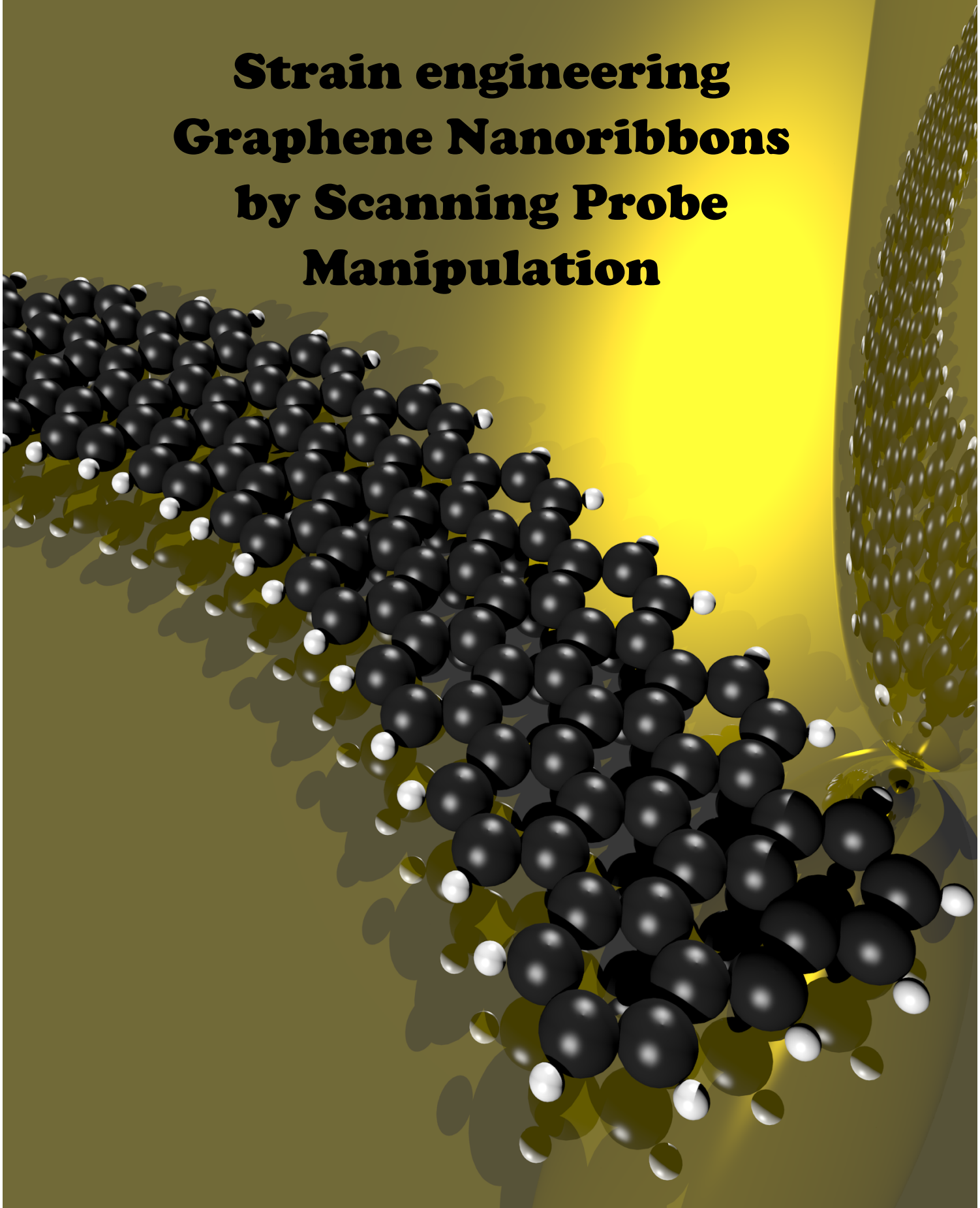
Peter Jacobse

Strain engineering Graphene Nanoribbons by Scanning Probe Manipulation



Strain engineering Graphene Nanoribbons by Scanning Probe Manipulation

Peter Herman Jacobse
peter.jacobse@gmail.com
Master Thesis
Nanomaterials: Chemistry and Physics
Utrecht University
Supervised by Drs. J. van der Lit
and Dr. I. Swart
August 2014, Utrecht

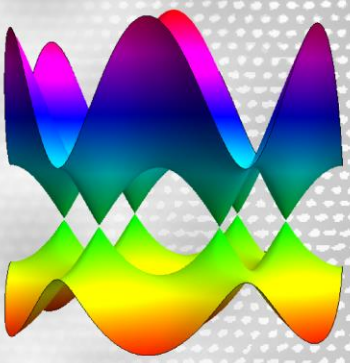
Universiteit Utrecht

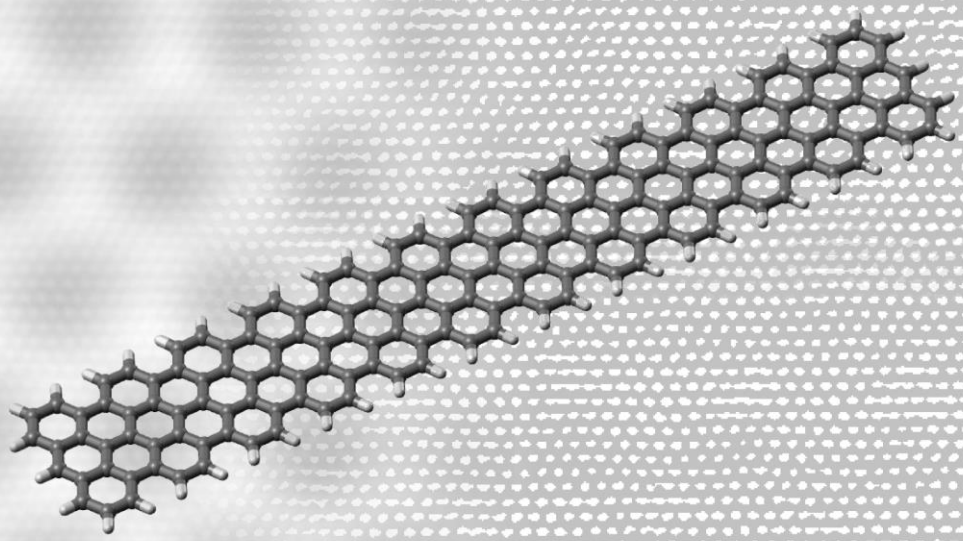
Contents

Abstract

The twodimensional material graphene features some unprecedented conductance properties, and is shaping up to become an important material in nanoelectronics. However, electronic components like transistors and diodes require a finite band gap to function; something that is absent in graphene. An approach to resolve this is to confine graphene into narrow strips called graphene nanoribbons. It has been found that strain and deformations can affect the electronic structure of graphene. Despite some theoretical efforts, it is not yet known experimentally what the impact of deformations on graphene nanoribbons is. We have used a scanning tunneling microscope to bend nanoribbons and find that the band gap of nanoribbons slightly diminishes with curvature, by approximately 1.5 % nm/^{o}.

Introduction





The recent hype on graphene should come as no surprise, given its many unique properties. Its characteristics like mechanical strength, electrical conductivity and impermeability - to name but a few - are unlike any materials discovered so far. Even before its discovery, many theoreticians have made promising predictions on the electronic properties of graphene, many of them verified after the first succesful synthesis by Geim and Novoselov[1].

Graphene is a single atom thick honeycomb-shaped network of carbon atoms. The essentially twodimensional material has only been discovered recently, but its supposed applicability in and impact on electronics industry has generated a huge incentive to perform extensive research[2]. One of the most important properties is graphenes extremely good electrical conduction, arising from its high charge carrier mobility.

Recently, the investigations on graphene extended to graphene nanosystems - structures with the same honeycomb building block, but with finite dimensions as opposed to the essentially twodimensional parent material. One of these is the graphene nanoribbon (GNR), a narrow strip of graphene, which is the main focus of this research. The main driving force for this is the opening of a band gap[3]. Similarly, some researches have investigated how a band gap can open up by straining graphene sheets. The influence of bending (a specific type of strain) on the electronic structure of graphene nanoribbons has not been been adressed experimentally yet.

GRAPHENE

The first isolation of graphene has been performed by peeling off some flakes from a piece of graphite (the material in pencil lead) with scotch tape multiple times - a method called the scotch tape method, or mechanical exfoliation[4]. The properties of graphene only arise after being free from the bulk graphite. Not long after that, it was discovered that chemical vapor deposition (CVD) can also be used for the growth of graphene[5]. Even though samples produced by the scotch tape method are often superior in quality than those grown by CVD[6], there is a large incentive to use CVD because of scaling issues[1][7].

Graphene is a material that features a very high conductivity[1]. This is mainly due to its unique atomic structure - a flat honeycomb lattice of carbon atoms - and the resulting electronic structure. It turns out that the band structure of graphene consists of a valence band and conductance band that exactly touch at the Fermi level, like two pencil tips. This makes graphene a zero-gap semiconductor, with vanishing density of states (DOS) at the Fermi energy. The DOS is furthermore linear at either side of the Fermi level. These features will become clear in the next chapter, where electronic structure calculations are carried out. A linear dispersion means that charge carriers in graphene can be thought to behave pretty much as photons! This translates to an exceedingly high conductivity. In fact, charge carriers are easily propelled up to relativistic speeds, and once graphene starts to be incorporated in nanoelectronics, this property will undoubtably play an important role in ultrafast electronics[1][8].

The electronic properties of graphene may be approximated in simple cases by the Drude model, some aspects of which are described in appendix A. The model describes diffusive electron transport, and is capable of yielding realistic values for resistivity and electron drift velocity. Within this framework, the conductivity of a material is proportional to its concentration of charge carriers, and has the material-specific mobility as a proportionality factor. The higher the mobility, the easier it is for electrons and holes to flow through the material. Geim and Novoselov found that the electron mobility in graphene was over $1.5 \cdot 10^4$ cm^2 V^{-1} s^{-1} [1], and values in excess of 10^5 cm^2 V^{-1} s^{-1} have already been reported[9] - close to the theoretical room temperature limit[10][11] and often higher than the best bulk metallic conductors[1]. The hole mobility was observed to be virtually the same as the electron mobility, which is not unexpected considering graphene's nearly symmetric dispersion around the Fermi energy. The room temperature carrier density is approximately 10^{12} cm^2[10]. Multiplying this charge carrier density with the experimental mobility already gives a value of the conductivity that is much larger than bulk metallic conductors. More transport properties of graphene can be appreciated in the very complete work of Das Sarma and coworkers[12].

To optimize the conductance it would be beneficial to have a nonzero density of states at the Fermi level, so that the charge carrier density would be higher. This can be done by shifting the Fermi energy, by means of introducing extrinsic carriers, or doping the material. Considering the fact that the Fermi level is determined by charge neutrality, introduction of donor impurities shifts the Fermi level up in the band structure, and vice versa for acceptor impurities. But in addition to affecting n_h and n_e, doping also affects μ_h and μ_e. The introduction of extrinsic carriers introduces scatterers, and the result is a reduced mean free path and compromised mobility[13][14].

An even more important feature is that the implementation of semiconductors in nanoelectronic components like transistors and diodes, requires a finite band gap[3][12]. Different approaches exist to open up the band gap in graphene. One of these is co-doping. As n-doping shifts the Fermi energy away from the VB and p-doping shifts it away from the CB, the net effect of co-doping donors and acceptors is to shift the bands apart. Co-doping of nitrogen and boron has been the subject of

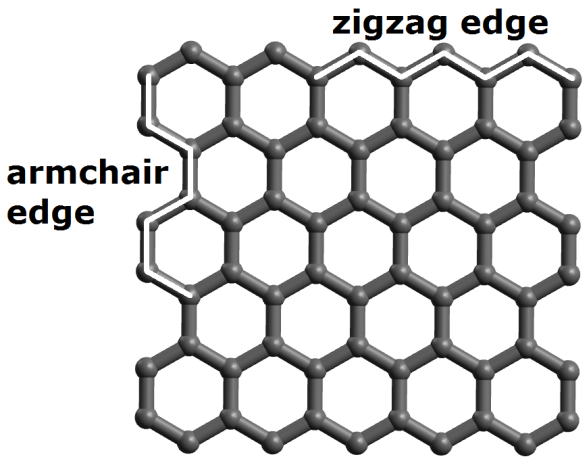


FIG. 1. The different edge structures - armchair and zigzag - that are encountered in graphene nanosystems. Hydrogen atoms are not shown.

theoretical[15] and experimental[16][17] research. Hexagonal boron nitride (hBN) can be seen as the most extreme case of co-doping, in which the total graphene A-sublattice has been exchanged with nitrogen atoms and the B-sublattice has been exchanged with boron atoms. hBN exhibits a significant band gap, but at the same time a low conductivity[18][19][20].

A second approach to generating narrow stripes of graphene, known as graphene nanoribbons (GNRs). These structures are "quantum confined" in one direction with respect to the parent material graphene, and depending on the width and cutting direction, the band gap can become nonzero. The conductance properties are again limited by scatterers. But as long as the nanoribbon edges are smooth enough, the scattering problem that hampered the co-doping method is absent in the ribbon method[14].

GRAPHENE NANORIBBONS

Many approaches have been used to synthesize different kinds of GNRs, most notably using laser cutting and chemical unzipping of carbon nanotubes[21][22]. Even scanning tunneling lithography has been used, where the current between tip and substrate cuts away carbon atoms[23]. A drawback of these methods is the lack of control over the precise direction of the nanoribbon, and the atomic defects in the edge, that give rise to scattering of free charge carriers[24]. In general, a graphene nanosystem can have zigzag edges and armchair edges, as shown in figure 1. The "cutting" direction has a large influence on the band structure of nanoribbons[25].

An alternative approach to synthesize nanoribbons has been used by Cai and coworkers[26]. They used an on-surface, metal-catalyzed polymerization reaction in ultra-high vacuum (UHV) conditions, utilizing 10,10'-dibromo-9,9'-bianthracene as a precursor. Their reaction allowed the production of atomically well-defined armchair graphene nanoribbons (AGNRs) with a width of seven atoms [27] (abbreviated 7-AGNR). The difference between this method and the methods proposed previously is the bottom-up fabrication - building up from molecular building blocks as opposed to breaking down graphene and carbon nanotubes. Throughout this research, the synthesis by Cai has been used to manufacture AGNRs. The resulting nanoribbons have a structure as displayed in figure 2.

Graphene nanoribbons, both the bottom-up and the top-down versions, have already made it to the first nanoelectronic components[28][29][30]. They are indeed only beginning to fulfill their promise as ultra-conductive finite gap semiconductors.

GRAPHENE SUBJECT TO STRESS

Now we want to look what happens when we take a patch of graphene, and start pulling on it. This introduces stress, and the graphene will deform accordingly. The strain in the pulling direction and transverse direction as a result of the applied stress are described by the Young's modulus and the Poisson ratio[31]. These can be related back to the elastic constants between the atoms in graphene[32][33]. This is an important point that I will come back to later in this thesis, but for now it suffices to note that graphene - like all materials - experiences some deformation when a stress is applied. Next, we can observe what happens to the electronic structure. In aromatic structures, significant changes in the electronic structure can in general be brought about by a simple conformational change[34]. We may also suspect that something will happen to the zero band gap in graphene. In graphene, strain reduces the lattice symmetry. The symmetry of the reciprocal space is altered accordingly, meaning that the dirac cones are also shifted

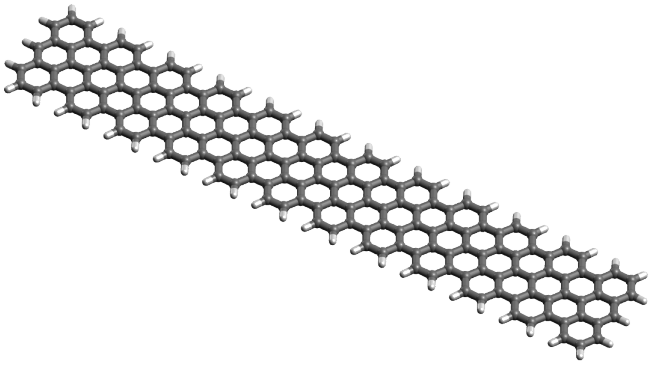


FIG. 2. A model of a 7-atom wide armchair graphene nanoribbon (7-AGNR).

from their positions[8]. It has been found theoretically that a band gap can open up after sufficient deformation, with threshold values for the applied tensile strain ranging from 0 %[35][36] to in excess of 20 %[37]. Transport characteristics are also altered[38]. Similar shifts of the high-symmetry point in the reciprocal space are reported in confined systems like carbon nanotubes[39] and graphene nanoribbons[40].

We can conclude that the band gap of graphene is tightly linked to its geometry, and can be tuned by confining the material or straining it. Now we would like to ask the question what would happen if we were to combine both quantum confinement and strain engineering. For graphene nanoribbons, several theoretical efforts have been carried out, for different kinds of deformations.

The geometry of nanoribbon deformations can be simulated using an atomistic molecular dynamics (or molecular mechanics) method. For example, Neek-Amal and Peeters used this method to simulate the out-of-plane buckling of graphene nanoribbons subject to compressive stress in the axial direction[41]. In general, molecular mechanics is capable of generating reasonable geometries for graphene and graphene nanosystems under strain, as long as the correct elastic parameters are used[31][32][33].

Using such a rippled nanoribbon as an input for a tight binding calculation, Costamagna and coworkers were able to derive a flattening of the band structure and an increase of the band gap with increasing rippling amplitude[42]. Using a non-equilibrium Green's function formalism (NEGF, a popular tool in ballistic transport calculations) they also calculated that the low-bias conductance should vanish, even for metallic ribbons, as the distortion is induced.

Li and coworkers considered both tensile and shear strain on nanoribbons using density functional theory (DFT) calculations[40]. Although the shear strain only brought about very mild variations in the band gap, the uniaxial strain had a very large effect. The simulations showed a shifting of the band gap in a zigzag pattern with approximately 100 meV per percent strain for armchair ribbons of all widths. The results for tensile strain were consistent with earlier work by Sun *et al.*, who used both DFT and tight binding[43], and Lu *et al.* [44].

Koskinen looked at a different kind of deformation, namely the in-plane bending of graphene nanoribbons[45]. Modeling the geometry itself with objective molecular dynamics[46] in large circular geometries. A tight-binding calculation was performed on a unit cell which was only a small part of this hoop. The electronic structure calculations were performed within the framework of density functional theory-based tight binding (DFTB), and could later be reproduced well with a nearest-neighbour tight binding model. He predicted the band gap for 7-AGNR to slightly decrease as a result of effective stretching of the covalent bonds in the molecule. The decrease is quadratic in the reciprocal of the radius of curvature, and gives band gap shifts of the order of 50 meV for a curvature of 10 ^{o} nm^{-1}.

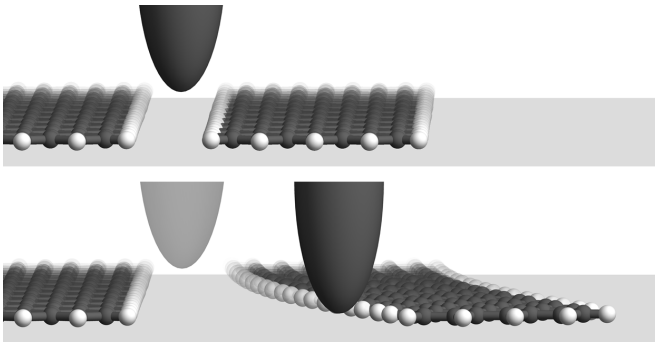


FIG. 3. A model of a nanoribbon that is bent with the tip of a scanning tunneling microscope

Despite the richness in theories, experimental work on straining graphene nanoribbons is still relatively unexplored territory. An experiment that approaches the desired bending of nanoribbons is the lifting of 7-AGNR off a metal surface with an STM tip by Koch *et al.* [47]. First, they carried out electronic structure calculations to derive a density of states. Pseudoballistic transport was modelled in terms of a transmission decay length β, which steps down as resonances enter the bias window. Then by pulling up nanoribbons, they found exponentially decaying transmission currents at different bias voltages, the values of which correspond well with their predicted values for $\beta(E)$. The out-of-plane bend that the nanoribbon makes to achieve its thermodynamic lowest state is claimed to lower the conductance somewhat, but its effect on the electronic structure was not established.

We are capable of manufacturing atomically precise 7-AGNR using the method of Cai *et al.* [26]. Using our scanning tunneling microscope, we can carefully monitor all electronic properties. These atomically well-defined systems have already been characterized well[48][49], so they provide an excellent system to investigate the impact of strain on electronic structure. In order to bring about strain, we would like to physically bend a ribbon over the surface with the tip of the scanning tunneling microscope, like in the model of figure 3.

There are two main reasons why strain in graphene nanoribbons is an important direction in GNR research. First, when GNR will be applied in nanoelectronics, they will inevitably experience some form of bending or strain. Second, it will give insight into the laws governing the relation between strain and electronic structure.

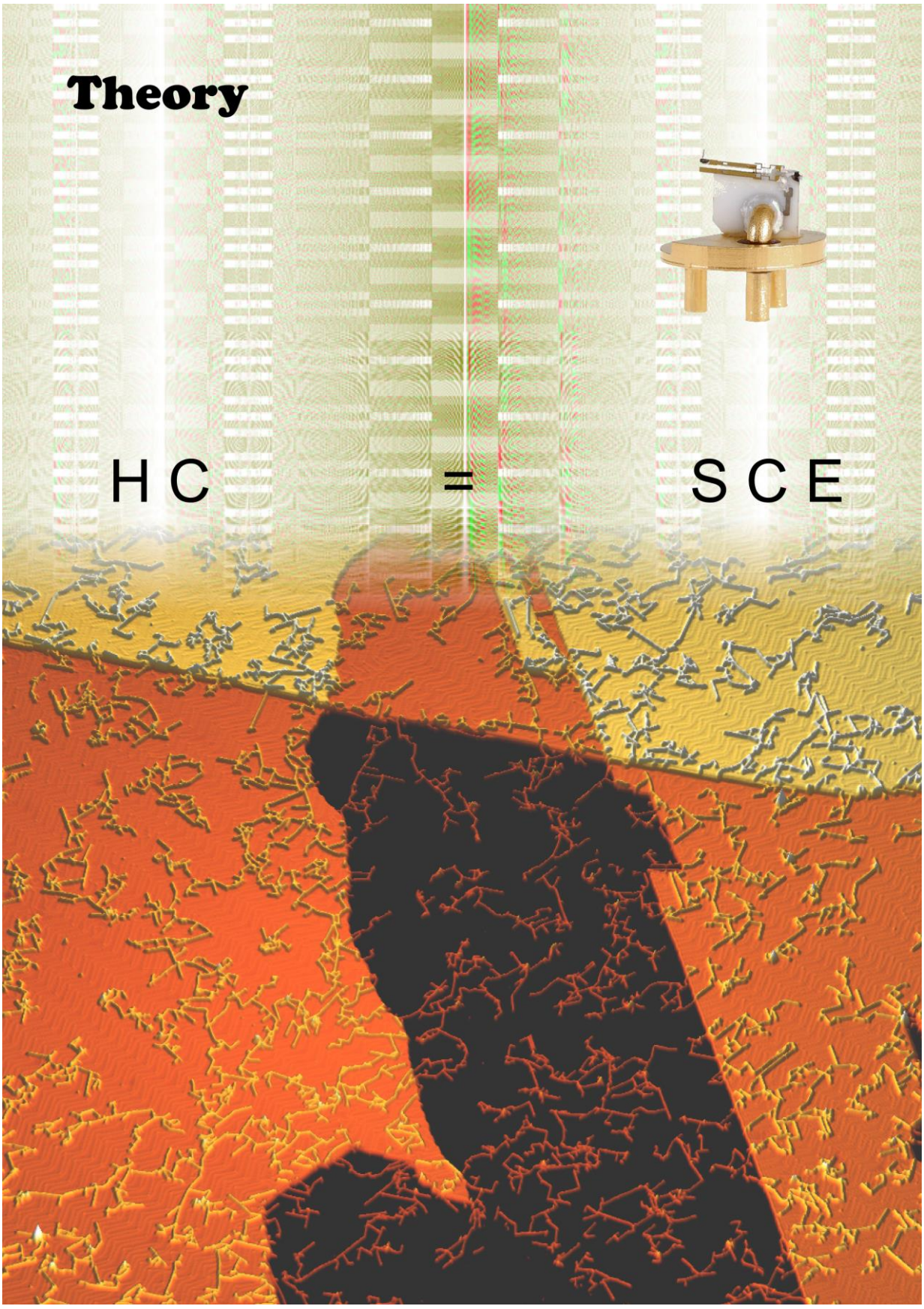
For this second reason, we would also like to include a theoretical part to the research. Therefore, tight binding calculations will be performed on graphene nanoribbons. Although not as accurate as (post) Hartree Fock (HF) or density functional (HF) methods, tight binding is encountered frequently in literature for electronic structure calculations on graphene and graphene

nanosystems[8][50]. This is because it is a computationally inexpensive and intuitively simple model, that still happens to give good results on graphene and its nanosystems. To simulate the deformations that we can bring about with the tip of a scanning tunneling microscope, a molecular mechanics method is used, similar to that used by Neek-Amal and Peeters [41].

In summary, this research will focus on the following questions: can we controllably bend graphene nanoribbons using SPM methods? Can we probe changes in the electronic structure? Can we generate a satisfying theoretical framework for the situation? And can we relate any possible changes to theoretical predictions?

