing PPC film in the window of the 3M scotch tape. This PPC can be gently heated by hovering it on top of a hot plate to smooth out wrinkles. The free-standing PPC is then directly put on top of the PDMS and transparent scotch tape. This completes the transfer slide.

F.2 Electron Beam Lithography Recipes

The electron beam writer of choice is the Raith 150 e-beam lithography system with which $20 - 30$ keV acceleration voltage is used. The lower voltage is used when a larger undercut is needed. For fine features, we use $10\,\mu$m aperture size on single layer PMMA 950K A6 resist spun at 3000 rpm and heated at 180°C for 2 mins and written with a dose of $470 - 500\,\mu$C/cm^2. This is a slight overdose to avoid liftoff problems. This recipe at 30 keV allows us to achieve a resolution of having 100 nm wide and $10\,\mu$m long lines that are spaced 150 nm apart over $10 \times 10\,\mu$m^2 large area. We typically use the same recipe for etch masks. This etch mask can sustain up to 10 mins of our BN etch recipe (see contact recipe section below). Coarse features are more susceptible to liftoff problems and so we use a bilayer resist that consists of co-polymer P(MMA-MAA) EL6 as the bottom layer and typically PMMA 495K A4 for the top layer. Both layers are spun at 3000 rpm and heated at 180°C for 2 mins. Due to the higher sensitivity of the co-polymer, this gives us a significant enough undercut to avoid liftoff problems. A typical dose for coarse features is $420 - 450\,\mu$C/cm^2. Finally, for ebeam writes associated with top contacts that do not include in-situ etching such as top contacts for carbon nanotubes, a dose of $600\,\mu$C/cm^2 on the single layer PMMA 950K A6 recipe is used. Note that the parameters mentioned is just a rough guideline. Parameters used will depend on the specific situation where we factor into write size, ebeam proximity effect and write time. The developer used is made of a IPA:H_2O mixture mixed with a $3:1$ ratio and cooled to 4°C. This developer is used because it

helps prevents cracks formed in the PMMA on BN upon development. The developer IPA/MIBK often created cracks at the sharp corners of features.

F.3 Metal Evaporation and Edge Contact Recipes

The safe thickness of metal that can be deposited with the above resist recipes is about 120 nm. The measured thickness of the above single and bilayer resist recipes is about 350 to 400 nm. A safe metal deposition thickness to avoid liftoff problems is supposedly $1/3$ of the resist thickness. For all our edge contact recipes, we evaporate with a rotating and 15° tilted stage. Typically $5, 15, 60$ nm of Cr, Pd, Au is thermally evaporated sequentially at $0.2, 0.3, 1$ Å/s. Cr is used as a sticking layer on SiO_2. Cr also has one of the finest grain sizes among evaporated metals, making them ideal for contacting individual atoms. Pd is used because it is strong. On atomically flat and smooth surfaces such as hBN, Cr does not act as a sticking layer as there are no oxygen dangling bonds for it to bond to. Pd instead is used to provide a strong infrastructure so that the metal leads (such as Au which is very soft) on top of the atomically flat surface do not bend and shift around. Au is used as a capping layer to provide a soft protection surface for wirebonding. For edge contacts, we perform an etch immediately before the evaporation. To etch hBN and Gr, we use a combination of CHF_3, O_2 and Ar gases to perform a dry RIE (reactive ion etch). The etch recipe name is akc_etch_BN. The time of etch depends on how thick the stack is. Typically $2 + 1$ mins is sufficient for a thorough etch. Initially we thought high RF power and low ICP power is critical for the success of edge contact as high RF power and low ICP power increases the anisotropy (directionality) of the etch. We would want to create a 45° edge profile for maximal contact chance during evaporation. But with some experimenting, the success window for edge contact is large. The above values are in some sense quite arbitrary. It is important to note that the etching of PMMA with anything that contains Fluorine will create a very insoluble Fluoride

polymer layer. This layer cannot be removed by solvents like Acetone or Chloroform. It becomes a danger as the Fluoride polymer can settle over the device. To circumvent this, after the Fluoride etch, a brief 30 secs O_2 plasma etch (akc_GentleO2Clean) is performed to get rid of this Fluoride polymer layer. All etches are done with the STS ICP RIE system. The etch recipes are detailed in the etch recipe section.

F.4 Atomic Layer Deposition Recipes

Atomic layer deposition (ALD) can be tricky on thick atomically flat flakes. This is because ALD relies on the existence of dangling bonds for the precursors to react and attach to. If no dangling bonds exist, which is the case of atomically flat flakes, the precursors will just get pumped away during the ALD process. To address this issue, a seeding layer is deposited first. This is done by performing a cold ALD step. Cold here means below the boiling temperature of water. This is so that when water is deposited, it won't evaporate easily and be pumped away but instead, would linger on the cold surface of the flake long enough so that it can grab hold the precursors introduced in the subsequent step. The ALD processes are done with the Cambridge Nanotech/Ultratech Savannah. For the cold ALD of Al_2O_3, the stage temperature is set to 80°C, the pulse time for both H_2O and TMA (Trimethylaluminum, the aluminum precursor) are both 0.015 sec with a pump time of 20 sec after each pulse. The deposition rate is roughly 1 Å/s and 100 cycles results in approximately 10 nm which is sufficient for the seed layer. After this a hot ALD is performed to create a higher quality film. The stage is heated to 250°C and the pulse time for both H_2O and TMA are still 0.015 sec but with a subsequent pump time of 10 sec. For HfO_2 deposition, depositions are performed exclusively at 80°C with a pulse time of 0.3 sec for TEMAH (Tetrakis(dimethylamino)hafnium, the hafnium precursor) and 0.015 sec for H_2O, with a subsequent pump time of 20 sec. This results in a growth rate of about 0.9 Å/s. Cold deposition is

tested to be superior than hot deposition comparing both breakdown fields and dielectric constants. The dielectric constants for Al_2O_3 and HfO_2 with the above recipes are: $\epsilon_{\text{AlO}} \approx 9$ and $\epsilon_{\text{HfO}} \approx 15$ with a breakdown field of about $0.3 - 0.5$ V/nm for both.

F.4.1 Substrates and Mechanical Exfoliation

$500\ \mu$m thick Phosphorous doped (P-doped, with resistivity $0.001 - 0.005\ \Omega$-cm) Silicon (Si) wafers that are thermally coated with 285 nm Silicon Dioxide (SiO_2), termed Si/SiO_2 wafers and bought from Nova Electronics Materials, are typically used as substrates where Van der Waals crystals are mechanically exfoliated onto. These wafers are chosen because the degenerately doped Si can be used as a electrical backgate needed to tune the chemical potential of the sample. The 285 nm SiO_2 is chosen because it yields the highest optical contrast for mono-layer graphene (7% per layer) which allows for easy identification when scanning for isolated graphene flakes. When identifying hexagonal-Boron Nitride (h-BN) that are below few layers thick, 90 nm-thick-SiO_2 coated Si wafers are used instead because this yields the highest optical contrast for these few-layered h-BN (2% per layer, right at the borderline of our naked eye detection ability). For some of our devices, $1\ \mu$m thick SiO_2 may be used to reduce the risk of gate leaks during wiring bonding and/or to reduce capacitance with the global Si backgate. Mechanical exfoliation is performed with the 3M scotch tape.

F.5 Carbon Nanotube Growth Chip Fabrication Recipe

—Target slit size: $50 - 80\ \mu$m by 3 mm

—Growth chip dimensions: 9 mm by 9 mm

—Front platform: 350 or 380 μm long, 3200 μm wide, $50 - 55\ \mu$m height

SiN Wafer from University Wafer:

—100 mm wafer

—Type N, Phosphorus Dopant

—$1 - 10\ \Omega$·cm Resistivity

—500 μm thick Si

—100 nm LPCVD Low-Stress SiN, coated and polished both sides

Back Window:

0.) Optional: Dehydrate wafer (bake wafer at 115°C for 5 mins)

1.) HDMS pre-coat, 3000 rpm, 30 secs, 115°C 1 min bake

2.) S1822, 3000 rpm, $45 - 60$ secs, 115°C 1 min bake

3.) OL-10 write: 160 dose (recipe: S1822, not S1822 3krpm), 405 nm laser (design: BackWindowFinal, use BackWindowSmallFinal for $20 - 50$ μm wide slits)

4.) CD-26, 2 min develop with agitation, rinse in H_2O

5.) Coat wafer edge with S1822. Bake on Tex wipe (to keep hot plate clean) at 115°C for 1 min

6.) RIE-8 etch (be sure to use vacuum oil for heat dissipation!): O_2 clean (akc_DeepO2Clean for at least 5 mins), SiN etch (akc_etch_SiN for 3 mins), Gentle O_2 clean (akc_DeepO2Clean for 2 mins)

7.) Soak wafer in Acetone (for an hour, at 70°C if resist is stubborn)

8.) Si wet etch: soak in KOH solution (Sigma Aldrich, 417661, 45 wt. % in H_2O) for $17 - 19$ hrs. 500 ml, 150°C in hot plate (corresponds to about $70 - 80^{\circ}$C solution), 100 rpm stirring, windows facing up and start with KOH solution at room temperature. Rinse in H_2O after. Remember to check about every 20 mins when nearing 17 hrs. When starting when solution has reached $70 - 80^{\circ}$C (usually because of etching right after another etch), it takes about $14 - 16$ hrs. Check every 20 mins after 10 hours.

Front Platform:

9.) Thermal tape the back side of wafer (to allow spinning of the wafer without breaking it). Make sure there is minimal bubbles and no creases

10.) HDMS pre-coat, 3000 rpm, 30 secs, 115°C 1 min bake

11.) AZ4620, 3000 rpm, $45 - 60$ secs, 115°C 10 min bake

12.) OL-10 write: 200 dose (recipe: AZ4620) (design: FrontPlatformWithMarksSmallFinal_Rev if slit width is smaller than 80 μm, FrontPlatformWithMarksFinal_Rev if slit width is smaller than 150 μm), use the slit in the center bottom left and right growth chips for alignment. The middle of the slit is at $(x, y) = (-4500, -4500), (4500, -4500)$ respectively.

13.) AZ400K:H_2O at $1:4$, 10 min develop with agitation, rinse in H_2O

14.) RIE-8 etch (be sure to use vacuum oil for heat dissipation!): O_2 clean (akc_DeepO2Clean for at least 5 mins), SiN etch (akc_etch_SiN for 3 mins)

15.) RIE-10 etch (be sure to use vacuum oil for heat dissipation! This oil here is important! The wafer will crack into pieces if not enough oil is applied): dry deep Si etch (akc_Si_etch, 30 loops for 52 μm)

16.) O_2 clean in RIE-8 (akc DeepO2Clean for 2 mins)

17.) AZ-100 Remover, and or Acetone/IPA clean

18.) Piranha clean (H_2SO_4:H_2O_2 at 3:1) for >20 mins. Absolutely have to make sure resist residue is gone

19.) RIE-8 etch (be sure to use vacuum oil for heat dissipation!): O_2 clean (akc_DeepO2Clean for at least 5 mins), O_2 clean the wafer (akc_DeepO2Clean for at 2 mins), SiN etch (akc_etch_SiN for 3 mins)

20.) Deposit SiO_2 on to wafer. CVD-11 TEOS SiO_2, 30 mins (for highest SiO_2 quality, preferred) or CVD-10 Wet/Dry Oxidation SiO_2, 1 hr Wet 1100°C (lower quality)

The design files are available upon reasonable request. The etch recipes are detailed in the following

section.

F.6 Etch Recipes

Reactive ion etching (RIE) is an ion plasma etch done primarily in RIE-8 (refers to STS ICP RIE by Surface Technology Systems). Deep anisotropic Si etch for a vertical profile where a Bosch process is implemented are all done in RIE-10 (refers to SPTS Rapier DRIE by SPTS Technologies). The deep Si etch is specific to CNT growth chips. Below are the recipes for the respective recipe names:

— akc_etch_BN:

10 mTorr base pressure, 2 sccm O_2, 10 sccm CHF_3, 10 sccm Ar, 13.56 MHz Coil: 30 W (ICP power), 13.56 MHz Platen: 10 W (RF power), 20°C stage temperature

— akc_GentleO2Clean:

10 mTorr base pressure, 10 sccm O_2, 13.56 MHz Coil: 30 W (ICP power), 13.56 MHz Platen: 10 W (RF power), 20°C stage temperature, 30 secs etch time

— akc_DeepO2Clean:

35 mTorr base pressure, 45 sccm O_2, 13.56 MHz Coil: 800 W (ICP power), 13.56 MHz Platen: 50 W (RF power), 20°C stage temperature

— akc_etch_SiN:

3 mTorr base pressure, 6 sccm Ar, 2 sccm CF_4, 50 sccm CHF_3, 13.56 MHz Coil: 600 W (ICP power), 13.56 MHz Platen: 125 W (RF power), 20°C stage temperature

— akc_Si_etch:

General parameters: 15 mtorr He back pressure, 150 Hz frequency, 10°C stage temperature

step 1: 55 mtorr base pressure, 300 sccm Ar, 100 sccm O_2, 100 sccm C_4F_8, 1 sccm SF_6, 1 sec process time, Platen power: 0 W (RF power), Source power: 2500 W (coil power)

step 2: 60 mtorr base pressure, 1 sccm O_2, 350 sccm C_4F_8, 1 sccm SF_6, 1.7 sec process time, Platen power: 0 W (RF power), Source power: 2500 W (coil power)

step 3: 30 mtorr base pressure, 1 sccm O_2, 1 sccm C_4F_8, 300 sccm SF_6, 1.7 sec process time, Platen power: 210 W (RF power), Source power: 2500 W (coil power)

step 4: 60 mtorr base pressure, 1 sccm O_2, 1 sccm C_4F_8, 660 sccm SF_6, 12 sec process time, Platen power: 40 W (RF power), Source power: 3000 W (coil power)

Repeat step 1 to 4 for 30 times.

F.7 Carbon Nanotube Growth Catalyst Recipe: CoMoCat

TEOS (Tetraethyl othosilicate) + Pluronic® F-127 + $Co(Ac)_2 + (MoAc_2)_2$ A typical growth catalyst is to deposit solid nanoparticles (such as Co or Fe) onto the growth chip but this requires ebeam evaporation and patterning. Another common growth catalyst is to swipe solution (such as $FeCl_5$) onto the growth chip and reduce the catalyst to its pure metal form through H_2 (reducing agent) anneal. Our catalyst is different and is advantageous for a few reasons. Our catalyst is a Co-Mo bi-metallic catalyst and comes in the form of liquid or gel and hence can be applied directly to the growth chip without evaporating a solid precursor. Because of the existence of Mo, the Co-Mo interaction inhibits the Co sintering that typically occurs at the high temperature required for CNT growth. When Co sinters, Co combine to form clusters or large Co particles and produces less desirable form of CNT(bundles or MWCNT). Because Co sintering is limited in our catalyst, Co clusters are small (i.e. <2 nm), SWCNT are more likely formed. Our catalyst is based on the CoMoCAT technique prepared by the sol-gel method (as opposed to the original impregnation method). The sol-gel method is more advantageous because it offers a higher surface area, narrower pore size distribution, higher support thermal stability than impregnated catalysts and higher resistance to metal particle sintering.

Chemicals required:

— TEOS(Tetraethyl orthosilicate) (from Aldrich, catalog # 86578)

— Pluronic®-F127 (from Sigma, catalog # P2443)

— Cobalt(II) Acetate Tetrahydrate 98+% (from Aldrich, catalog # $40,302-4$)

— Molybdenum Acetate Dimer 98% (from Sigma, catalog # 232076)

— 0.1M HCl (made in Harvard CNS)

— Ethyl Alcohol, Pure, 200 Proof, ACS Reagent, $>$ 99.5% (from Sigma Aldrich)

Catalyst Preparation:

— Catalyst preparation: 1.04 g tetraethyl orthosilicate (TEOS), 6.3 mL ethanol, 0.5 mL 0.1M HCl was mixed by stirring and heated at 70 degree Celsius for 1 hr to make solution A.

— 4.5 g pluronic F-127, 99 mg cobalt acetate tetrahydrate and 66 mg molybdenum acetate was mixed with 50 g EtOH by sonication for 1 hr to make solution B.

— Mix 5.5 g of solution B with all of solution A to get a transparent solution C. Age solution C at room temperature for 2 weeks, which would turn into a gel. Use toothpick to apply this gel onto chips for growth of SWNTs. Solution B and C should appear pink and faint pink respectively.

— Solution B can also be used directly as a more concentrated and liquid form of catalyst.