[1] A. K. Geim and K. S. Novoselov, Nature materials **6**, 183 (2007).
[2] A. K. Geim, Science (New York, N.Y.) **324**, 1530 (2009).
[3] F. Schwierz, Nature Nanotechnology **5**, 487 (2010).
[4] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and a. a. Firsov, Science (New York, N.Y.) **306**, 666 (2004).
[5] A. Nagashima, K. Nuka, H. Itoh, and T. Ichinokawa, Surface Science **291**, 93 (1993).
[6] A. Venugopal, J. Chan, X. Li, C. W. Magnuson, W. P. Kirk, L. Colombo, R. S. Ruoff, and E. M. Vogel, Journal of Applied Physics **109**, (2011).
[7] H. K. Yu, K. Balasubramanian, K. Kim, J.-L. Lee, M. Maiti, C. Ropers, J. Krieg, K. Kern, and A. M. Wodtke, ACS nano , 0 (2014).
[8] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).
[9] K. I. Bolotin, K. J. Sikes, Z. Jiang, M. Klima, G. Fudenberg, J. Hone, P. Kim, and H. L. Stormer, Solid State Communications **146**, 351 (2008).
[10] E. V. Castro, H. Ochoa, M. I. Katsnelson, R. V. Gorbachev, D. C. Elias, K. S. Novoselov, A. K. Geim, and F. Guinea, Phys. Rev. Lett. **105**, 266601 (2010).
[11] A. Akturk and N. Goldsman, Journal of Applied Physics **103**, 053702 (2008).
[12] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, Reviews of Modern Physics **83**, 407 (2011).
[13] J.-H. Chen, C. Jang, S. Adam, M. S. Fuhrer, E. D. Williams, and M. Ishigami, Nature Physics **4**, 377 (2008), 0708.2408 [cond-mat.other].
[14] X. Wang, Y. Ouyang, L. Jiao, H. Wang, L. Xie, J. Wu, J. Guo, and H. Dai, Nature Nanotechnology **6**, 563 (2011).
[15] S. Yu, W. Zheng, C. Wang, and Q. Jiang, ACS nano **4**, 7619 (2010).
[16] Z. Jin, J. Hao, C. Kittrell, and J. M. Tour, ACS Nano , 4112 (2011).
[17] R. Lv, Q. Li, A. R. Botello-Méndez, T. Hayashi, B. Wang, A. Berkdemir, Q. Hao, A. L. Elías, R. Cruz-Silva, H. R. Gutiérrez, Y. A. Kim, H. Muramatsu, J. Zhu, M. Endo, H. Terrones, J.-C. Charlier, M. Pan, and M. Terrones, Scientific reports **2**, 586 (2012).
[18] A. Abdellaoui, A. Bath, B. Bouchikhi, and O. Baehr, Materials Science and Engineering **47**, 257 (1997).
[19] V. Solozhenko, A. Lazarenko, J.-P. Petitet, and A. Kanaev, Journal of Physics and Chemistry of Solids **62**, 1331 (2001).
[20] N. Ooi, a. Rairkar, L. Lindsley, and J. B. Adams, Journal of Physics: Condensed Matter **18**, 97 (2006).
[21] M. Terrones, A. R. Botello-Méndez, J. Campos-Delgado, F. López-Urías, Y. I. Vega-Cantú, F. J. Rodríguez-Macías, A. L. Elías, E. Muñoz Sandoval, A. G. Cano-Márquez, J.-C. Charlier, and H. Terrones, Nano Today **5**, 351 (2010).
[22] X. Wang, Y. Ouyang, L. Jiao, H. Wang, L. Xie, J. Wu, J. Guo, and H. Dai, Nature nanotechnology **6**, 563 (2011).
[23] L. Tapasztó, G. Dobrik, P. Lambin, and L. P. Biró, Nature Nanotechnology **3**, 397 (2008).
[24] H. Xu and T. Heinzel, Journal of physics. Condensed matter : an Institute of Physics journal **24**, 455303 (2012).
[25] V. Barone, O. Hod, and G. E. Scuseria, Nano letters **6**, 2748 (2006).
[26] J. Cai, P. Ruffieux, R. Jaafar, M. Bieri, T. Braun, S. Blankenburg, M. Muoth, A. P. Seitsonen, M. Saleh, X. Feng, K. Müllen, and R. Fasel, Nature **466**, 470 (2010).
[27] The synthesis is described in appendix E.
[28] J. G. Son, M. Son, K.-J. Moon, B. H. Lee, J.-M. Myoung, M. S. Strano, M.-H. Ham, and C. a. Ross, Advanced materials **25**, 4723 (2013).
[29] Y. Ye, L. Gan, L. Dai, H. Meng, F. Wei, Y. Dai, Z. Shi, B. Yu, X. Guo, and G. Qin, Journal of Materials Chemistry **21**, 11760 (2011).
[30] P. B. Bennett, Z. Pedramrazi, A. Madani, Y.-C. Chen, D. G. de Oteyza, C. Chen, F. R. Fischer, M. F. Crommie, and J. Bokor, Applied Physics Letters **103**, 253114 (2013).
[31] F. Scarpa, S. Adhikari, and a. Srikantha Phani, Nanotechnology **20**, 065709 (2009).
[32] P. Zhao and G. Shi, Tech Science Press **5**, 49 (2011).
[33] K. Alzebdeh, International Journal of Mechanics and Materials in Design **8**, 269 (2012).
[34] R. Kroemer and K. Liedl, Journal of the American Chemical Society , 12573 (1998).
[35] G. Gui, J. Li, and J. Zhong, Phys. Rev. B **78**, 075435 (2008).
[36] Z. H. Ni, T. Yu, Y. H. Lu, Y. Y. Wang, Y. P. Feng, and Z. X. Shen, ACS Nano **2**, 2301 (2008).
[37] V. Pereira, a. Castro Neto, and N. Peres, Physical Review B **80**, 045401 (2009).
[38] Y. Xu, H. Gao, M. Li, Z. Guo, H. Chen, Z. Jin, and B. Yu, Nanotechnology **22**, 365202 (2011).
[39] L. Yang and J. Han, Phys. Rev. Lett. **85**, 154 (2000).
[40] Y. Li, X. Jiang, Z. Liu, and Z. Liu, Nano Research **3**, 545 (2010).
[41] M. Neek-Amal and F. M. Peeters, Physical Review B **82**, 085432 (2010).
[42] S. Costamagna, O. Hernandez, and a. Dobry, Physical Review B **81**, 115421 (2010).
[43] L. Sun, Q. Li, H. Ren, H. Su, Q. W. Shi, and J. Yang, The Journal of chemical physics **129**, 074704 (2008).
[44] Y. Lu and J. Guo, Nano Research **3**, 189 (2010).
[45] P. Koskinen, Physical Review B **85**, 205429 (2012).
[46] T. Dumitric and R. D. James, Journal of the Mechanics and Physics of Solids **55**, 2206 (2007).
[47] M. Koch, F. Ample, C. Joachim, and L. Grill, Nature

nanotechnology **7**, 713 (2012).
[48] P. Ruffieux, J. Cai, N. C. Plumb, L. Patthey, D. Prezzi, A. Ferretti, E. Molinari, X. Feng, K. Müllen, C. A. Pignedoli, and R. Fasel, ACS Nano **6**, 6930 (2012).
[49] J. van der Lit, M. P. Boneschanscher, D. Vanmaekelbergh, M. Ijäs, A. Uppstu, M. Ervasti, A. Harju, P. Liljeroth, and I. Swart, Nature Communications **4**.
[50] K. Wakabayashi, K.-i. Sasaki, T. Nakanishi, and T. Enoki, Science and Technology of Advanced Materials **11**, 054504 (2010).


Theory
H C = S C E

QUANTUM THEORY

A quantum system is fully described by its wavefunction, which is a solution to the Schrödinger equation (SE). As the SE cannot be solved analytically for systems with more than two mutually interacting particles, a number of approximations and assumptions should be made. Different approximation schemes have been used in the fields of solid state physics and quantum chemistry to deal with this problem in the most efficient way. It turns out that atomistic tight binding is a perfect tool for performing electronic structure calculations on graphene nanosystems. It can furthermore be combined with molecular mechanics, which allows the simulation of strained nanoribbon geometries.

All measurements in this research have been performed using scanning probe techniques. In scanning probe microscopy (SPM) an extremely sharp metal tip is brought into close proximity of a flat surface. The surface is then scanned, line by line. In scanning tunneling microscopy, the tunnel current (the current that flows through the vacuum gap) is recorded[1], whereas in atomic atomic force microscopy, the force on the tip is measured[2]. Scanning probe microscopy is one of the few ways to atomically resolve the structure of a surface, or structures lying on top of a surface. Techniques based on photonic imaging can not provide atomic resolution as this is far below the diffraction limit. Techniques like scanning electron microscopy (SEM) and transmission electron microscopy (TEM) use energetic electrons, rather than photons, and can achieve much higher resolution. However, the resolution is still larger than what can be achieved in SPM.

Scanning probe techniques have the added advantage that they allow spectroscopic measurements to be performed. Apart from their imaging capabilities, they allow the possibility of measuring observables like the local density of states and tip-sample interactions. Furthermore, the SPM tip can be used to controllably manipulate atoms and molecules on a surface. This chapter tries to explain a couple of the powerful capacities of SPM after carefully explaining the background of the techniques.

Scattering of free electrons

As derived in appendix B, the wave function of an electron with energy E in a constant potential $V < E$ is given by

$$\psi(\mathbf{r},t) = c_1 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + c_2 e^{i(-\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

with c_1 and c_2 the amplitudes of the complex electron waves, ω the angular frequency and $\mathbf{k}$ the wave vector, the magnitude of which is determined by the parabolic

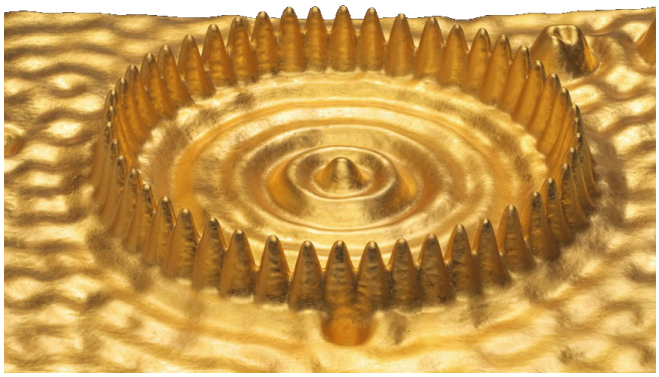


FIG. 1. "The Well (Quantum Corral)", by artist Julian Voss-Andreae, exhibiting data from Crommie, Lutz and Eigler[3], who used scanning tunneling microscopy to make the confined standing electron waves visible.

dispersion relation

$$E = V + \frac{\hbar^2 k^2}{2m} \;\Leftrightarrow\; k^2 = \frac{2m(E-V)}{\hbar^2}$$

An electron scattering elastically on an interface of the potential well to the vacuum will generate a standing wave as a result of the interference of the incident and reflected parts of the envelope. For the sake of argument, let's imagine a surface at $x = 0$, with a constant potential on the right and an infinite potential on the left. ψ cannot penetrate into an infinite potential, but should be continuous, so that there is a boundary condition $\psi \to 0$ at the surface. This is ensured by setting

$$\psi_{inc} = ce^{i(kx-\omega t)} \qquad \psi_{refl} = -ce^{i(-kx-\omega t)}$$

$$\psi = \psi_{inc} + \psi_{refl} = c\sin(kx)e^{-i\omega t}$$

In general, electrons form standing waves around potential energy humps and "walls". This is the basis of so-called surface/edge/end/Tamm/Schottky states and defect/scattering states, the former of which will be encountered in the case of graphene nanoribbons.

Quantum confinement

If a second wall is added at $x = x_0$, the same argument can be invoked to find a standing wave solution for the SE

$$\psi = c\sin(k(x-x_0))e^{-i\omega t}$$

For these wave functions to interconnect, the condition

$$kx = k(x-x_0) + n\pi \Rightarrow k = \frac{n\pi}{x_0}$$

should be satisfied for integer values of n. The constructed wave functions should be orthonormal, meaning that their overlap integral evaluates to one for a

single wave function product and to zero for a product of wave functions with different quantum numbers ($\langle\psi_i|\psi_j\rangle = \delta_{ij}$). So in general, the wave functions of a small system are standing waves that can be thought of as superpositions of plane waves bouncing back and forth. A beautiful experiment that visualized the formation of these standing-wave states was carried out by Crommie, Lutz and Eigler[3], where they could carefully position iron atoms in a ring on a copper surface with a scanning tunneling microscope (STM). The standing electron waves result from scattering from the potential energy humps that are cast by the atoms in the ring. An STM scan reveals the the circular wave pattern, as seen in figure 1.

A standing wave of the form $\psi = c\sin(kx)e^{-i\omega t}$ is also an eigenfunction of the constant-potential Hamiltonian. Its eigenenergy is identical to the plane wave energy, featuring the same quadratic dispersion. As long as the electron is in a constant potential, its dispersion is quadratic, but k is quantized to the values $k = n\pi/x_0$ (for integer values of n). The lowest energy is achieved for setting up a wave with a single node: $n = 1$.

Now if N electrons reside in the volume between the two boundaries, and their mutual repulsion is ignored, they will pair up and decay to the lowest $N/2$ energy levels. The Fermi wave number $k_F = N\pi/(2x_0)$ is the wave number of the highest energy electrons, and the corresponding energy $E_F = V + \hbar^2 k_F^2/(2m)$ is known as the Fermi energy. For any energy we have

$$E - V = \frac{\hbar^2 k_F^2}{2m} = \frac{\hbar^2 N^2 \pi^2}{8mx_0^2}$$

The density of states - the number of states per unit energy interval - is found straightforwardly by differentiating N with respect to E.

$$\mathrm{DOS}(E) = \frac{\partial N}{\partial E} = \frac{\partial}{\partial E}\frac{2\sqrt{2m}x_0}{\hbar\pi}\sqrt{E-V}$$

$$= \frac{\sqrt{2m}x_0}{\hbar\pi\sqrt{E-V}}$$

The density of states of a one-dimensional system is thus a function that is strongly peaked (contains a singularity) at $E = V$, and decays toward zero for higher energies.

The energy difference between the nth and the $n+1$th eigenstate is

$$\Delta E = \frac{\hbar^2\pi^2[(n+1)^2 - n^2]}{2mx_0^2} = \frac{\hbar^2\pi^2[2n+1]}{2mx_0^2}$$

$$\approx \frac{n\hbar^2\pi^2}{mx_0^2}$$

So when x_0 is made small, the energy differences between the levels increase. This effect is called quantum confinement. A visual example of this effect can be seen in the

FIG. 2. A quantum dot rainbow. Each subsequent vial contains quantum dots with increasing particle size, decreasing HOMO-LUMO gap and increasing the fluorescence emission wavelength. Courtesy of prof. Andries Meijerink, Condensed Matter and Interfaces, Utrecht University.

quantum dots displayed in figure 2, where the increasing particle size in the subsequent vials cause the fluorescence emission wavelength to shift to larger valued. To examine this confinement somewhat more in detail, we compare a wire of material of length x_0 and a wire of the same material of length $2x_0$. The energy difference between the nth and $n+1$th eigenstate is obviously 4 times larger for the smallest wire. However, the number of electrons that go inside is proportional to the length of the wire,

$$n_F = \frac{\lambda}{2}x_0$$

with λ the number of electrons per unit length of the wire. The band gap is the energy difference between n_F and $n_F + 1$, so

$$\Delta E \approx \frac{n_F\hbar^2\pi^2}{mx_0^2} \approx \frac{\lambda\hbar^2\pi^2}{2mx_0}$$

This is another important result, that will be used in the simulations in the next chapter. As long as there is a quadratic dispersion, the band gap energy should go up as x_0^{-1} upon confining a system.

Now it is time to look what happens to the energy levels of a free-electron system upon placing additional boundaries in the y-dimension, at $y = 0$ and $y = y_0$. The time-independent wave equation that satisfies the boundary scattering condition becomes

$$\psi = ce^{i\mathbf{k}\cdot\mathbf{r}} - ce^{-i\mathbf{k}\cdot\mathbf{r}} = c\sin k_x x \sin k_y y$$

with $\mathbf{k}$ quantized as

$$\mathbf{k} = \pi\left(\frac{n_x}{x_0}\hat{\mathbf{x}} + \frac{n_y}{y_0}\hat{\mathbf{y}}\right)$$

where n_x, and n_y are positive integers. As long as the dispersion relation is parabolic, the general solution for few-dimensional systems scale as

$$\mathrm{DOS}_{1D}(E) = V_0 + \frac{c}{\sqrt{E}}, \quad \mathrm{DOS}_{2D}(E) = V_0 + c$$

$$\mathrm{DOS}_{3D}(E) = V_0 + c\sqrt{E}$$

where the constants like mass, Planck's constant and potential energy are absorbed into the constants c. When the background potential is not constant anymore, there is the possibility of the formation of multiple bands. These will all show up individually in the density of states, at different characteristic energy onsets. Another point is that the density of states is particularly peaked for the one-dimensional case. The point where DOS goes to infinity, is known as a van Hove-singularity[4].

Post free-electron models

The problems that we are interested in - the electronic structure of graphene and graphene nanoribbons - are many-body problems, with an inhomogeneous potential energy background cast by the nuclei and electron-electron repulsion. The free-electron model always returns a parabolic dispersion and can not give any material-specific structure in the density of states. It will be shown that the free-electron model is really poor in the case of graphene and graphene nanoribbons, and it cannot account for their excellent transport properties. This problem calls for refined models.

The wavefunction of the molecule is inherently a function of the coordinates of all electrons and all nuclei. The Hamiltonian operator contains contributions from the internuclear repulsion, electron-nuclear attraction, kinetic energy and electron-electron repulsion

$$\hat{\mathrm{H}} = \hat{\mathrm{V}}_{\mathrm{Ne}} - \frac{\hbar^2}{2m}\nabla^2 + \hat{\mathrm{V}}_{\mathrm{ee}} + \mathrm{V}_{\mathrm{NN}}$$

Obviously, it is impossible to solve the SE for thousands of particles. Therefore, the following common approximations and simplifications will be made.

- The nuclei are considered stationary on the timescale of electronic motion (Born-Oppenheimer (BO) approximation). The BO approximation decouples the nuclear wave function from the electronic wave function.
- The electronic wave function is treated as a product of single-electron wave functions (independent particle model).
- Electron-electron repulsion is ignored.

In the nearly-free electron model, the constant potential background V is replaced by a perturbed potential $V(\mathbf{r})$ that varies across the unit cell. As shown in appendix B, the introduction of Fourier expansions for the wave function and potential energy landscape yields the central equations.

$$\left(E_i - \frac{\hbar^2 k^2}{2m_e}\right)\psi(\mathbf{k}) = \sum_{\mathbf{Q}} V_{\mathbf{Q}}\psi(\mathbf{k}-\mathbf{Q})$$

where $\{\mathbf{Q}\}$ is the set of reciprocal lattice vectors. These equations are related to the secular equations shown later in this chapter.

The explicit introduction of the lattice may have a profound impact on the dispersion and density of states of a system. Later in this chapter, atomistic methods are used which also take the graphene lattice into account, and they are indeed found to give the correct dispersion for graphene. Nearly-free electron models can also give a band gap - something that does not exist in free electron models. Another way to resolve the featureless density of states of graphene is by taking relativistic effects into account. The relativistic Klein-Gordon equation is treated succinctly in appendix B, and also allows different regimes in the dispersion relation.

Electron-electron repulsion, exchange and correlation

One of the most severe assumptions in the free and nearly-free electron models is the complete negligence of electron-electron interaction. A crude model of electron repulsion is the mean field Coulomb repulsion. A single-electron wave function ϕ establishes a charge density

$$\rho(\mathbf{r}_1) = -e|\phi(\mathbf{r}_1)|^2$$

and the interaction of this charge density with an electron (2) is found by integrating over the entire distribution of electron (1)

$$E_{12} = e^2 \iiint \frac{\rho(\Delta\mathbf{r})d\mathbf{r}_1}{4\pi\epsilon_0|\Delta\mathbf{r}|}$$

$$= e^2 \iiint \frac{|\phi(\Delta\mathbf{r})|^2 d\mathbf{r}_1}{4\pi\epsilon_0|\Delta\mathbf{r}|}$$

This second electron is not a point charge, but is also diffuse. So the real mean-field repulsion is the integral over both distributions

$$E_{12} = e^2 \iiint\iiint \frac{|\phi_1(\Delta\mathbf{r})|^2|\phi_1(\Delta\mathbf{r})|^2 d\mathbf{r}_1 d\mathbf{r}_2}{4\pi\epsilon_0|\Delta\mathbf{r}|}$$

$$= \frac{e^2}{4\pi\epsilon_0}\left\langle \phi_1 \left| \frac{|\phi_2\rangle\langle\phi_2|}{|\Delta\mathbf{r}|} \right| \phi_1 \right\rangle$$

This so-called mean field approximation overestimates electron-electron interaction for the simple reason that electrons are found to move in a correlated, mutually avoiding fashion. This happens in such a way that the chance to find an individual electron at a certain point in space in space is still given by $|\phi|^2$, but the chance to find two electrons close together is smaller than $|\phi_1|^2|\phi_2|^2$. Repulsion and correlation have large consequences for computational effort in most models.

Another important point is that van der Waals attraction is basically a correlation effect. It is explained in terms of an induced-dipole induced-dipole attraction, where the dipoles are created through correlation between the two electrons[5][6]. Empirically, this gives an attractive energy between the two dipoles that scales as

$$E_{vdW} \propto r^{-6}$$

where r is the separation. This interaction happens to be very important in atomic force microscopy, as described in appendix G.

The total electronic wave function of a two-electron system may be represented as the product of single-electron wave functions

$$\Psi_e = |1(\mathbf{1}_\alpha)\rangle|1(\mathbf{2}_\beta)\rangle$$

that is, electron 1 with position **1** and spin α in orbital $|1\rangle$ and electron 2 with position **2** and spin β in orbital $|1\rangle$. However, electrons are indistinguishable, so that the wave function with interchanged spins is just as valid. The actual wave function is therefore a superposition of wave function products. Another apparently equivalent situation is encountered when the two particles are exchanged. By the Pauli exchange principle, the wave function should be antisymmetric for this operation

$$\Psi_e(1,2) = -\Psi_e(2,1)$$

so that the true wave function in the ground state is represented as a Slater determinant.

$$\Psi_e = \mathrm{N} \begin{vmatrix} |1(\mathbf{1}_\alpha)\rangle & |1(\mathbf{1}_\beta)\rangle \\ |1(\mathbf{2}_\alpha)\rangle & |1(\mathbf{2}_\beta)\rangle \end{vmatrix}$$

Eigenstates of the Hamiltonian should always be orthonormal, so that the normalization constant is $N = 1/\sqrt{n!}$, with n the number of single-electron wave functions. As a means of convenient shorthand notation, an antisymmetrization operator $\hat{\mathrm{A}}$ can be invoked from this point on to turn a many-orbital product into the antisymmetrized Slater determinant (where $\tilde{\Psi}$ is the ground state Slater matrix).

$$\hat{\mathrm{A}}\Psi = \frac{1}{\sqrt{n!}}\det\tilde{\Psi}$$

Operating a Hamiltonian with a mean-field electron-electron repulsion on a Slater determinant wave function automatically gives a Coulomb term and an exchange term.

Exchange lies on the basis of Pauli repulsion. If two electrons with equal spin are forced into the same region of space, they would render the Slater determinant 0, and the total wave function would not be normalizable. Therefore, electrons that are "pushed together" have to develop nodes in their wave function to remain orthonormal. This increases their energy. Although the functional form of Pauli repulsion strongly depends on the type of system, it is more or less proportional to electron density. Empirically, the Pauli repulsion energy is most often modeled as

$$E_P \propto r^{-12}$$

The Pauli interaction is the most important contribution in AFM contrast formation, as described in appendix G.

SCANNING TUNNELING MICROSCOPY

In STM, an atomically sharp tip can be raster scanned in the x, y and z-direction with great precision thanks to piezoelectric elements. The tip is biased and positioned close to a conductive sample - on the order of a few nm - to create a so-called tunnel barrier[1]. In the next section, the tunnel effect is explained. After that, it is explained how this effect is exploited in STM when performing measurements on graphene nanoribbons.

Tunneling

The difference between the vacuum energy and the Fermi energy is the work function.

$$V_{vac} - E_F = W$$

A shortcoming of the free-electron model up to now is that it thinks of the potential energy well of the solid as infinitely deep. But by definition of the work function, the depth of the well with respect to the vacuum is

$$V_{well} - V_{vac} = W + E_F$$

As shown in appendix B, the wave function for an electron in vacuum with too little energy to overcome the potential is

$$\psi = c_1 e^{\mathbf{k}\cdot\mathbf{r}} + c_2 e^{-\mathbf{k}\cdot\mathbf{r}}$$

As long as these exponential solutions are only part of the complete wavefunction, it may remain normalizable. This means that electrons may effectively spill over into the vacuum, past their classical turning points.

As an example we take a wall at $x = 0$ that separates the potential $V = 0$ on the left and $V = V > E$ on the right. The wave function is written by tying together the left and right solutions at $x = 0$ (omitting the exponentially rising solution as it is unnormalizable).

$$\psi_l = c_1 e^{ikx} + c_2 e^{-ikx} \qquad \psi_r = c_3 e^{-k'x}$$

By choosing $c_3 = c_1 + c_2$, the wave function becomes continuous.

$$\psi_l(0) = c_1 e^0 + c_2 e^0 = c_3 = c_3 e^0 = \psi_r(0)$$

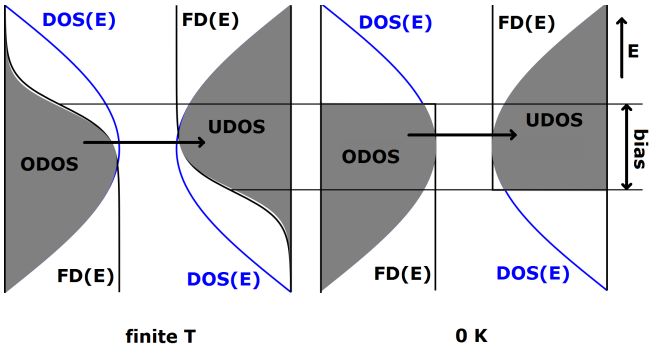


FIG. 3. A schematic energy diagram of a tunnel barrier at high temperature (broad Fermi-Dirac distribution) and low temperature (Fermi-Dirac distribution becomes step function). The black curves are the Fermi-Dirac functions, the blue curves the density of states, and the gray areas the occupied density of states and unoccupied density of states - the integral over FD(E)DOS(E)) and (1-FD(E))DOS(E), respectively. The bias potential is indicated on the right.

And momentum is conserved if $\langle p_x \rangle_l = \langle p_x \rangle_r$. This requires that the total function is smooth across the boundary, so that

$$\left[\frac{\partial \psi_l}{\partial x}\right]_0 = ikc_1 - ikc_2 = \left[\frac{\partial \psi_r}{\partial x}\right]_0 = -k'c_3$$

$$ik(c_1 - c_2) = -k'(c_1 + c_2) \Rightarrow \frac{c_2}{c_1} = \frac{ik + k'}{ik - k'}$$

If the vacuum is only a small gap between two allowed regions, the free-electron wavefunctions may connect through the vacuum. This means that an electron incident on a vacuum barrier has a finite probability to cross the gap. This phenomenon is the tunnel effect.

If we have a tunnel barrier, and a potential difference is applied on the two sides, electrons may preferably tunnel from the high energy side to the low energy side. So a net current from high side to low side will occur, the strength of which is dependent both on the size of the gap and the integral over the occupied density of states (ODOS) on the left to the unoccupied density of states on the right (UDOS). A schematic picture is shown in figure 3.

A tunnel barrier is modeled as a square potential energy hump between $x = 0$ and $x = x_0$ with $V > E$, so that $k_2 = (V - E)/(2m\hbar^2)$, where E is the energy of the electrons that are incident on the barrier. On either side the potential energy is $V < E$ so that plane waves incident on the barrier spill over. The wave vector on the left is $k_1 = (E - V)/(2m\hbar^2)$ and on the right is $k_3 = k_1$. The total wave function is defined in pieces as

$$\psi = \left\{ \begin{array}{c} \psi_1 = Ae^{ik_1 x} + Be^{-ik_1 x} \mid x < 0 \\ \psi_2 Ce^{-k_2 x} + De^{k_2 (x - x_0)} \mid 0 \leq x \leq x_0 \\ \psi_3 Ee^{ik_3 (x - x_0)} + Fe^{-ik_3 (x - x_0)} \mid x > x_0 \end{array} \right\}$$

Application of continuity and smoothness at 0 and x_0 easily leads to a set of four coupled equations, which are solved for the expansion coefficients as

$$\begin{pmatrix} -1 & -1 & 1 & u & 0 & 0 \\ -in_{12} & in_{12} & -1 & u & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & u & -1 & in_{32} & -in_{32} \\ 0 & 0 & u & 1 & -1 & -1 \end{pmatrix} \mathbf{c} = \mathbf{0}$$

with $u = e^{-k_2 x_0}$, $n_{12} = k_1/k_2$ and $n_{32} = k_3/k_2$, and where the trivial equations $\mathbf{0} \cdot \mathbf{c} = \mathbf{0} \cdot (A, B, C, D, E, F) = \mathbf{0}$ are added to make the matrix square. The allowed $\mathbf{c}$ thus span the nullspace of this boundary condition matrix. Now it turns out that there are two solution vectors - the system is overdetermined. Obviously, any superposition of these is also a solution. The solution space reduces to one dimension when another constraint is added. One can require that there are no ingoing waves from the right by simply setting F to zero as an extra boundary condition. The solutions to the tunnel barrier problem are depicted in figure 4. This picture shows electron waves that are incident from the left and right, and inter-

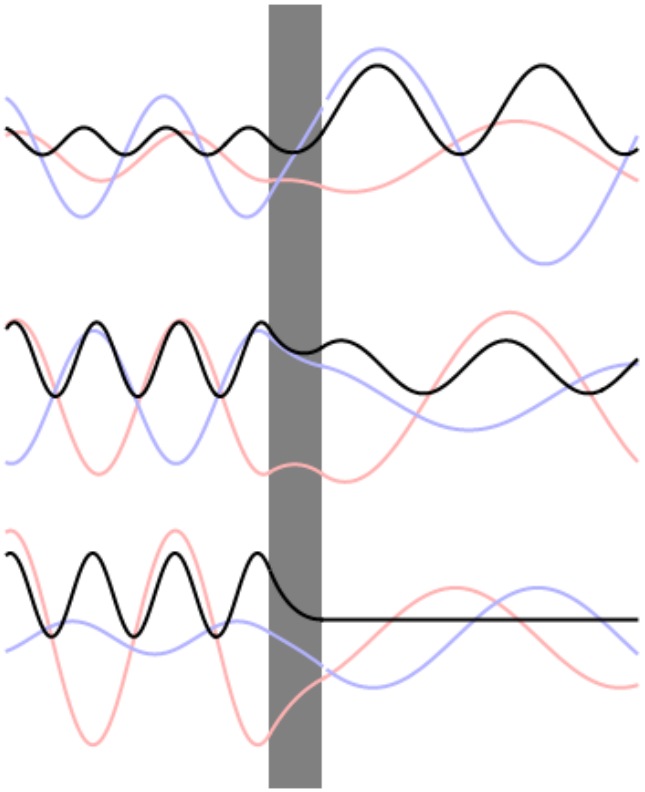


FIG. 4. Two solutions for the tunnel barrier problem ($k_1 = k_2 = k_3$, $x_0 = 1/k_2$) and their linear superposition solution upon setting the additional boundary condition $F = 0$ (no incoming waves from the right). Blue = real, red = imaginary, black = probability density amplitude (absolute squared wave function).

connect in the barrier region. The third situation has no incoming waves from the right. The situation is now an electron wave from the left that transmits a small part to the right and has a large part reflecting on the barrier. The incident and reflected waves interfere to generate a standing wave pattern.

The charge density ρ of an electron wave ψ is given by $\rho = e|\psi|^2$. The continuity relation can be used to find the current density

$$\nabla \cdot \mathbf{J} = -e\frac{\partial|\psi|^2}{\partial t}$$

The solution in one dimension is

$$J = \frac{\hbar e}{2mi}\left(\psi^*\frac{d\psi}{dx} - \psi\frac{d\psi^*}{dx}\right) = e\left(\frac{\psi^*\hat{p}\psi - \psi\hat{p}\psi^*}{2m}\right)$$

For the left and right regions respectively, the current evaluates to

$$J_{left} = \hbar k e\frac{|A|^2 - |B|^2}{m} = J_{right} = \hbar k e\frac{|E|^2 - |F|^2}{m}$$

For no incoming waves from the right, $|F|^2 = 0$, so the transmitted current from the left must be $J = \hbar k|E|^2/m$. This means that the coefficients must obey

$$|A|^2 - |B|^2 = |E|^2$$

This is analogous to the Fresnel equations for a simple wave transmission problem, where $|A|^2$ is the incident intensity, $|B|^2$ is the reflected intensity, and $|E|^2$ is the transmitted intensity. Dividing all intensities through $|A|^2$ gives

$$\frac{|B|^2}{|A|^2} + \frac{|B|^2}{|A|^2} = R + T = 1$$

where R and T are the reflection and transmission coefficient, respectively.

The transmission coefficient is proportional to

$$T \propto e^{-2k_2x_0}$$

demonstrating the extreme dependence of tunnel current with vacuum gap size. For a vacuum gap of 5 eV - approximately the work function of copper - the decay constant k_2 is 11.5 nm^{-1}, meaning that a change in gap length of a single Angstrom results in a 10-fold increase or decrease in tunnel current!

It would be desirable to see how the probability current depends on bias potential. If the left reservoir is biased with respect to the right one

$$V_1 = +eV/2,\ V_2 = -eV/2$$

the wave vectors at the Fermi level k_1 and k_3 will start to differ ($k_1 = \sqrt{2m(E - V_1)}/\hbar < k_3$), which makes sense, since a wave function of equal energy in higher potential has a lower kinetic energy. Bardeen treated this situation by splitting the wave function contributions on either side of the barrier in the following two waves[7]:

$$\psi_A = \left\{ \begin{array}{c} \psi_1 = Ae^{ik_1x} + Be^{-ik_1x} \mid x < 0 \\ \psi_2 Ce^{-k_2x} \mid 0 \le x \le x_0 \end{array} \right\}$$

$$\psi_B = \left\{ \begin{array}{c} \psi_2 = De^{k_2(x-x_0)} \mid 0 \le x \le x_0 \\ \psi_3 Ee^{ik_3(x-x_0)} + Fe^{-ik_3(x-x_0)} \mid x > x_0 \end{array} \right\}$$

he started with wave function ψ_A, which satisfies the Hamiltonian on the left and in the middle of the gap, but not on the right side. The trick is to consider the Hamiltonian on the right side of the gap as a perturbation to H on the left side. Since ψ_b satisfies the Hamiltonian in the gap region and on the right, it is a solution to the perturbation. The tunneling rate can then be found in terms of Fermi's golden rule

$$\Gamma_{A\to B} = \frac{1}{h}|\langle\psi_B|\hat{\mathrm{H}}|\psi_A\rangle|^2\delta_{A,occ}(E - E_0)\delta_{B,un}(E - E_0)$$

In the case of a continuum, $\delta(E - E_0) = \rho(E - E_0)$. For the one-dimensional case, the matrix element $\langle\psi_B|\mathrm{H}|\psi_A\rangle$ reduces to the formula of the probability current, and therefore implicitly contains the factor e^{-2k_2x}. Additionally, there is the condition that the tranfer rate to an occupied state or from an empty state is zero. The occupied density of states on the positively biased "left" electrode becomes ODOS = $\rho_A\mathrm{FD}(E + eV/2)$ and the unoccupied density of states on the negatively biased "right" electrode becomes UDOS = $\rho_B(1 - \mathrm{FD}(E - eV/2))$, where FD is the Fermi-Dirac distribution. These distribution functions introduce the so-called bias window

$$\Gamma = \frac{1}{h}\int|\langle\psi_B|\hat{\mathrm{H}}|\psi_A\rangle|^2\rho_A\mathrm{FD}(E + eV/2)\times$$

$$\rho_B(1 - \mathrm{FD}(E - eV/2))dE$$

So the tunnel current is proportional to the overlap integral over the UDOS and UDOS, as shown schematically in figure 3.

The current is the surface integral of the transfer rate. Only the wave function matrix element is a function of surface integration.

$$I = \frac{2e}{h}\int_{-\infty}^{\infty}\rho_A\rho_B|M_{AB}|^2\ \mathrm{FD}(1 - \mathrm{FD})dE$$

with $|M_{AB}|^2 = \int\langle\psi_B|\hat{\mathrm{H}}|\psi_A\rangle dA$ and where the bias window is written succinctly as FD(1 − FD). The factor 2 comes from the fact that both spin-up and spin-down electrons contribute.

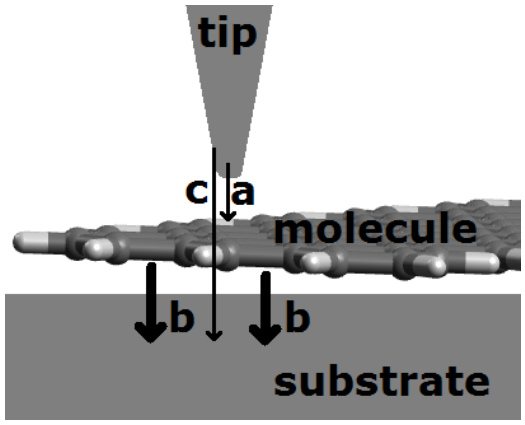


FIG. 5. 1) The double tunnel barrier formed by tip, molecule and substrate. The different tunneling processes are indicated by the letters a, b and c. a: (rate-dependent) tunneling from tip to molecule. b: fast tunneling from molecule to substrate. c: background tunneling from tip to substrate.

Microscopy

Now we want to analyse the situation in which tunneling first occurs from tip to GNR, and then from GNR to substrate. This two-step tunneling process is the situation that is encountered when scanning over nanoribbons and performing spectroscopy experiments on them. In addition to the two-step tunneling process, tunneling directly from tip to substrate without passing the molecule may occur. The three processes in the double-barrier tunnel junction are pictured in figure 5. The most common mode in which STM is performed is constant current mode, in which the measured tunnel current is used as a feedback parameter controlling the tip height. Therefore, the tip height is constantly adjusted to keep the tunnel current more or less constant. Doing this dynamically while scanning across the surface allows for the visualization of the "height" profile of the surface. Of course, since tunneling depends on the electronic structure of the materials on either side of the junction, the observed height is actually a convolution of real height and the local density of states within the bias window. Nevertheless, it may be expected that STM images give a good impression of the height profile.

An issue with the current formula so far is that the bias window has edges with a width of order kT through the Fermi-Dirac distributions. This means that the electronic spectra of the molecules are smeared out by roughly 26 meV at room temperature. This thermal broadening comes on top of the natural broadening of the energy levels through coupling with the substrate (and tip)[8]. We would like to probe the electronic structure of our molecules as accurately as possible, so our experiments are performed at low temperatures (4K). There are additional arguments for going to 4K - the reduction of thermal vibrations of the sample, reduction of piezo drift, reduced adsorbate mobility (including GNR) - all of which will aid in creating a stable tunnel junction. The Fermi-Dirac distributions at low temperatures can be approximated as step functions, so that the new tunnel current formula becomes

$$I = \frac{2e}{h} \int_{E_F - \frac{eV}{2}}^{E_F + \frac{eV}{2}} \rho_A \rho_B |M_{AB}|^2 dE$$

The rate of tunneling from either electrode to the molecule is roughly proportional to the coupling. Weak coupling means that an electron can reside on the molecule for a sufficient time. Strong coupling causes the electronic structure of the molecule to become significantly distorted, and electrons can be localized in orbitals on both the tip and the molecule at the same time. The time scale for tunneling is so short that the uncertainty in energy becomes quite large. This causes lifetime broadening. So in order to "see" a molecule in a relatively unperturbed state, strong coupling is detrimental. One of the important reasons to choose a gold substrate is the fact that it is a noble metal, and as such establishes only weak coupling[9]. The region between the ribbon and the gold can be truly regarded as a second tunnel barrier[10].

There is a second kind of coupling: capacitative coupling. When taking both electronic and capacitative coupling into account, the double barrier junction can be regarded as an electronics problem. This is shown in figure 7. Now both the source and drain (the tip and substrate)

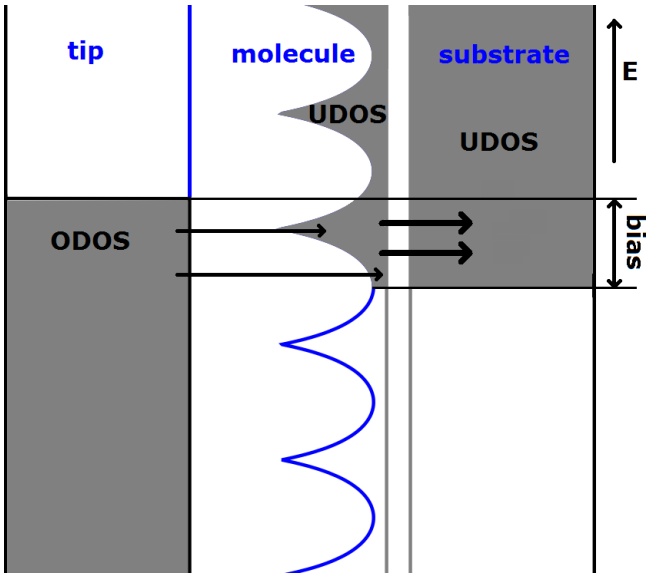


FIG. 6. A schematic energy diagram of a double tunnel barrier, where a two-step tunneling process takes place. The strongly asymmetric capacitative coupling "ties" the molecular density of states to the substrate voltage. The tunnel rate is determined by tunneling from the ODOS of the tip to the UDOS of the molecule (or vice versa for negative bias).

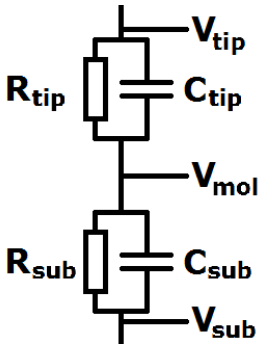


FIG. 7. 1) An equivalent electronic circuit describing the transport and capacitative coupling in the double tunnel barrier.

are connected to the GNR via a resistor (R_{tip} and R_{sub}) and a capacitor (C_{tip} and C_{sub}). The capacities are determined by geometric factors - the surface area and the distance between molecule and metal. Since $C = Q/\Delta V$, the charge on the molecule can be written as

$$Q_{mol} = C_{tip}(V_{tip} - V_{mol}) = C_{sub}(V_{sub} - V_{mol})$$

and the bias voltage is

$$V_{mol} = \frac{C_{sub}V_{sub} + C_{tip}V_{tip}}{C_{sub} + C_{tip}}$$

The electrostatic potential is thus determined by the "capacitance-weighted" average of the bias of both tip and substrate. The quantum mechanical structure - the orbitals at their respective energy levels - comes on top of this constant potential. Since GNR are large, flat molecules, it is assumed from now on that $C_{sub} >> C_{tip}$ so that $V_{mol} \approx V_{sub}$. The capacitative effect of the tip is thus ignored. So the molecule does not change appreciably in electrostatic energy upon changing the bias potential.

There are now two bias windows for each tunneling step. The bias window between molecule and substrate is much smaller than between tip and molecule. Since the molecule is "tied" to the voltage of the substrate, one can define the effective bias window, where the electrostatic potential of molecule and substrate is set to a reference value of zero. Now the bias window over the tip-molecule junction runs from E_F to $E_F + eV$. Furthermore, the tunnel barrier from the molecule to the substrate is assumed to be much smaller than to the tip ($R_{tip} >> R_{sub}$), so that the current from tip to molecule becomes the rate-determining step for the total current, and the electronic structure of the substrate is not important for the two-step tunnel process (although it does show up in the background tunneling process shown in figure 5).

$$I = \frac{G_0}{e}\int_{E_F}^{E_F+eV} \rho_m\rho_t|M_{m,t}|^2 dE$$

where G_0 is the conductance quantum and the subscripts m, s and t refer to molecule, substrate and tip respectively. When the tip has a relatively flat density of states over a significant bias range, ρ_t can be taken outside of the integral.

$$I = G_0\rho_t\int_{E_F/e}^{E_F/e+V} \rho_m|M_{m,t}|^2 dV$$

Whether the tip is clean or not can be probed by sweeping the bias voltage when the tip is in tunnel contact with the gold surface.

When the clean tip is brought on top of a molecule and the bias is increased, the current formula dictates that the current steps up each time a molecular level (or "transmission channel") is passed. The derivative of the current with respect to bias voltage, or differential conductance, directly probes the density of states.

$$\frac{dI}{dV} = G_0\rho_t\rho_m(V)|M_{m,t}|^2 \propto \rho_m(V)|\langle\psi_m|\psi_t\rangle|^2$$

Here, a result by Tersoff and Haman is used. They argued that, when the tip is atomically sharp, the tunnel matrix element is approximately proportional to the square of the overlap integral with the molecular wave function[11] ($|M_{m,t}|^2 \propto |\langle\psi_m|\psi_t\rangle|^2$). Figure 6 shows an idealized situation of the problem at low temperature. On the left side, the tip is depicted, with a flat density of states, so that the occupied density of stated (ODOS) is proportional to the Fermi-Dirac "step-function". The molecular states are tied capacitatively to the bias potential of the substrate. By opening the bias window, the tunnel current will increase proportional to the current integral, which contains the unoccupied density of states on the molecule. The differential conductance then returns the density of states on the molecule. Of course, the sign of the bias can be changed, changing the direction of the current from molecular occupied states to unoccupied states on the tip.

A molecule, like a GNR, features a series of discrete energy levels. However, due to lifetime broadening and electronic interaction with the substrate, these levels are broadened. The DOS can now be expanded as a sum of Lorentzians.

$$\rho_m(V) = \frac{1}{\pi}\sum_i \frac{\gamma}{(V - \frac{E_i}{e})^2 + (\frac{\gamma}{2})^2}$$

where E_i are the molecular eigenenergies and γ is the homogeneous broadening. The differential conductance is then written as

$$\frac{dI}{dV}(x,y) \propto \sum_i \frac{\gamma|\psi_{m,i}(x,y)|^2}{(V - \frac{E_m}{e})^2 + (\frac{\gamma}{2})^2}$$

Hence, the differential conductance can teach us two things. In a bias voltage sweep experiment, one can probe

the local density of states (the density of states in which each energy level is weighed with a factor $|\psi^2|$). Second, a scan of the differential conductance at a bias voltage in resonance with a single energy level E_i will provide an image of the orbital $|\psi_i|^2$.

Scanning tunneling spectroscopy

In an STM experiment, the tip can be positioned at a certain point on the surface, and a scanning tunneling spectroscopy (STS) measurement can be performed. As the bias voltage is swept, the resulting tunneling current is measured and the differential conductance can be calculated by differentiating the current with respect to the bias voltage. But in numerical derivation, noise in I is amplified to a large extent. To obtain a better signal to noise ratio, the technique of choice is to use a lock-in amplifier.

The tunnel current can be written in terms of the equivalent resistance of the tunnel junction $R = \int_{E_F/e}^{V_B} dI/dV \; dV$. In the lock-in amplifier, a harmonic oscillation is added to the bias voltage over the tunnel junction.

$$V = V_B + V_{osc} \; e^{i\omega t}$$

The result is a harmonic current through the junction

$$I = \frac{V_B}{R} + \frac{V_{osc}}{R} e^{i\omega t}$$

The current is fed back to the lock-in amplifier. First, a band-pass filter discards the constant tunnel offset and most of the noise. Then, the current oscillation is multiplied with the output signal, and the product signal is integrated over a few cycles. The output voltage is basically the Fourier integral of the a.c. tip current over the output signal of the lock-in amplifier. This output voltage is proportional to the differential conductance.

Differential conductance mapping

A problem with STM is that one cannot really visualize individual eigenstates after the HOMO or LUMO. To do so, it is really necessarry to be able to look at the differential conductance spatially. Intuitively, this technique can be seen as a method to "subtract" the first orbital from the image containing both orbitals. Using a harmonic oscillation generated by a lock-in amplifier, we can bring a single molecular level in and out of resonance of the bias window. Now, a scan can be made of the differential conductance, returning an image of the molecular orbital.

In general, energy levels are broadened by such an amount that they may overlap significantly. Nevertheless, by choosing a small voltage oscillation around a single level, it may often be possible to single out one orbital[12]. Of course, when energy levels become closely spaced, the scan still returns a superposition of orbitals. Correct differential conductance maps are therefore created by summing over all molecular orbitals, weighed by their Lorentzian factors integrated over the entire voltage oscillation. An example of this procedure is given in figure 8. The differential conductance first shows that there are two levels close in energy in the density of states. Mapping out the differential conductance spatially at a bias level where the resonances overlap, or where the bias oscillation - indicated by the horizontal arrows - spreads out over multiple peaks, results in a pattern that is a superposition of the individual eigenstates.

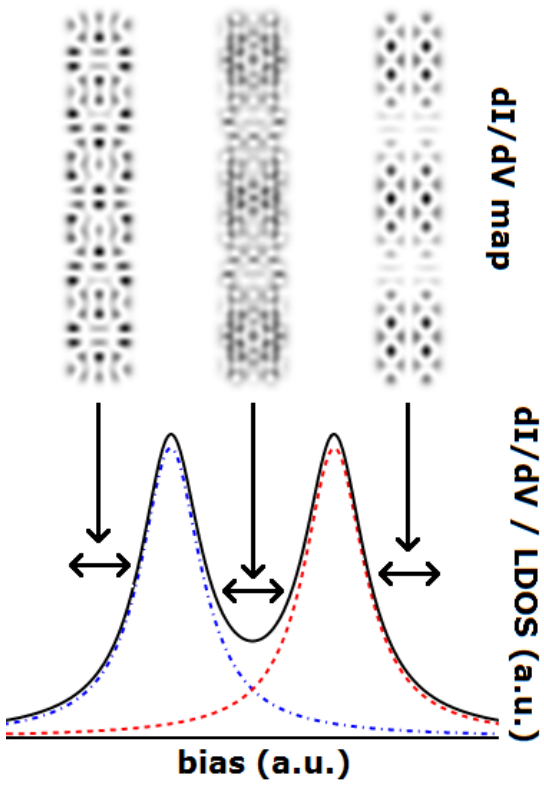


FIG. 8. An example of eigenstate superpositions in differential conductance mapping. For the sake of argument, two random orbitals from the short nanoribbon were used that are close in energy ($\Delta E = 4$ meV). The real density of states is much more convoluted with multiple eigenstates in the vicinity, so this should only be regarded as a simple model.

Manipulation

So far, many SPM techniques have been covered, but the focus has been on probing. The extremely precise control - on picometer scale - also allows for the precise manipulation of molecules and atoms on surfaces. The first success with STM atom manipulation came by Eigler

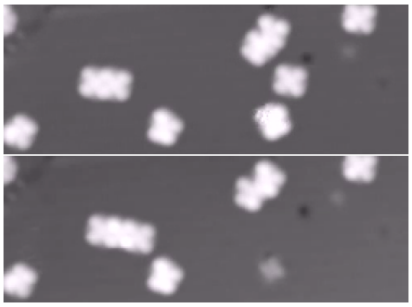


FIG. 9. An experiment, in which a molecule was shifted over a sodium chloride layer and contacted with another molecule by means of SPM manipulation. Experiment performed by Swart *et al.* [14]

and Schweizer, who managed to push xenon atoms into the correct positions to have them spell out "IBM", the name of their company[13].

Most STM manipulation experiments described in literature involve the lateral manipulation of atoms or complete molecules[15][14] (an example of which is given in figure 9), picking up atoms and molecules with an STM tip[16][17] and even giving voltage pulses to bring about a reaction[18][19]. In this research, we will focus on inducing conformational changes in a nanoribbon. The details of the nanoribbon manipulation experiments will be treated in the next chapter.

ATOMISTIC QUANTUM METHODS

In contrast with the nearly free-electron model, where the background potential is treated to first order as being constant, and only in higher order acquires a perturbation due to the lattice, the atomistic approach uses the exact Coulomb interactions of all nuclei to build up the potential energy landscape. The atomistic approach was originally devised for nonperiodic, molecular systems, and as such is used chiefly with space-localized basis functions rather than plane waves[20]. Nevertheless, extensions to periodic solids are used frequently nowadays[21][22].

The molecular single-electron wave functions are usually approximated as linear combinations of atomic orbitals χ_i (LCAO approximation) in a molecule (although some additional functions may be chosen that do not really relate to atomic orbitals).

$$\phi = \sum_i c_i \chi_i \quad \text{or} \quad \phi = \mathbf{c}.\vec{\chi}$$

with c_i the expansion coefficient of atomic orbital i, and $\mathbf{c}$ and $\vec{\chi}$ the coefficient vector and atomic orbital vector, respectively. The complete set of basis functions that describes the atomic orbitals - for example, cartesian Gaussians or Slater-type orbitals of different size and quantum numbers - is called the basis set. The projection of the basis set onto the atomic positions defines the complete atomic orbital vector $\vec{\chi}$. The problem of finding the molecular single-electron wave function is now reduced to finding the correct expansion coefficients on each atomic orbital.

The variational principle states that any wave function estimate ϕ will have a higher energy than the ground-state wave function ϕ_0 (which is an eigenfunction of the Hamiltonian), unless it is equivalent to the ground state wave function, in which case the energies are equivalent

$$\langle E \rangle = \left\langle \phi \left| \hat{\mathrm{H}} \right| \phi \right\rangle \geq \frac{\hat{\mathrm{H}}\phi_0}{\phi_0}$$

$$\phi = \phi_0 \Rightarrow \frac{\partial \langle \mathrm{E} \rangle}{\partial c_i} = \frac{\partial \left\langle \mathbf{c}.\vec{\chi} \left| \hat{\mathrm{H}} \right| \mathbf{c}.\vec{\chi} \right\rangle}{\partial c_i} = 0 \; \forall \; c_i$$

This problem is analytically solved under the orthonormality constraint (using the method of Lagrange multipliers) to give the secular equations

$$\mathrm{HC} = \mathrm{SCE}$$

C is the matrix containing the expansion coefficient vectors for each molecular wave function. E is a matrix containing the electronic energies on its diagonal. H is the Hamiltonian matrix, and S is the overlap matrix. Their elements are

$$\mathrm{H}_{ij} = \left\langle \chi_i \left| \hat{\mathrm{H}} \right| \chi_j \right\rangle \qquad \mathrm{S}_{ij} = \langle \chi_i | \chi_j \rangle$$

By virtue of the linearity of the Hamiltonian operator, the Hamiltionan matrix H can be separated into (or built up from) a single-electron Hamiltonian matrix and a multi-electron Hamiltonian matrix

$$\mathrm{H} = \mathrm{H}_1 + \mathrm{H}_2$$

The secular equations are a special kind of eigenvalue problem, and can be solved using the machinery of linear algebra. In the case that overlap is ignored, so that S = I, they can be written as a trivial eigenvalue equation

$$\mathrm{C}^{-1}\mathrm{HC} = \mathrm{E}$$

so that the diagonal eigenenergy matrix E contains the eigenvalues and C the eigenvectors of H. Each eigenvalue E_{ii} is the solution to the characteristic polynomial

$$\det(\mathrm{H} - \mathrm{E}_{ii}\mathrm{I}) = 0$$

and the expansion coefficients are the elements of the eigenvector $\mathbf{c}$ from the matrix C of that specific eigenvalue.

$$\mathrm{H}\mathbf{c} = \mathrm{E}_{ii}\mathbf{c}$$

This last equation is the secular equations in single-electron basis.

Although the addition of the overlap matrix means that HC = SCE cannot be calculated anymore by a simple eigendecomposition, the secular equations can still be solved in the single-electron form

$$(\mathrm{H} - \mathrm{E}_{ii}\mathrm{S})\mathbf{c} = 0$$

Eigendecomposition of the Hamiltonian matrix is thus seen to involve solving the n-dimensional characteristic polynomial for the eigenenergies and subsequent eigenvector formation by finding the nullspace (reduced: without the trivial zero element) of all matrices $\mathrm{H}_1 - \mathrm{E}_{ii}\mathrm{S}$, defined as

$$N(\mathrm{H} - \mathrm{E}_{ii}\mathrm{S}) = \{\mathbf{c} | (\mathrm{H} - \mathrm{E}_{ii}\mathrm{S})\mathbf{c} = \mathbf{0}, \mathbf{c} \neq \mathbf{0}\}$$

For degenerate states, the eigenenergies are identical, so that the nullspace contains multiple vectors for that state. In general, the degeneracy level g is the dimensionality of (reduced) nullspace

$$g(\mathrm{E}_{ii}) = \mathrm{span}(N(\mathrm{H} - \mathrm{E}_{ii}\mathrm{S}))$$

Two degenerate eigenstates have the property that they can be mixed

$$\mathrm{H}(\mathbf{c}_1 + \mathbf{c}_2) = \mathrm{H}\mathbf{c}_1 + \mathrm{H}\mathbf{c}_2 = \mathrm{SE}\mathbf{c}_1 + \mathrm{SE}\mathbf{c}_2 = \mathrm{SE}(\mathbf{c}_1 + \mathbf{c}_2)$$

This means that the superposition eigenstate $\mathbf{c}_3 = \mathbf{c}_1 + \mathbf{c}_2$ is also an eigenstate of the Hamiltonian matrix. The entire plane in state space spanned by vectors $\mathbf{c}_1$ and $\mathbf{c}_2$ contains solutions to the Hamiltonian, and the set of normalized states span a circle. Two degenerate orthonormal eigenstates can thus be mapped onto two other degenerate orthonormal eigenstates by using a rotational transformation

$$\begin{pmatrix} \mathbf{c}_1' \\ \mathbf{c}_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}$$

or, equivalently,

$$\begin{pmatrix} \psi_1' \\ \psi_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

An example of a rotational transformation will be encountered in the tight binding calculations in the next chapter.

We would like to solve the secular equations directly for all orbitals, without going to the single electron basis and applying all energy eigenvalues. The trick is to orthogonalize the atomic basis set and find the square root of the overlap matrix so that the secular equations can be rewritten to a simple eigenvalue problem.

- The (diagonal) eigenvalue matrix Λ and its corresponding eigenvector matrix X of S are computed by an eigendecomposition

$$\Lambda = \mathrm{X}^{-1}\mathrm{SX}$$

The matrix X is thus an expansion coefficient matrix that diagonalizes S. When X is applied to the atomic orbitals $\vec{\chi}$, the resulting molecular orbitals are mapped to orthogonal basis functions.

- Since Λ is diagonal, the square root of this matrix is simply calculated as the matrix containing the square root of all diagonal entries.

$$\Lambda^{\frac{1}{2}} = \sqrt{\Lambda}$$

- $\Lambda^{\frac{1}{2}}$ is used to calculate the square root of S.

$$(\mathrm{S}^{\frac{1}{2}})(\mathrm{S}^{\frac{1}{2}}) = \mathrm{S} = \mathrm{X}\Lambda\mathrm{X}^{-1} = \mathrm{X}\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}\mathrm{X}^{-1}$$

$$= (\mathrm{X}\Lambda^{\frac{1}{2}}\mathrm{X}^{-1})(\mathrm{X}\Lambda^{\frac{1}{2}}\mathrm{X}^{-1}) \Rightarrow \mathrm{S}^{\frac{1}{2}} = \mathrm{X}\Lambda^{\frac{1}{2}}\mathrm{X}^{-1}$$

- The inverse of $\mathrm{S}^{\frac{1}{2}}$ is calculated by a matrix inversion algorithm.

$$\mathrm{S}^{-\frac{1}{2}} = \left(\mathrm{S}^{\frac{1}{2}}\right)^{-1}$$

- Using these square root matrices, the secular equations are written as

$$\mathrm{HC} = \mathrm{SCE} = \mathrm{H}(\mathrm{S}^{-\frac{1}{2}}\mathrm{S}^{\frac{1}{2}})\mathrm{C} = \mathrm{S}(\mathrm{S}^{-\frac{1}{2}}\mathrm{S}^{\frac{1}{2}})\mathrm{CE}$$

$$\mathrm{HS}^{-\frac{1}{2}}(\mathrm{S}^{\frac{1}{2}}\mathrm{C}) = \mathrm{S}^{\frac{1}{2}}(\mathrm{S}^{\frac{1}{2}}\mathrm{C})\mathrm{E}$$

$$(\mathrm{S}^{-\frac{1}{2}}\mathrm{HS}^{-\frac{1}{2}})(\mathrm{S}^{\frac{1}{2}}\mathrm{C}) = (\mathrm{S}^{\frac{1}{2}}\mathrm{C})\mathrm{E}$$

$$\mathrm{H}'\mathrm{C}' = \mathrm{C}'\mathrm{E}$$

The last equations are the orthogonalized secular equations. They have the form of an eigenvalue equation, which we know how to solve (or let the computer solve). Therefore, initially the matrix $\mathrm{H}' = \mathrm{S}^{-\frac{1}{2}}\mathrm{HS}^{-\frac{1}{2}}$ is calculated and subsequently eigendecomposed.