F.8 Carbon Nanotube Growth Procedure

General parameters:

— Ar (Ultra High Purity 5.0 Grade): 200 sccm

— H_2 (Ultra High Purity 5.0 Grade): 200 sccm

— Ar through EtOH bubbler (Pure, 200 Proof, ACS Reagent, >99.5%): 100 sccm

— Growth temperature (as shown on furnace temperature controller): 910°C

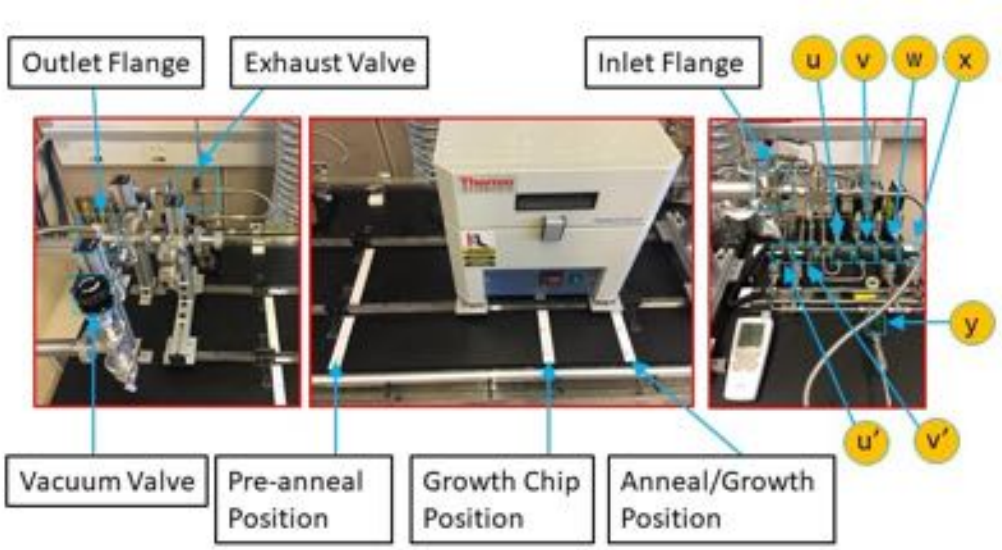


Figure F.4: The carbon nanotube growth furnace. The furnace is composed of three sections highlighted in red. Left: outlet for vacuum pumping to evacuate the tube furnace and for growth gas exhaust. Middle: moving furnace for fast heating and cooling. Right: inlet that includes four mass/gas flow meters for hydrogen and argon, two bubblers for ethanol and distilled water, and a hygrometer.

— Growth period: $5 - 10$ min

Notes:

— Ar and H_2 tanks should be changed before reaching 500 psi

— To reduce amorphous carbon buildup, bake the quartz tube in air for an hour at 1000°C after a couple uses

— Close and open valves such that pressure is not built up in closed sections

— Recommended to change quartz tube after every $6 - 8$ uses

— Recommended to always be flushing the furnace with Ar at about 80 sccm even when not in use

Fast heating and cooling method:

1.) Wear gloves, respect the furnace and keep it clean.

2.) Sonicate catalyst (Co Mo mix) for at least 10 mins (DO NOT HEAT THE CATALYST!)

3.) Apply catalyst onto growth chip using a new stick with fresh surface

4.) Check gas valves and pressure gauges in the gas cabinet and wetlab manifold. Make sure gases can be readily delivered to the furnace

5.) Valves (u), (v), (w) and (x) should be closed. If Ar was left flowing, then close valve (y) first then close valve (w) to pressurize the gas line with Ar. Alternatively, Ar can be left flowing in which case valve (y) and (w) is left opened.

6.) Remove connection at the inlet and outlet flange. Using the clean PTFE rod (rod must be non-metallic because metal particles left in the tube act as growth catalyst. Make sure rod is clean, if it's not, wipe it down with ethanol or IPA), push the quartz boat out of the tube. Use a clean petri dish (ethanol or IPA cleaned) to hold onto the quartz boat. Use clean (non-metallic) tweezers to hold onto the quartz boat.

7.) Put the growth chips onto the clean quartz boat while sitting on the clean petri dish (be sure the boat does not come in contact to the table, gloves or other foreign surfaces.

8.) Push the quartz boat back into the quartz tube using the PTFE rod (make sure PTFE is clean, if it's not, wipe it down with ethanol or IPA). Push it to the growth chip position.

9.) With the inlet and outlet flange still removed such that the tube is open, burn the growth chips in air at 400°C for 30 mins. This is to burn out the polymer/hydrocarbons.

10.) If valves (w) and (y) are not open, open them now so that Ar is flowing.

11.) Replace the inlet and outlet flanges connection to begin flowing Ar into the furnace. Check the exhaust bubbler to see if the vacuum oil is bubbling. If it is not, there is a pinch somewhere.

12.) Set the furnace temperature to zero and move furnace back to pre-anneal position. You can open the furnace to let it cool faster.

13.) Once at room temperature, first close the exhaust valve. Check if vacuum pump is on, if it's not

then switch it on. Open the vacuum valve very gently and slowly. If the vacuum valve is opened too fast, the growth chips will move out of place. Finally, close the Ar gas flow by closing valve (w). Pump down the furnace for at least 30 mins to remove residue gases. Exhaust valve must be closed first before pumping else the oil at the exhaust bubbler will be sucked into the tube furnace.

14.) After pumping for 30 mins, open valve (w) to introduce Ar gas flow first. Then close the vacuum valve. Read the pressure gauge. Wait till pressure is above 1atm=0psig. Then open exhaust valve gently and slowly. If exhaust valve is opened too fast the bubbler will bubble very aggressively.

15.) Open valves (x) to introduce H_2. Then open valve (v), (u) and then valve (v'), (u') to flow Ar through EtOH and water respectively. Doing in this order pressurizes the bubbler and prevents backflow of gas into the bubbler.

16.) The furnace should be in the pre-anneal position. Heat up the furnace to 1000°C. This will take approximately 18mins.

17.) Fast move the furnace to the growth position and then set the temperature of the furnace to 910°C. The temperature reading of the furnace will reach around 840°C.

18.) Grow for desired amount of time (recommended time is 7 mins for single walled CNT with good spacing between CNT). Timing begins the moment the furnace is moved over to the growth position.

19.) After 7 mins, turn off valve (v), (u) then valve (v'), (u'). Valves (w), (x), (y) remain open.

20.) Move furnace away to pre-anneal position and then turn the temperature of furnace to zero.

21.) When the furnace temperature reading is around 400°C (approximately 45mins), you can safely open the furnace to make it cool faster.

22.) Close valve (x) to stop H_2 flow.

23.) When the furnace is cool enough (below 100°C, takes about 30 mins more. Note that PTFE does not sustain above 250°C) you can close valve (y) then close valve (w). Alternatively, let Ar flow at 80

sccm.

24.) Open inlet and outlet flanges and push samples out with PTFE rod (make sure PTFE is clean, if not, wipe it down with ethanol or IPA). Remember to keep quartz boat clean.

25.) Return the furnace back to it default configuration, i.e. quartz boat is back inside the furnace, inlet and outlet is closed with their respective flanges with either 80 sccm Ar gas flowing or no gas flowing at all and that the area is clean. Switch off the scroll pump to prolong tip seal lifetime.

F.8.1 Furnace Cleaning

Over time because of outside contamination or carbon buildup due to growths, the quartz tube and quartz boat will become dirty. The quartz tube or boat glowing orange red when heated to high temperatures is a sign that they are dirty and needs to be cleaned. To clean the quartz tube or boat, heat the quartz tube and boat in air at 1000°C for over 6 hours. This is to burn out the carbon built up in the tube or boat due to the many growth cycles. If dirt is stubborn and does not go away after annealing (because it is not carbon and is not burnt away with oxygen), the quartz tube or boat must be replaced with a new one.

The PTFE rod will easily become dirty and so it should be wiped down with ethanol or IPA. Note that using IPA as opposed to ethanol may introduce more semiconducting tubes relative to metallic tubes as suggested by literature.

F.9 Janis

The electrical connection to and within the cryostat can be viewed as a recipe. The grounding and wiring are all done very carefully and requires much planning. The wiring of the probe uses quad-twist phospher bronze wires. The wires are all sufficiently long to allow the wires (and hence electrons)

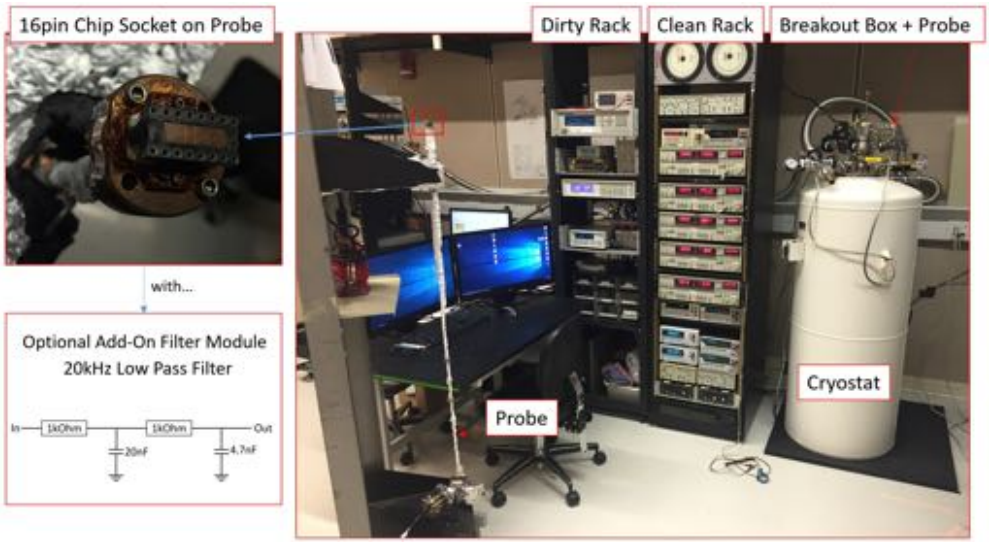


Figure F.5: The Janis setup.

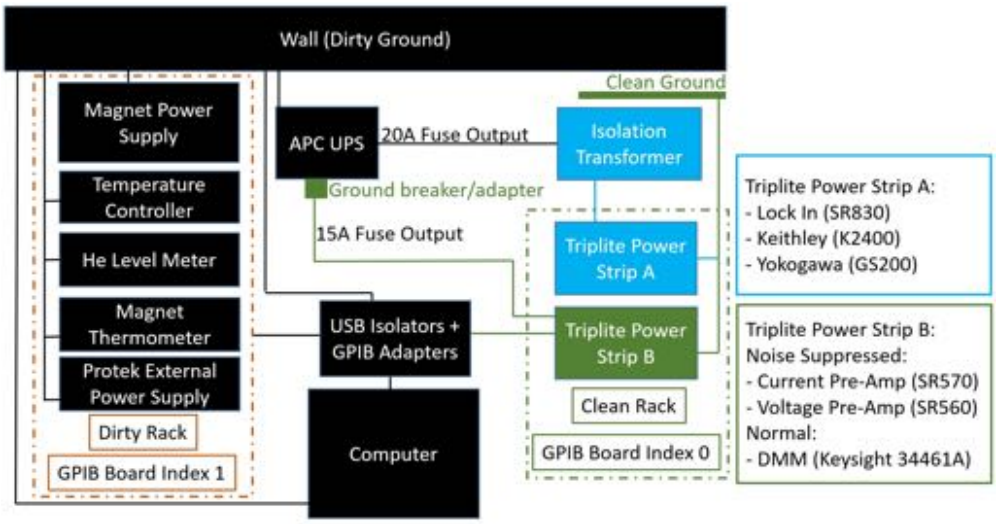


Figure F.6: Grounding of Janis.

to thermalize to base temperature (1.6 K). The wires are wound in clockwise and counter clockwise equal number of times along the probe to limit electromagnetic flux noise. GPIB cables are equipped with ferrite cores to reduce electromagnetic noise. Racks are separated physically and electrically into

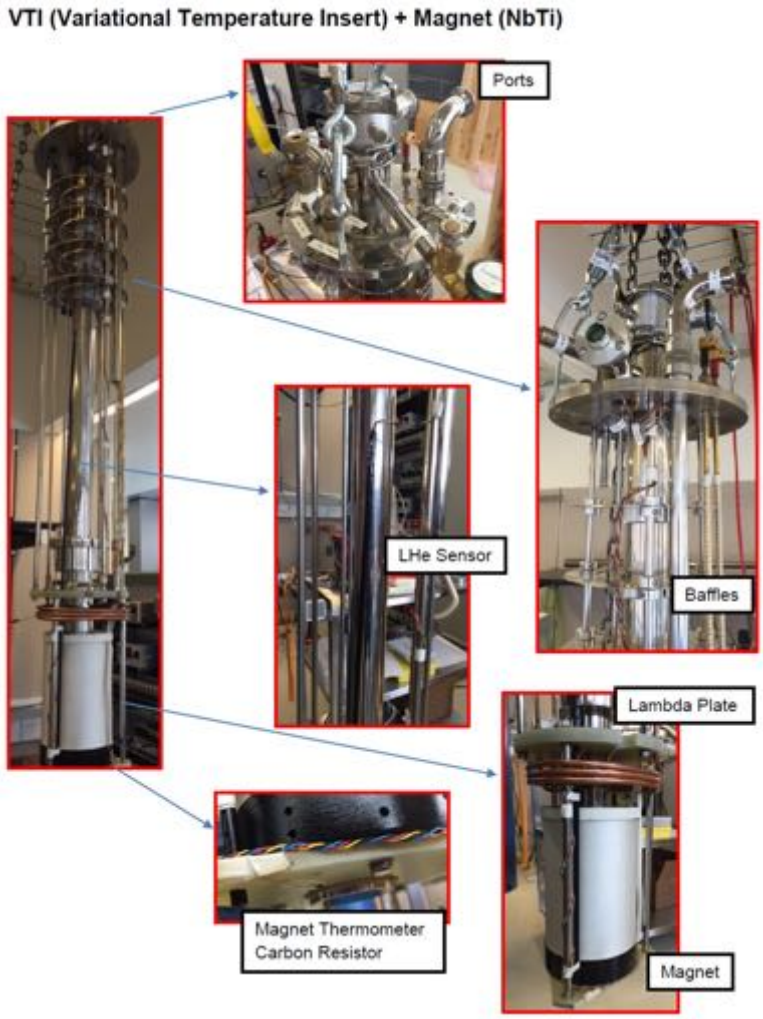


Figure F.7: Variational temperature insert (VTI) and magnet

dirty (high power instruments not directly measuring the device) and clean (sensitive instruments used to measure and excite the device) racks. The instruments in the clean rack are all decoupled from everything else by the isolation transformer and is grounded only to the clean ground. GPIB isolators are avoided as they are measured to give a lot of noise. Instead, USB isolators are used.

G

Additional Supporting Data

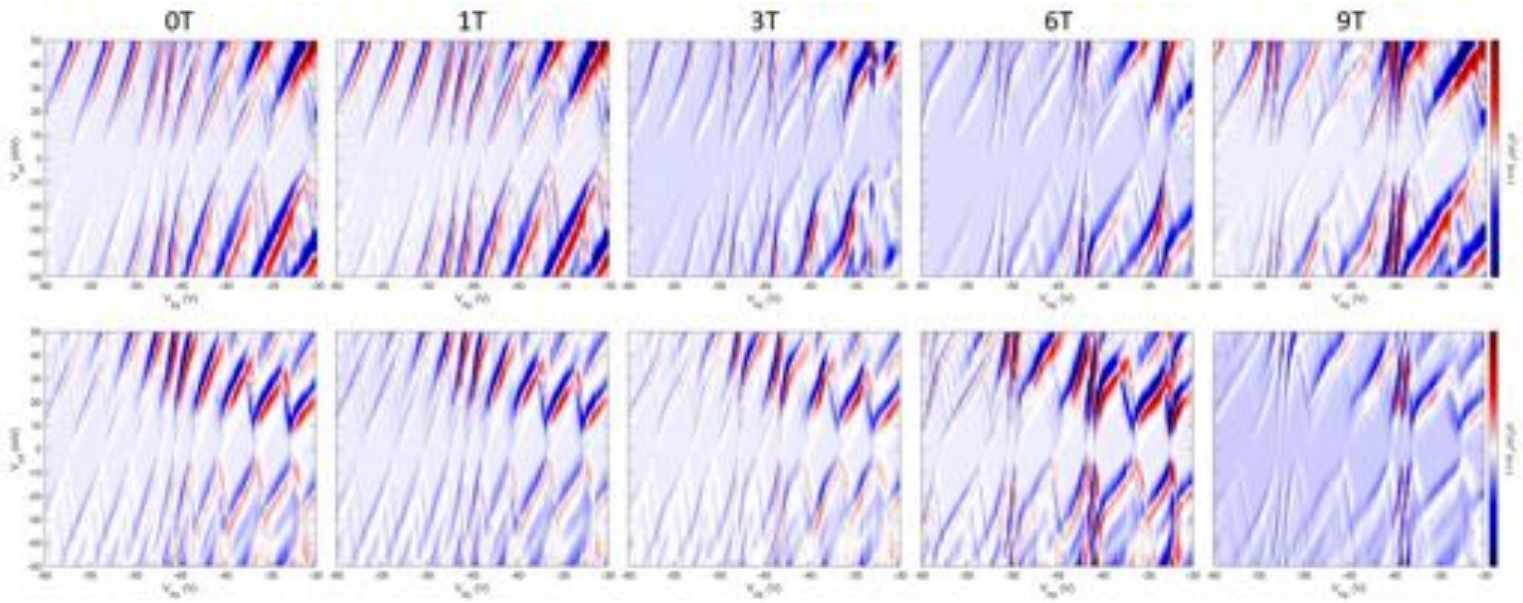


Figure G.1: Additional plots of magnetic field dependence of a carbon nanotube conductance on graphene. Plotted is the derivative of the differential conductance of the carbon nanotube as a function of the global Si backgate (V_{bg}) and source drain bias (V_{sd}) at fixed magnetic fields. See figure 5.3 for device reference. The two rows of measurements belong to different graphene potentials. The first row is taken with $V_{\mathrm{Gr}} = 1\mathrm{V}$ and the second row is taken with $V_{\mathrm{Gr}} = 2\mathrm{V}$. Regardless of the graphene potential, the charging energy is not affected by the magnetic field and is $\sim 20\mathrm{meV}$. The variation of the diamond widths along the backgate axis is due to the screening from Landau levels in Gr. The diamonds at 1T is completely equivalent to the diamonds at 0T. See chapter 5 for more details on Landau level screening as well as the insensitivity to small magnetic fields.