- The resultant eigenvector matrix $\mathrm{C}' = \mathrm{S}^{\frac{1}{2}}\mathrm{C}$ is premultiplied by $\mathrm{S}^{-\frac{1}{2}}$ to give C.

Hartree Fock and Density functional Theory

Quantum chemical methods usually work in a couple of steps, as follows. First, single electron integrations over atomic orbitals are carried out and the resulting values are stored in the Hamiltonian matrix and the overlap matrix. For example

$$\mathrm{H}_{ij,1} = \left\langle \chi_i \left| \hat{\mathrm{V}}_{Ne} - \frac{\hbar^2}{2m}\nabla^2 \right| \chi_j \right\rangle$$

The functions χ themselves may be expanded in so-called primitive functions (contractions), and the basis set may be extended from the minimal case (diffuse orbitals, polarization functions, etc.). This increases the accuracy of the computation but also the number of integrals to obtain $\mathrm{H}_{ij,1}$. Evaluation of the multi-electron Hamiltonian matrix terms is more involved, and differs for various quantum chemical methods. In general, the multi-electron Hamiltonian is adjusted several times during a computation. The result after all molecular integral evaluations is an overlap matrix S and a Hamiltonian matrix H, and at this point the secular equations can be solved.

The largest difference in quantum chemical methods is how to account for the multiparticle term $\hat{\mathrm{V}}_{ee}$.

- Using the mean-field repulsions between electron densities set up by the respective atomic orbitals (the two-electron Coulomb repulsion as set up by the respective electron wave function densities). This is done in the Hartree-Fock (HF) method. Electron correlation effects can be accounted for by using post-HF methods, albeit at a large computational cost that scales increasingly worse for large systems.

- Using an empirical function(al), that depends on the total electron density and that correctly accounts for correlated repulsion. This is used in the Kohn-Sham implementation of density functional theory (DFT)[23].

In both cases, the calculated coefficients are a function of the electron density through the multi-electron Hamiltonian. However, the electron density in turn is related to the coefficients of occupied orbitals, which is the output of the calculation. The problem is usually solved by taking a guess density and iterating the calculation until the input and output density are consistent. This procedure is known as the self-consistent field (SCF) approach. Whether or not a computation will actually converge is a key issue in SCF computations. Various methods (like density matrix damping, DIIS and level shifting) are used to help convergence, but it may remain a problem in some computations.

Density functional theory is the most popular quantum chemical method nowadays, as the computational cost is comparable to Hartree Fock but its results are often more accurate[24]. It is important to note that both methods require the input of explicit functions, the basis set, and the Hamiltonian and overlap matrix elements are obtained by explicitly integrating over these functions.

Symmetry, sparsity and separability

The orbitals in graphene and graphene nanosystems are derived from the s and p atomic orbitals of the carbon atoms. The s-orbitals and two of the p-orbitals, lie in the molecular plane. They span the so-called σ-framework. The remaining p-orbitals protrude on either side of the molecular plane, and combine to form the so-called π-system. They will be denoted $p_\perp$, indicating their perpendicularity with respect to the framework orbitals.

The orbitals spanning the π-system and σ-framework are of different parity. $p_\perp$-orbitals are, in contrast to the other AO's, antisymmetric with respect to reflection in the z-plane.

$$\phi_{2pz}(-z) = -\phi_{2pz}(z), \quad \phi_{2px}(-z) = \phi_{2px}(z)$$

The product of a symmetric and an antisymmetric function is antisymmetric

$$\phi_{2pz}\phi_{2px}(-z) = -\phi_{2pz}\phi_{2px}(z)$$

This property holds for every AO pair, regardless of (relative) position in the xy-plane. An overlap matrix element is

$$\mathrm{S}_{ij} = \langle\phi_{2px}|\phi_{2pz}\rangle = \iint \left[\int_{-\infty}^{\infty} \phi_{2px}\phi_{2pz}dz\right] dx\ dy$$

where the integral within brackets is expanded

$$[] = \int_{-\infty}^{0} \phi_{2px}\phi_{2pz}(z)dz + \int_{0}^{\infty} \phi_{2px}\phi_{2pz}(z)dz$$

and the antisymmetry is applied to the first term

$$[] = -\int_{0}^{\infty} \phi_{2px}\phi_{2pz}(z)dz + \int_{0}^{\infty} \phi_{2px}\phi_{2pz}(z)dz = 0$$

So the overlap of orbitals with different parity vanishes. Especially in a high-symmetry system like graphene, many elements in both the single-electron Hamiltonian matrix and the overlap matrix integrate to zero, simply because of the orthogonality of the basis functions. Also - this is completely general - the magnitudes of matrix elements that are not orthogonal still approaches zero quite quickly when AO's become separated by a couple of nm. These effects render the single-electron Hamiltonian matrix and overlap matrix very sparse. In addition to being convenient for linear algebraic diagonalization algorithms, these matrices can become block diagonal, e.g.

$$\mathrm{H}_1 = \begin{pmatrix} \mathrm{A} & \mathbf{0} \\ \mathbf{0} & \mathrm{B} \end{pmatrix}$$

The eigenvalues of H_1 are the solutions to the characteristic polynomial

$$\det(\mathrm{H}_1 - \mathrm{E}_{ii}\mathrm{I}) = 0$$

In the block-diagonal case this is - by definition of the determinant - identical to

$$\det(\mathrm{H}_1 - \mathrm{E}_{ii}\mathrm{I}) = \det(\mathrm{A} - \mathrm{E}_{ii}\mathrm{I})\det(\mathrm{B} - \mathrm{E}_{ii}\mathrm{I})$$

This polynomial equation is solved if

$$\Rightarrow \det(\mathrm{A} - \mathrm{E}_{ii}\mathrm{I}) = 0 \ \mathit{or} \ \det(\mathrm{B} - \mathrm{E}_{ii}\mathrm{I}) = 0$$

The eigenvalues of the combined system are simply those of A and B, and the eigenvectors are also completely unperturbed by the mutual presence. So, the wave functions are orthogonal, the states do not mix, and the eigenvalues and -vectors are simply those of A and B. The subsystems are "invisible" to each other! In the particular case of graphene, the one-electron Hamiltonian matrix can be written as a block-diagonal matrix containing σ-framework and π-system Hamiltonian submatrices.

$$\mathrm{H}_1 = \begin{pmatrix} \Sigma & \mathbf{0} \\ \mathbf{0} & \Pi \end{pmatrix}$$

where Σ is the Hamiltonian of the σ-framework and Π the Hamiltonian of the π-system. Since we concluded that these subsystems are "invisible" to each other, they are solved by simply putting the subsystems into the secular equations, one by one.

$$\Sigma\mathrm{C}_\sigma = \mathrm{S}_\sigma\mathrm{C}_\sigma\mathrm{E}_\sigma, \ \Pi\mathrm{C}_\pi = \mathrm{S}_\pi\mathrm{C}_\pi\mathrm{E}_\pi$$

Now we can invoke the Hückel approximation and claim that the interesting physics of aromatic molecules is in the π-system[25]. The approach would then be to not care about the equations of the σ-framework and only solve the π-SE. In comparison with the total secular equations for all AO's, a significant reduction in basis orbitals is made (only the $p_\perp$-orbitals are left), and the Hamiltonian is drastically reduced in size. The computational effort of solving the secular equations has become a much smaller and a much quicker one. In summary, orthogonality causes the π-system, which is composed of $p_\perp$-orbitals to be entirely decoupled from the σ framework.

Unfortunately, the multi-electron Hamiltonian does not benefit from symmetry because of the fact that the energy is proportional to the squared wave functions. Squared AO's are always positive and can never be anisymmetric with respect to an inversion or with respect to each other. Although H_1 and S may be block-diagonal, H_2 is not block-diagonal. Hence, H is not block-diagonal, ruining the $\sigma - \pi$-separability. Electron-electron interactions cause a coupling between the σ-framework and π-system, increasing the computational effort in atomistic methods tremendously.

However, this annoying feature does not have to damp all hopes to benefit from symmetry. Might it be possible to tweak the π-framework Hamiltonian (Π) in such a way that the effect of σ-interaction is accurately represented, even when disregarding the σ-lattice itself? The answer is: yes, but at a slight decrease of computational accuracy. In the framework of graphene tight-binding, it is normally assumed that the $\sigma - \pi$-subsystem coupling is weak and translationally invariant, and it is implicitly taken into account in assigning the values of the Hamiltonian matrix for the π-system. The translational invariance at edges can be fixed in an empirical way, as will be shown later.

It will be clear from this discussion that for the efficient evaluation of quantum mechanical systems, sparsity is our friend. The computer needs to find the characteristic polynomial for the Hamiltonian matrix by calculating the determinant, the cost of which scales with the number of off-diagonal elements. The same goes for the overlap matrix, as will be shown later. A final point to be made is that the input orbitals (generally the basis set projected on the atoms) are not constrained to be atomic orbitals, and may be LCAO expansions themselves. It is even possible to solve the eigensystem for a set of molecular orbitals and have them mix to form other molecular orbitals (as long as the input set is linearly independent, otherwise, degrees of freedom are lost). A popular method to achieve even sparser Hamiltonian matrices is to employ the symmetry of a system, by first forming symmetry-adapted linear combinations (SALCs) of atomic orbitals and then performing the calculation on this basis. SALCs work when they respect certain symmetry elements of the system, for example (anti)symmetry with respect to inversion. In addition, the formation of SALCS puts constraints on expansion coefficient for each type, so that the number of "degrees of freedom" is reduced. Symmetry is commonly implemented in quantum chemical software, like ADF and ORCA, which were used in this research.

The tight binding method

Tight binding, or Hückel theory, is a relatively simple method for finding the electronic structure of molecules, by empirical parametrization of molecular integrals in the secular equations and invoking symmetry arguments to minimize the computational cost. Despite the severity of the assumptions used, surprisingly accurate results have been achieved, and nowadays it has become a very popular tool in electronic structure calculations of graphene and graphene nanosystems[26][27][28][29]. $\pi - \sigma$-separability is employed and only the π-system has to be taken into account. Another helpful aspect is that these materials tend to be weakly correlated[30]. The weak correlation of graphenes arises from the fact that the relative electron-electron separation is quite large, especially in the π system. Since the electron density of different molecular orbitals is also very evenly distributed, it deviates only very slightly for electrons in different orbitals.

Another assumption is that the potential energy landscape on each carbon atom is relatively invariant from one atom to the next. Of course, this is true for the bulk, and the potential energy wells only deviate near the edges. The problem is solving the secular equations for

the π system exclusively based on empirical values of the Hamiltonian and overlap matrix elements. By choosing them correctly, for example by fitting it to other atomistic methods, interaction effects with the lattice can be slightly recreated within the π-Hamiltonian. TB - although arguably somewhat less sophisticated than DFT or HF methods - is much more efficient because of the following factors

- Single parametrization of integrals in the Hamiltonian matrix without specifying a basis set
- Lack of four-electron integrals (HF) or exchange correlation interaction (DFT) integrals
- No need for SCF iterations
- Disregard of orbitals of a certain symmetry and decoupling from interesting orbitals, in this case the entire σ-framework

As a simple system that is easily extrapolated to graphene and graphene nanoribbons, the ethylene molecule is taken as an example. The Hamiltonian matrix for ethylene within the tight-binding framework is constructed by only treating the $p_\perp$-orbitals on both carbon atoms. These are referred to as $|1\rangle$ and $|2\rangle$. By the hermiticity of the Hamiltonian operator

$$\left\langle \chi_i \left| \hat{\mathrm{H}} \right| \chi_j \right\rangle = \overline{\left\langle \chi_j \left| \hat{\mathrm{H}} \right| \chi_i \right\rangle} \; \forall \; \chi_i, \; \chi_j$$

the Hamiltonian matrix should be symmetric and the interaction integral between the two $p_\perp$-orbitals is equivalent (called $\beta = \mathrm{H}_{ij}$). By symmetry, the carbon atom self energies, (called $\alpha = \mathrm{H}_{ii}$) should also be equal. The matrix is

$$\mathrm{H} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$$

H is separated as

$$\mathrm{H} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix} = \alpha \mathrm{I} + \mathrm{H}_\beta$$

Ignoring overlap, so that $\mathrm{S} = \mathrm{I}$, the secular equations are written as a simple eigenvalue equation

$$\mathrm{E} = \mathrm{C}^{-1}\mathrm{H}\mathrm{C} = \mathrm{C}^{-1}(\alpha \mathrm{I} + \mathrm{H}_\beta)\mathrm{C}$$

$$= \alpha \mathrm{I}\mathrm{C}^{-1}\mathrm{C} + \mathrm{C}^{-1}\mathrm{H}_\beta \mathrm{C} = \alpha \mathrm{I} + \mathrm{C}^{-1}\mathrm{H}_\beta \mathrm{C}$$

which has a trivial solution

$$\mathrm{C} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad \mathrm{E} = \alpha \mathrm{I} + \begin{pmatrix} \beta & 0 \\ 0 & -\beta \end{pmatrix}$$

The two row vectors that make up C denote the bonding and antibonding combinations, respectively

$$\phi_g = \frac{\sqrt{2}}{2}\left(|1\rangle + |2\rangle\right) \qquad \phi_u = \frac{\sqrt{2}}{2}\left(|1\rangle - |2\rangle\right)$$

the subscripts denote *gerade* (bonding) and *ungerade* (antibonding), by concensus. Their respective energies are the diagonal elements in the eigenenergy matrix

$$E_g = \alpha + \beta \qquad E_u = \alpha - \beta$$

By symmetry, there should be no other combinations than a symmetric and antisymmetric one.

The aim of the decomposition of the Hamiltonian matrix into two matrices is to show that if all diagonal terms are the same ($\mathrm{H}_{ii} = \alpha \; \forall \; i$), the effect on the energy levels is a uniform displacement of all energy levels by this amount (by virtue of the factor αI in the eigenenergy matrix). So tuning α shifts the MO diagram or band structure as a whole. For a system with one electron per carbon atom, this self-energy parameter also happens to coincide with the Fermi level of the molecule ($E_F = \alpha$). In most calculations, the Hamiltonian matrix diagonal is simply set to zero so that automatically the Fermi level is set to the reference value of 0. It is also easily observed that the off-diagonal interaction strength β scales the energy spacing between successive energy levels. It is seen empirically that, even in large calculations, the parameter β can be tuned to uniformly increase energy level spacings.

At this point it is instructive to compare our matrix tight-binding approach to another conventional form of tight-binding, since latter notation is commonly used in physics literature. Although its appearance is completely different, a completely equivalent form of writing down the tight-binding Hamiltonian is

$$\hat{\mathrm{H}}_{ij} = \sum_{ij} t_{ij} c_i^\dagger c_j + h.c.$$

with t_{ij} the interaction, or "hopping" integral, *h.c.* standing for Hermitian conjugate, $c^\dagger$ the creation operator and c the annihilation operator, which operate on an implicit minimal basis set. The sum runs over all atom pairs ij. The key thought of this formulation of tight binding is that the interaction integral of different atomic orbitals is equivalent to the energy needed to have an electron "hop" from site i to site j, or vice versa. This hopping is described by the simultaneous annihilation of an electron from state j and creation of it in state i. The hopping parameter t is equivalent to the value β or H_{ij} that we would like to put into the Hamiltonian matrix. Minimization of wave function energies as requested by the variational principle now comes down to minimization of the total energy expectation value of the TB operator on all electron pairs.

In the ethylene problem, there is only a single nearest neighbour interaction constant t, and for simplicity, we can neglect overlap. The Hamiltonian can be written as

$$\hat{\mathrm{H}} = t \sum_{ij} |i\rangle\langle j| + h.c.$$

We apply this Hamiltonian on a yet undetermined ethylene orbital ϕ, defined as

$$|\phi\rangle = c_1|1\rangle + c_2|2\rangle$$

This results in

$$\hat{\mathrm{H}}|\phi\rangle = t(c_1|1\rangle\langle 2|1\rangle + c_2|1\rangle\langle 2|2\rangle$$

$$+c_1|2\rangle\langle 1|1\rangle + c_2|2\rangle\langle 1|2\rangle)$$

$$= tc_2|1\rangle + tc_1|2\rangle$$

$$= E|\phi\rangle = E(c_1|1\rangle + c_2|2\rangle)$$

This requires $tc_2 - Ec_1 = 0$ and $tc_1 - Ec_2 = 0$. So

$$\begin{pmatrix} -E & t \\ t & -E \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This is again the same eigenvalue equation $(\mathrm{H} - \mathrm{IE_{ii}})\mathbf{c} = \mathbf{0}$, with β disguised as t! So this description leads to the same matrix equations. For practical calculations, however, it is sufficient to start off by formulating a matrix Hamiltonian.

In absence of overlap ($s = 0$), all energy levels are symmetric with respect to the Fermi energy. Now we set $\langle 1|2\rangle = h.c. = s$ so that

$$\mathrm{S} = \begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix}$$

We can calculate HC

$$\mathrm{HC} = \mathrm{H}_\beta\mathrm{C} = \frac{\sqrt{2}}{2}\begin{pmatrix} \beta & -\beta \\ \beta & \beta \end{pmatrix}$$

and SCE

$$\mathrm{SCE} = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & s \\ s & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} E_g & 0 \\ 0 & E_u \end{pmatrix}$$

$$= \frac{\sqrt{2}}{2}\begin{pmatrix} E_g(1+s) & E_u(1-s) \\ E_g(1+s) & -E_u(1-s) \end{pmatrix}$$

The secular equations are solved in terms of energies by termwise comparison of matrix components of HC and SCE.

$$E_g(1+s) = \beta \Rightarrow E_g = \frac{\beta}{1+s}$$

$$E_u(1+s) = -\beta \Rightarrow E_u = -\frac{\beta}{1-s}$$

Equivalently, in the single electron basis, only one vector is taken from the matrix C, and it was shown that the secular equations are solved when

$$(\mathrm{H} - \mathrm{E_{ii}S})\mathbf{c} = \mathbf{0}$$

$$\begin{pmatrix} -E & \beta - sE \\ \beta - sE & -E \end{pmatrix}\mathbf{c} = \mathbf{0}$$

The characteristic quadratic equation again lead to the same two eigenvalues

$$(\beta - sE)^2 = E^2 \Rightarrow E(1 \pm s) = \pm\beta$$

and the eigenvectors are again the symmetric and antisymmetric superposition. Overlap effectively antisymmetrizes, or skews the energy levels. Energy levels above the Fermi level move away from the Fermi level and energy levels below move closer towards it. This is a result that will carry over into larger calculations, as will be seen in the next chapter.

THE STRUCTURE OF GRAPHENE

The electronic magic of graphene comes from its flat honeycomb structure. The formation of this structure can be explained from examining the atomic orbitals of carbon. A carbon atom contains a pair of electrons in its core shell, and four electrons in its valence shell. If carbon were a free atom, these electrons would doubly occupy the $2s$ orbital and singly occupy both the perpendicular $2p_x$ and $2p_y$ orbitals by the Aufbau principle. This leaves an empty p_z orbital. Inside a crystal lattice, the energies of these orbitals are displaced and it may become more energetically favourable for electrons to occupy orbitals which are mixtures of $2s$ and $2p$ orbitals. The orbitals hybridize depending on their chemical environment, or ligand field. This causes carbon to exist in different allotropes, e.g. sp^2-hybridized carbon will form graphene or graphite lattices and sp^3-hybridized carbon will form diamond lattices.

The wavefunction for the $2s$ orbital is, to a good approximation

$$\phi_{2s} = \left[\frac{Z^{3/2}}{2\sqrt{2\pi}}\right](2 - Zr)e^{-\frac{Zr}{2}}$$

where the term in brackets is a normalization constant and Z is the effective nuclear charge, which is lower than

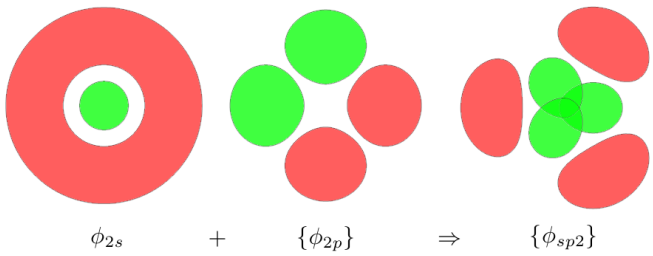


FIG. 10. Hybridization of the carbon $2s$ atomic orbital with the $2p_x$ and $2p_y$ orbitals gives sp^2 hybrids.

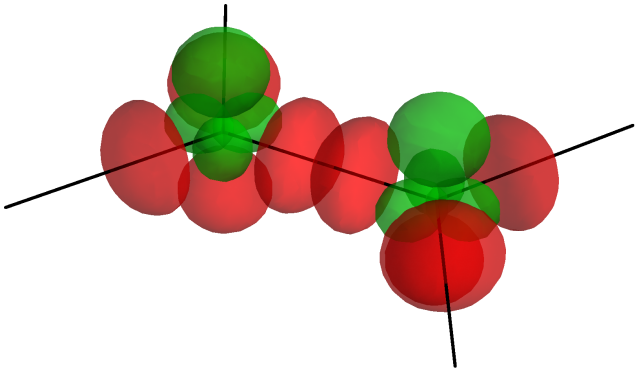


FIG. 11. sp^2 hybrid orbitals (shown on adjacent atomic sites) can be seen as the basic building blocks of the graphene honeycomb σ lattice. They will mix to form the molecular orbitals in the plane. The remaining, singly occupied $p_\perp$ orbital mix to form the π-system.

6 because of the shielding effect of the $1s$ orbital electron pair. The wavefunction for the $2p_x$ orbital is

$$\phi_{2p_x} = \left[\frac{Z^{5/2}}{4\sqrt{2\pi}}\right] xe^{-Zr}$$

and similarly for $2p_y$ and $2p_z$. The sp^2 hybrid orbitals are then the orthonormal linear combinations

$$\phi_{sp^2,1} = \frac{\phi_{2s} - \sqrt{2}\phi_{2p_x}}{\sqrt{3}}$$

$$\phi_{sp^2,2} = \frac{1}{\sqrt{3}}\left(\phi_{2s} + \frac{\phi_{2p_x}}{\sqrt{2}} + \sqrt{\frac{3}{2}}\phi_{2p_y}\right)$$

$$\phi_{sp^2,3} = \frac{1}{\sqrt{3}}\left(\phi_{2s} + \frac{\phi_{2p_x}}{\sqrt{2}} - \sqrt{\frac{3}{2}}\phi_{2p_y}\right)$$

The hybridized orbitals protrude at 120 degree angles relative to each other, perpendicular to the the remaining unhybridized p_z orbital. A cross section is shown in figure 10. In the framework of valence bond theory, one can now say that carbon atom in graphene form σ-type bonds between its sp^2 orbital and a sp^2 orbital of a neighbour atom, leaving the bonding carbon-carbon bond doubly occupied. When surrounded by carbon atoms, each carbon-carbon bond is thus doubly occupied. The 120 degree angle between σ bonds results in the honeycomb lattice. The singly occupied $2p_z$ orbital, which from now on is called the $p_\perp$ orbital, protrudes perpendicularly on every carbon atom of the lattice. Two graphene lattice positions with sp^3-hybridized basis functions are shown in figure 11. The honeycomb framework of σ-type bonds arising from sp^2 orbitals are now referred to as the σ lattice, and the $p_\perp$ orbitals span the π system on both sides of the graphene plane11.

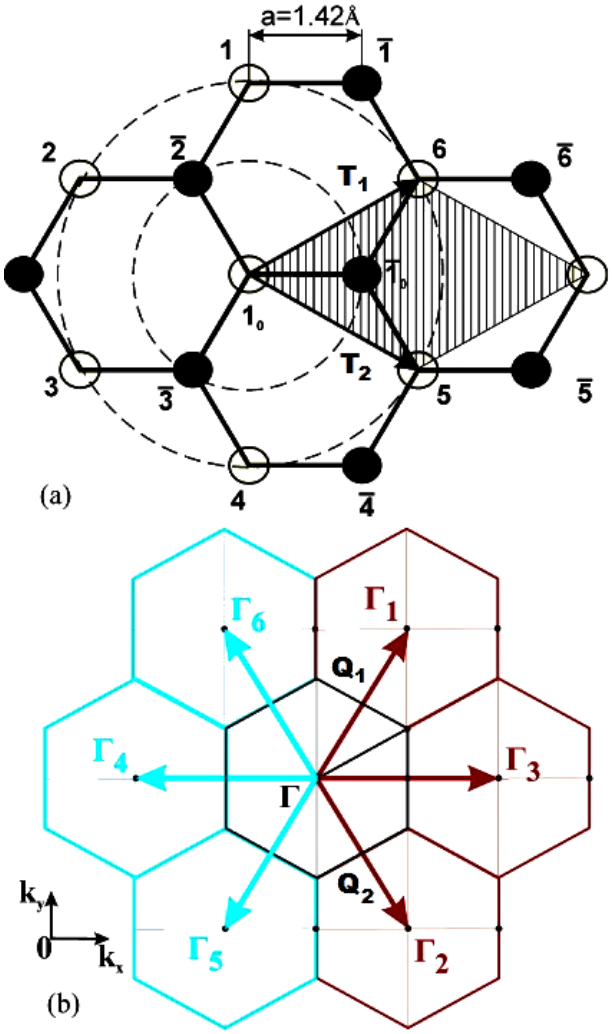


FIG. 12. a: The reciprocal lattice, with vectors $\mathbf{Q}_1$ and $\mathbf{Q}_2$. b: the graphene lattice. The lattice vectors $\mathbf{T}_1$ and $\mathbf{T}_2$ are shown. Atoms in the hexagonal A-sublattice are white, those in the B-sublatice are black (and their numbers overlined). The unit cell, spanned by the lattice vectors, is shown in gray. Nearest neigbour vectors span from 1 to $\bar{1}$, to $\bar{2}$, and to $\bar{3}$. The circles have radii $r_0 = 1.42$Å and $\sqrt{3}\, r_0 = 2.46$Å. (Image taken from D. L. Nika and A. A. Balandin, *J. Phys. Cond. Matter* 24233203 (2012)

The honeycomb lattice can be seen as two interpenetrating hexagonal sublattices A and B (the black and white atoms in figure 12). The unit cell is a rhombus containing a basis of one carbon atom from both the A and B sublattice. The crystal is obtained by translation of this basis by any translation vector $\mathbf{T}$ in the twodimensional set

$$\{\mathbf{T}\} = \{k\mathbf{T}_1 + l\mathbf{T}_2 | k, l \in \mathbb{N}\}$$

where the translation basis vectors are

$$\mathbf{T}_1 = \sqrt{3}\, r_0 \left(\frac{\sqrt{3}}{2}\hat{\mathbf{x}} + \frac{1}{2}\hat{\mathbf{y}}\right)$$

$$\mathbf{T}_2 = \sqrt{3}\, r_0 \left(\frac{\sqrt{3}}{2}\hat{\mathbf{x}} - \frac{1}{2}\hat{\mathbf{y}} \right)$$

with

$$||\mathbf{T}_1|| = ||\mathbf{T}_2|| = \sqrt{3}\, r_0 = 2r_0 \cos 30^o$$

$r_0 \approx 1.42$Å is the graphene carbon-carbon bond length, and $\sqrt{3}\, r_0 \approx 2.46$Å is the lattice constant. The area of the unit cell is

$$\mathrm{A} = ||\mathbf{T}_1 \times \mathbf{T}_2|| = \frac{3\sqrt{3}\, r_0^2}{2}$$

The translation vectors from a carbon atom in the A-sublattice to a nearest neighbour atom in the B sublattice is

$$\{\mathbf{T}_{AB}/r_0\} =$$

$$\left\{ \hat{\mathbf{x}}, -\frac{\hat{\mathbf{x}}}{2} + \frac{\sqrt{3}}{2}\hat{\mathbf{y}}, -\frac{\hat{\mathbf{x}}}{2} - \frac{\sqrt{3}}{2}\hat{\mathbf{y}} \right\}$$

The graphene lattice, with lattice vectors and coordinates of interest is displayed in figure 12 b.

In order to be able to analyse waves - be it electronic, photonic or phononic - running through the graphene lattice, the reciprocal lattice of graphene will have to be defined. Wavevector space (also called reciprocal space, k-space or Fourier space) contains two particular regions of interest

- Lattice vectors in the reciprocal lattice, described by the set of vectors $\{\mathbf{Q}\}$, describe short-wavelength waves with an integer number of wavelengths between atomic planes in the real lattice. Vectors from $\{\mathbf{Q}\}$ will become useful in Fourier problems, like the nearly free electron model.
- Wavevectors corresponding to waves of wavelength (much) longer than the lattice distance are also important. The set of these long-wavelength waves $\{\mathbf{k}\}$ defines the first Brillouin zone.

The reciprocal lattice vectors are defined so that

$$\mathbf{T}_i \cdot \mathbf{Q}_j = 2\pi\delta_{ij}$$

They reciprocal lattice vectors can be calculated as

$$\mathbf{Q}_1 = 2\pi \frac{\mathbf{T}_2 \times \hat{\mathbf{z}}}{\mathrm{A}} = -\frac{2\pi}{3r_0}\left(\hat{\mathbf{x}} + \sqrt{3}\hat{\mathbf{y}} \right)$$

$$\mathbf{Q}_2 = 2\pi \frac{\mathbf{T}_1 \times \hat{\mathbf{z}}}{\mathrm{V}} = \frac{2\pi}{3r_0}\left(\hat{\mathbf{x}} - \sqrt{3}\hat{\mathbf{y}} \right)$$

The entire reciprocal lattice is then defined - analogous to the real space lattice - as

$$\{\mathbf{Q}\} = \{k\mathbf{Q}_1 + l\mathbf{Q}_2 | k, l \in \mathbb{N}\}$$

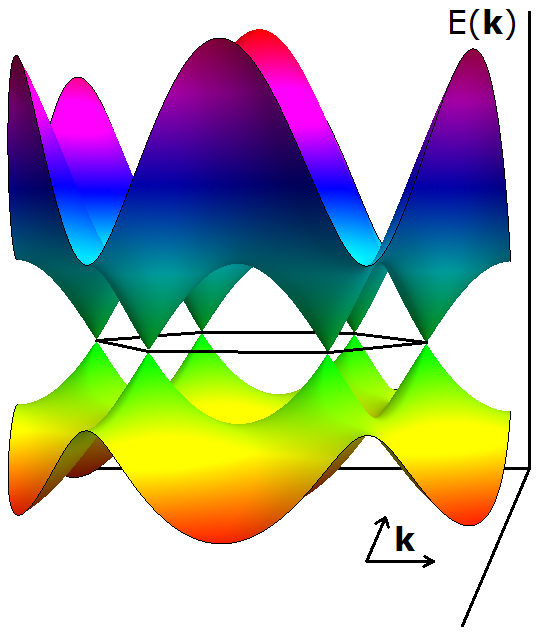


FIG. 13. The dispersion relation E($\mathbf{k}$) for the entire two-dimensional Brillouin zone of graphene, as calculated with a nearest-neighbour periodic tight binding model. The band gap is maximum for electron waves with zero wave vector, and reduces to zero for electrons at the Fermi level, at the Dirac cones.

All points of interest in the reciprocal lattice are displayed in figure 12.

We can now use the tight binding model to derive the band structure of graphene, in the lines of the pioneering work by Wallace[31]. In graphene, a basis of two $p_\perp$-orbitals on the carbon atoms is repeated in a hexagonal lattice - essentially the periodic variant of the ethylene molecule. The basis functions are Bloch waves originating from the atomic positions, that contain a localized part on either sublattice.

$$\chi_A = c_A|A\rangle e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_A)}, \quad \chi_B = c_B|B\rangle e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_B)}$$

In the nearest neighbour-approximation, the coupling

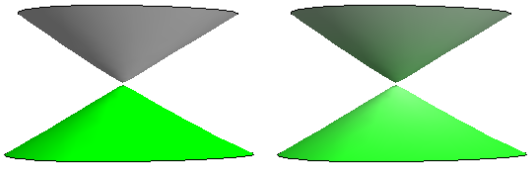


FIG. 14. A zoom-in on Dirac cones of the graphene dispersion. The occupation probability of holes (gray) and electrons (green) is given by the FD-distribution FD(E) and the carrier concentration by the integrals over FD(E)DOS(E). Close to absolute zero, this is a sharp cut-off (left) and the carrier concentration is close to zero. At higher temperatures, some electrons reside in the CB and holes in the VB (right).

and overlap is given by contributions from all three neighbours on the other sublattice, the magnitude of which is determined by the Bloch phase difference. The overlap matrix is

$$\mathrm{S} = \begin{pmatrix} 1 & s\sqrt{f} \\ s\sqrt{f} & 1 \end{pmatrix}$$

where $s = \langle A|B\rangle$ and $\sqrt{f}$ is the sum of phase factors from all neighbours

$$\sqrt{f} = e^{-i\mathbf{k}(\cdot\mathbf{r}-\mathbf{r}_A)} \sum_B e^{i\mathbf{k}\cdot\mathbf{r}-\mathbf{r}_B} = \sum_B e^{i\mathbf{k}\cdot\mathbf{T}_{AB}}$$

This phase arises in the same way in the Hamiltonian matrix. The nearest-neighbour vectors were shown in chapter 1. The sum becomes

$$\sqrt{f} = e^{i\mathbf{k}\cdot r_0\hat{\mathbf{x}}} + e^{i\mathbf{k}\cdot r_0(-\hat{\mathbf{x}}+\sqrt{3}\hat{\mathbf{y}})/2} + e^{i\mathbf{k}\cdot r_0(-\hat{\mathbf{x}}-\sqrt{3}\hat{\mathbf{y}})/2}$$

and can be written in terms of the translation vectors as

$$f = 3 + 2\cos\mathbf{k}\cdot\mathbf{T}_1 + 2\cos\mathbf{k}\cdot\mathbf{T}_2 + 2\cos\mathbf{k}\cdot(\mathbf{T}_2 - \mathbf{T}_1)$$

The secular equations in single-electron form are

$$(\mathrm{H} - \mathrm{S}E_{ii})\mathbf{c} = \begin{pmatrix} -E & (\beta - sE)\sqrt{f} \\ (\beta - sE)\sqrt{f} & -E \end{pmatrix}\mathbf{c} = \mathbf{0}$$

The determinant gives the characteristic parabolic equation $E^2 = (\beta - sE)^2 f$, which is solved to give

$$E \mp sE = \mp\beta\sqrt{f} \Rightarrow E = \frac{\mp\beta\sqrt{f}}{1 \mp s\sqrt{f}}$$

and the eigenvectors are

$$\mathbf{c}_g = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mathbf{c}_u = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

which are the same eigenenergies and orbital coefficient expansions as ethene, except for the function $\sqrt{f(\mathbf{k})}$, which introduces a band structure. The full band structure is plotted in figure 13. It is shown in appendix D that this dispersion relation is indeed linear in the K-points at the Fermi level.

Figure 14 shows a zoom on the Dirac cones, which are the regions in the band structure where electron transport takes place. As described in appendix A, the conductivity is a function of charge carrier mobility and charge carrier concentration. The density of states can formally be calculated by integrating the number of $\mathbf{k}$-vectors for each energy value interval and subsequently taking the derivative with respect to energy

$$E(\mathbf{k}) \Leftrightarrow \mathbf{k}(E) \qquad N(E) = \iint \mathbf{k}(E)dE$$

$$\mathrm{DOS}(E) = \frac{\partial N}{\partial E}$$

Around the Fermi level, the density of states - like the dispersion - turns out to be linear. The charge carrier concentration can now be found (analytically) by integrating the density of states over the thermal occupancy probability (given by the Fermi-Dirac distribution). The occupations are shown schematically for 0 K and for elevated temperature in figure 14, and the procedure for evaluating the conductivity integral is shown in appendix A. Using these tools, it can now be appreciated that the vanishing density of states at the Fermi level gives a conductivity that approaches zero at zero temperature.

THE MOLECULAR MECHANICS METHOD

Molecular dynamics (MD) is a computational method to evaluate the dynamical properties of molecules and between molecules by numerically integrating the forces they experience in time[32]. These forces are approximated by some predefined interactions, like quadratic bond potentials and inverse quadratic Coulomb interactions. Molecular mechanics (MM) can be seen as a subset of molecular dynamics, focused on computing or obtaining molecular geometries by running a molecular dynamics simulation using mainly intramolecular forces. There is no real distinction between the two terms, but in general, the term MD is used more often for intermolecular problems and MM more for intramolecular problems.

Although MD methods are inherently approximations of the real dynamic nature of molecules - because they simply don't capture the quantum nature of molecules - they are known to work well for large systems[33]. The use of MM has been particularly popular in calculating equilibrium geometries of macromolecules like proteins, and numerous successes have been achieved using this method[34][35]. Using an inherently classical method on molecular systems seems rather crude, but quantum chemical simulations may actually give potential energy landscapes which often turn out to be very similar to those used in MM. Alternatively, molecular mechanics force field can also be calibrated with quantum chemical simulations[36].

A classical parametrization of intramolecular interaction energies consists of contributions from bond lengths, bond angles, dihedral angles and electrostatic interactions. Given an equilibrium bond length r_0, the energy of a covalent bond between atoms i and j is given by

$$E = \frac{1}{2}k_b(r_{ij} - r_0)^2$$

If atom i and atom k are both covalently bonded to atom j, they will experience an interaction energy when the angle θ_{ijk} deviates from the equilibrium bond angle. The quadratic potential for bond angle variations takes the form of

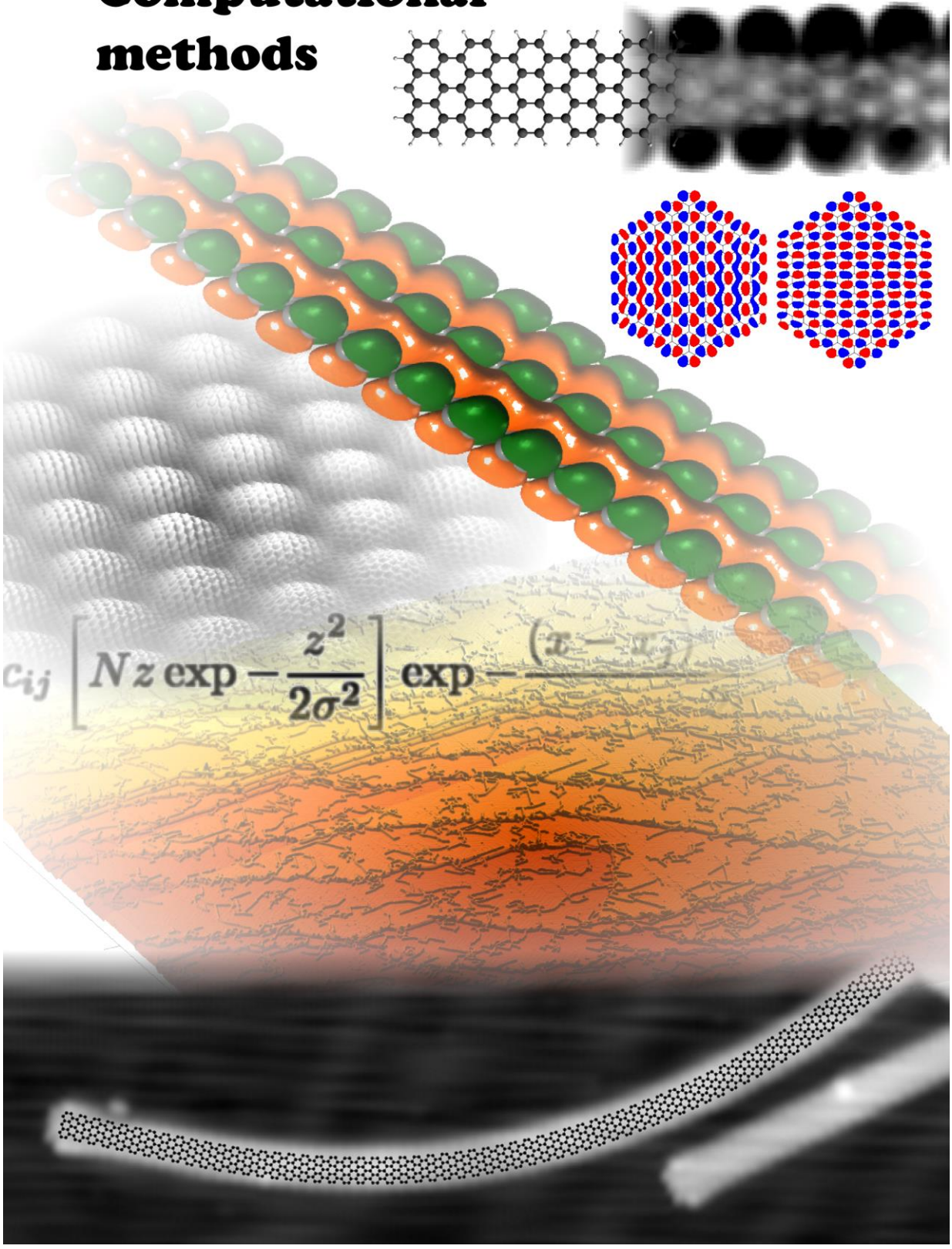
$$E = \frac{1}{2} k_\phi (\theta_{ijk} - \theta_0)^2$$

If the MM model is extended further, dihedral angles should also be taken into account. The dihedral angle ϕ_{ijkl} is the angle between the plane spanned by atoms i, j and k and the atoms j, k and l. Additionally, electrostatic terms and van der Waals terms for non-bonded atoms are taken into account. As the subscripts suggest, the number of atoms to take into account increases strongly by taking into account dihedral contributions. However, for this research, we only considered in-plane bending, conveniently rendering all dihedral angles zero.

It will be shown in the next chapter how a molecular mechanics "force field" - basically the set of forces resulting from the aforementioned interaction energies - can be implemented computationally. Then, this method is used to create bent nanoribbon geometries. By combining the two methods - performing tight binding calculations on geometries prepared using the molecular mechanics method - we will be able to make predictions on the changes in the electronic structure of graphene nanoribbons when they are bent.

[1] G. Binnig, Applied Physics Letters **40**, 178 (1982).
[2] G. Binnig, "Atomic force microscope and method for imaging surfaces with atomic resolution," (1988), uS Patent 4,724,318.
[3] M. Crommie, C. Lutz, and D. Eigler, Science **262**, 218 (1993).
[4] L. van Hove, Physical Review **89**, 1189=1193 (1952).
[5] R. Eisenschitz and F. London, Zeitschrift fur Physik **60**, 491 (1930).
[6] F. London, Trans. Faraday Soc. **33**, 8b (1937).
[7] J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).
[8] U. Fano, Physical Review **124**, 1866 (1961).
[9] W.-H. Soe, C. Manzano, A. De Sarkar, N. Chandrasekhar, and C. Joachim, Phys. Rev. Lett. **102**, 176102 (2009).
[10] J. van der Lit, M. P. Boneschanscher, D. Vanmaekelbergh, M. Ijäs, A. Uppstu, M. Ervasti, A. Harju, P. Liljeroth, and I. Swart, Nature communications **4**.
[11] J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985).
[12] W.-H. Soe, C. Manzano, H. S. Wong, and C. Joachim, Journal of physics. Condensed matter : an Institute of Physics journal **24**, 354011 (2012).
[13] D. K. Schweizer and E. K. Eigler, Nature **344**, 524 (1990).
[14] I. Swart, T. Sonnleitner, J. Niedenführ, and J. Repp, Nano letters **12**, 1070 (2012).
[15] G. Meyer, J. Repp, S. Zphel, K.-F. Braun, S. W. Hla, S. Flsch, L. Bartels, F. Moresco, and K. H. Rieder, Single Molecules **1**, 79 (2000).
[16] L. Lafferentz, F. Ample, H. Yu, S. Hecht, C. Joachim, and L. Grill, Science **323**, 1193 (2009).
[17] M. Koch, F. Ample, C. Joachim, and L. Grill, .
[18] M. Alemani, M. V. Peters, S. Hecht, K.-H. Rieder, F. Moresco, and L. Grill, Journal of the American Chemical Society **128**, 14446 (2006).
[19] Y. Jiang, Q. Huan, L. Fabris, G. C. Bazan, and W. Ho, Nature chemistry **5**, 36 (2013).
[20] R. Pariser and R. G. Parr, The Journal of Chemical Physics **21**, 466 (1953).
[21] S. Hammes-Schiffer and H. C. Andersen, The Journal of Chemical Physics **101** (1994).
[22] P. J. Hasnip, K. Refson, M. I. J. Probert, J. R. Yates, S. J. Clark, and C. J. Pickard, Royal Society of London Philosophical Transactions Series A **372**, 30270 (2014).
[23] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[24] P. Bou, Journal of Computational Chemistry **21**, 8 (2000).
[25] E. Hückel, Zeitschrift fur Physik **70**, 204 (1931).
[26] Y. Li, X. Jiang, Z. Liu, and Z. Liu, Nano Research **3**, 545 (2010).
[27] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejón, Physical Review B **66**, 035412 (2002).
[28] A. H. Castro Neto, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, "The electronic properties of graphene," (2009).
[29] H. Raza and E. Kan, Physical Review B , 1 (2008).
[30] P. Abbamonte, J. P. Reed, B. Uchoa, Y. I. Joe, E. Fradkin, and D. Casa, in *APS Meeting Abstracts* (2010) p. 21001.
[31] P. R. Wallace, Phys. Rev. **71**, 622 (1947).
[32] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **31**, 459 (1959).
[33] B. Schwenzer, S. N. Kerisit, and M. Vijayakumar, RSC Adv. **4**, 5457 (2014).
[34] M. Karplus and J. A. McCammon, Nature structural biology **9**, 646 (2002).
[35] X. Xu, L. F. C. Pereira, Y. Wang, J. Wu, K. Zhang, X. Zhao, S. Bae, C. Tinh Bui, R. Xie, J. T. L. Thong, B. H. Hong, K. P. Loh, D. Donadio, B. Li, and B. Özyilmaz, Nature communications **5**, 3689 (2014).
[36] T. Zhu, X. Xiao, C. Ji, and J. Z. H. Zhang, Journal of Chemical Theory and Computation **9**, 1788 (2013).

Computational methods



MOLECULAR MECHANICS ON GRAPHENE NANORIBBONS

The molecular mechanics interaction energies shown in the previous chapter - parabolic wells for the bending angle and bond length - look trivial, but knowing the energies is not sufficient to implement a MM simulation. Rather, they need to be transformed to vectorial forces on all atoms. A molecular mechanics force field was written to prepare bent and strained geometries for AGNRs.

The vector from atom i to atom j is

$$\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$$

and its corresponding unit vector is

$$\mathbf{e}_{ij} = \frac{\mathbf{r}_{ij}}{||\mathbf{r}_{ij}||}$$

The bond length (spring) force on atom i due to atom j is in the direction from i to j and can be calculated as

$$\mathbf{F}_i(j) = k_b(||\mathbf{r}_{ij}|| - r_0)\mathbf{e}_{ij}$$

Due to momentum conservation, the same force is experienced by atom j due to atom i

$$\mathbf{F}_j(i) = -\mathbf{F}_i(j) = -k_b(||\mathbf{r}_{ij}|| - r_0)\mathbf{e}_{ij}$$

For bond angles, the force field is slightly more involved. θ is taken to be the bond angle between $\mathbf{e}_{ji}$ and $\mathbf{e}_{jk}$, centered on atom j. It is calculated according to

$$\cos(\theta) = \frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{kj}}{||\mathbf{r}_{ij}||||\mathbf{r}_{kj}||} = \mathbf{e}_{ij} \cdot \mathbf{e}_{kj}$$

So the magnitude of the force is given by

$$||\mathbf{F}_{i/k}|| = -||\nabla E|| = -k_\theta||\nabla(\theta - \theta_0)^2||$$

$$= \frac{k_\theta}{r}\frac{\partial(\theta - \theta_0)^2}{\partial\theta} = \frac{2k_\theta(\arccos[\mathbf{e}_{ij} \cdot \mathbf{e}_{kj}] - \theta_0)}{r_{ij/kj}}$$

Atom i rotates around atom j in a direction perpendicular to $\mathbf{e}_{ji}$. The rotation of atom i and k is in the direction defined by a unit vectors tangent to the "circle" on which the atoms rotate. The unit vectors for two rotational directions (clockwise and couterclockwise) are given by

$$\mathbf{e}'_i = \pm[(\mathbf{e}_{ji} \cdot \hat{\mathbf{x}})\hat{\mathbf{y}} - (\mathbf{e}_{ji} \cdot \hat{\mathbf{y}})\hat{\mathbf{x}}]$$

and similarly for $\mathbf{e}_{jk}$. Whether the rotational direction should be clockwise or counterclockwise for i (and vice versa for k) is now accounted for by taking the sign of the cross product between $\mathbf{e}_{ij}$ and $\mathbf{e}_{kj}$, that is, the orientation of its normal vector

$$s_{ijk} = \frac{\mathbf{e}_{ij} \times \mathbf{e}_{kj}}{||\mathbf{e}_{ij} \times \mathbf{e}_{kj}||}$$

The forces on i and k are now calculated by combining latter three equations

$$\mathbf{F}_i = ||\mathbf{F}_{i,k}||s_{ijk}\mathbf{e}'_i, \quad \mathbf{F}_k = -||\mathbf{F}_{i,k}||s_{ijk}\mathbf{e}'_k$$

To conserve momentum, the force experienced by j by the bending of i and k should be

$$\mathbf{F}_j = -\mathbf{F}_i - \mathbf{F}_k$$

This ensures that, if atom i and k rotate towards each other, atom j rocks back slightly, keeping the center of mass of the three atoms at the same position.

After loading an initial geometry, the forces on all atoms are calculated. A small timestep δt is chosen, after which the forces, velocity changes and coordinate changes are calculated. Numerical time integrations are subject to numerical noise, caused by the finite timestep size and rounding errors. The leapfrog algorithm was devised to numerically integrate MD/MM problems in time, with an increased stability with respect to the Euler algorithm[1]. After setting an initial geometry $\mathbf{r}_0$ and setting $\mathbf{v}_0 = 0$, the order in which the forces, velocities and positions are updated is

- The velocities are updated a half time step using the previous accelerations (which at time zero are 0, but in general, at time t_0 are given by $\mathbf{a}(t_0)$).

$$\mathbf{v}(t_0 + \frac{1}{2}\delta t) = \mathbf{v}(t_0 - \frac{1}{2}\delta t) + \mathbf{a}(t_0)\delta t$$

- The positions are updated to time $t_0 + \delta t$ using the velocities.

$$\mathbf{r}(t_0 + \delta t) = \mathbf{r}(t_0) + \mathbf{v}(t_0 + \frac{1}{2}\delta t)\delta t$$

- The accelerations are computed at a time $t_0 + \delta t$ (j are the force components due to all strains and bends of all atoms around atom i).

$$\mathbf{a}_i(t_0 + \delta t) = \sum_j \frac{\mathbf{F}_{ij}(\mathbf{r}_i(t_0 + \delta t))}{m}$$

This algorithm effectively ensures that the velocities are always determined by the average force between time t_0 and time $t_0 + \delta t$, so they are approximately determined by $\mathbf{a}(t_0 + \frac{1}{2}\delta t)$.

$$\Rightarrow \mathbf{v}(t_0 + \delta t) = \mathbf{v}(t_0) + \frac{\mathbf{a}(t_0) + \mathbf{a}(t_0 + \delta t)}{2}\delta t$$

The input geometry in the nanoribbon problems is the straight, relaxed hexamer ribbon. The position of the center carbon atoms on both termini are constrained: forced to move slowly towards each other. During this process of "pinching", the ribbon will wiggle and bounce.

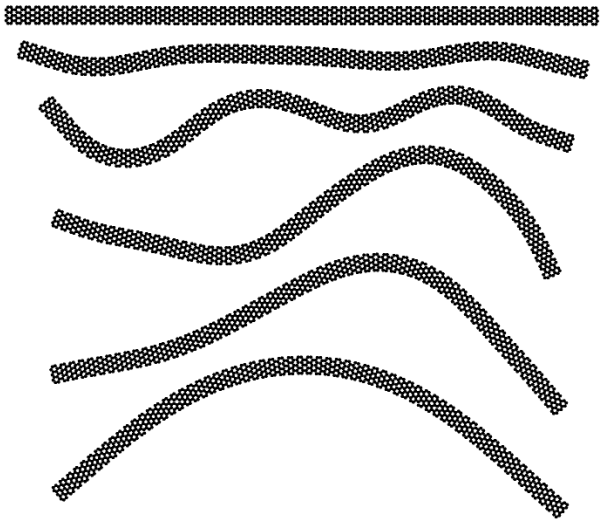


FIG. 1. Stills from a molecular mechanics simulation. Two atoms on either side of the ribbon are constrained and pinched towards each other. The ribbon is stressed, then it bends, wiggles, and relaxes.

Pinching introduces a steady increase in potential energy, which is quickly absorbed as kinetic energy of the atoms. In a proper MM simulation, a molecule will keep wiggling and bouncing forever. Therefore, the ribbon is forced to relax by perturbing the position update with a damping term

$$\mathbf{r}(t_0 + \delta t) = \mathbf{r}(t_0) + \mathbf{v}(t_0 + \frac{1}{2}\delta t)\delta t + 0.1\mathbf{a}(t_0)\delta t dt$$

where dt is the time corresponding to a step of unity in atomistic unit. Also, every number of cycles, the atomic velocity vectors are scaled by a factor between 0.95 and 0.99. Using these damping mechanisms, pinching and waiting for long enough, bent geometries have been obtained for tight binding simulations. A pinching simulation is shown in figure 1.

For molecular mechanics on graphene, it is generally found that bonds are much stiffer than bond angles[2][3][4]. Good values of k_b and k_θ appear to be around

$$k_b = 60 \text{ nN Å}^{-1}$$

$$k_\theta = 8 \text{ nN Å rad}^{-2}$$

where only the ratio of the parameters matters for the equilibrium result. Figure 2 shows how bending can in general be compensated in different ways, proving the importance of a correct geometrical description. The following parameter ratios were used in the generation of these geometries, respectively: $k_b/k_\theta = 100$ Å^{-2} rad^{-2} and $k_b/k_\theta = 0.01$ Å^{-2} rad^{-2}. In the angle-preserving regime, bonds are severely shortened on the inside of the bend, and elongated on the outside. In the bond length-preserving regime, the angles are deformed in such a way that the benzene rings on the inside are almost pushed into rectangles, whereas the outer benzene rings stretch out. According to the parameters obtained from molecular mechanics on bulk graphene, latter scenario might be more realistic.

FIG. 2. Bending can be compensated by angles, bonds or a combination of the two. These rather extreme examples show the limiting scenarios, generated by performing a MM simulation with parameter ratios of 0.01 and 100 Å^{-2} rad^{-2} respectively.

TIGHT BINDING ON GRAPHENE NANOSYSTEMS

Parametrization

The first issue to address is how to account for the bending-induced distortions in the nanoribbon in tight-binding. Obviously, assigning fixed parameters for the nearest and next-nearest neigbour interaction integrals cannot capture any electronic structure changes in bending. The parameters themselves should somehow be a function of the interatomic separation.

Since the interaction parameter is related to the interaction integral between two $p_\perp$-orbitals (explicitly so in Hartree Fock and Density Functional Theory), it should have a similar distance dependence as the actual integral. The same goes for the overlap integral. At large distances, the values should for example approach zero. A natural choice to model this function is to use the exponential decay[6][7]. An additional feature of such a function is that it should also be able to give reasonable values for the next- and second next-nearest neighbour parameters. For his work on strained graphene sheets, Pereira used the function

$$\mathrm{H}_{ij} = \mathrm{H}_0 e^{-\kappa\left(\frac{r_{ij}}{r_0}-1\right)}$$

with $\kappa = 3.37$, an "interaction range parameter"[7]. In stead of defining parameters for multiple neighbour distances, there are now only two parameters. Subsequent work by Ribeiro *et al.* was performed to look at κ in more detail. They carried out DFT calculations on strained graphene sheets, where the strain was applied in multiple directions, and tried to find the correct tight-binding parameters to replicate the distortions[8]. The interaction parameter was then fitted to the exponentially decaying function, giving a value of κ in the range $2.6-4$ for multiple directions. Although neither Pereira nor Ribeiro took overlap into account, it is assumed here that the overlap scales in the same way.

$$\mathrm{S}_{ij} = \mathrm{S}_0 e^{-\kappa\left(\frac{r_{ij}}{r_0}-1\right)}$$

For S_0 around 0.1, this appears to be in relatively good agreement with one of the parameter sets used by Hancock *et al.*[9] and Kundu *et al.*[10]. The parameters were first chosen to be in relatively good agreement with earlier works, meaning for example that the value H_0 should not differ too much from -2.7 meV to -3.1 meV. Tuning the parameters within this range was found to have only minor effects on the DOS around the band gap, although the antisymmetrizing effect of the overlap is accompanied by a small increase in band gap. The effect in H_0 is linear in the band gap (the band gap is roughly $\mathrm{H}_0/2$) and the interaction range parameter κ does not affect the band gap too much.