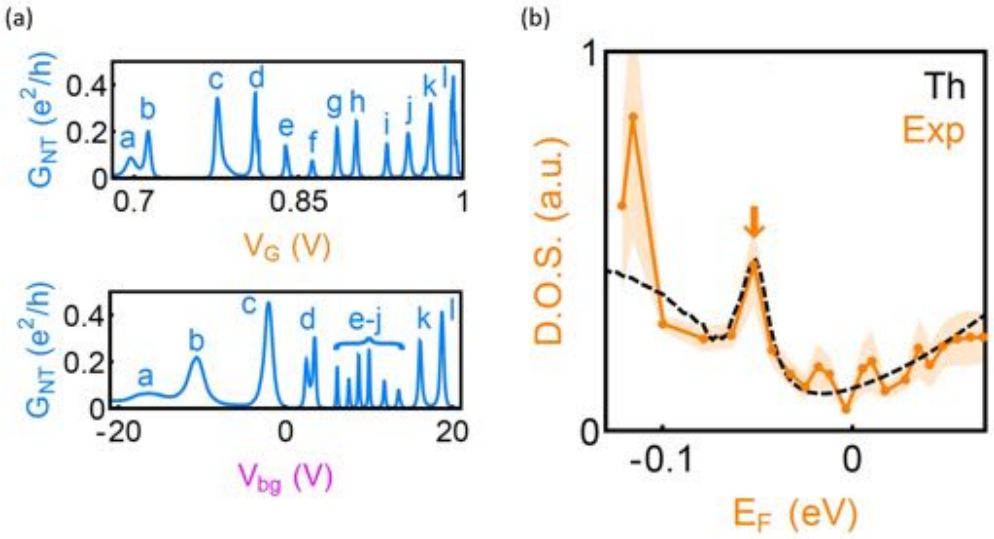


Figure G.2: Bound state in 6 nm device A

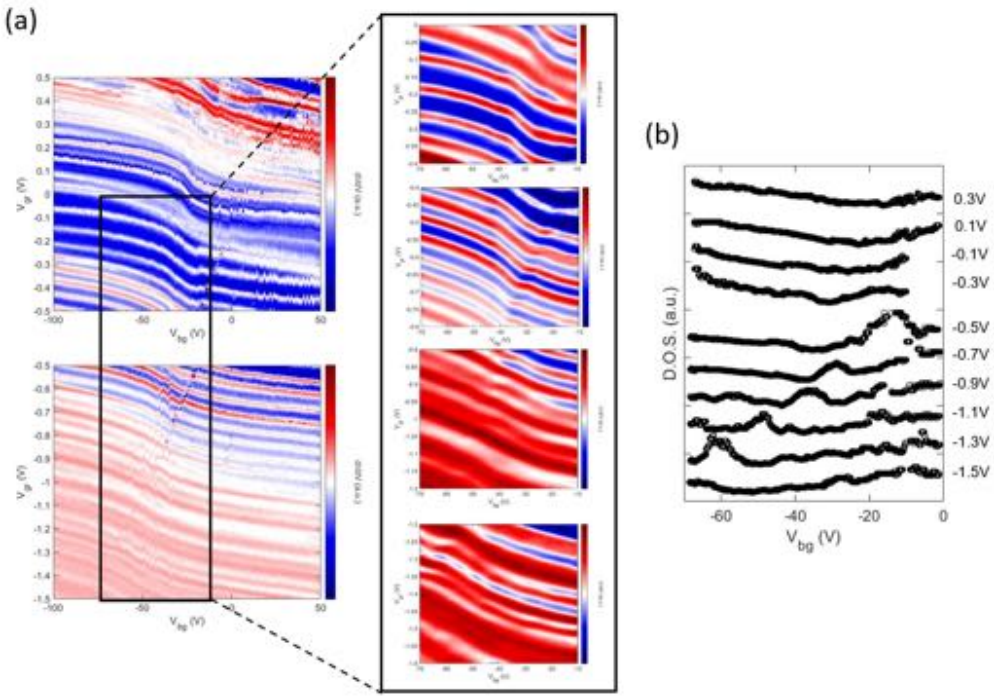


Figure G.3: Bound state in 6 nm device B

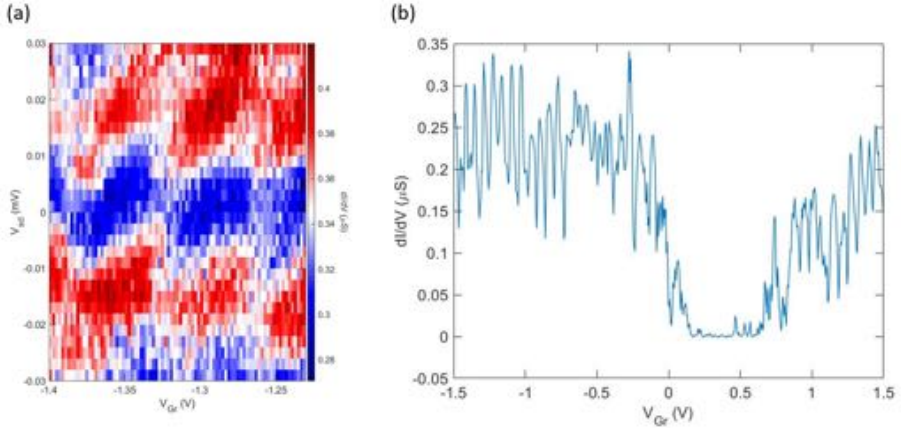


Figure G.4: Coulomb diamonds of 4 nm device.

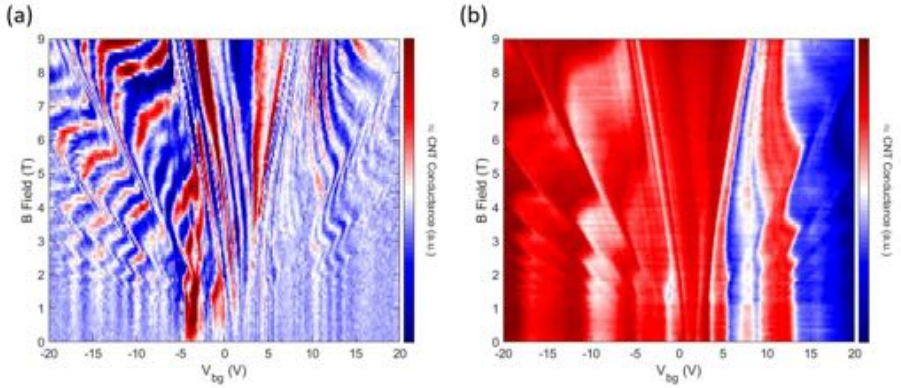


Figure G.5: Additional fan diagrams.

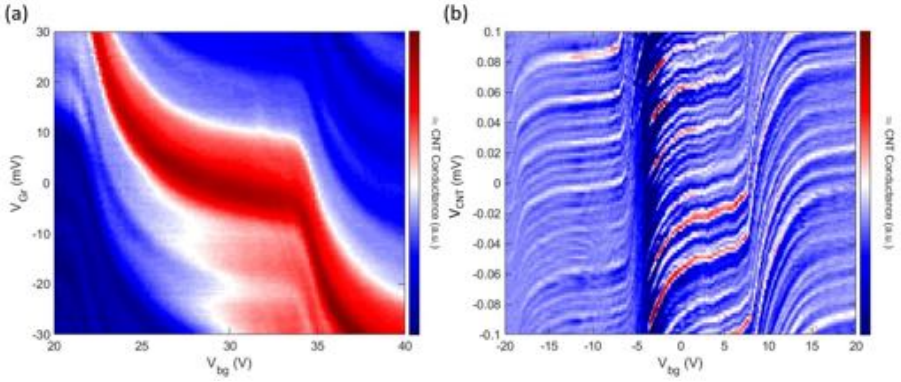


Figure G.6: Kink zoom of compressibility measurement in quantum Hall regime. Note that on the right the CNT potential is swept and not the graphene.

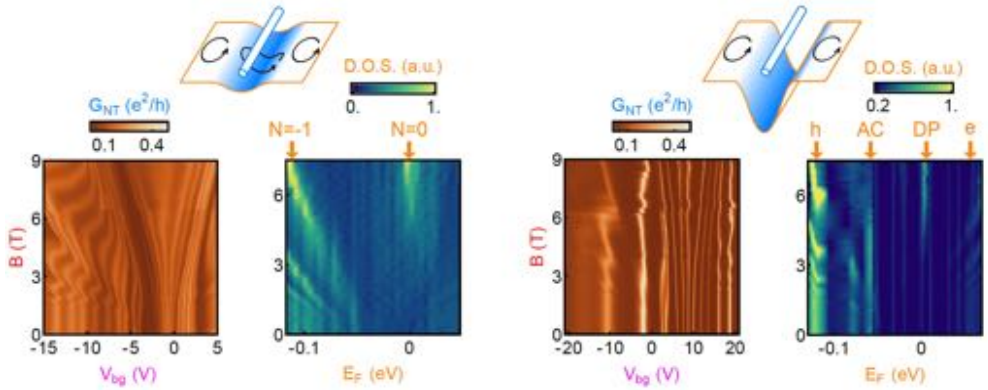


Figure G.7: Additional fan diagrams showing Landau level collapse in device A.

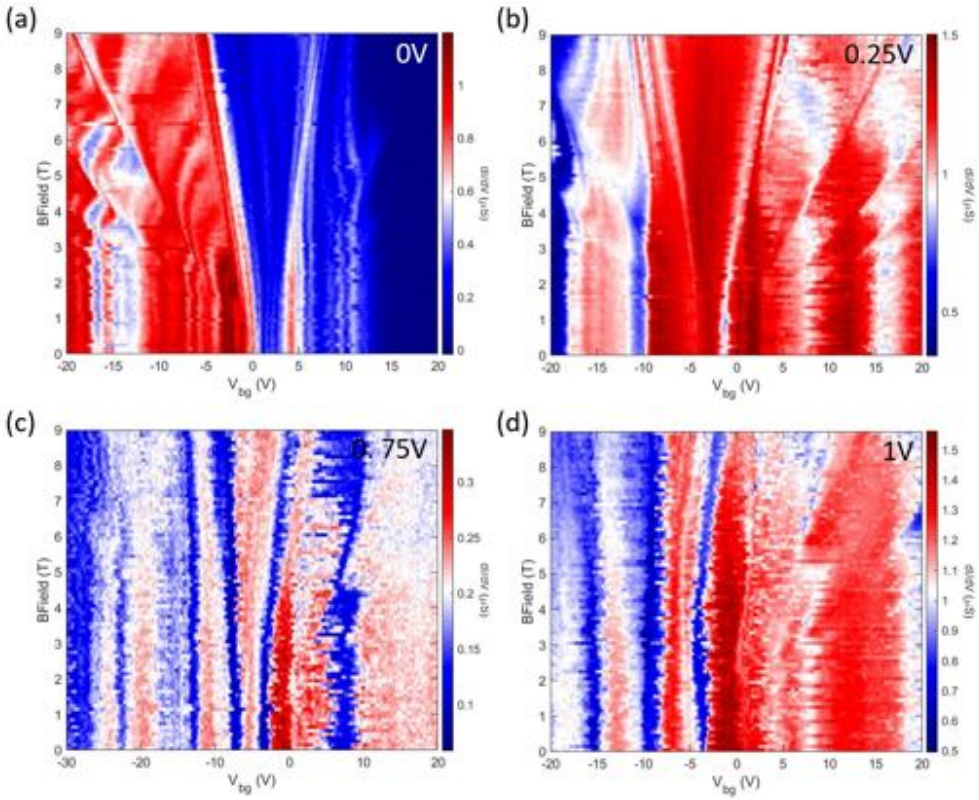


Figure G.8: Additional fan diagrams showing Landau level collapse in device B.

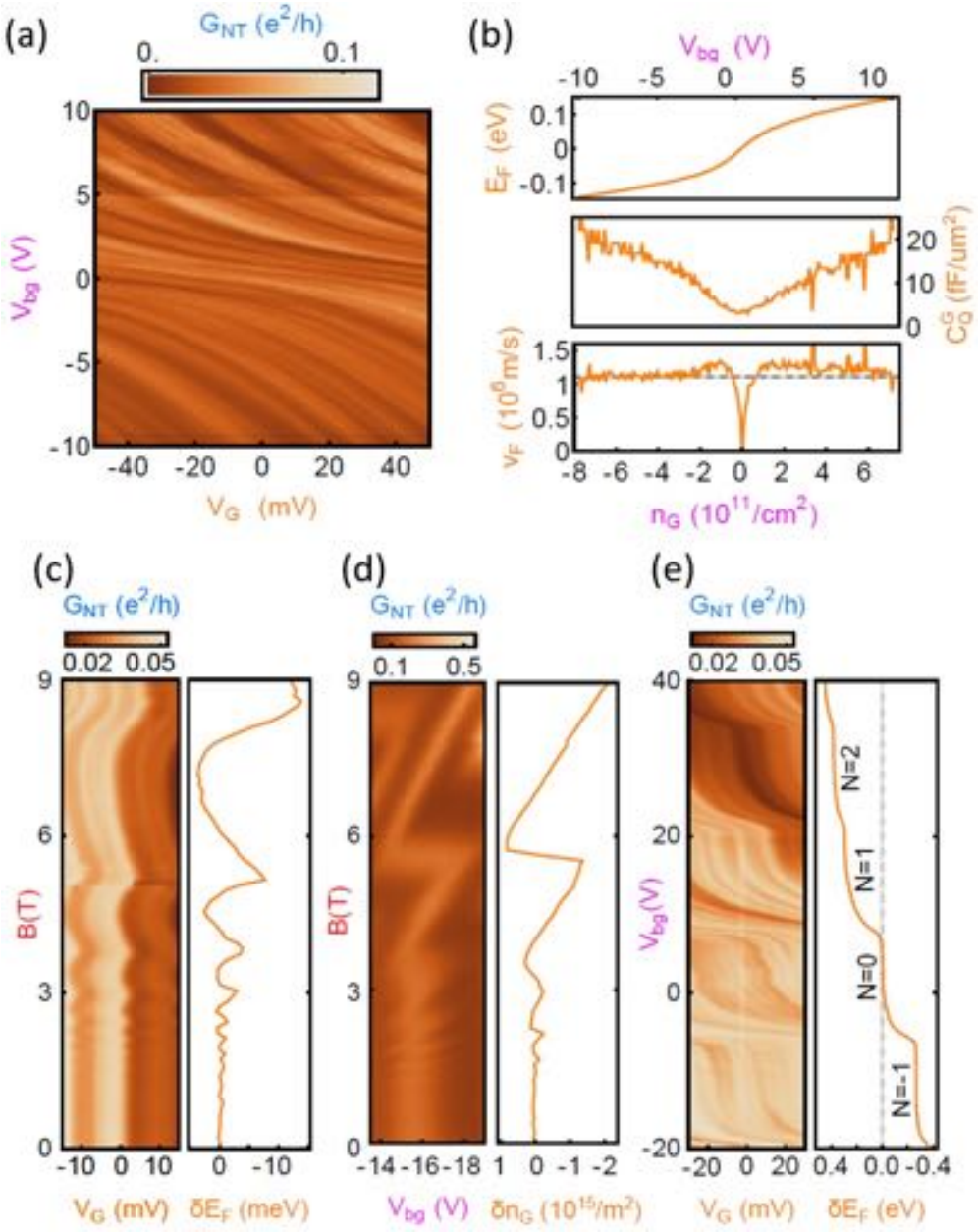


Figure G.9: Measurements of carbon nanotube conductance as functions of different parameters and the associated extracted quantities regarding graphene.

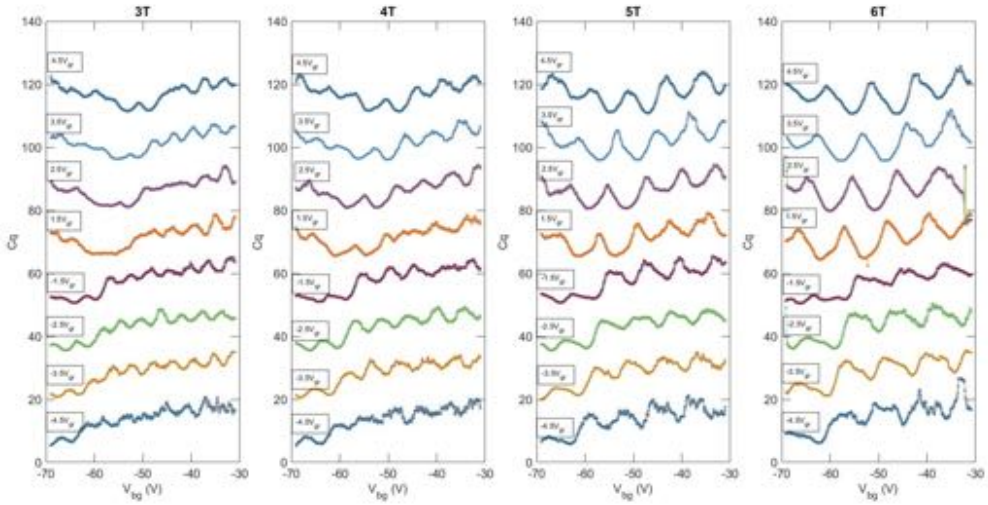


Figure G.10: Quantum capacitance line cuts at different magnetic fields showing Landau level collapse.

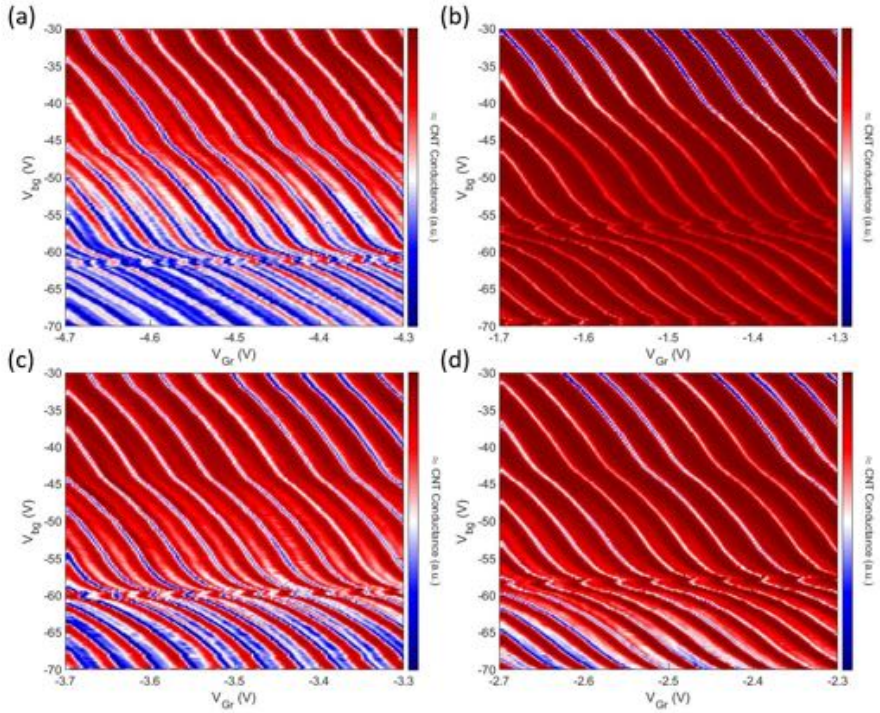


Figure G.11: Negative compressibility at 9 T implied by the sharp kink with positive slope. Negative compressibility is known to be a consequence of many-body effects.

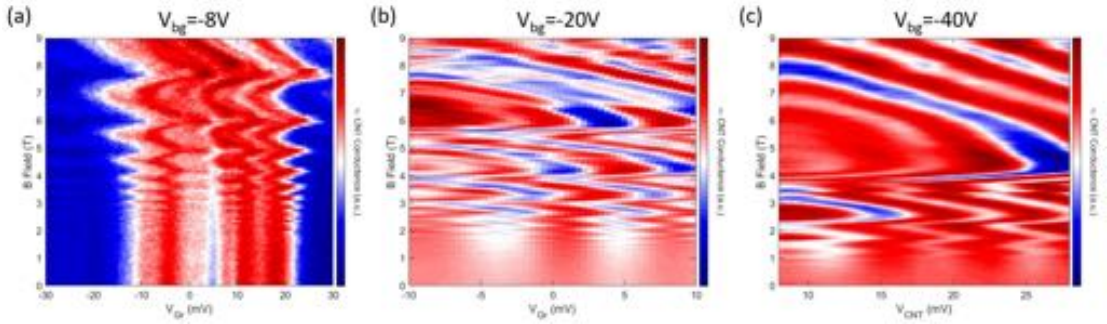


Figure G.12: Carbon nanotube conductance as a function of B field and graphene potential. The oscillations correspond to the Landau levels in graphene. Graphene density is fixed for each sweep.

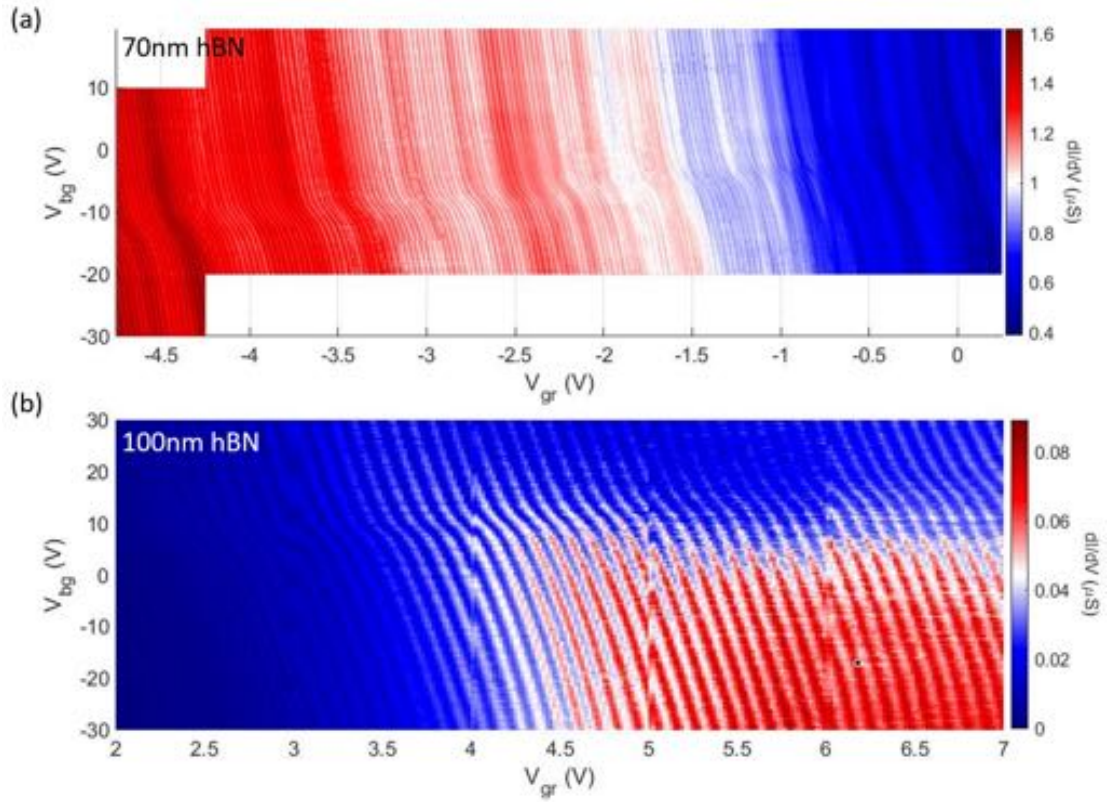


Figure G.13: Local compressibility measurements of (a) 70 nm and (b) 100 nm thick BN devices. No signs of resonances.

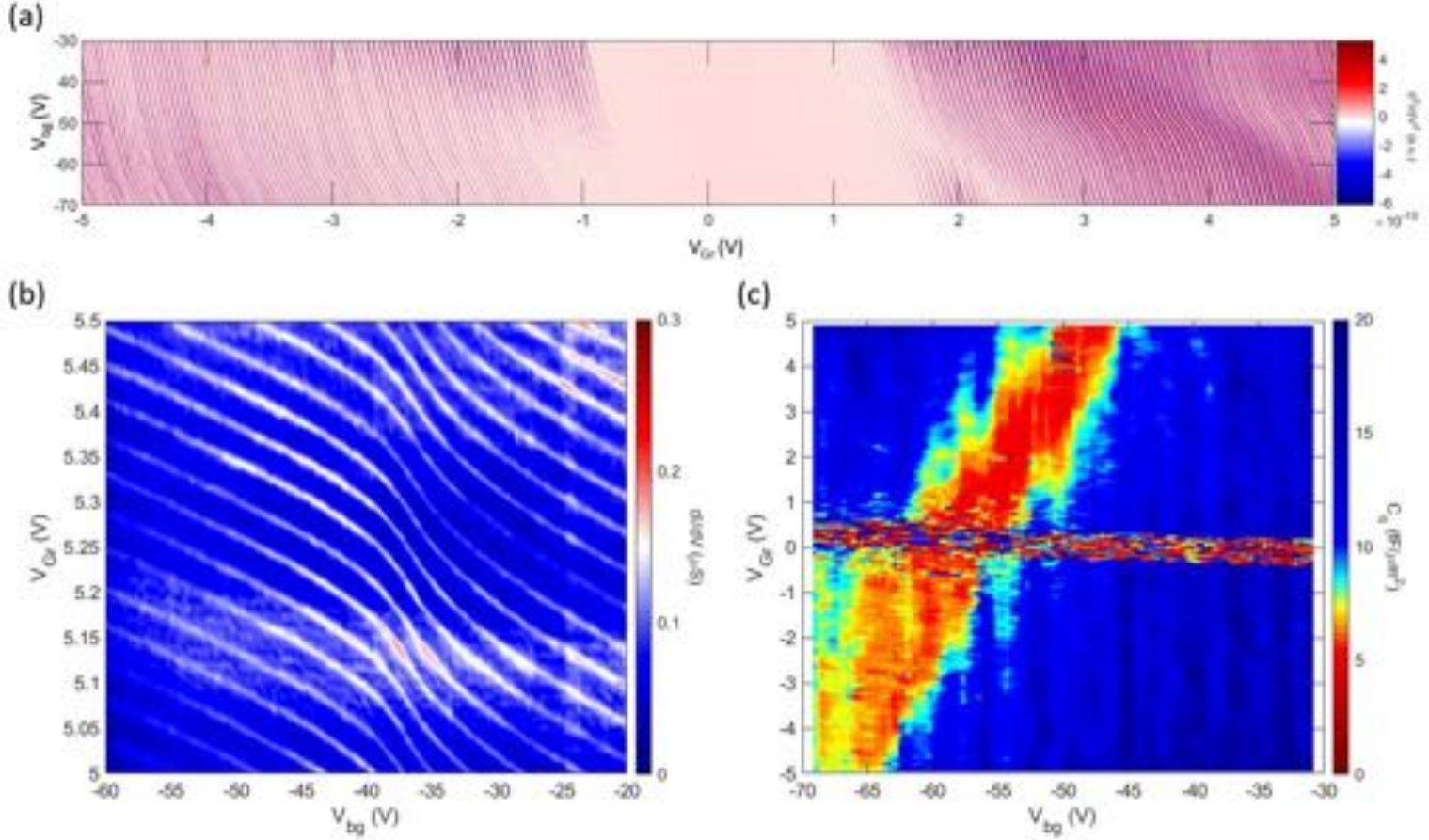


Figure G.14: Local compressibility measurement and quantum capacitance extraction of the 10 nm hBN device. No signs of resonance can be seen in this 10 nm hBN device.

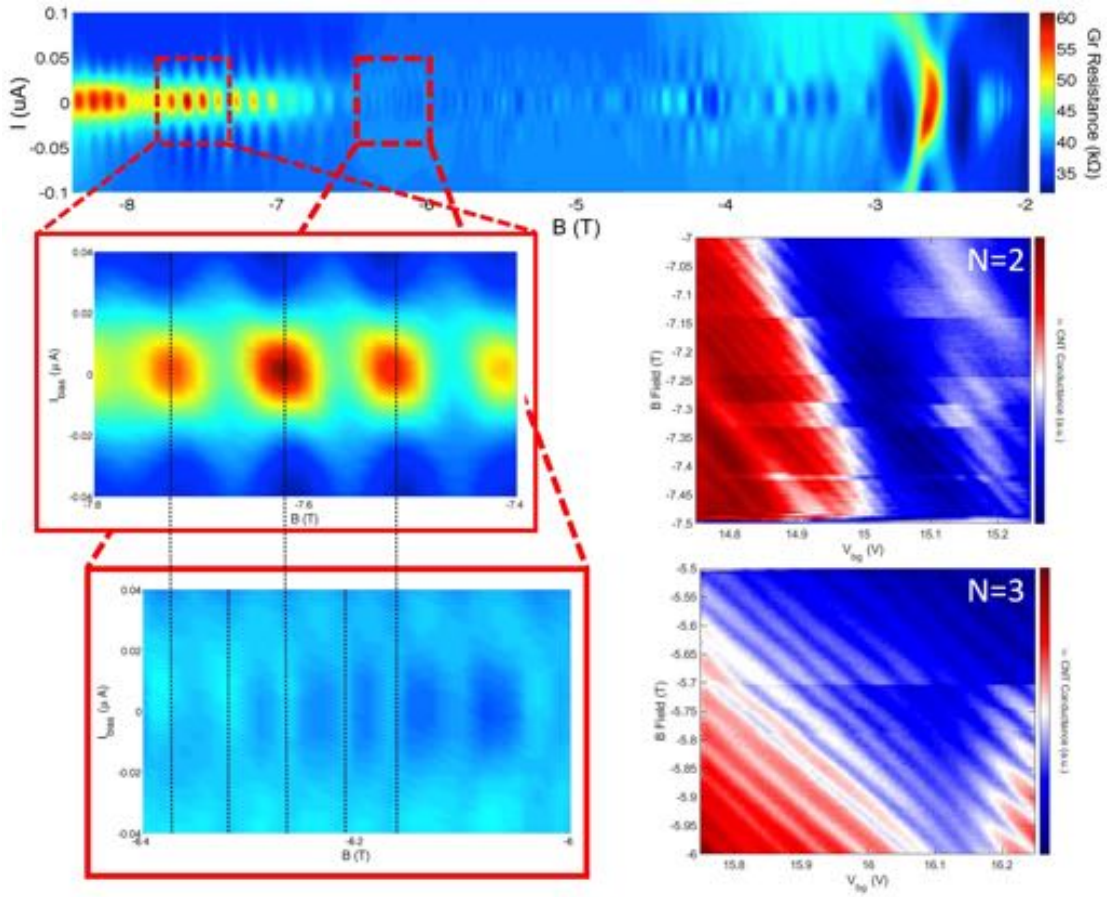


Figure G.15: Aharonov Bohm effect in graphene caused by carbon nanotube gating and observation of period doubling depending on Landau level index.

References

[1] David Larousserie. Graphene - the new wonder material. *The Guardian*, November 2013. URL https://www.theguardian.com/science/2013/nov/26/graphene-molecule-potential-wonder-material.

[2] 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, October 2004. doi: 10.1126/science.1102896. URL http://science.sciencemag.org/content/306/5696/666.

[3] Elton J. G. Santos and Efthimios Kaxiras. Electric-field dependence of the effective dielectric constant in graphene. *Nano Letters*, 13(3):898–902, Mar 2013. ISSN 1530-6984. doi: 10.1021/nl303611v. URL https://doi.org/10.1021/nl303611v.