Implementation

The tight-binding machinery as described in the previous chapter, has been implemented in a tight-binding program, which was written in Mathematica. As input, both a molecular geometry or a molecular dynamics file containing geometry "frames" can be read. First the program detects which atoms constitute a frame, distinguishes carbon atoms from hydrogen, finds and enforces the molecular plane and centers the molecule at the origin. Then the program loops over all carbon atom pairs, while assigning the correct values of H and S as a function of distance. The orthogonal transformation is performed to give H', and the secular equations are solved by diagonalization of H'. The result is the matrix C', which is backtransformed to C, and E, the diagonal eigenenergy matrix. A check ensures whether the secular equations are indeed solved, by looking at the residual matrix HC-SCE[11]. The program outputs the list of energy eigenvalues $\vec{\epsilon}$ and writes the coefficient matrix to a file. The clean geometry and energies are written to a file as well.

For molecular orbital visualization purposes, a minimal single-Gaussian basis set is used. The orbital number i and a grid size are requested, after which the program calculates $|\phi_i(\mathbf{r})|^2$ over all grid positions $\mathbf{r}$ in 3D space. The details of this process will be discussed in the section on AFM/STM experiment simulations later in this chapter. The values $|\phi_i(\mathbf{r})|^2$ can be written to a Gaussian cube file, after which it can be imported into orbital visualization programs, like Avogadro or Molekel.

Before carrying out a tight binding simulation, a DFT calculation was performed on a hexamer nanoribbon[12]1. The Perdew Becke Ernzerhof exchange correlation functional (PBE) was used within the generalized gradient approach (GGA-PBE), and the basis set was double zeta valence (DZV) set. A geometry optimization was performed in D2h symmetry, with the molecule in the yz-plane.

DFT finds a band gap of 2.06 eV between a doubly occupied state and a doubly unoccupied state. Halfway, two almost degenerate states are found: the HOMO and the LUMO. An orbital plot of the HOMO and LUMO shows a symmetric and an antisymmetric state with respect to the mirror plane halfway the ribbon, respectively, localized on both sides of the nanoribbon. Obviously, these will never be visible in periodic calculations - an infinite nanoribbon has no ends. All states belonging to the π-system were extracted from the molecular orbital energy levels by choosing the levels with the correct antisymmetry with respect to the yz-plane. These are the Au, B1g, B2g and B3u irreducible representations respectively. The resulting eigenenergies are broadened into a pDOS (density of states spanned by the $p_\perp$-orbitals). The pDOS and the orbital plots of the "gap states" are shown in figure . Notice that the gap states are not re-

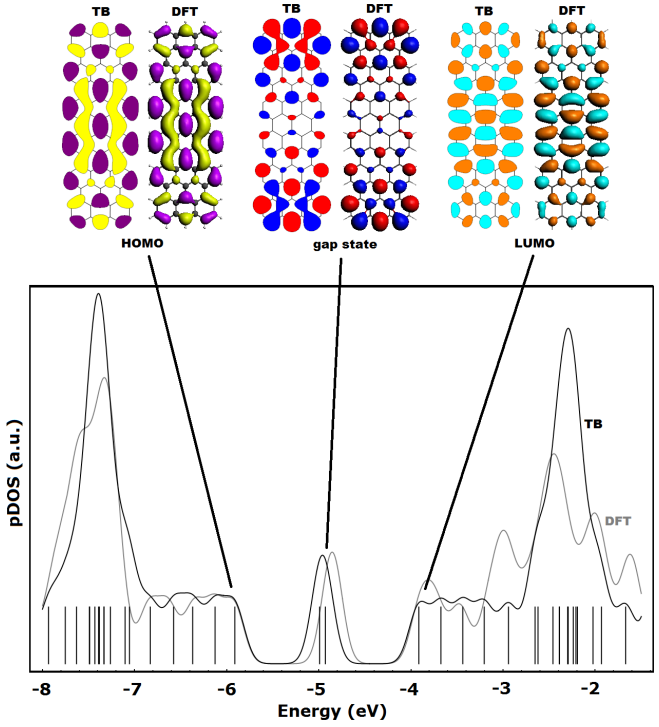


FIG. 3. The density of states and the frontier orbitals of a hexamer nanoribbon, calculated with tight binding and density functional theory

n-mer	1	2	3	6	12
band gap (eV)	3.70	2.58	2.01	1.47	1.37
end state gap (eV)	0.876	0.203	0.059	0.002	$\approx 10^{-5}$

TABLE I. Energy gaps in graphene nanoribbons as a function if length, calculated with tight binding

ferred to as HOMO and LUMO. The reason for doing so is explained later.

Now a tight bindind calculation is carried out on the same hexamer ribbon. It was found that the orbitals of the states around the band gap are visually in extremely good agreement with DFT. The same near-degenerate symmetric and antisymmetric end-localized states are encountered, and the orbitals of the first states on either side of the gap also compare very well. For this calculation, the DOS was also generated by broadening the energy levels, and a fair agreement was found to the DFT pDOS. The tight binding parameters were now calibrated with the DFT results by comparing the overlap of the two plots of the DOS. The best agreement was obtained by setting $H_0 = -2.95$ eV, $S_0 = 0.085$ and $\kappa = 3.9$. The orbitals and DOS from both tight binding and DFT are displayed in figure .

More tight binding calculations were performed on ribbons with different lengths. The band gap is defined as the energy difference between the eigenstates on either side of the energy gap, not counting the end-localized states. Table I shows the band gap, as well as the energy gap between the two end states, as function of length. The band gap converges to a value around 1.3 eV as the ribbon size increases. The gap between the end states quickly becomes negligible. The Fermi-Dirac distribution tells us that if the energy difference of the gap states is small compared to the thermal energy, they may well both be singly occupied. In fact, this may actually be favored by spin interaction, as is well-known from the Aufbau principle (and for example high-spin inorganic systems).

It was noted in the previous chapter that a pair of degenerate eigenstates can be rotationally mixed to a new pair of orthonormal eigenstates. Such a rotational transformation makes sense when the theory that fails to account for effects that would benefit from such a mixing. In the tight binding model, no explicit electron-electron repulsion term was present (in fact, no two-electron matrix at all). Assuming that the degeneracy of the end states is good enough for ribbons of sufficient length, a rotation by 45 degrees can be performed. The rotational transformation then gives

$$|\text{gap}_1\rangle' = \frac{1}{\sqrt{2}}\left(|\text{gap}_1\rangle + |\text{gap}_2\rangle\right)$$

$$|\text{gap}_2\rangle' = \frac{1}{\sqrt{2}}\left(|\text{gap}_1\rangle - |\text{gap}_2\rangle\right)$$

It turns out that $|\text{gap}_1\rangle'$ and $|\text{gap}_2\rangle'$ are localized on one side of the ribbon only (the left and right side respectively). These states therefore leave the electrons on either side of the ribbon - far away from each other, which is favourable with respect to electron-electron interaction.

The symmetric and antisymmetric combinations were found in DFT calculations as well. However, this is

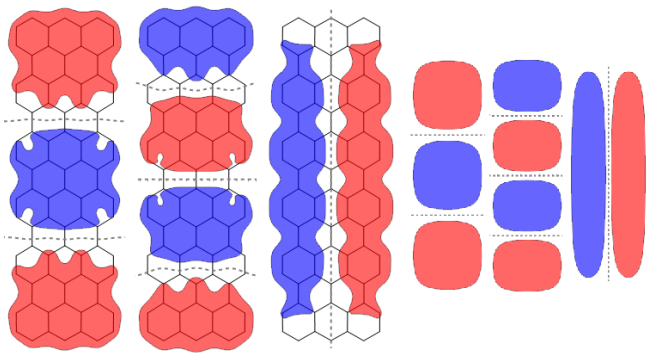


FIG. 4. Left: The 3rd, 4th and 5th lowest energy eigenfunctions on a 3-mer AGNR, as calculated with the TB method. The nodal patterns compare extremely well to the 3rd, 4th and 5th lowest standing waves in a "particle-in-a rectangular 2D box" model, as shown on the right

an artefact of the applied symmetry - states localized on either side of a nanoribbon simply are not symmetric with respect to a reflection to the ribbon's other end. Upon performing the calculation with symmetry turned off and a singlet ground state, physical separation of the end states was achieved. In later tight binding simulations on longer systems, the single end-localized states were sometimes encountered automatically. The energy difference is probably lower than the numerical precision of the calculation, returning either the symmetric and antisymmetric gap states or the end-localized states. The presence of end-localized states on the zigzag termini of nanoribbons is well established in literature[13][14][15][16][17].

Figure 4 shows the 3rd, 4th and 5th eigenstate of π-electrons on a graphene nanoribbon. Interestingly enough, there is an extremely good comparison to the 3rd, 4th and 5th lowest standing waves in a "particle-in-a rectangular 2D box" model. This should come as no surprise. Remember that the Bloch waves in the solid state are $u(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$, where the spatial component $u(\mathbf{r})$ was found to correspond to (a combination of) atomic orbitals. In the atomistic approach, the atomic $\vec{\chi}$ orbitals were put into the model (either explicitly or implicitly). Now the expansion coefficients $\mathbf{c}$ should fulfill the role of a standing waves across the molecule.

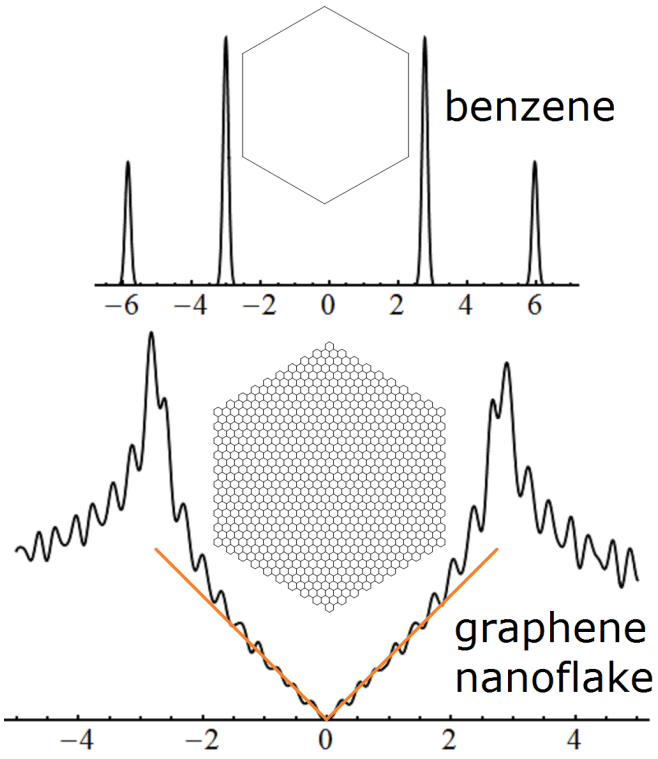


FIG. 5. The calculated density of states for a benzene molecule, and a graphene nanoflake. Upon increasing the system size, the density of states automatically obtains the characteristic linear dispersion regime.

Benchmarking

To gain more confidence in the tight binding model, a few calculations were performed to see whether the model accurately describes well-known results from experiment and theory. The focus has for the most part been on size effects. These calculations are shown in this paragraph.

The smallest "graphene nanosystem" is arguably benzene. A tight binding calculation yields six π states, two pairs of which are degenerate. Upon extending the nanosystem towards a graphene nanoflake, an increasing amount of energy levels is obtained, and by broadening them into a density of states, plot is obtained where a "valley" is flanked by two peaks at the positions of the degenerate states. If the system is increased sufficiently enough, the valley turns out to get a linear shape, as seen in figure 5. The appearance of the DOS quickly begins to resemble the DOS of periodic graphene[8]. This suggests that the finite tight-binding model extrapolates well to periodic systems.

As explained in the previous chapter, the band gap energy for a parabolic dispersion material is inversely proportional to its dimensions

$$\Delta E \propto x_0^{-1}$$

For an isotropic twodimensional material, the confinement in a perpendicular direction must have the same proportionality and

$$\Delta E \propto A^{-1}$$

where A is the area. Any fingerprint for a linear dispersion would then surely show up as a deviation in this proportionality. Tight binding calculations were performed As seen in figure 6, the power dependence was fitted from

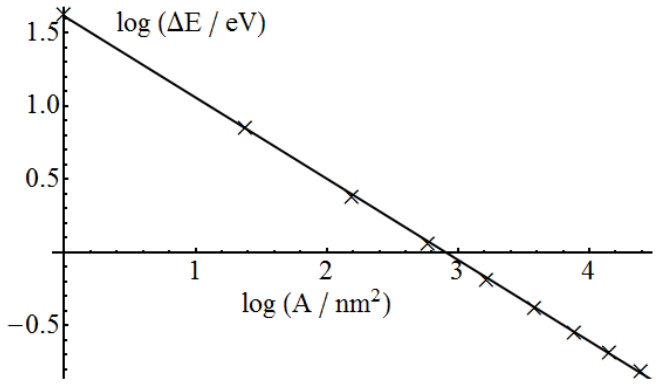


FIG. 6. The band gap of nanoflakes as a function of their area.

the tight binding results as

$$\Delta E = cA^n \Rightarrow \frac{d\log \Delta E}{d\log A} = n$$

It was found that the band gap scales roughly with the square root of the area: $n = 0.555$. This result compares well with scanning tunneling microscopy experiments on graphene nanoflakes, showing the applicability of the Klein-Gordon equation[18].

It was stated in the introduction that the electronic structure of graphene nanoribbons is tightly linked to its atomic structure. The width and "cutting direction" have a large influence on the band gap. To appreciate this effect, we should once again have a look at the twodimensional dispersion relation of graphene. Quantum confining graphene to one dimension means singling out a single direction of the band structure, while discretizing the perpendicular direction. This creates a discrete set of bands, that are mapped onto the one-dimensional band structure[16][19]. Quantum confining graphene into armchair edge nanoribbons often discards all K- points in the discretization of the first Brillouin zone, whereas the zigzag edge nanoribbons do maintain the characteristic touching of the valence band and conductance band. This peculiar result means that the zigzag graphene nanoribbons are metallic, and armchair nanoribbons can be either semiconducting or metallic.

In GNR, lateral confinement alters the electronic structure in a non-uniform way. The band gap of n atom-wide ribbons exist in three families. The $3n$ and $3n+1$ atom ribbons have a finite band gap, that uniformly decreases to zero as n becomes larger. But $3n+2$ atom wide ribbons have a vanishingly small band gap[20][16][21][13]. This result is not unique to tight binding calculations: it

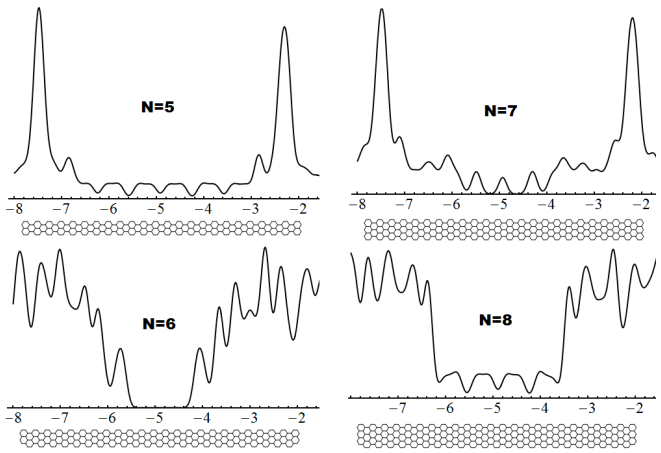


FIG. 7. The density of states as obtained by TB calculations on finite nanoribbons of various widths. The DOS is nonzero at the Fermi level (which is set to -5 eV) in the N=5 and N=8 AGNRs, meaning that the band gap is zero or close to zero. Furthermore, it is seen that the in-gap states do not exist on N=6 AGNRs.

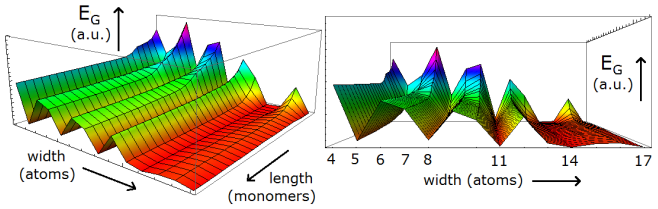


FIG. 8. The band gap of graphene nanoribbons as function of width and length. For short ribbons, the extra longitudinal confinement increases the band gap. As the length increases, the band gaps quickly approach the periodic result.

has been reproduced with DFT as well[22][15]. They are effectively metallic! This strange behaviour was reproduced in the calculations, as shown for example in the density of states plots for ribbons of different widths in figure 7. For the 5- and 8-atom wide GNR, the density of states does not vanish in the middle of the band. In addition, longitudinal confinement can increase the gap even further, but it does so in a uniform way. A plot of calculated band gap versus length and width is shown in figure 8. The longitudinal confinement is lost relatively quickly, and the pattern semiconducing - semiconducting - metallic remains for the "quasi-infinite" case.

Simulating AFM and STM experiments

It was explained in the previous chapter how differential conductance mapping can generate images of molecular orbitals in real space, as long as a correct bias voltage and oscillation voltage is chosen that singles out a state. The tight-binding theory allows the molecular orbitals to be calculated, and can therefore also provide a simulation of STM experiments. After a calculation, the molecular geometry is ammended with the eigenenergy matrix E and the coefficient matrix C. These are the three ingredients to simulate the STM maps, using the Tersoff-Haman approach[23][13].

Every wave function was described by an eigenvector in C as $\psi_i = \mathbf{c}_i \cdot \vec{\chi}$. As basis functions, the minimal Gaussian basis set is again employed, with

$$\chi(x,y,z) = Nz\exp\left(-\frac{x^2+y^2+z^2}{2\sigma^2}\right)$$

with N a normalization factor and σ the orbital size. The wave function is

$$\psi_i = \sum_j c_{ij}\left[Nz\exp\left(-\frac{z^2}{2\sigma^2}\right)\right]\times$$

$$\exp\left(-\frac{(x-x_j)^2+(y-y_j)^2}{2\sigma^2}\right)$$

where the z-dependence can be split out of the Gaussian by its cartesian separability. The summation runs over

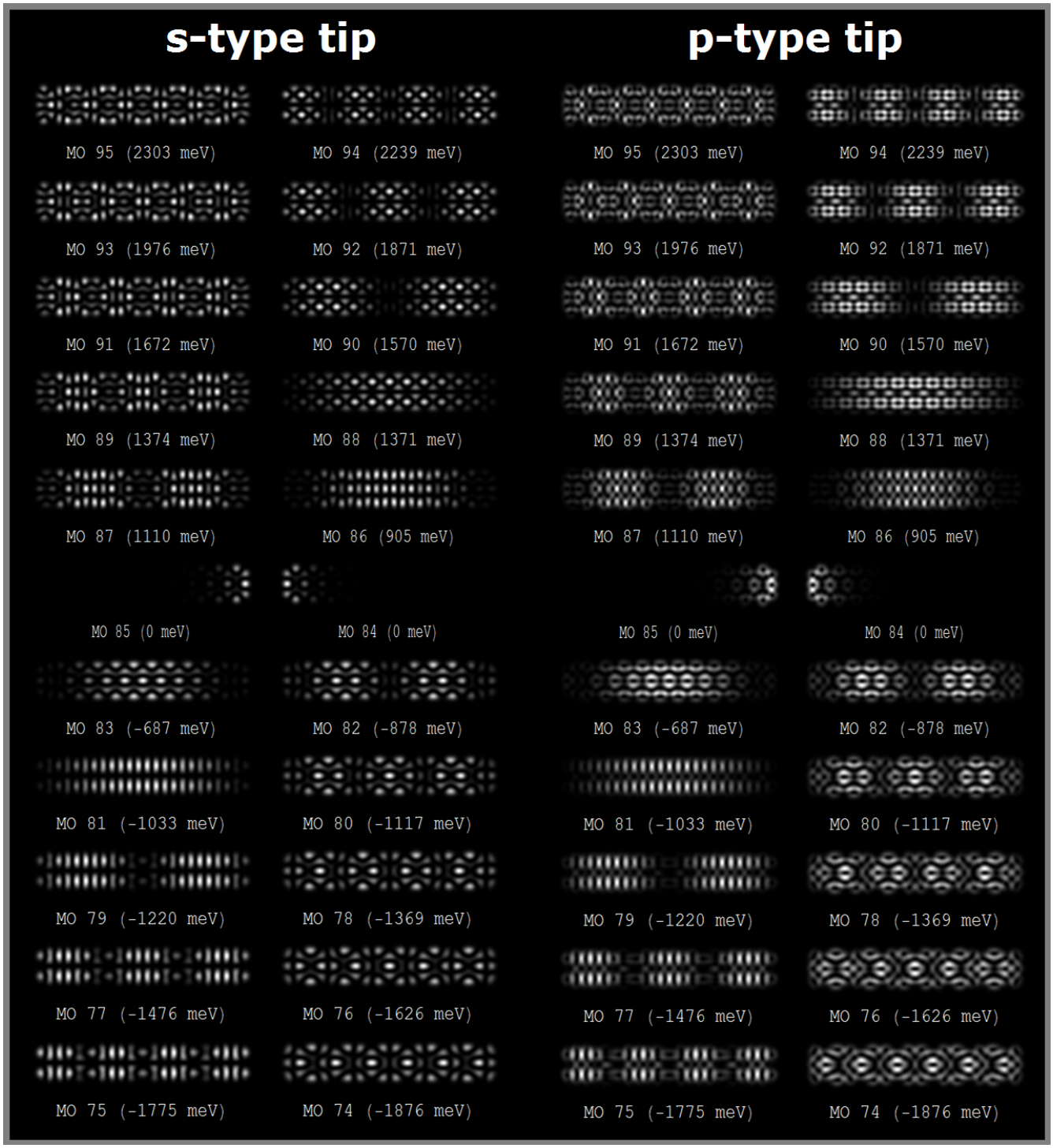


FIG. 9. STM simulations of nanoribbon eigenstates with a sharp metallic tip (s-type) and a carbon monoxide terminated tip (p-type). Parameters: H_0 = -2.8 eV, $S_0 = 0.2$, $\kappa = 3$, $\sigma = 0.8$ nm.

all atomic orbitals j, and their corresponding z_j-values are 0 since the molecule lies in the xy plane. This means that the resulting factor $Nz\exp(-z^2/2\sigma^2)$ can be taken outside of the summation.

$$\psi = \left[Nz\exp-\frac{z^2}{2\sigma^2}\right]\times$$

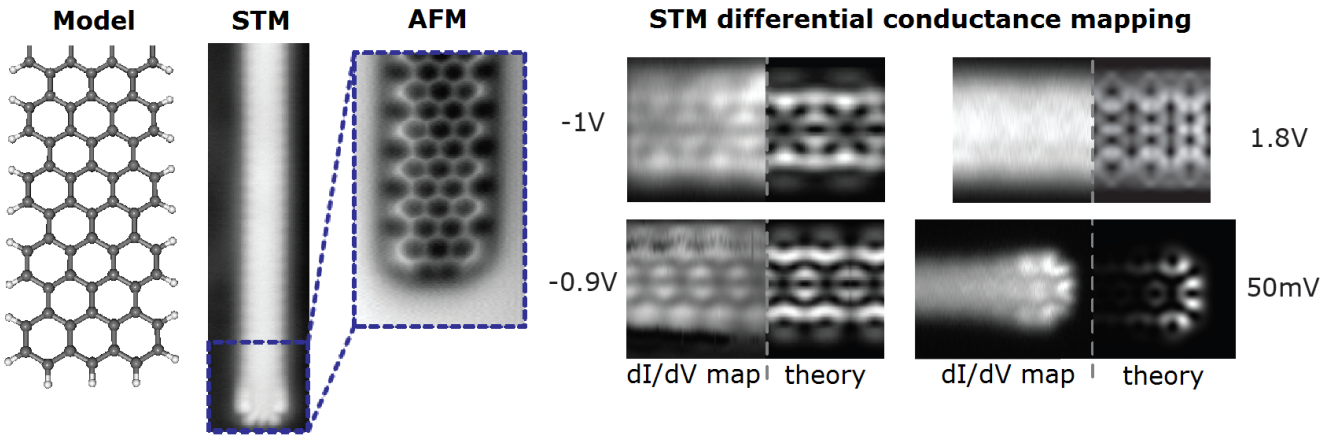


FIG. 10. An overview of characterizations on graphene nanoribbons, where the differential conductance maps have been amended with predictions from theory. Experiments by van der Lit *et al.* [24]

$$\left[\sum_j c_{ij} \exp\left(-\frac{(x-x_j)^2+(y-y_j)^2}{2\sigma^2}\right)\right]$$

Now in order to calculate the orbitals, $\psi(x,y,z_0)$ is calculated on a grid (x,y), where z_0 is conveniently chosen such that $N^2 z_0^2 = \exp z_0^2/\sigma^2$. This choice renders the z-dependent prefactor equal to 1, so that

$$\psi(x,y,z_0) = \left[\sum_j c_{ij} \exp\left(-\frac{(x-x_j)^2+(y-y_j)^2}{2\sigma^2}\right)\right]$$

The three-dimensional wave function is then constructed by setting

$$\psi(x,y,z) = \psi(x,y,z_0)\left[Nz \exp -\frac{z^2}{2\sigma^2}\right]$$

The electron density of this wave function is found by

$$\rho(x,y,z)_i = |\psi(x,y,z)_i|^2 = \psi(x,y,z_0)^2 \times$$

$$\left[N^2 z^2 \exp\left(-\frac{z^2}{\sigma^2}\right)\right]$$

The phase of the orbitals is respected by explicitly taking the sign of ψ and z into account. Some constant height-STM simulations for the individual eigenstates on a short ribbon are shown in figure 9, where the height is simply z_0.

When tunneling proceeds through a $p-$orbital of the carbon monoxide, the overlap is given by the maximum overlap of this antisymmetric wavefunction with the molecular wavefunction, and as a consequence, effectively probes the gradient of the orbitals on the molecules, as shown by Gross *et al.*[25]

$$|M_{AB}|^2 \propto \left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2$$

This formula is used to calculate STM images for p-terminated tips, again by first calculating

$$\psi(x,y,z_0) = \left[\sum_j c_{ij} \exp\left(-\frac{(x-x_j)^2+(y-y_j)^2}{2\sigma^2}\right)\right]$$

on a grid. Then the numerical derivatives with respect to x and y are calculated, squared and added. The results are shown in figure 9. The predictions can already be compared with earlier work. Figure 10 shows differential conductance map on the HOMO, end state and another occupied state at -1 V. The measurements were performed by van der Lit and coworkers, and the calculated STM maps have been produced by Ijäs *et al.*[24][13]. The calculated eigenstates seem to be in perfect agreement with the theoretical and experimental results. The orbital pattern at -1 V is no directly obtained as an eigenstate, but it is explained as a superposition of multiple valence band states with patterns like molecular orbital 80 and 81 in figure 9.

The AFM contrast in the repulsive regime mainly comes from the Pauli repulsion, and as such it should be rougly proportional to the total electron density[26]. Using the coefficient matrix C, we can now try to calculate the electron density, and start to compare theory with an AFM experiment.