[4] Joaquin. F. Rodriquez-Nieva. *Novel Electronic Behaviors in Graphene Nanostructures*. PhD thesis, Massachusetts Institute of Technology, 2016. URL http://hdl.handle.net/1721.1/107041.

[5] D A Bandurin, I Torre, R Krishna Kumar, M Ben Shalom, A Tomadin, A Principi, G H Auton, E Khestanova, K S Novoselov, I V Grigorieva, L A Ponomarenko, A K Geim, and M Polini. Negative local resistance caused by viscous electron backflow in graphene. *Science*, 351:1055–1058, March 2016. doi: 10.1126/science.aad0201. URL http://science.sciencemag.org/content/351/6277/1055.

[6] J. Crossno, J. K. Shi, K. Wang, X. Liu, A. Harzheim, A. Lucas, S. Sachdev, P. Kim, T. Taniguchi, K. Watanabe, T. A. Ohki, and K. C. Fong. Observation of the dirac fluid and the breakdown of the wiedemann-franz law in graphene. *Science*, 351:1058–1061, March 2016. doi: 10.1126/science.aad0343. URL http://science.sciencemag.org/content/351/6277/1058.

[7] P. J. W. Moll, P. Kushwaha, N. Nandi, B. Schmidt, and A. P. Mackenzie. Evidence for hydrodynamic electron flow in pdcoo2. *Science*, 351:1061–1064, March 2016. doi: 10.1126/science.aac8385. URL http://science.sciencemag.org/content/351/6277/1061.

[8] L.W. Molenkamp and M.J.M. de Jong. Observation of knudsen and gurzhi transport regimes in a two-dimensional wire. *Solid-State Electronics*, 37:551–553, April 1994. doi: 10.1016/0038-1101(94)90244-5. URL http://www.sciencedirect.com/science/article/pii/0038110194902445.

[9] Markus Müller, Jörg Schmalian, and Lars Fritz. Graphene: A nearly perfect fluid. *Phys. Rev. Lett.*, 103:025301, Jul 2009. doi: 10.1103/PhysRevLett.103.025301. URL https://link.aps.org/doi/10.1103/PhysRevLett.103.025301.

[10] Sumio Iijima. Helical microtubules of graphitic carbon. *Nature*, 354:56–58, November 1991. doi: 10.1038/354056a0. URL http://www.nature.com/nature/journal/v354/n6348/abs/354056a0.html.

[11] Mikhail I. Katsnelson. Massless dirac fermions in graphene. In *Graphene: Carbon in Two Dimensions*, chapter 1, page 11. Cambridge University Press, 2012. ISBN 978-0-521-19540-9.

[12] Marc Bockrath, David H. Cobden, Jia Lu, Andrew G. Rinzler, Richard E. Smalley, Leon Balents, and Paul L. McEuen. Luttinger-liquid behaviour in carbon nanotubes. *Nature*, 397:598 EP –, Feb 1999. URL https://doi.org/10.1038/17569.

[13] Zhiwen Shi, Xiaoping Hong, Hans A. Bechtel, Bo Zeng, Michael C. Martin, Kenji Watanabe, Takashi Taniguchi, Yuen-Ron Shen, and Feng Wang. Observation of a luttinger-liquid plasmon in metallic single-walled carbon nanotubes. *Nature Photonics*, 9:515 EP –, Jul 2015. URL https://doi.org/10.1038/nphoton.2015.123.

[14] Vikram V. Deshpande and Marc Bockrath. The one-dimensional wigner crystal in carbon nanotubes. *Nature Physics*, 4:314 EP –, Mar 2008. URL https://doi.org/10.1038/nphys895. Article.

[15] Edward A. Laird, Ferdinand Kuemmeth, Gary A. Steele, Kasper Grove-Rasmussen, Jesper Nygård, Karsten Flensberg, and Leo P. Kouwenhoven. Quantum transport in carbon nanotubes. *Reviews of Modern Physics*, 87:703–764, July 2015. doi: 10.1103/RevModPhys.87.703. URL https://link.aps.org/doi/10.1103/RevModPhys.87.703.

[16] G. A. Steele, F. Pei, E. A. Laird, J. M. Jol, H. B. Meerwaldt, and L. P. Kouwenhoven. Large spin-orbit coupling in carbon nanotubes. *Nature Communications*, 4:1573 EP –, Mar 2013. URL https://doi.org/10.1038/ncomms2584. Article.

[17] Yongyou Zhang, Qingyun Zhang, and Udo Schwingenschlögl. Spin-charge separation in finite length metallic carbon nanotubes. *Nano Letters*, 17(11):6747–6751, Nov 2017. ISSN 1530-6984. doi: 10.1021/acs.nanolett.7b02880. URL https://doi.org/10.1021/acs.nanolett.7b02880.

[18] N. Tombros, S. J. van der Molen, and B. J. van Wees. Separating spin and charge transport in single-wall carbon nanotubes. *Phys. Rev. B*, 73:233403, Jun 2006. doi: 10.1103/PhysRevB.73.233403. URL https://link.aps.org/doi/10.1103/PhysRevB.73.233403.

[19] Jie Jiang, Jinming Dong, and D. Y. Xing. Zeeman effect on the electronic spectral properties of carbon nanotubes in an axial magnetic field. *Phys. Rev. B*, 62:13209–13215, Nov 2000. doi: 10.1103/PhysRevB.62.13209. URL https://link.aps.org/doi/10.1103/PhysRevB.62.13209.

[20] David H. Cobden, Marc Bockrath, Paul L. McEuen, Andrew G. Rinzler, and Richard E. Smalley. Spin splitting and even-odd effects in carbon nanotubes. *Phys. Rev. Lett.*, 81:681–684, Jul 1998. doi: 10.1103/PhysRevLett.81.681. URL https://link.aps.org/doi/10.1103/PhysRevLett.81.681.

[21] Neil W. Ashcroft and N. David. Mermin. Thomas-fermi theory of screening. In *Solid State Physics*, chapter 17, page 342. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[22] S. Das Sarma, Shaffique Adam, E. H. Hwang, and Enrico Rossi. Electronic transport in two-dimensional graphene. *Rev. Mod. Phys.*, 83:407–470, May 2011. doi: 10.1103/RevModPhys.83.407. URL https://link.aps.org/doi/10.1103/RevModPhys.83.407.

[23] Robert F. Pierret. Holes. In *Semiconductor Device Fundamentals*, chapter 2, page 50. Addison-Wesley Publishing, 1996. ISBN 0-201-54393-1.

[24] T. Fang, A. Konar, H. Xing, and D. Jena. Carrier statistics and quantum capacitance of graphene sheets and ribbons. *Applied Physics Letters*, 91:092109, August 2007. doi: 10.1063/1.2776887. URL http://aip.scitation.org/doi/10.1063/1.2776887.

[25] 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, November 2008. doi: 10.1038/nphys781. URL http://dx.doi.org/10.1038/nphys781.

[26] Jiamin Xue, Javier Sanchez-Yamagishi, Danny Bulmash, Philippe Jacquod, Aparna Deshpande, K. Watanabe, T. Taniguchi, Pablo Jarillo-Herrero, and Brian J. LeRoy. Scanning tunnelling microscopy and spectroscopy of ultra-flat graphene on hexagonal boron nitride. *Nature Physics*, 10:282–285, February 2011. doi: 10.1038/nmat2968. URL http://dx.doi.org/10.1038/nmat2968.

[27] Jilin Xia, Fang Chen, Jinghong Li, and Nongjian Tao. Measurement of the quantum capacitance of graphene. *Nature Nanotechnology*, 4:505–509, August 2009. doi: 10.1038/nnano.2009.177. URL http://www.nature.com/articles/nnano.2009.177.

[28] Jiale Liang, Deji Akinwande, and H.-S. Philip Wong. Carrier density and quantum capacitance for semiconducting carbon nanotubes. *Journal of Applied Physics*, 104:064515, September 2008. doi: 10.1063/1.2986216. URL http://aip.scitation.org/doi/10.1063/1.2986216.

[29] S. J. Tans, M. H. Devoret, H. Dai, Andreas Thess, R. E. Smalley, L. J. Geerlings, and C. Dekker. Individual single-wall carbon nanotubes as quantum wires. *Nature*, 386:474–477, April 1997. doi: 10.1038/386474a0. URL https://www.nature.com/articles/386474a0.

[30] W. Liang, M. Bockrath, and H. Park. Shell filling and exchange coupling in metallic single-walled carbon nanotubes. *Physical Review Letters*, 88:126801, March 2002. doi: 10.1103/PhysRevLett.88.126801. URL https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.88.126801.

[31] M. Bockrath, D. H. Cobden, P. L. McEuen, N. G. Chopra, A. Zettl, A. Thess, and R. E. Smalley. Single-electron transport in ropes of carbon nanotubes. *Science*, 275:1922–1925, March 1997. doi: 10.1126/science.275.5308.1922. URL http://science.sciencemag.org/content/275/5308/1922.

[32] Seyoung Kim, Insun Jo, D. C. Dillen, D. A. Ferrer, B. Fallahazad, Z. Yao, S. K. Banerjee, and E. Tutuc. Direct measurement of the fermi energy in graphene using a double-layer heterostructure. *Phys. Rev. Lett.*, 108:116404, Mar 2012. doi: 10.1103/PhysRevLett.108.116404. URL https://link.aps.org/doi/10.1103/PhysRevLett.108.116404.

[33] Wei-Hung Chiang and R. Mohan Sankaran. Linking catalyst composition to chirality distributions of as-grown single-walled carbon nanotubes by tuning nixfe1−x nanoparticles. *Nature Materials*, 8:882–886, September 2009. doi: 10.1038/nmat2531. URL https://www.nature.com/articles/nmat2531.

[34] Rahul Rao, David Liptak, Tonya Cherukuri, Boris I. Yakobson, and Benji Maruyama. In situ evidence for chirality-dependent growth rates of individual carbon nanotubes. *Nature Materials*, 11:213–216, January 2012. doi: 10.1038/nmat3231. URL https://www.nature.com/articles/nmat3231.

[35] Sayangdev Naha and Ishwar K. Puri. A model for catalytic growth of carbon nanotubes. *The Journal of Physics D: Applied Physics*, 41:065304, February 2008. doi: 10.1088/0022-3727/41/6/065304. URL http://stacks.iop.org/0022-3727/41/i=6/a=065304.

[36] Mukul Kumar. Carbon nanotube synthesis and growth mechanism. In *Carbon Nanotubes - Synthesis, Characterization, Applications*, chapter 8, pages 147–170. InTechOpen Limited, 2011. ISBN 978-953-307-497-9.

[37] Limin Huang, Xiaodong Cui, Brian White, and Stephen P. O'Brien. Long and oriented single-walled carbon nanotubes grown by ethanol chemical vapor deposition. *The Journal of Physical Chemistry B*, 108:16451–16456, September 2004. doi: 10.1021/jp0474125. URL https://cdn-pubs.acs.org/doi/10.1021/jp0474125.

[38] B. Kitiyanan, W. E. Alvarez, J. H. Harwell, and D. E. Resasco. Controlled production of single-wall carbon nanotubes by catalytic decomposition of co on bimetallic co–mo catalysts. *Chemical Physical Letters*, 317:497–503, February 2000. doi: 10.1016/S0009-2614(99)01379-2. URL https://www.sciencedirect.com/science/article/pii/S0009261499013792.

[39] W. Z. Li, J. G. Wen, Y. Tu, and Z. F. Ren. Effect of gas pressure on the growth and structure of carbon nanotubes by chemical vapor deposition. *Applied Physics A*, 73:259–264, May 2001. doi: 10.1007/s003390100916. URL https://link.springer.com/article/10.1007/s003390100916.

[40] Banghua Peng, Yagang Yao, and Jin Zhang. Effect of the reynolds and richardson numbers on the growth of well-aligned ultralong single-walled carbon nanotubes. *The Journal of Physical Chemistry C*, 114:12960–12965, July 2010. doi: 10.1021/jp103731p. URL https://pubs.acs.org/doi/abs/10.1021/jp103731p.

[41] Y Zhang, J Gregoire, and A van Dover R Hart. Ethanol-promoted high-yield growth of few-walled carbon nanotubes. *The Journal of Physical Chemistry C*, 114:6389––6395, March 2010. doi: 10.1021/jp100358j. URL https://pubs.acs.org/doi/abs/10.1021/jp100358j.

[42] Shaoming Huang, Mike Woodson, Richard Smalley, and Jie Liu. Growth mechanism of oriented long single walled carbon nanotubes using "fast-heating" chemical vapor deposition process. *Nano Letters*, 4:1025–1028, May 2004. doi: 10.1021/nl049691d. URL https://cdn-pubs.acs.org/doi/10.1021/nl049691d.

[43] Kenji Hata, Don N. Futaba, Kohei Mizuno, Tatsunori Namai, Motoo Yumura, and Sumio Iijima. Water-assisted highly efficient synthesis of impurity-free single-walled carbon nanotubes. *Science*, 306:1362–1364, November 2004. doi: 10.1126/science.1104962. URL http://science.sciencemag.org/content/306/5700/1362.

[44] Don N. Futaba, Kenji Hata, Takeo Yamada, Kohei Mizuno, Motoo Yumura, and Sumio Iijima. Kinetics of water-assisted single-walled carbon nanotube synthesis revealed by a time-evolution analysis. *Physical Review Letters*, 95:056104, July 2005. doi: 10.1103/PhysRevLett.95.056104. URL https://link.aps.org/doi/10.1103/PhysRevLett.95.056104.

[45] Y Choi and S Lim. Selective oxidation of amorphous carbon layers without damaging embedded single wall carbon nanotube bundles. *Japanese Journal of Applied Physics*, 52:115101, October 2013. doi: 10.7567/JJAP.52.115101. URL http://iopscience.iop.org/article/10.7567/JJAP.52.115101.

[46] Guangyu Zhang, David Mann, Li Zhang, Ali Javey, Yiming Li, Erhan Yenilmez, Qian Wang, James P. McVittie, Yoshio Nishi, James Gibbons, and Hongjie Dai. Ultra-high-yield growth of vertical single-walled carbon nanotubes: Hidden roles of hydrogen and oxygen. *Proceedings of the National Academy of Sciences*, 102:16141–16145, May 2005. doi: 10.1073/pnas.0507064102. URL https://www.pnas.org/content/102/45/16141.

[47] S.B. Sinnott, R. Andrews, D. Qian, A.M. Rao, Z. Mao, E.C. Dickey, and F. Derbyshire. Model of carbon nanotube growth through chemical vapor deposition. *Chemical Physical Letters*, 315: 25–30, July 1999. doi: 10.1016/S0009-2614(99)01216-6. URL http://www.sciencedirect.com/science/article/pii/S0009261499012166.

[48] M.S. Dresselhaus, G. Dresselhaus, R. Saitoc, and A. Joriod. Raman spectroscopy of carbon nanotubes. *Physics Reports*, 409:47–99, March 2005. doi: 10.1016/j.physrep.2004.10.006. URL https://www.sciencedirect.com/science/article/pii/S0370157304004570.

[49] Zheng Yi Zhang. *Optical and Electrical Properties of Single-walled Carbon Nanotubes with Known Chiralities*. PhD thesis, Columbia University, 2014. URL https://doi.org/10.7916/D8F769PF.

[50] Matthew J. Baker, Caryn S. Hughes, and Katherine A. Hollywood. *Raman Spectroscopy*. Morgan and Claypool Publishers, 2016. ISBN 978-1-6817-4007-2.

[51] Tony F. Heinz. *Rayleigh Scattering Spectroscopy*, pages 353–369. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-72865-8. doi: 10.1007/978-3-540-72865-8_11. URL https://doi.org/10.1007/978-3-540-72865-8_11.

[52] M. F. Lin. Optical spectra of single-wall carbon nanotube bundles. *Phys. Rev. B*, 62:13153–13159, Nov 2000. doi: 10.1103/PhysRevB.62.13153. URL https://link.aps.org/doi/10.1103/PhysRevB.62.13153.

[53] David J. Griffiths. Absorption and dispersion. In *Introduction to Electrodynamics*, chapter 9, page 403. Addison-Wesley, 2015. ISBN 0-13-805326-X.

[54] Ermin Malić, Matthias Hirtschulz, Frank Milde, Yang Wu, Janina Maultzsch, Tony F. Heinz, Andreas Knorr, and Stephanie Reich. Theory of rayleigh scattering from metallic carbon nanotubes. *Phys. Rev. B*, 77:045432, Jan 2008. doi: 10.1103/PhysRevB.77.045432. URL https://link.aps.org/doi/10.1103/PhysRevB.77.045432.

[55] R. Saito, G. Dresselhaus, and M. S. Dresselhaus. Trigonal warping effect of carbon nanotubes. *Phys. Rev. B*, 61:2981–2990, Jan 2000. doi: 10.1103/PhysRevB.61.2981. URL https://link.aps.org/doi/10.1103/PhysRevB.61.2981.

[56] Stéphane Berciaud, Christophe Voisin, Hugen Yan, Bhupesh Chandra, Robert Caldwell, Yuyao Shan, Louis E. Brus, James Hone, and Tony F. Heinz. Excitons and high-order optical transitions in individual carbon nanotubes: A rayleigh scattering spectroscopy study. *Phys. Rev. B*, 81:041414, Jan 2010. doi: 10.1103/PhysRevB.81.041414. URL https://link.aps.org/doi/10.1103/PhysRevB.81.041414.

[57] Feng Wang, Matthew Y. Sfeir, Limin Huang, X. M. Henry Huang, Yang Wu, Jaehee Kim, James Hone, Stephen O'Brien, Louis E. Brus, and Tony F. Heinz. Interactions between individual carbon nanotubes studied by rayleigh scattering spectroscopy. *Phys. Rev. Lett.*, 96:167401, Apr 2006. doi: 10.1103/PhysRevLett.96.167401. URL https://link.aps.org/doi/10.1103/PhysRevLett.96.167401.

[58] Matthew Y. Sfeir, Feng Wang, Limin Huang, Chia-Chin Chuang, J. Hone, Stephen P. O'Brien, Tony F. Heinz, and Louis E. Brus. Probing electronic transitions in individual carbon nanotubes by rayleigh scattering. *Science*, 306:1540–1543, November 2004. doi: 10.1126/science.1103294. URL http://science.sciencemag.org/content/306/5701/1540.

[59] S. Maruyama. Kataura plot for resonant raman, 2002-01-18. URL http://www.photon.t.u-tokyo.ac.jp/~maruyama/kataura/kataura.html.

[60] Md Shamsul Arefin. Empirical equation based chirality (n, m) assignment of semiconducting single wall carbon nanotubes from resonant raman scattering data. *Nanomaterials*, 3:1–21, March 2013. doi: 10.3390/nano3010001. URL http://www.mdpi.com/2079-4991/3/1/1.

[61] G. R. Ahmed Jamal and S. M. Mominuzzaman. Beyond the basic tight binding model to calculate optical transition energies of swcnt. *Journal of Nanoscience and Nanoengineering*, 1:84–95, August 2015. URL http://files.aiscience.org/journal/article/html/70270014.html.

[62] G. R. Ahmed Jamal and S. M. Mominuzzaman. Universal empirical formula for optical transition energies of semiconducting single-walled carbon nanotubes. *AIP Advances*, 6:1–6, January 2016. doi: 10.1063/1.4940761. URL https://aip.scitation.org/doi/full/10.1063/1.4940761.

[63] Kaihui Liu, Jack Deslippe, Fajun Xiao, Rodrigo B. Capaz, Xiaoping Hong, Shaul Shaul Aloni, Alex Zettl, Wenlong Wang, Xuedong Bai, Steven G. Louie, Enge Wang, and Feng Wang. An atlas of carbon nanotube optical transitions. *Nature Nanotechnology*, 7:325–329, April 2012. doi: 10.1038/nnano.2012.52. URL https://www.nature.com/articles/nnano.2012.52.

[64] L. Wang, I. Meric, P. Y. Huang, Q. Gao, Y. Gao, H. Tran, T. Taniguchi, K. Watanabe, L. M. Campos, D. A. Muller, J. Guo, P. Kim, J. Hone, K. L. Shepard, and C. R. Dean. One-dimensional electrical contact to a two-dimensional material. *Science*, 342:614–617, November 2013. doi: 10.1126/science.1244358. URL http://science.sciencemag.org/content/342/6158/614.

[65] Jhao-Wun Huang, Cheng Pan, Son Tran, Bin Cheng, Kenji Watanabe, Takashi Taniguchi, Chun Ning Lau, and Marc Bockrath. Superior current carrying capacity of boron nitride encapsulated carbon nanotubes with zero-dimensional contacts. *Nano Letters*, 15:6836––6840, September 2015. doi: 10.1021/acs.nanolett.5b02716. URL https://pubs.acs.org/doi/abs/10.1021/acs.nanolett.5b02716.

[66] Charles A. Brau. Putting a spin on qft. In *Modern Problems in Classical Electrodynamics*, chapter 1, pages 29–72. Oxford University Press, 2004. ISBN 0-19-514665-4.

[67] Tom Lancaster and Stephen J. Blundell. Feynman's interpretation of the negative energy states. In *Quantum Field Theory for the Gifted Amateur*, chapter 6, pages 61–63. Oxford University Press, 2014. ISBN 978-0-19-969932-2.

[68] Charles Kittel. Holes. In *Introduction to Solid State Physics*, chapter 8, pages 194–197. John Wiley and Sons, 2005. ISBN 978-0-471-41526-8.

[69] Walter Greiner, B. Müller, and J. Rafelski. The klein paradox. In *Quantum Electrodynamics of Strong Fields: With an Introduction into Modern Relativistic Quantum Mechanics*, chapter 5, pages 112–121. Springer Publishing, 1985. ISBN 978-3-642-82272-8.

[70] P. E. Allain and J. N. Fuchs. Klein tunneling in graphene: optics with massless electrons. *The European Physical Journal B*, 83:301, October 2011. doi: 10.1140/epjb/e2011-20351-3. URL https://doi.org/10.1140/epjb/e2011-20351-3.

[71] Vadim V. Cheianov and Vladimir I. Fal'ko. Selective transmission of dirac electrons and ballistic magnetoresistance of $n-p$ junctions in graphene. *Phys. Rev. B*, 74:041403, Jul 2006. doi: 10.1103/PhysRevB.74.041403. URL https://link.aps.org/doi/10.1103/PhysRevB.74.041403.

[72] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim. Chiral tunnelling and the klein paradox in?graphene. *Nature Physics*, 2:620 EP –, Aug 2006. URL https://doi.org/10.1038/nphys384. Article.

[73] E. B. Sonin. Effect of klein tunneling on conductance and shot noise in ballistic graphene. *Phys. Rev. B*, 79:195438, May 2009. doi: 10.1103/PhysRevB.79.195438. URL https://link.aps.org/doi/10.1103/PhysRevB.79.195438.

[74] Andrea F. Young and Philip Kim. Quantum interference and klein tunnelling in graphene heterojunctions. *Nature Physics*, 5:222 EP –, Feb 2009. URL https://doi.org/10.1038/nphys1198.

[75] V. P. Neznamov and I. I. Safronov. New method for solving the z > 137 problem and determining hydrogen-like energy levels. *Uspekhi Fizicheskikh Nauk (UFN) Journal*, 57:189–193, February 2014. doi: 10.3367/ufne.0184.201402i.0200. URL https://doi.org/10.3367%2Fufne.0184.201402i.0200.

[76] I. Pomeranchuk and Y. Smorodinsky. On the energy levels of systems with z>137. *Journal of Physics*, 9:97–100, June 1945. URL ufn.ru/ufn13/ufn13_4/Russian/pomeranchuk.pdf.

[77] Ya. B. Zeldovich and V. S. Popov. Electronic structure of superheavy atoms. *Soviet Physics Uspekhi*, 14:673–694, June 1972. doi: 10.1070/pu1972v014n06abeh004735. URL https://doi.org/10.1070%2Fpu1972v014n06abeh004735.

[78] Walter Greiner, B. Müller, and J. Rafelski. Solutions of the dirac equation for a coulomb potentialand a scalar 1/r potential. In *Quantum Electrodynamics of Strong Fields: With an Introduction into Modern Relativistic Quantum Mechanics*, chapter 4, pages 83–86. Springer Publishing, 1985. ISBN 978-3-642-82272-8.

[79] Walter Greiner, B. Müller, and J. Rafelski. Resonant states in supercritical fields. In *Quantum Electrodynamics of Strong Fields: With an Introduction into Modern Relativistic Quantum Mechanics*, chapter 6, pages 122–137. Springer Publishing, 1985. ISBN 978-3-642-82272-8.

[80] Walter Greiner, B. Müller, and J. Rafelski. The charged vacuum. In *Quantum Electrodynamics of Strong Fields: With an Introduction into Modern Relativistic Quantum Mechanics*, chapter 1, pages 1–7. Springer Publishing, 1985. ISBN 978-3-642-82272-8.

[81] Paul Jerabek, Bastian Schuetrumpf, Peter Schwerdtfeger, and Witold Nazarewicz. Electron and nucleon localization functions of oganesson: Approaching the thomas-fermi limit. *Phys. Rev. Lett.*, 120:053001, Jan 2018. doi: 10.1103/PhysRevLett.120.053001. URL https://link.aps.org/doi/10.1103/PhysRevLett.120.053001.

[82] Y. Wang, D. Wong, A. Shytov, V. 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, May 2013. doi: 10.1126/science.1234320. URL http://science.sciencemag.org/content/340/6133/734.

[83] N. M. R. Peres and E. V. Castro. Algebraic solution of a graphene layer in transverse electric and perpendicular magnetic fields. *Journal of Physics: Condensed Matter*, 19:406231, September 2007. doi: 10.1088/0953-8984/19/40/406231. URL https://doi.org/10.1088%2F0953-8984%2F19%2F40%2F406231.

[84] Vinu Lukose, R. Shankar, and G. Baskaran. Novel electric field effects on landau levels in graphene. *Phys. Rev. Lett.*, 98:116802, Mar 2007. doi: 10.1103/PhysRevLett.98.116802. URL https://link.aps.org/doi/10.1103/PhysRevLett.98.116802.

[85] B.-Y. Jiang and M. M. Fogler. Electronic response of graphene to linelike charge perturbations. *Phys. Rev. B*, 91:235422, Jun 2015. doi: 10.1103/PhysRevB.91.235422. URL https://link.aps.org/doi/10.1103/PhysRevB.91.235422.

[86] R. R. Hartmann, N. J. Robinson, and M. E. Portnoi. Smooth electron waveguides in graphene. *Phys. Rev. B*, 81:245431, Jun 2010. doi: 10.1103/PhysRevB.81.245431. URL https://link.aps.org/doi/10.1103/PhysRevB.81.245431.

[87] Y. Jompol, C. J. B. Ford, J. P. Griffiths, I. Farrer, G. A. C. Jones, D. Anderson, D. A. Ritchie1, T. W. Silk, and A. J. Schofield. Probing spin-charge separation in a tomonaga-luttinger liquid. *Science*, 325:597–601, July 2009. doi: 10.1126/science.1171769. URL http://science.sciencemag.org/content/325/5940/597.

[88] A. Yacoby, H. L. Stormer, Ned S. Wingreen, L. N. Pfeiffer, K. W. Baldwin, and K. W. West. Nonuniversal conductance quantization in quantum wires. *Phys. Rev. Lett.*, 77:4612–4615, Nov 1996. doi: 10.1103/PhysRevLett.77.4612. URL https://link.aps.org/doi/10.1103/PhysRevLett.77.4612.

[89] Alexander S. Mayorov, Roman V. Gorbachev, Sergey V. Morozov, Liam Britnell, Rashid Jalil, Leonid A. Ponomarenko, Peter Blake, Kostya S. Novoselov, Kenji Watanabe, Takashi Taniguchi, and A. K. Geim. Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano Letters*, 11(6):2396–2399, Jun 2011. ISSN 1530-6984. doi: 10.1021/nl200758b. URL https://doi.org/10.1021/nl200758b.

[90] A. N. Grigorenko, M. Polini, and K. S. Novoselov. Graphene plasmonics. *Nature Photonics*, 6:749 EP –, Nov 2012. URL https://doi.org/10.1038/nphoton.2012.262. Review Article.

[91] F. Bonaccorso, Z. Sun, T. Hasan, and A. C. Ferrari. Graphene photonics and optoelectronics. *Nature Photonics*, 4:611 EP –, Aug 2010. URL https://doi.org/10.1038/nphoton.2010.186. Review Article.

[92] G. X. Ni, A. S. McLeod, Z. Sun, L. Wang, L. Xiong, K. W. Post, S. S. Sunku, B.-Y. Jiang, J. Hone, C. R. Dean, M. M. Fogler, and D. N. Basov. Fundamental limits to graphene plasmonics. *Nature*, 557(7706):530–533, 2018. ISSN 1476-4687. doi: 10.1038/s41586-018-0136-9. URL https://doi.org/10.1038/s41586-018-0136-9.

[93] Arthur W. Barnard, Alex Hughes, Aaron L. Sharpe, Kenji Watanabe, Takashi Taniguchi, and David Goldhaber-Gordon. Absorptive pinhole collimators for ballistic dirac fermions in graphene. *Nature Communications*, 8:1541, May 2017. doi: 10.1126/science.1171769. URL https://doi.org/10.1038/ncomms15418.

[94] Peter Rickhaus, Péter Makk, Ming-Hao Liu, Klaus Richter, and Christian Schönenberger. Gate tuneable beamsplitter in ballistic graphene. *Applied Physics Letters*, 107:251901, December 2015. doi: doi:10.1063/1.4938073. URL https://aip.scitation.org/doi/10.1063/1.4938073.

[95] Ming Liu, Xiaobo Yin, Erick Ulin-Avila, Baisong Geng, Thomas Zentgraf, Long Ju, Feng Wang, and Xiang Zhang. A graphene-based broadband optical modulator. *Nature*, 474:64 EP –, May 2011. URL https://doi.org/10.1038/nature10067.

[96] Peter Rickhaus, Romain Maurand, Ming-Hao Liu, Markus Weiss, Klaus Richter, and Christian Schönenberger. Ballistic interferences in suspended graphene. *Nature Communications*, 4:2342 EP –, Aug 2013. doi: 10.1038/ncomms3342. URL https://doi.org/10.1038/ncomms3342. Article.

[97] V. V. Cheianov, V. Fal'ko1, and B. L. Altshuler. The focusing of electron flow and a veselago lens in graphene p-n junctions. *Science*, 315:1252–1255, March 2007. doi: 10.1126/science.1138020. URL http://science.sciencemag.org/content/315/5816/1252.

[98] Shaowen Chen, Zheng Han, Mirza M. Elahi, K. M. Masum Habib, Lei Wang, Bo Wen, Yuanda Gao, Takashi Taniguchi, Kenji Watanabe, James Hone, Avik W. Ghosh, and Cory R. Dean. Electron optics with p-n junctions in ballistic graphene. *Science*, 353:1522–1525, September 2016. doi: 10.1126/science.aaf5481. URL http://science.sciencemag.org/content/353/6307/1522.

[99] Peter Rickhaus, Péter Makk, Ming-Hao Liu, Endre Tóvári, Markus Weiss, Romain Maurand, Klaus Richter, and Christian Schönenberger. Snake trajectories in ultraclean graphene p-n junctions. *Nature Communications*, 6:6470 EP –, Mar 2015. doi: 10.1038/ncomms7470. URL https://doi.org/10.1038/ncomms7470. Article.

[100] M. T. Allen, O. Shtanko, I. C. Fulga, A. R. Akhmerov, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, L. S. Levitov, and A. Yacoby. Spatially resolved edge currents and guided-wave electronic states in graphene. *Nature Physics*, 12:128 EP –, Nov 2015. doi: 10.1038/nphys3534. URL https://doi.org/10.1038/nphys3534.

[101] Peter Rickhaus, Ming-Hao Liu, Péter Makk, Romain Maurand, Samuel Hess, Simon Zihlmann, Markus Weiss, Klaus Richter, and Christian Schönenberger. Guiding of electrons in a few-mode ballistic graphene channel. *Nano Letters*, 15(9):5819–5825, Sep 2015. ISSN 1530-6984. doi: 10.1021/acs.nanolett.5b01877. URL https://doi.org/10.1021/acs.nanolett.5b01877.

[102] J. R. Williams, Tony Low, M. S. Lundstrom, and C. M. Marcus. Gate-controlled guiding of electrons in graphene. *Nature Nanotechnology*, 6:222 EP –, Feb 2011. doi: 10.1038/nnano.2011.3. URL https://doi.org/10.1038/nnano.2011.3.

[103] Clevin. J. Handschin. *Quantum Transport in Encapsulated Graphene P-N Junctions*. PhD thesis, University of Basel, 2017. URL https://edoc.unibas.ch/57468/.

[104] Viktor G Veselago. The electrodynamics of substances with simultaneously negative values of ϵ and μ. *Soviet Physics Uspekhi*, 10:509–514, April 1968. doi: 10.1070/pu1968v010n04abeh003699. URL https://doi.org/10.1070%2Fpu1968v010n04abeh003699.