Every wave function was described by an eigenvector in C as $\psi_i = \mathbf{c}_i \cdot \vec{\chi}$. As basis functions, the minimal Gaussian basis set is again employed, with

$$\chi(x,y,z) = Nz \exp\left(-\frac{x^2+y^2+z^2}{2\sigma^2}\right)$$

with N a normalization factor and σ the orbital size. The electron density of this wave function is found by

$$\rho_i = |\psi_i^2| = \left[\sum_j c_{ij}\chi_i\right]^2$$

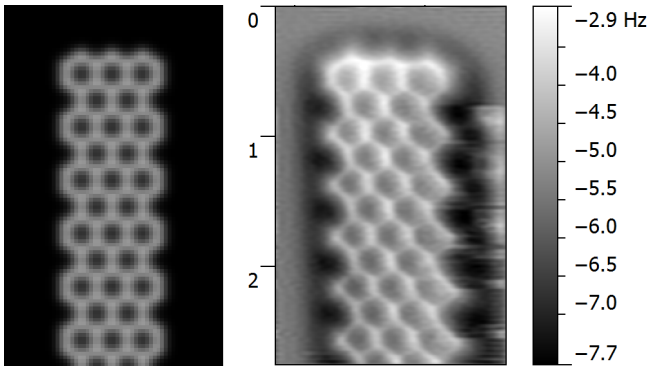


FIG. 11. Left: A simulated AFM image, using the sum of single electron densities in the pi-system only. Right: An AFM image recorded with a passivating carbon monoxide tip.

where j runs over all occupied states. Again, χ_i is decomposed using the cartesian separability of Gaussian functions, so that only the in-plane contribution has to be calculated. It is approximated that

$$\rho_{total} = \sum_i \rho_i = \sum_i \left[\sum_j c_{ij} \chi_i \right]^2$$

ignoring any cross terms between different orbitals. So basically, all orbital images of filled orbitals are superimposed to generate the total electron density. Going from the twodimensional cross section to the threedimensional plot is again just a matter of multiplying the cross section with the height-dependent term. The resulting density cross sections are first-order models for AFM images. Such a calculated AFM image is shown in figure 11, together with a constant height AFM experiment with carbon monoxide tip on a graphene nanoribbon, carried out by Joost van der Lit[24]. Interestingly enough, although all molecular orbitals seem to have quite random patterns, they happen to add up to a total electron density in the π system that perfectly shows the molecular structure. Better models can be generated from the electron density by a few mathematical operations[27][28].

COMBINING TIGHT BINDING AND MOLECULAR MECHANICS

Now we would like to see what happens to the band gap when a nanoribbon is bent. To do this, a series of nanoribbon geometries is prepared in various curvatures, and the tight binding calculation is carried out with the exponential parametrization of overlap and Hamiltonian elements. As a first calculation, we decided to take a hexamer nanoribbon and bend it to over 100 degrees. For theses calculations, the following spring constant was used erroneously: $k_b = 600$ nN Å^{-1}. According to tight binding, the band gap increases slightly up to around 90 degrees, after which the band gap decreases again. It seems that the HOMO on the ribbon overtakes the HOMO-1, resulting in a change in the HOMO identity. The result is backed up by DFT single-point calculations on a few geometries in the series. The results are shown in figure 14, where the vertical red arrows in the tight binding plot indicate the magnitude of the band gap at two angle values. Even though the quantitative features differ slightly for the methods, there seems to be an agreement in the trend, and DFT also predicts a transition from increasing to decreasing band gap. Although these optimistic calculations give some confidence as both theories again seem to give similar results, the extreme curvatures are not really attainable in experiments.

To obtain a better model of a nanoribbon, we determined to recreate a model of a nanoribbon that was actually bent in an STM experiment (this ribbon will again be encountered in the next chapter). Starting from this geometry, the ribbon was relaxed. The DOS was again obtained by a Lorentzian expansion of the eigenenergies. A geometry was prepared that emulated the experimental STM image of a bent nanoribbon. This required the restraint of two atoms in the end and one in the middle. Since STS was also performed on this ribbon when it was still straight, another geometry was made in which all atoms were relaxed and the "digital ribbon" was allowed to straighten out. This ribbon had a length of 28

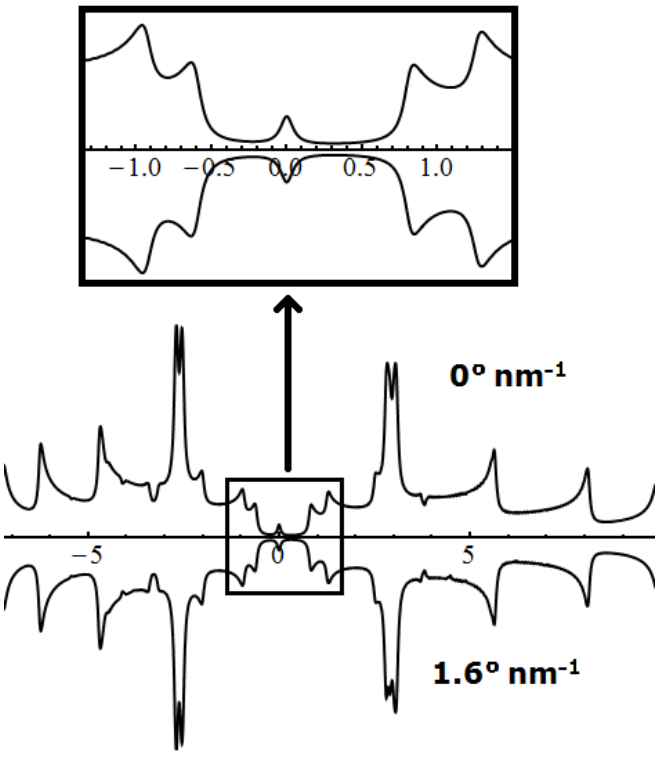


FIG. 12. The calculated density of states for a straight and bent graphene nanoribbon. Molecular mechanics parameter ratio: 100 Å^{-2} rad^{-2}. Tight binding parameters: H_0 = -2.8 eV. $S_0 = 0.15$. $\kappa = 3$.

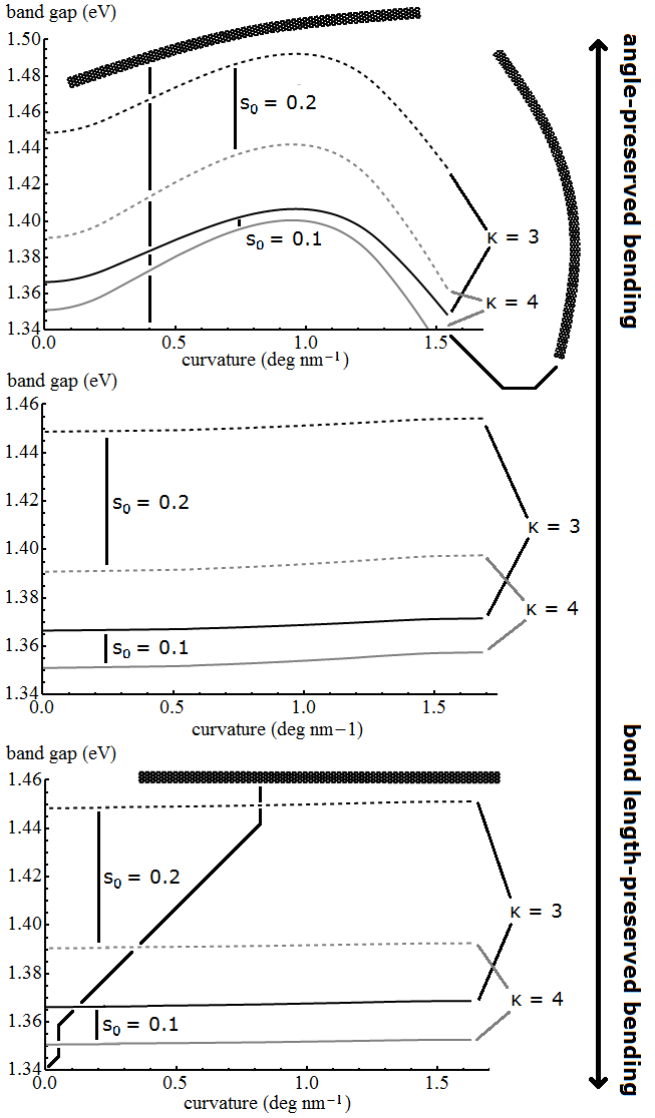


FIG. 13. Influence of the molecular mechanics parameters on the trend in the band gap versus curvature. MM parameter ratios are 0.01, 1 and 100 Å^{-2} rad^{-2}, respectively.

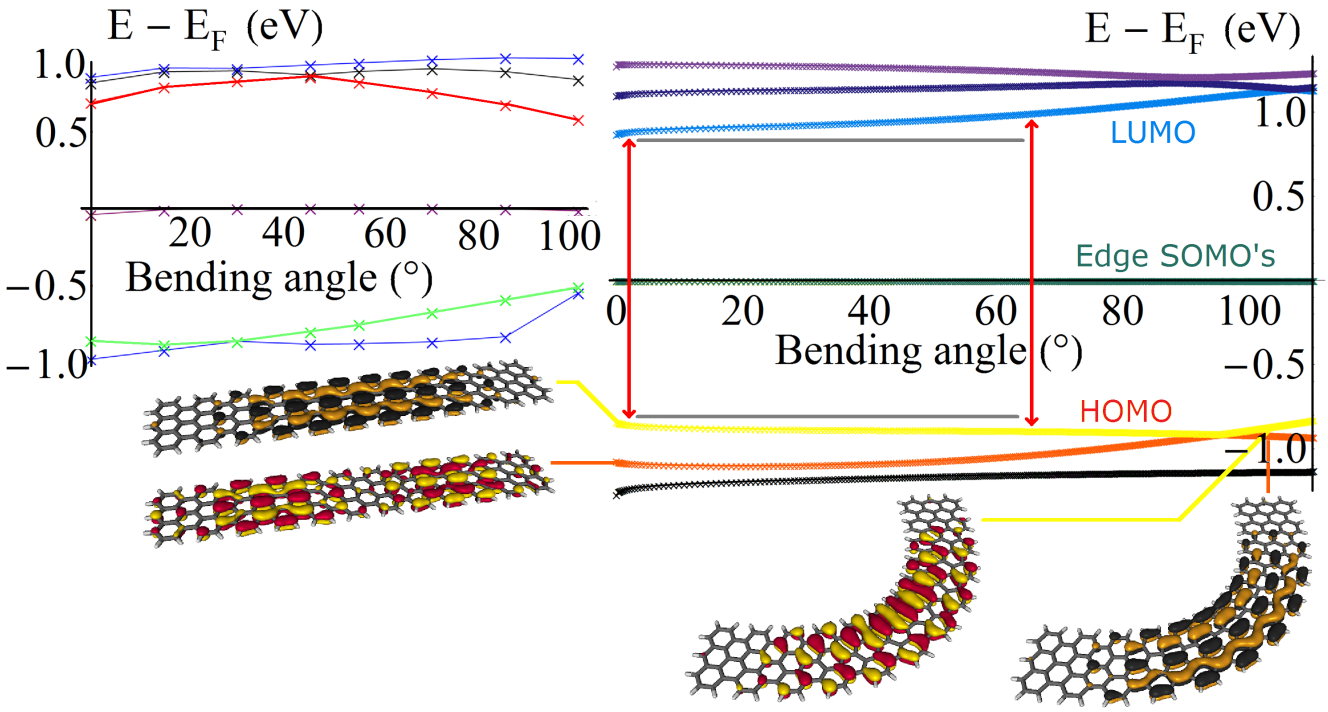


FIG. 14. The energy eigenvalues of a few frontier orbitals of graphene nanoribbons as a function of their bending angle, for DFT and tight binding. The images are orbital plots of the HOMO and HOMO-1 from the tight binding calculations.

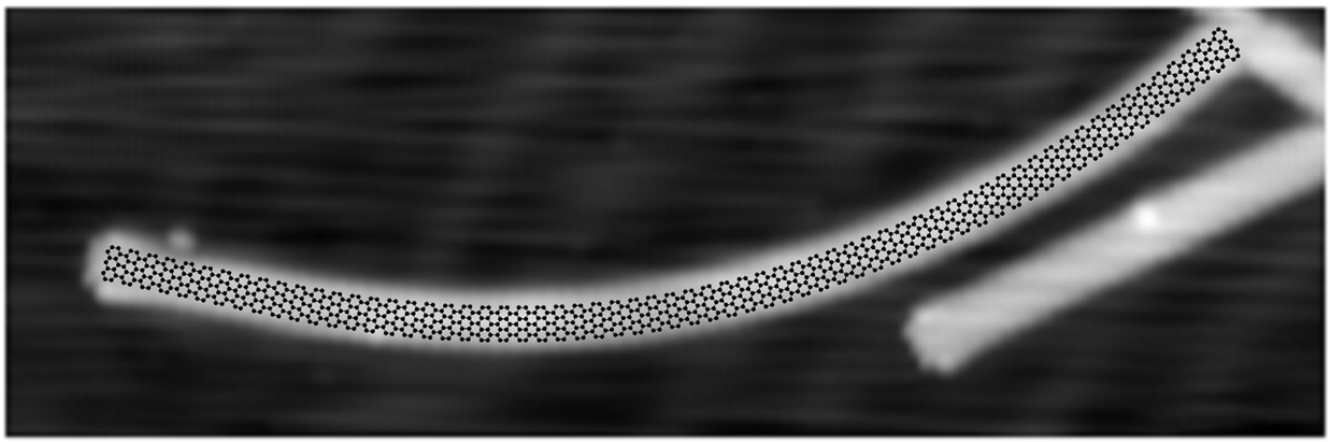


FIG. 15. An STM scan of a bent nanoribbon, with a model from molecular mechanics overlayed.

nm, much longer than what was looked at earlier.

First, let's look at the density of states. The multitude of eigenstate is seen to create a spectrum that looks much more like the periodic result than the density of states for the short ribbon in figure . Many "spikes", or van Hove-singularities can be observed in the DOS, reflecting the quasi-onedimensionality of the system - see figure 12. The calculated band gap of a graphene nanoribbon is of the order of 1.5 eV. It turns out that this is a lot less than the experimental band gap, measured with STS, which is around 2.4 eV. Also, a peculiar result that was seen earlier in STS was the fact that the conductance band peak is always twice as high as the valence band peak, but the theory suggests that there should be no difference. Although we did not look into these effects, we think that these might be due to the nanoribbon being charged.

Now upon comparing the DOS between the curved ribbon and the straight ribbon - see for example figure 12 - no differences could be seen whatsoever! When comparing the band gap for a bent ribbon with respect to a straight ribbon, an increase of only 3 meV could be seen in the bent case on a total calculated gap of 1.4 V, a difference of hardly 0.2%! This is ascribed to the stiffness of

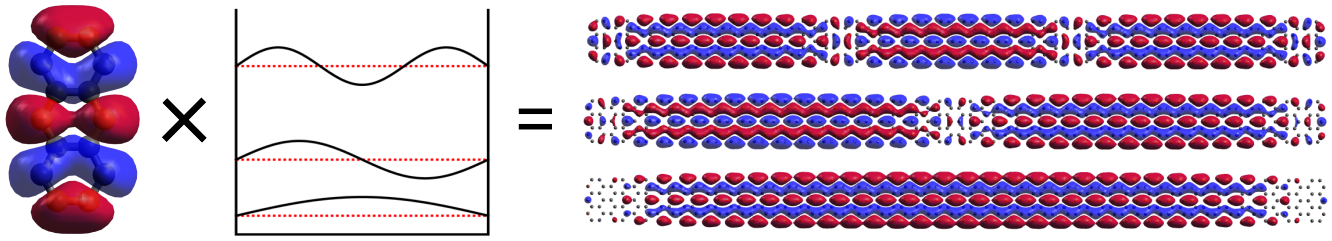


FIG. 16. The nanoribbon orbitals are found to be anthracene orbitals modulated with one-dimensional standing wave envelopes.

the bonds in the molecular mechanics simulations. Using the literature parameters, any curvature is almost completely directed by bond angle deformation, leaving the nearest-neighbour distances almost unperturbed. The most important contributions in the Hamiltonian matrix are however nearest-neighbour interactions. The next-nearest neighbour interactions do change considerably, but the effect on the eigenenergies is of course much less pronounced.

This appears to be in contradiction with the results by Koskinen, who found a decrease in band gap upon bending[5]. It turns out that the discrepancy arises from the anharmonic spring constants that were used in his calculations. These effectively elongate the average carbon-carbon distance upon bending. The change in band gap then turns out to be dictated mainly by the effective stretching in the same way as uniaxial stretching.

As a simple experiment, I decided to desert the bulk graphene parameters and prepare the bent and straight geometry again with stiff angles and weak bonds. Surprisingly, when comparing the band gaps again, a slight decrease was found in the bent case! So the mere act of shifting the rigidness between bonds and angles in the MM simulation has a large impact on the trend in the band gap as a function of curvature! This is shown in figure 13, where a series of geometries was prepared based on the long model ribbon before, ranging from completely straight to the experimental curvature where the model was based on. This was done for three molecular mechanics parameter ratios: $k_b/k_\theta = 0.01$, 1 and 100 Å^{-2} rad^{-2}, respectively. Then for these sets of geometries the band gap was calculated. Now the real power of tight binding is experienced: althought there are 896 carbon atoms in this ribbon, and therefore 802816 entries in the Hamiltonian and overlap matrices (almost 1 MB each in terms of computer memory), the calculation for each conformation is completed within 5 seconds, completing an entire band gap versus curvature graph in just over a minute. The bottleneck is actually the molecular mechanics simulation. 4 lines are plotted, for different values of the overlap S_0, and interaction range parameter κ, showing that these parameters do not really affect the trend as much as the molecular mechanics.

Last, we determined to look at the orbitals again, and found something that went completely unnoticed in calculations on smaller ribbons. The HOMO-1 state was found to be completely identical in nodal pattern to the HOMO, except for an additional nodal plane in the middle of the nanoribbon, as shown in figure 16. The HOMO-2 features an additional node, etc. The same was true for the LUMO, LUMO+1, etc. Then when orbitals with energies within the next peak in the DOS were mapped, it was found that they had a completely different nodal pattern, but again, the HOMO-1 had one more nodal plane than the HOMO, and so forth. So each orbital pattern casts a pseudoband of states with an integer number of nodes. Effectively, this means that every state in a band can be described as a one-dimensional particle-in-a-box state. So every molecular state on the monomer - antracene - creates a band of states of the form

$$|\psi\rangle = |n\rangle|\psi_{anthr}\rangle$$

where $|n\rangle$ is the nth one-dimensional particle-in-a-box state and $|\psi_{anthr}\rangle$ is the anthracene molecular orbital. Of course, this must be the case, since the model must somehow extrapolate to the Bloch waves in the periodic case. The Hamiltonian can be split linearly as $\hat{\mathrm{H}} = \hat{\mathrm{H}}_0 + \hat{\mathrm{H}}_{\mathrm{latt}}$, where $\hat{\mathrm{H}}_0$ is the constant potential Hamiltonian, of which the particle-in-a-box waves are eigenfunctions, and $\hat{\mathrm{H}}_{\mathrm{latt}}$ is the lattice perturbation, to which - to a first approximation - the anthracene molecular orbitals are eigenfunctions. Then

$$\hat{\mathrm{H}}|\psi\rangle = \left(\hat{\mathrm{H}}_0 + \hat{\mathrm{H}}_{\mathrm{latt}}\right)|n\rangle|\psi_{anthr}\rangle$$

$$\approx \left(c_1\frac{\hbar^2k^2}{2m} + c_2E_{anthr}\right)|\psi\rangle$$

where c_1 and c_2 are coupling terms. We know that the quadratic dispersion for a one-dimensional system gives a density of states proportional to $E^{-1/2}$, so our nanoribbon system will have a density of states characterized with these peaked one-dimensional particle-in-a-box states derived from anthracene orbitals with a certain offset given by c_2E_{anthr} and a certain band width (and direction) given by c_1.

We can use the standing wave envelopes to calculate the local density of states on several points of the ribbon semianalytically. After setting $E_n = n^2$ and $|\psi_n|^2(x) =$

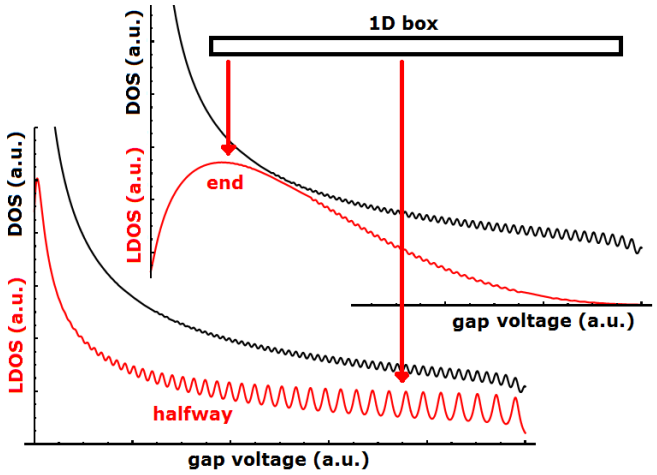


FIG. 17. Model calculation of the LDOS halfway and near the end of a quasi-one-dimensional box.

$\sin^2 n\pi x/x_0$, the LDOS was calculated in the usual way, by Lorentzian broadening the substituent eigenenergies. For $x = \frac{1}{2}x_0$, there are equal contributions from states with odd n - since they have an antinode in the middle - and no contributions from states with even n - since they have a node in the middle. Closer to the end, the differential conductance is blind to the low wave vector states, since they have most density on the middle, but really catch on for higher values as the wave spreads more evenly over the nanoribbon. These scenarios are shown in figure 17. These predictions are in excellent agreement with experimental work on short, conjugated molecules. It was found by Repp and coworkers that the band onset in oligothiophenes (small, conjugated molecules) could be described well with this particle-in-a-box theory[29].

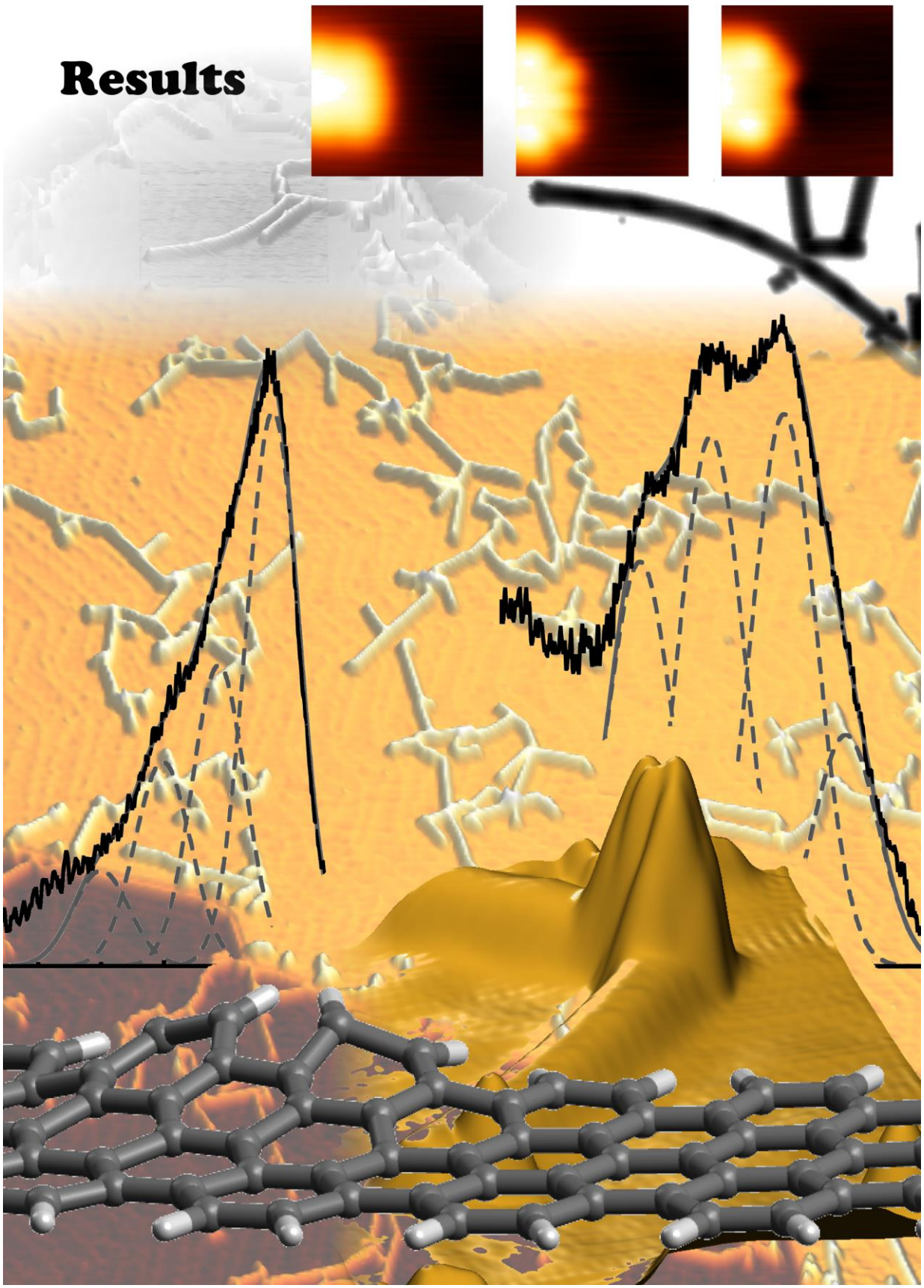
If the tight-binding theory is appropriate for the decription of nanoribbons, we should be able to find the modulations in the band onsets by measuring the LDOS at various points over the length of the nanoribbon. The band onsets should be the sharpest on the middle of the ribbon, and become flat when going further towards the end. Additionally, the variations in the LDOS means that we have to be careful with positioning the tip if we want to extract proper values of valence band and conductance band onset. In addition to the fact that variations in height offset can give variations in the broadening through electronic coupling and possibly a nonnegligible capacitative effect, the peak position is also dependent on the position of the tip on the ribbon. Therefore, the VB and CB onset values can best be obtained by positioning the tip exactly halfway the ribbon, at the same feedback setpoint every time. Of course, positioning the tip in a node of a certain orbital makes it blind to the corresponding pseudoband, so by positioning the tip carefully on the armchair edge we want to counteract this as much as possible. Here there is as little cancellation of orbitals with different signs as possible with an s-type tip[30].

Now that we have a good understanding of the electronic structure of graphene nanoribbons and formulated a theoretical framework, it is time to carry out the experiments.

[1] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles, New York: McGraw-Hill, 1981* (1981).
[2] C. Li and T.-W. Chou, International Journal of Solids and Structures **40**, 2487 (2003).
[3] P. Zhao and G. Shi, Tech Science Press **5**, 49 (2011).
[4] K. Alzebdeh, International Journal of Mechanics and Materials in Design **8**, 269 (2012).
[5] P. Koskinen, Physical Review B **85**, 205429 (2012).
[6] D. A. Papaconstantopoulos, M. J. Mehl, and S. C. Erwin, *Tight-binding hamiltonians for carbon and silicon* (1998) p. 221.
[7] V. Pereira, A. Castro Neto, and N. Peres, Physical Review B **80**, 045401 (2009).
[8] R. M. Ribeiro, V. M. Pereira, N. M. R. Peres, P. R. Briddon, and A. H. Castro Neto, New Journal of Physics **11**, 115002 (2009).
[9] Y. Hancock, A. Uppstu, K. Saloriutta, A. Harju, and M. J. Puska, Phys. Rev. B **81**, 245402 (2010).
[10] R. Kundu, Modern Physics Letters B , 1 (2011).
[11] The large matrices can be represented in a visual way by turning them into images, the pixel value of which is proportional to the value of the matrix elements. The top part of the title page background for the theory chapter actually shows such a representation for the identical matrix products HC and SCE, side by side, for a graphene nanoribbon.
[12] The terminology is chosen such that an n-mer refers to a polymer with n repeating subunits. In the synthesis, the precursors are bianthracene molecules, so n attached precursors give rise to a $2n$-mer.
[13] M. Ijäs, M. Ervasti, A. Uppstu, P. Liljeroth, J. van der Lit, I. Swart, and A. Harju, Phys. Rev. B **88**, 75429 (2013).
[14] L. Talirz, H. Söde, J. Cai, P. Ruffieux, S. Blankenburg, R. Jafaar, R. Berger, X. Feng, K. Müllen, D. Passerone, R. Fasel, and C. a. Pignedoli, Journal of the American Chemical Society **135**, 2060 (2013).
[15] P. Shemella, Y. Zhang, M. Mailman, P. M. Ajayan, and S. K. Nayak, Applied Physics Letters **91**, 042101 (2007).
[16] H. Zheng, Z. Wang, T. Luo, Q. Shi, and J. Chen, Physical Review B **75**, 165414 (2007).
[17] P. Ruffieux, J. Cai, N. C. Plumb, L. Patthey, D. Prezzi, A. Ferretti, E. Molinari, X. Feng, K. Müllen, C. A. Pignedoli, and R. Fasel, ACS Nano **6**, 6930 (2012).
[18] Z. Sun, M. P. Boneschanscher, A. Uppstu, M. Ij, A. Harju, and P. Liljeroth, Physical Review Letters (2011).
[19] Y. Li, X. Jiang, Z. Liu, and Z. Liu, Nano Research **3**, 545 (2010).
[20] M. Ezawa, Phys. Rev. B **73**, 45432 (2006).
[21] H. Raza and E. Kan, Physical Review B , 1 (2008).
[22] V. Barone, O. Hod, and G. E. Scuseria, Nano letters **6**, 2748 (2006).
[23] J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805

(1985).
[24] J. van der Lit, M. P. Boneschanscher, D. Vanmaekelbergh, M. Ijäs, A. Uppstu, M. Ervasti, A. Harju, P. Liljeroth, and I. Swart, Nature communications **4**.
[25] L. Gross, N. Moll, F. Mohn, A. Curioni, G. Meyer, F. Hanke, and M. Persson, Physical Review Letters **107**, 086101 (2011).
[26] L. Gross, F. Mohn, N. Moll, P. Liljeroth, and G. Meyer, Science **325**, 1110 (2009).
[27] F. Giessibl and H. Bielefeldt, Annalen der Physik , 1 (2001).
[28] N. Moll, L. Gross, F. Mohn, A. Curioni, and G. Meyer, New Journal of Physics **14**, 083023 (2012).
[29] J. Repp, P. Liljeroth, and G. Meyer, Nature Physics **6**, 975 (2010).
[30] N. Pavliček, I. Swart, J. Niedenführ, G. Meyer, and J. Repp, Physical Review Letters **110**, 136101 (2013).