[105] D. Gloge. Weakly guiding fibers. *Appl. Opt.*, 10(10):2252–2258, Oct 1971. doi: 10.1364/AO.10.002252. URL http://ao.osa.org/abstract.cfm?URI=ao-10-10-2252.

[106] John A. Buck. *Fundamental of Optical Fibers*. John Wiley and Sons, 1995. ISBN 978-0471221913.

[107] K. Akkaravarawong, O. Shtanko, and L. Levitov. Ballistic guided electron states in graphene. *Arxiv*, December 2015. URL https://arxiv.org/abs/1512.04185v1. arXiv:1512.04185v1 [cond-mat.mes-hall].

[108] S. Samaddar, I. Yudhistira, S. Adam, H. Courtois, and C. B. Winkelmann. Charge puddles in graphene near the dirac point. *Phys. Rev. Lett.*, 116:126804, Mar 2016. doi: 10.1103/PhysRevLett.116.126804. URL https://link.aps.org/doi/10.1103/PhysRevLett.116.126804.

[109] C. W. J. Beenakker, R. A. Sepkhanov, A. R. Akhmerov, and J. Tworzydło. Quantum goos-hänchen effect in graphene. *Phys. Rev. Lett.*, 102:146804, Apr 2009. doi: 10.1103/PhysRevLett.102.146804. URL https://link.aps.org/doi/10.1103/PhysRevLett.102.146804.

[110] C. A. Downing and M. E. Portnoi. One-dimensional coulomb problem in dirac materials. *Phys. Rev. A*, 90:052116, Nov 2014. doi: 10.1103/PhysRevA.90.052116. URL https://link.aps.org/doi/10.1103/PhysRevA.90.052116.

[111] H. Chau Nguyen, M. Tien Hoang, and V. Lien Nguyen. Quasi-bound states induced by one-dimensional potentials in graphene. *Phys. Rev. B*, 79:035411, Jan 2009. doi: 10.1103/PhysRevB.79.035411. URL https://link.aps.org/doi/10.1103/PhysRevB.79.035411.

[112] Yang Wang, Victor W. Brar, Andrey V. Shytov, Qiong Wu, William Regan, Hsin-Zon Tsai, Alex Zettl, Leonid S. Levitov, and Michael F. Crommie. Mapping dirac quasiparticles near a single coulomb impurity on graphene. *Nature Physics*, 8:653 EP –, Jul 2012. URL https://doi.org/10.1038/nphys2379.

[113] Jan Zaanen. Electrons go with the flow in exotic material systems. *Science*, 351(6277):1026–1027, 2016. ISSN 0036-8075. doi: 10.1126/science.aaf2487. URL https://science.sciencemag.org/content/351/6277/1026.

[114] Andrew Lucas and Kin Chung Fong. Hydrodynamics of electrons in graphene. *Journal of Physics: Condensed Matter*, 30(5):053001, jan 2018. doi: 10.1088/1361-648x/aaa274. URL https://doi.org/10.1088%2F1361-648x%2Faaa274.

[115] R N Gurzhi. Hydrodynamic effects in solids at low temperature. *Soviet Physics Uspekhi*, 11: 255–270, February 1968. doi: 10.1070/pu1968v011n02abeh003815. URL https://doi.org/10.1070%2Fpu1968v011n02abeh003815.

[116] M. J. M. de Jong and L. W. Molenkamp. Hydrodynamic electron flow in high-mobility wires. *Phys. Rev. B*, 51:13389–13402, May 1995. doi: 10.1103/PhysRevB.51.13389. URL https://link.aps.org/doi/10.1103/PhysRevB.51.13389.

[117] Moshe Kaveh and Nathan Wiser. Electron-electron scattering in conducting materials. *Advances in Physics*, 33(4):257–372, 1984. doi: 10.1080/00018738400101671. URL https://doi.org/10.1080/00018738400101671.

[118] A. I. Berdyugin, S. G. Xu, F. M. D. Pellegrino, R. Krishna Kumar, A. Principi, I. Torre, M. Ben Shalom, T. Taniguchi, K. Watanabe, I. V. Grigorieva, M. Polini, A. K. Geim, and D. A. Bandurin. Measuring hall viscosity of graphene's electron fluid. *Science*, 364(6436):162–165, 2019. doi: 10.1126/science.aau0685. URL https://science.sciencemag.org/content/364/6436/162.

[119] R. Krishna Kumar, D. A. Bandurin, F. M. D. Pellegrino, Y. Cao, A. Principi, H. Guo, G. ?. H. Auton, M. Ben Shalom, L. A. Ponomarenko, G. Falkovich, K. Watanabe, T. Taniguchi, I. ?. V. Grigorieva, L. S. Levitov, M. Polini, and A. ?. K. Geim. Superballistic flow of viscous electron

fluid through graphene constrictions. *Nature Physics*, 13:1182 EP –, Aug 2017. URL https://doi.org/10.1038/nphys4240.

[120] Iacopo Torre, Andrea Tomadin, Andre K. Geim, and Marco Polini. Nonlocal transport and the hydrodynamic shear viscosity in graphene. *Phys. Rev. B*, 92:165433, Oct 2015. doi: 10.1103/PhysRevB.92.165433. URL https://link.aps.org/doi/10.1103/PhysRevB.92.165433.

[121] Francesco M. D. Pellegrino, Iacopo Torre, Andre K. Geim, and Marco Polini. Electron hydrodynamics dilemma: Whirlpools or no whirlpools. *Phys. Rev. B*, 94:155414, Oct 2016. doi: 10.1103/PhysRevB.94.155414. URL https://link.aps.org/doi/10.1103/PhysRevB.94.155414.

[122] Leonid Levitov and Gregory Falkovich. Electron viscosity, current vortices and negative nonlocal resistance in graphene. *Nature Physics*, 12:672 EP –, Feb 2016. URL https://doi.org/10.1038/nphys3667.

[123] Gregory Falkovich and Leonid Levitov. Linking spatial distributions of potential and current in viscous electronics. *Phys. Rev. Lett.*, 119:066601, Aug 2017. doi: 10.1103/PhysRevLett.119.066601. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.066601.

[124] Denis A. Bandurin, Andrey V. Shytov, Leonid S. Levitov, Roshan Krishna Kumar, Alexey I. Berdyugin, Moshe Ben Shalom, Irina V. Grigorieva, Andre K. Geim, and Gregory Falkovich. Fluidity onset in graphene. *Nature Communications*, 9(1):4533, 2018. doi: 10.1038/s41467-018-07004-4. URL https://doi.org/10.1038/s41467-018-07004-4.

[125] D. Svintsov. Hydrodynamic-to-ballistic crossover in dirac materials. *Phys. Rev. B*, 97:121405, Mar 2018. doi: 10.1103/PhysRevB.97.121405. URL https://link.aps.org/doi/10.1103/PhysRevB.97.121405.

[126] Mani Chandra, Gitansh Kataria, Deshdeep Sahdev, and Ravishankar Sundararaman. Hydrodynamic and ballistic ac transport in two-dimensional fermi liquids. *Phys. Rev. B*, 99:165409, Apr 2019. doi: 10.1103/PhysRevB.99.165409. URL https://link.aps.org/doi/10.1103/PhysRevB.99.165409.

[127] Andrey Shytov, Jian Feng Kong, Gregory Falkovich, and Leonid Levitov. Particle collisions and negative nonlocal response of ballistic electrons. *Phys. Rev. Lett.*, 121:176805, Oct

2018. doi: 10.1103/PhysRevLett.121.176805. URL https://link.aps.org/doi/10.1103/PhysRevLett.121.176805.

[128] Markus Müller and Subir Sachdev. Collective cyclotron motion of the relativistic plasma in graphene. *Phys. Rev. B*, 78:115419, Sep 2008. doi: 10.1103/PhysRevB.78.115419. URL https://link.aps.org/doi/10.1103/PhysRevB.78.115419.

[129] D. Svintsov, V. Vyurkov, S. Yurchenko, T. Otsuji, and V. Ryzhii. Hydrodynamic model for electron-hole plasma in graphene. *Journal of Applied Physics*, 111(8):083715, Apr 2012. ISSN 0021-8979. doi: 10.1063/1.4705382. URL https://doi.org/10.1063/1.4705382.

[130] S. S. Apostolov, A. Levchenko, and A. V. Andreev. Hydrodynamic coulomb drag of strongly correlated electron liquids. *Phys. Rev. B*, 89:121104, Mar 2014. doi: 10.1103/PhysRevB.89.121104. URL https://link.aps.org/doi/10.1103/PhysRevB.89.121104.

[131] B. N. Narozhny and A. Levchenko. Coulomb drag. *Rev. Mod. Phys.*, 88:025003, May 2016. doi: 10.1103/RevModPhys.88.025003. URL https://link.aps.org/doi/10.1103/RevModPhys.88.025003.

[132] J. I. A. Li, T. Taniguchi, K. Watanabe, J. Hone, A. Levchenko, and C. R. Dean. Negative coulomb drag in double bilayer graphene. *Phys. Rev. Lett.*, 117:046802, Jul 2016. doi: 10.1103/PhysRevLett.117.046802. URL https://link.aps.org/doi/10.1103/PhysRevLett.117.046802.

[133] T. J. Gramila, J. P. Eisenstein, A. H. MacDonald, L. N. Pfeiffer, and K. W. West. Mutual friction between parallel two-dimensional electron systems. *Phys. Rev. Lett.*, 66:1216–1219, Mar 1991. doi: 10.1103/PhysRevLett.66.1216. URL https://link.aps.org/doi/10.1103/PhysRevLett.66.1216.

[134] Karsten Flensberg, Ben Yu-Kuang Hu, Antti-Pekka Jauho, and Jari M. Kinaret. Linear-response theory of coulomb drag in coupled electron systems. *Phys. Rev. B*, 52:14761–14774, Nov 1995. doi: 10.1103/PhysRevB.52.14761. URL https://link.aps.org/doi/10.1103/PhysRevB.52.14761.

[135] N. P. R. Hill, J. T. Nicholls, E. H. Linfield, M. Pepper, D. A. Ritchie, A. R. Hamilton, and G. A. C. Jones. Frictional drag between parallel two-dimensional electron gases in a perpendicular magnetic field. *Journal of Physics: Condensed Matter*, 8:L557–L562, September 1996.

doi: 10.1088/0953-8984/8/39/001. URL https://doi.org/10.1088%2F0953-8984%2F8%2F39%2F001.

[136] X. G. Feng, S. Zelakiewicz, H. Noh, T. J. Ragucci, T. J. Gramila, L. N. Pfeiffer, and K. W. West. Negative electron drag and holelike behavior in the integer quantum hall regime. *Phys. Rev. Lett.*, 81:3219–3222, Oct 1998. doi: 10.1103/PhysRevLett.81.3219. URL https://link.aps.org/doi/10.1103/PhysRevLett.81.3219.

[137] Karsten Flensberg. Coulomb drag of luttinger liquids and quantum hall edges. *Phys. Rev. Lett.*, 81:184–187, Jul 1998. doi: 10.1103/PhysRevLett.81.184. URL https://link.aps.org/doi/10.1103/PhysRevLett.81.184.

[138] R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy, I. V. Grigorieva, A. H. MacDonald, S. V. Morozov, K. Watanabe, T. Taniguchi, and L. A. Ponomarenko. Strong coulomb drag and broken symmetry in double-layer graphene. *Nature Physics*, 8:896–901, October 2012. doi: 10.1038/nphys2441. URL https://www.nature.com/articles/nphys2441.

[139] A. F. Croxall, K. Das Gupta, C. A. Nicoll, M. Thangaraj, H. E. Beere, I. Farrer, D. A. Ritchie, and M. Pepper. Anomalous coulomb drag in electron-hole bilayers. *Phys. Rev. Lett.*, 101:246801, Dec 2008. doi: 10.1103/PhysRevLett.101.246801. URL https://link.aps.org/doi/10.1103/PhysRevLett.101.246801.

[140] Pauline Simonet, Szymon Hennel, Hiske Overweg, Richard Steinacher, Marius Eich, Riccardo Pisoni, Yongjin Lee, Peter Märki, Thomas Ihn, Klaus Ensslin, Mattias Beck, and Jérôme Faist. Anomalous coulomb drag between bilayer graphene and a gaas electron gas. *New Journal of Physics*, 19:103042, Nov 2017. doi: 10.1088/1367-2630/aa887e. URL https://doi.org/10.1088%2F1367-2630%2Faa887e.

[141] Seyoung Kim, Insun Jo, Junghyo Nah, Z. Yao, S. K. Banerjee, and E. Tutuc. Coulomb drag of massless fermions in graphene. *Phys. Rev. B*, 83:161401, Apr 2011. doi: 10.1103/PhysRevB.83.161401. URL https://link.aps.org/doi/10.1103/PhysRevB.83.161401.

[142] Kayoung Lee, Jiamin Xue, David C. Dillen, Kenji Watanabe, Takashi Taniguchi, and Emanuel Tutuc. Giant frictional drag in double bilayer graphene heterostructures. *Phys. Rev. Lett.*, 117:046803, Jul 2016. doi: 10.1103/PhysRevLett.117.046803. URL https://link.aps.org/doi/10.1103/PhysRevLett.117.046803.

[143] A G Rojo. Electron-drag effects in coupled electron systems. *Journal of Physics: Condensed Matter*, 11:R31–R52, January 1999. doi: 10.1088/0953-8984/11/5/004. URL https://doi.org/10.1088%2F0953-8984%2F11%2F5%2F004.

[144] Markus Müller, Lars Fritz, and Subir Sachdev. Quantum-critical relativistic magnetotransport in graphene. *Phys. Rev. B*, 78:115406, Sep 2008. doi: 10.1103/PhysRevB.78.115406. URL https://link.aps.org/doi/10.1103/PhysRevB.78.115406.

[145] Alex Kamenev and Yuval Oreg. Coulomb drag in normal metals and superconductors: Diagrammatic approach. *Phys. Rev. B*, 52:7516–7527, Sep 1995. doi: 10.1103/PhysRevB.52.7516. URL https://link.aps.org/doi/10.1103/PhysRevB.52.7516.

[146] Ben Yu-Kuang Hu and Karsten Flensberg. Electron-electron scattering in linear transport in two-dimensional systems. *Phys. Rev. B*, 53:10072–10077, Apr 1996. doi: 10.1103/PhysRevB.53.10072. URL https://link.aps.org/doi/10.1103/PhysRevB.53.10072.

[147] M. Pustilnik, E. G. Mishchenko, L. I. Glazman, and A. V. Andreev. Coulomb drag by small momentum transfer between quantum wires. *Phys. Rev. Lett.*, 91:126805, Sep 2003. doi: 10.1103/PhysRevLett.91.126805. URL https://link.aps.org/doi/10.1103/PhysRevLett.91.126805.

[148] M. Yamamoto, M. Stopa, Y. Tokura, Y. Hirayama, and S. Tarucha1. Negative coulomb drag in a one-dimensional wire. *Science*, 313:204–207, July 2006. doi: 10.1126/science.1126601. URL http://science.sciencemag.org/content/313/5784/204.

[149] D. Laroche1, G. Gervais1, M. P. Lilly, and J. L. Reno. 1d-1d coulomb drag signature of a luttinger liquid. *Science*, 343:631–634, February 2014. doi: 10.1126/science.1244152. URL http://science.sciencemag.org/content/343/6171/631.

[150] D. Nandi, A. D. K. Finck, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Exciton condensation and perfect coulomb drag. *Nature*, 488:481–484, August 2012. doi: 10.1038/nature11302. URL https://www.nature.com/articles/nature11302.

[151] X. Liu, K. Watanabe, T. Taniguchi, B. I. Halperin, and P. Kim. Quantum hall drag of exciton condensate in graphene. *Nature Physics*, 13:746–750, May 2017. doi: 10.1038/nphys4116. URL https://www.nature.com/articles/nphys4116.

[152] J. Li, T. Taniguchi, K. Watanabe, J. Hone, and C. Dean. Quantum hall drag of exciton condensate in graphene. *Nature Physics*, 13:751–755, May 2017. doi: 10.1038/nphys4140. URL https://www.nature.com/articles/nphys4140.

[153] M. Pustilnik, E. G. Mishchenko, L. I. Glazman, and A. V. Andreev. Coulomb drag by small momentum transfer between quantum wires. *Phys. Rev. Lett.*, 91:126805, Sep 2003. doi: 10.1103/PhysRevLett.91.126805. URL https://link.aps.org/doi/10.1103/PhysRevLett.91.126805.

[154] Thomas Fuchs, Rochus Klesse, and Ady Stern. Coulomb drag between quantum wires with different electron densities. *Phys. Rev. B*, 71:045321, Jan 2005. doi: 10.1103/PhysRevB.71.045321. URL https://link.aps.org/doi/10.1103/PhysRevB.71.045321.

[155] Wang-Kong Tse, Ben Yu-Kuang Hu, and S. Das Sarma. Theory of coulomb drag in graphene. *Phys. Rev. B*, 76:081401, Aug 2007. doi: 10.1103/PhysRevB.76.081401. URL https://link.aps.org/doi/10.1103/PhysRevB.76.081401.

[156] K. Muraki, J. G. S. Lok, S. Kraus, W. Dietsche, K. von Klitzing, D. Schuh, M. Bichler, and W. Wegscheider. Coulomb drag as a probe of the nature of compressible states in a magnetic field. *Phys. Rev. Lett.*, 92:246801, Jun 2004. doi: 10.1103/PhysRevLett.92.246801. URL https://link.aps.org/doi/10.1103/PhysRevLett.92.246801.

[157] Felix von Oppen, Steven H. Simon, and Ady Stern. Oscillating sign of drag in high landau levels. *Phys. Rev. Lett.*, 87:106803, Aug 2001. doi: 10.1103/PhysRevLett.87.106803. URL https://link.aps.org/doi/10.1103/PhysRevLett.87.106803.

[158] Xiaomeng Liu, Lei Wang, Kin Chung Fong, Yuanda Gao, Patrick Maher, Kenji Watanabe, Takashi Taniguchi, James Hone, Cory Dean, and Philip Kim. Frictional magneto-coulomb drag in graphene double-layer heterostructures. *Phys. Rev. Lett.*, 119:056802, Aug 2017. doi: 10.1103/PhysRevLett.119.056802. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.056802.

[159] J. C. W. Song, D. A. Abanin, and L. S. Levitov. Coulomb drag mechanisms in graphene. *Nano Letters*, 13:3631––3637, July 2013. doi: 10.1021/nl401475u. URL https://pubs.acs.org/doi/abs/10.1021/nl401475u.

[160] Justin C. W. Song and Leonid S. Levitov. Energy-driven drag at charge neutrality in graphene. *Phys. Rev. Lett.*, 109:236602, Dec 2012. doi: 10.1103/PhysRevLett.109.236602. URL https://link.aps.org/doi/10.1103/PhysRevLett.109.236602.

[161] Herbert B. Callen. The application of onsager's reciprocal relations to thermoelectric, thermomagnetic, and galvanomagnetic effects. *Phys. Rev.*, 73:1349–1358, Jun 1948. doi: 10.1103/PhysRev.73.1349. URL https://link.aps.org/doi/10.1103/PhysRev.73.1349.

[162] Justin C. W. Song and Leonid S. Levitov. Hall drag and magnetodrag in graphene. *Phys. Rev. Lett.*, 111:126601, Sep 2013. doi: 10.1103/PhysRevLett.111.126601. URL https://link.aps.org/doi/10.1103/PhysRevLett.111.126601.

[163] Lars Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, Feb 1931. doi: 10.1103/PhysRev.37.405. URL https://link.aps.org/doi/10.1103/PhysRev.37.405.

[164] Lars Onsager. Reciprocal relations in irreversible processes. ii. *Phys. Rev.*, 38:2265–2279, Dec 1931. doi: 10.1103/PhysRev.38.2265. URL https://link.aps.org/doi/10.1103/PhysRev.38.2265.

[165] Charles Kittel and Herbert Kroemer. Kinetic theory. In *Thermal Physics*, chapter 14, page 406. W. H. Freeman and Company, 1980. ISBN 978-0-7167-1088-2.

[166] H. B. G. Casimir. On onsager's principle of microscopic reversibility. *Rev. Mod. Phys.*, 17:343–350, Apr 1945. doi: 10.1103/RevModPhys.17.343. URL https://link.aps.org/doi/10.1103/RevModPhys.17.343.

[167] Seyoung Kim, Insun Jo, Junghyo Nah, Z. Yao, S. K. Banerjee, and E. Tutuc. Coulomb drag of massless fermions in graphene. *Phys. Rev. B*, 83:161401, Apr 2011. doi: 10.1103/PhysRevB.83.161401. URL https://link.aps.org/doi/10.1103/PhysRevB.83.161401.

[168] Seyoung Kim and E. Tutuc. Coulomb drag and magnetotransport in graphene double layers. *Solid State Communications*, 152:1283–1288, August 2012. doi: 10.1016/J.SSC.2012.04.032. URL https://doi.org/10.1016/j.ssc.2012.04.032.

[169] A. S. Price, A. K. Savchenko, B. N. Narozhny, G. Allison, and D. A. Ritchie. Giant fluctuations of coulomb drag in a bilayer system. *Science*, 316:1055–1058, April 2007. doi: 10.1126/science.1139227. URL http://science.sciencemag.org/content/316/5821/99/tab-pdf.

[170] Lars Fritz, Jörg Schmalian, Markus Müller, and Subir Sachdev. Quantum critical transport in clean graphene. *Phys. Rev. B*, 78:085416, Aug 2008. doi: 10.1103/PhysRevB.78.085416. URL https://link.aps.org/doi/10.1103/PhysRevB.78.085416.

[171] Youngwoo Nam, Dong-Keun Ki, David Soler-Delgado, and Alberto F. Morpurgo. Electron-hole collision limited transport in charge-neutral bilayer graphene. *Nature Physics*, 13:1207 EP –, Aug 2017. URL https://doi.org/10.1038/nphys4218. Article.

[172] Dmitri K. Efetov and Philip Kim. Controlling electron-phonon interactions in graphene at ultrahigh carrier densities. *Phys. Rev. Lett.*, 105:256805, Dec 2010. doi: 10.1103/PhysRevLett.105.256805. URL https://link.aps.org/doi/10.1103/PhysRevLett.105.256805.

[173] W. Chen, A. V. Andreev, and A. Levchenko. Boltzmann-langevin theory of coulomb drag. *Phys. Rev. B*, 91:245405, Jun 2015. doi: 10.1103/PhysRevB.91.245405. URL https://link.aps.org/doi/10.1103/PhysRevB.91.245405.

[174] Kaihui Liu, Xiaoping Hong, Qin Zhou, Chenhao Jin, Jinghua Li, Weiwei Zhou, Jie Liu, Enge Wang, Alex Zettl, and Feng Wang. High-throughput optical imaging and spectroscopy of individual carbon nanotubes in devices. *Nature Nanotechnology*, 8:917 EP –, Nov 2013. doi: 10.1038/nnano.2013.227. URL https://doi.org/10.1038/nnano.2013.227.

[175] Elizabeth Gibney. How 'magic angle' graphene is stirring up physics. *Nature*, January 2019. URL https://www.nature.com/articles/d41586-018-07848-2.

[176] Yuan Cao, Valla Fatemi, Ahmet Demir, Shiang Fang, Spencer L. Tomarken, Jason Y. Luo, Javier D. Sanchez-Yamagishi, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, Ray C. Ashoori, and Pablo Jarillo-Herrero. Correlated insulator behaviour at half-filling in magic-angle graphene superlattices. *Nature*, 556:80 EP –, Mar 2018. URL https://doi.org/10.1038/nature26154.

[177] Yuan Cao, Valla Fatemi, Shiang Fang, Kenji Watanabe, Takashi Taniguchi, Efthimios Kaxiras, and Pablo Jarillo-Herrero. Unconventional superconductivity in magic-angle graphene superlattices. *Nature*, 556:43 EP –, Mar 2018. URL https://doi.org/10.1038/nature26160. Article.

[178] Mikhail I. Katsnelson. *Graphene: Carbon in Two Dimensions*. Cambridge University Press, 2012. ISBN 978-0-521-19540-9.

[179] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Reviews of Modern Physics*, 81:109–162, January 2009. doi: 10.1103/RevModPhys.81.109. URL https://link.aps.org/doi/10.1103/RevModPhys.81.109.

[180] H. S. Philip Wong and Deji Akinwande. *Carbon Nanotubes and Graphene Device Physics*. Cambridge University Press, 2011. ISBN 978-0-521-51905-2.

[181] Neil W. Ashcroft and N. David. Mermin. *Solid State Physics*. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[182] Neil W. Ashcroft and N. David. Mermin. General remarks about bloch's theorem. In *Solid State Physics*, chapter 8, page 139. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[183] Neil W. Ashcroft and N. David. Mermin. The reciprocal lattice is a bravais lattice. In *Solid State Physics*, chapter 5, page 86. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[184] Piers Coleman. Landau fermi-liquid theory. In *Introduction to Many-Body Physics*, chapter 6, page 127. Cambridge University Press, 2015. ISBN 978-0-521-86488-6.

[185] Neil W. Ashcroft and N. David. Mermin. General remarks about bloch's theorem. In *Solid State Physics*, chapter 8, page 134. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[186] Neil W. Ashcroft and N. David. Mermin. Perturbation theory and weak periodic potentials. In *Solid State Physics*, chapter 9, pages 152–162. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[187] Andrea F. Young, Yuanbo Zhang, and Philip Kim. Experimental manifestation of berry phase in graphene. In *Physics of Graphene*, chapter 1, pages 3–27. Springer International Publishing, 2013. ISBN 978-3-319-02632-9.

[188] Tsuneya Ando, Takeshi Nakanishi, and Riichiro Saito. Berry's phase and absence of back scattering in carbon nanotubes. *Journal of the Physical Society of Japan*, 67(8):2857–2862, 1998. doi: 10.1143/JPSJ.67.2857. URL https://doi.org/10.1143/JPSJ.67.2857.

[189] Paul L. McEuen, Marc Bockrath, David H. Cobden, Young-Gui Yoon, and Steven G. Louie. Disorder, pseudospins, and backscattering in carbon nanotubes. *Physical Review Letters*, 83: 5098–5101, Dec 1999. doi: 10.1103/PhysRevLett.83.5098. URL https://link.aps.org/doi/10.1103/PhysRevLett.83.5098.

[190] Neil W. Ashcroft and N. David. Mermin. Measuring the fermi surface. In *Solid State Physics*, chapter 14, page 272. Brooks-Cole Publishing, 1976. ISBN 978-0-03-083993-1.

[191] H. Kataura, Y. Kumazawa, Y. Maniwa, I. Umezu, S. Suzuki, Y. Ohtsuka, and Y. Achiba. Optical properties of single-wall carbon nanotubes. *Synthetic Metals*, 103:2555–2558, June 1999. doi: 10.1016/S0379-6779(98)00278-1. URL https://doi.org/10.1016/S0379-6779(98)00278-1.

[192] R. Zhang, Y. Zhang, Q. Zhang, H. Xie, W. Qian, and F. Wei. Growth of half-meter long carbon nanotubes based on schulz–flory distribution. *ACS Nano*, 7:6156–6161, June 2013. doi: 10.1021/nn401995z. URL https://pubs.acs.org/doi/10.1021/nn401995z.

[193] H. S. Philip Wong and Deji Akinwande. Cnt brillouin zone. In *Carbon Nanotubes and Graphene Device Physics*, chapter 4, page 83. Cambridge University Press, 2011. ISBN 978-0-521-51905-2.

[194] J. W. Mintmire and C. T. White. Universal density of states for carbon nanotubes. *Phys. Rev. Lett.*, 81:2506–2509, September 1998. doi: 10.1103/PhysRevLett.81.2506. URL https://link.aps.org/doi/10.1103/PhysRevLett.81.2506.

[195] H. S. Philip Wong and Deji Akinwande. Cnt equilibrium properties. In *Carbon Nanotubes and Graphene Device Physics*, chapter 5, page 107. Cambridge University Press, 2011. ISBN 978-0-521-51905-2.

[196] Deji Akinwande, Yoshio Nishi, and H. S.Philip Wong. An analytical derivation of the density of states, effective mass, and carrier density for achiral carbon nanotubes. *IEEE Transactions on Electron Devices*, 55:289–297, January 2008. doi: 10.1109/TED.2007.911078. URL https://ieeexplore.ieee.org/document/4408769.

[197] Ying Tian, Hua Jiang, Patrik Laiho, and Esko I. Kauppinen. Validity of measuring metallic and semiconducting single-walled carbon nanotube fractions by quantitative raman spectroscopy. *Analytical Chemistry*, 90:2517–2525, 2018. doi: 10.1021/acs.analchem.7b03712. URL https://doi.org/10.1021/acs.analchem.7b03712.

[198] A. Grüneis, R. Saito, Ge. G. Samsonidze, T. Kimura, M. A. Pimenta, A. Jorio, A. G. Souza Filho, G. Dresselhaus, and M. S. Dresselhaus. Inhomogeneous optical absorption around the k point in graphite and carbon nanotubes. *Phys. Rev. B*, 67:165402, Apr 2003. doi: 10.

1103/PhysRevB.67.165402. URL https://link.aps.org/doi/10.1103/PhysRevB.67.165402.

[199] C. W. J. Beenakker. Theory of coulomb-blockade oscillations in the conductance of a quantum dot. *Phys. Rev. B*, 44:1646–1656, Jul 1991. doi: 10.1103/PhysRevB.44.1646. URL https://link.aps.org/doi/10.1103/PhysRevB.44.1646.

[200] G. Breit and E. Wigner. Capture of slow neutrons. *Phys. Rev.*, 49:519–531, Apr 1936. doi: 10.1103/PhysRev.49.519. URL https://link.aps.org/doi/10.1103/PhysRev.49.519.

[201] W. Liang, M. Bockrath, D. Bozovic, J. H. Hafner, M. Tinkham, and H. Park. Fabry - perot interference in a nanotube electron waveguide. *Nature*, 411:665–669, June 2001. doi: 10.1038/35079517. URL https://www.nature.com/articles/35079517.

[202] J. N. Fuchs, F. Piéchon, M. O. Goerbig, and G. Montambaux. Topological berry phase and semiclassical quantization of cyclotron orbits for two dimensional electrons in coupled band models. *The European Physical Journal B*, 77:351–362, October 2010. doi: 10.1140/epjb/e2010-00259-2. URL https://doi.org/10.1140/epjb/e2010-00259-2.

[203] Yisong Zheng and Tsuneya Ando. Hall conductivity of a two-dimensional graphite system. *Phys. Rev. B*, 65:245420, June 2002. doi: 10.1103/PhysRevB.65.245420. URL https://link.aps.org/doi/10.1103/PhysRevB.65.245420.

[204] K. S. Novoselov, Z. Jiang, S. V. Zhang, Y.and Morozaov, H. L. Stormer, U. Zeitler, J. C. Maan, G. S. Boebinger, P. Kim, and A. K. Geim. Room-temperature quantum hall effect in graphene. *Science*, 315:1379, March 2007. doi: 10.1126/science.1137201. URL http://science.sciencemag.org/content/315/5817/1379.

[205] Mark O. Goerbig. The quantum hall effect in graphene – a theoretical perspective". *Comptes Rendus Physique*, 12(4):369 – 378, 2011. doi: 10.1016/j.crhy.2011.04.012. URL http://www.sciencedirect.com/science/article/pii/S1631070511000958.

[206] Yuanbo Zhang, Yan-Wen Tan, Horst L. Stormer, and Philip Kim. Experimental observation of the quantum hall effect and berry's phase in graphene. *Nature*, 438:201–204, November 2005. doi: 10.1038/nature04235. URL https://www.nature.com/articles/nature04235.

[207] M. Mendoza, H. J. Herrmann, and S. Succi. Hydrodynamic model for conductivity in graphene. *Scientific Reports*, 3:1052, December 2013. doi: 10.1038/srep01052. URL http://www.nature.com/articles/srep01052.

[208] Derek Y. H. Ho, Indra Yudhistira, Nilotpal Chakraborty, and Shaffique Adam. Microscopic theory for electron hydroydnamics in monolayer and bilayer graphene. *Arxiv*, October 2017. URL https://arxiv.org/abs/1710.10272v1. arXiv:1710.10272v1 [cond-mat.mes-hall].

[209] Derek Y. H. Ho, Indra Yudhistira, Nilotpal Chakraborty, and Shaffique Adam. Theoretical determination of hydrodynamic window in monolayer and bilayer graphene from scattering rates. *Phys. Rev. B*, 97:121404, Mar 2018. doi: 10.1103/PhysRevB.97.121404. URL https://link.aps.org/doi/10.1103/PhysRevB.97.121404.

[210] Jesse Crossno, Xiaomeng Liu, Thomas A. Ohki, Philip Kim, and Kin Chung Fong. Development of high frequency and wide bandwidth johnson noise thermometry. *Applied Physics Letters*, 106(2):023121, 2015. doi: 10.1063/1.4905926. URL https://doi.org/10.1063/1.4905926.

Thank you Mom and Dad! This thesis would not have been possible without you. Thank you for being alongside me every step of the way and cheering for me. The above image was created using nanotechnology. This thesis was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. A template that can be used to format a PhD dissertation with this look & feel has been released under the permissive AGPL license, and can be found online at github.com/suchow/Dissertate or from its lead author, Jordan Suchow, at suchow@post.harvard.edu.