Results



We have performed electronic structure calculations on graphene nanoribbons to obtain a reasonable hypothesis for the effects that we expect to see in bending nanoribbons. It was shown how scanning probe will be the work horse in manipulation and characterization of graphene nanoribbons, and how it can be used to prove - or debunk - the theoretical efforts. Now we are finally in the position to bring everything together and perform the experiments.

A multitude of experiments was carried out during this research. During the course of the experiments, our understanding of the situation kept improving. The progressive understanding led to better theories and better experiments, which eventually led to many useful results.

SCANNING

The GNR sample was prepared by sputtering/annealing an Au(111) crystal twice, before depositing precursor molecules on the surface. The sample was kept at a constant temperature of approximately 200 ^{o}C through resistive heating, while preheating the precursor molecule evaporator resistively for 10 minutes behind a shutter. Now the shutter was opened for 60 seconds. Afterwards, the sample was annealed at approximately 400 ^{o}C for 10 minutes.

The sample is brought to the low temperature STM (Omicron LT AFM/STM with qPlus tip). An overview scan of the ribbons on the gold surface is shown in figure 3. Feedback settings vary, but they are all in-gap (50-100 mV) with a feedback setpoint of around 5 to 50 pA. In all scans, the Au(111) herringbone reconstruction is easily identified[1]. A regional scan is shown in figure 4, where a line profile shows the apparent height of the reconstuction, nanoribbons and step edges. The fact that the herringbone is unaffected by the presence of graphene nanoribbons is strong evidence for weak electronic coupling to the ribbon. Any coupling would surely perturb

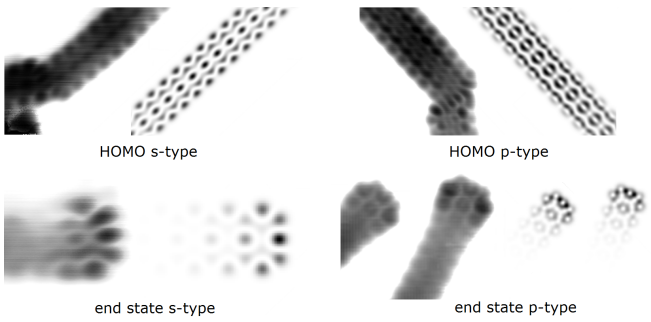


FIG. 1. Measured and simulates frontier orbitals. The end states are measured at 100 mV. The HOMO is measured at -800 mV. Measurements are performed in STM feedback.

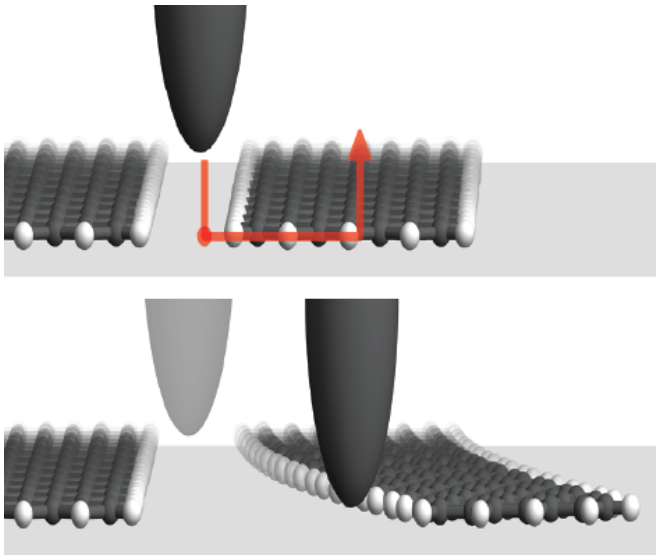


FIG. 2. A model of the nanoribbon bending scheme.

the weak surface energies associated with surface reconstructions. We found that our synthesis did not succeed in generating free ribbons, but all ribbons seem interconnected into networks. It is not known yet why free ribbons were not obtained, but presumably the reaction conditions were not ideal (maybe a too high/low synthesis temperature). Nevertheless, the free ends of the ribbons appear atomically defect-free.

We can be sure that the ribbon structures are actually graphene nanoribbons. Although the apparent width is influenced by the sharpness of the STM tip, we find it to be just under 1.5 nm. Together with an apparent height of 0.2 nm at in-gap bias voltage, this corresponds well with literature[2][3]. Scans also reveals "cat-paw" like structures at the termini of the ribbons, in agreement with the expected end states[4][2][5][3]. The tip is conditioned on the gold surface to give a relatively flat density of states within the regions where the characteristic peaks of the GNR DOS are expected. Then a GNR free end is characterized in STS. An oscillation voltage of 20 mV rms at a frequency of 2999 Hz is used, with an integration time of 50 ms per point. Most scans are made from a tip position corresponding to a feedback setpoint of 800 pA at 2.0 V or, for the valence band spectra, 500 pA at -0.5 V. As can be seen in figure 5, differential conductance spectroscopy with a clean metal tip (flat density of states on the substrate) reveals a density of states on the ribbon with a valence band onset at roughly -700 mV and a conductance band onset of +1600 mV, in perfect agreement with Ruffieux *et al.*[6].

We performed STM with an atomically sharp tip and with a carbon monoxide tip. Scans are made at low bias voltage (100 mV), and in resonance with the HOMO (-800 mV). The results are compared to the calculated STM images in figure 1. Although the images were mea-

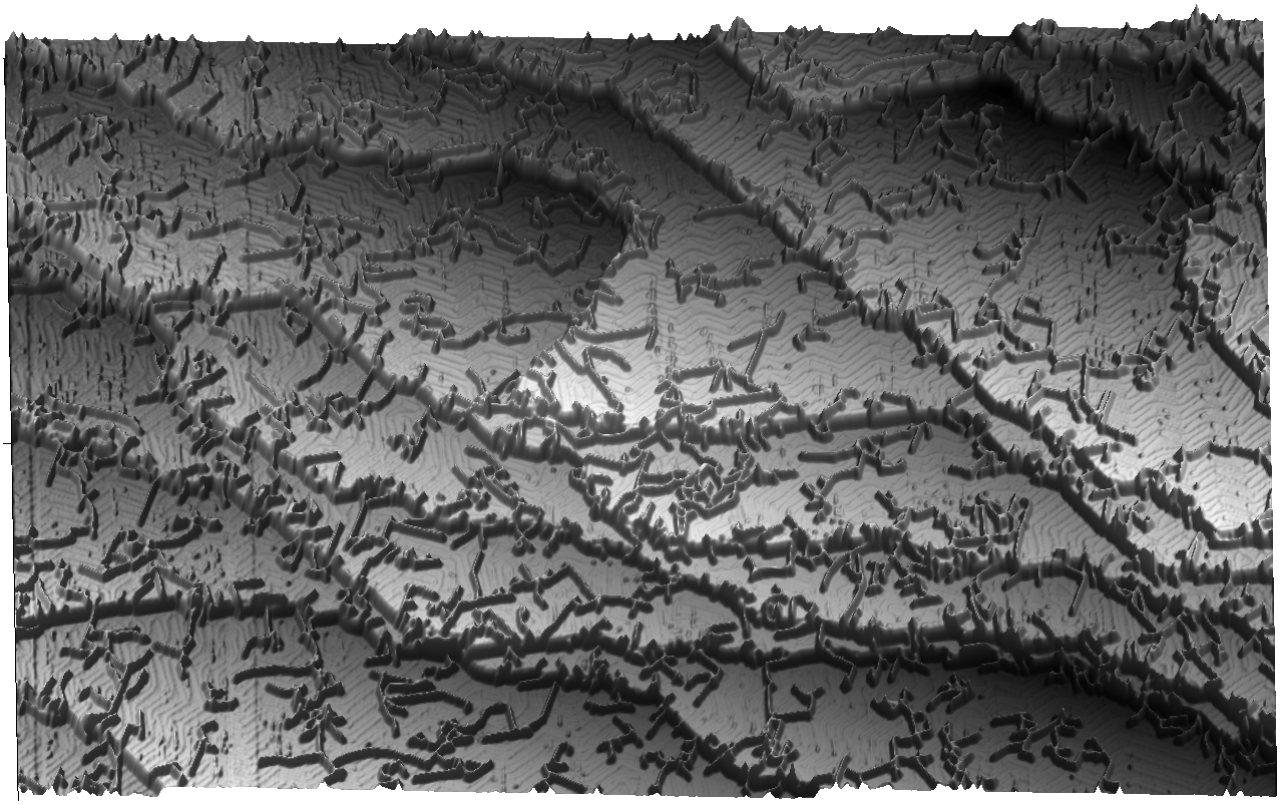


FIG. 3. A typical STM overview scan in a threedimensional representation, and with artificial lighting, giving an impression of the relative corrugations of the nanoribbons, step edges and herringbone reconstruction. Recorded at 100 mV 20 pA feedback.

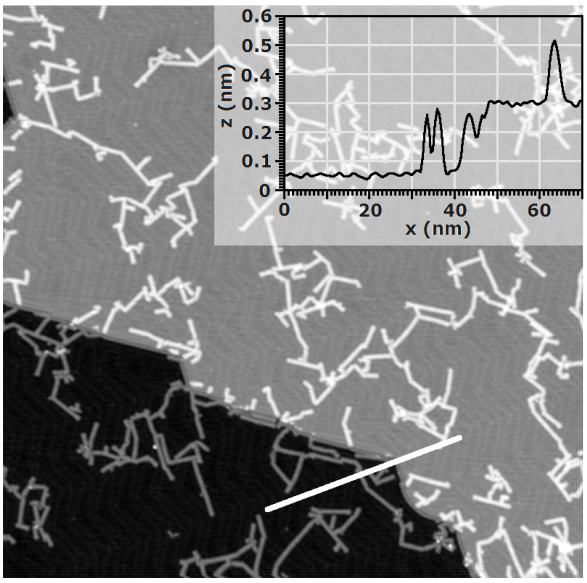


FIG. 4. A typical regional STM feedback scan. The inset is the height profile of the white line trace. The trace crosses the herringbone, nanoribbons and a step edge.

sured in feedback and calculated at constant height, this does not make a significant differance for the nodal pattern.

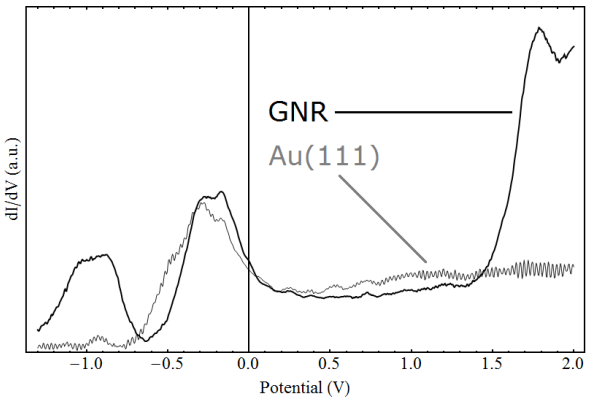


FIG. 5. A typical STS differential conductance spectrum recorded on a graphene nanoribbon and the Au(111) surface. The states between $V = -0.5$ V and $V = 0.2$ V are characteristic of the Au(111). The states that are visible on either side of the spectrum, in the nanoribbon only, define the onset of the valence band and the conductance band.

BENDING

Bending is achieved using the following steps

- From an intermediate STM current feedback setpoint at 0.1 V of 1.0 nA, the feedback is discon-

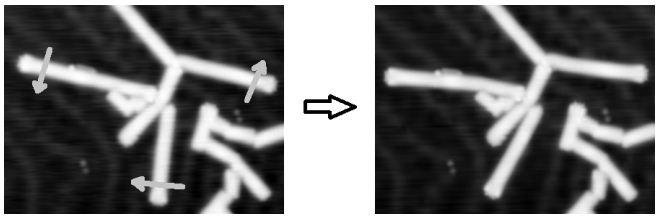


FIG. 6. An experiment, in which three ribbon ends were bent.

nected to go to constant height mode.

- The tip is lowered to the surface by 3 to 4 Å.
- The tip is moved slowly across the surface with a speed of 5 to 10 nm/s. Note that when the feedback is disabled, the tip should still follow the gold surface without crashing into it or losing its proximity. For this reason, an empty gold terrace is selected as the reference for defining a z-plane for constant-height mode.
- The tip is retracted by approximately 1 nm and the STM current feedback is reinitiated.

this scheme is shown in figure 2. Figure 6 shows an experiment where this scheme was applied to three different nanoribbon ends in close proximity.

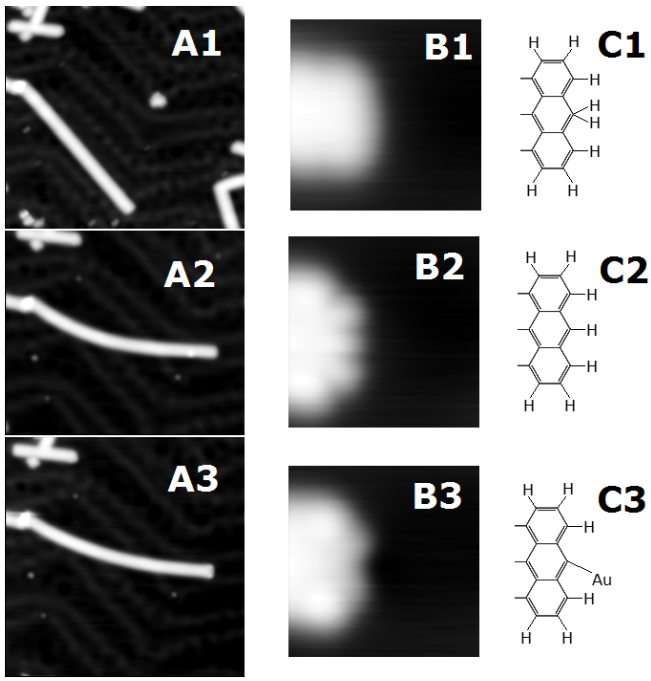


FIG. 7. A: An experiment, in which bending (from A1 to A2) was followed by pinning (A3). B: Scans on the ribbons terminus before (B1) pinning, after a single pinning step (B2) and after two pinning steps (B3). C: The molecular models of the termini in their pinned and unpinned state, with C1 corresponding to B1, C2 to B2 and C3 to B3.

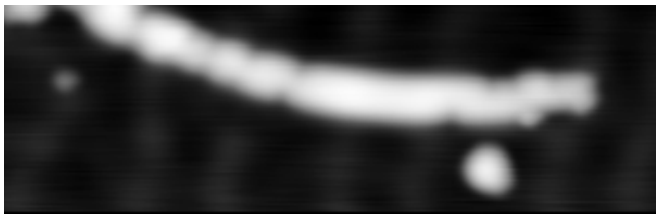


FIG. 8. A bent and pinned ribbon, which was severely dehydrogenated by a voltage pulse

To investigate the mechanism of manipulation, a few attempts have been done to pull, rather than push a graphene nanoribbon. This was performed by physically approaching the tip adjacent to the ribbon end and moving it in the direction opposite of the nanoribbons. These attempts have all been unsuccessful, indicating that it is more likely that the actual mechanism is pushing. Further evidence is given by the fact that a nanoribbon can be bent further - by one nanoribbon width - than expected by a pulling mechanism.

The succes rate of nanoribbon bending depends on the distance over which the ribbon is attempted to be curved. For relatively small angles, bending is almost always succesful. For larger angles, the nanoribbons presumably "whip back". It turned out that nanoribbons are most easily bent around defects, where the number of C-C bonds is lower. This is for example the case in regions where the ribbon is attached to another ribbon by one or two covalent bonds. Well-defined nanoribbons are somewhat stiffer, but they can still be bent easily with an STM tip. Empirically, it was found that for defect-free ribbons, the maximum obtainable curvature was around $2\ ^o\ \mathrm{nm}^{-1}$.

To achieve more control over the nanoribbons, the manipulation "toolbox" was amended with the contacting method, or pinning, which was developed by Repp and coworkers[7], and later applied to GNR by van der Lit *et al.*[3]. In this method, a carbon-hydrogen bond is broken by an inelastic tunneling proces, and a carbon-gold bond is formed subsequently. To prepare for pinning, first a zoom on the nanoribbon free end is made. The terminus in the in-gap image most often happens to have the resemblence of a cat's paw, as can be seen in the figure 7:B2. The tip is positioned on the "middle finger", and a voltage pulse is applied. To controllably give the pulse, a voltage sweep is set up to rapidly go to 3 to 4 V and back again to the feedback voltage. While ramping the bias, the current is recorded, and any jump in the spectrum is an indication for a change. When imaging the "paw" again, it is observed that the middle finger is lost and the ribbon slightly lifts off from the surface on either side (figure 7:B3). This is ascribed to breaking of the carbon-hydrogen bond and formation of a covalent carbon-gold bond (as in the model in figure 7:C3), with subsequent geometric relaxation. Therefore, pinning is really spot-

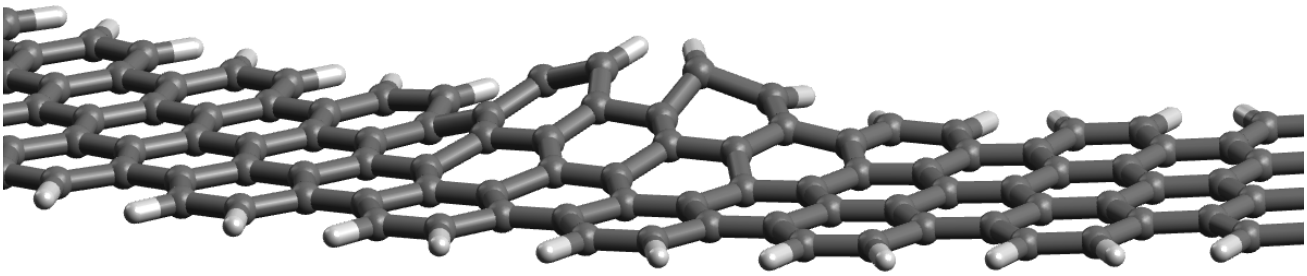


FIG. 9. A molecular model of a ribbon, where buckling starts to appear on the inside

welding of nanoribbons on the atomic scale! Sometimes, a "blunt" nanoribbon end is encountered, as in figure 7:B1. Here, the initial voltage pulse turns the blunt end into the cat paw. Presumably, the terminus has two hydrogens on top of each other, meaning that the middle carbon atom is sp^3 hybridized[4]. An example of bending, followed by pinning is given in figure 7:A1-3. It should be noted, that voltage pulse-induced dehydrogenation is not specific for the terminus of a ribbon. In one case, tip conditioning utilising voltage pulses of around 8 V was performed too close to a previously bent ribbon. The result was a ribbon that was severely dehydrogenated and attached to the gold at all these points (figure 8). The fact that ribbons are found to chemically bind to the substrate after STM-induced carbon-hydrogen bond dissociation seems in good agreement with predictions by Yuan *et al.*[8].

It was hypothesized that the elastic "whip-back" of nanoribbons could be counteracted by firmly securing them to the gold surface, using the pinning method. The question that remained was how to get these ribbons into place before applying the pulse in the first place. This was resolved in the course of the experiment by giving an additional Δz-step (approximately 800 pm down coming from a setpoint of 100 mV and 20 pA) in the manipulation, crashing the tip into the surface and leaving behind an atom or a small cluster of atoms. The cluster or atom could be moved relatively easily over the surface. Manipulation was then performed by sweeping the cluster against the nanoribbon and bending it. The humps of atoms prevent whip-back, and the ribbon can be bent further than previously.

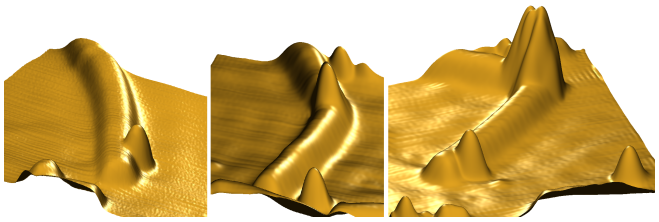


FIG. 10. Three-dimensional plots of three STM feedback scans of the experiment shown in figure 12.

BUCKLING

In this new range of obtainable ribbon curvatures, a strange effect was found to occur. After sufficient bending, a hump started to appear around which the nanoribbon "buckles". This hump increased in apparent height to around 0.7 nm. This is explained as the ribbon starting to lift off the substrate to slightly relieve the in-plane stress. This results in a situation where the bending is out-of-plane in stead of in-plane[9]. Figure 10 gives an impression of the relative heights of the buckling point with respect to the flat ribbon. These images are part of the experiment shown in figure 12. Figure 9 displays a model of a ribbon, where the threshold for buckling has just been crossed.

In some early experiments a small "pivot point" was sometimes used in addition to the anchor point (a similar atom or cluster of atoms). The nanoribbon was bent around this and the pivot point appeared to grow in size. Assuming that this hump was still due to the pivot point, it was attempted a couple of times to wipe the presumed cluster of atoms away. In one case, the buckling point itself was shifted over the ribbon (see figure 11). In another case, where it was attempted to wipe the supposed "atoms" over the ribbon, the ribbon actually snapped in two pieces - see figure 13. We have managed to break strong covalent bonds within a graphene nanoribbon simpy by pushing against it with a biased tip of an STM!

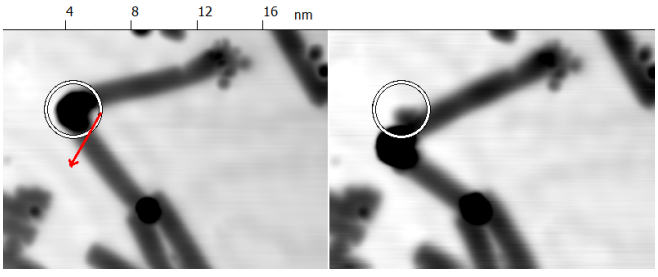


FIG. 11. An experiment, in which the buckling point was moved over the ribbon. The circle indicates the initial position of the buckling point. The arrow shows the manipulation.

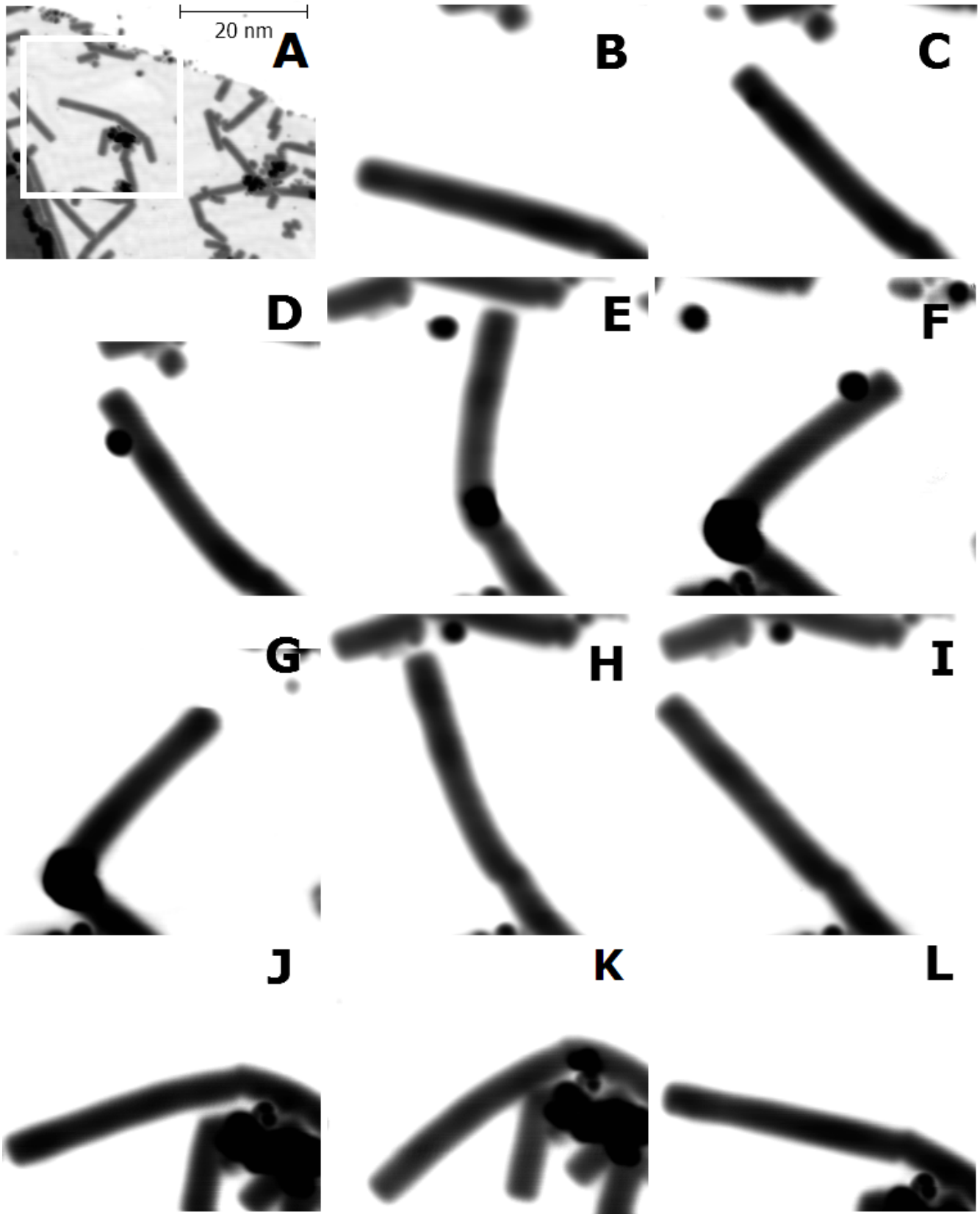


FIG. 12. An experiment, in which a ribbon was bent in two directions, and buckling was observed in either way. A: regional scan. B: the ribbon in its initial state. C: Bent upwards. D: Bent further. E: Bent even further - buckling is observed. F: Maximally buckled. G: Anchor point removed. H: Straightened. I: Straightened further. J: Bent downwards. K. Bent further down - buckling is observed. L: Straightened again.

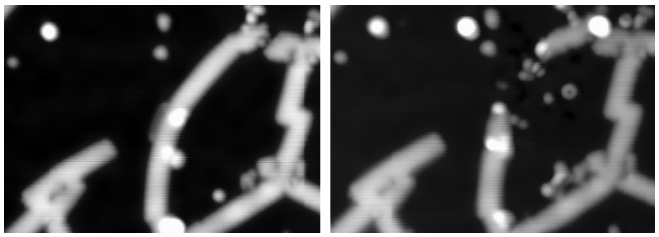


FIG. 13. An experiment, in which excessive pushing against the buckling point of a pinned ribbon resulted in its complete destruction

In some experiments buckling was followed by pinning. Then, the anchor point could be wiped away from next to the ribbon with some minor effort, and the ribbon was kept into place in a buckled way through its contact point with the surface. Attempts to bend the ribbon back were now unsuccesful, indicating that the ribbon is completely locked into place by the pinning procedure.

The onset of buckling has been carefully resolved. As seen in figure 14, the inside edge of a ribbon strained to just below the buckling onset shows some extremely subtle variation in observed height. This patern is only seen in resonance with the valence band, but does not differ for different bias voltages, indicating that this is a geometric rather than electronic effect. Now an in-gap scan was made at low bias (100 mV) and high feedback (3 nA) and the tip was found to "bounce" up and down exactly on these points. The tip itself remained stable. The explanation is that the ribbon already starts to "ripple" on the inside, partially releaving strain. The attraction to the gold surface of the higher points is compromised and its position is now determined by a trade-off between surface attraction energy and partial strain relieve. At high current feedback and low gap voltage, the tip may approach closely enough to disturb this balance. The

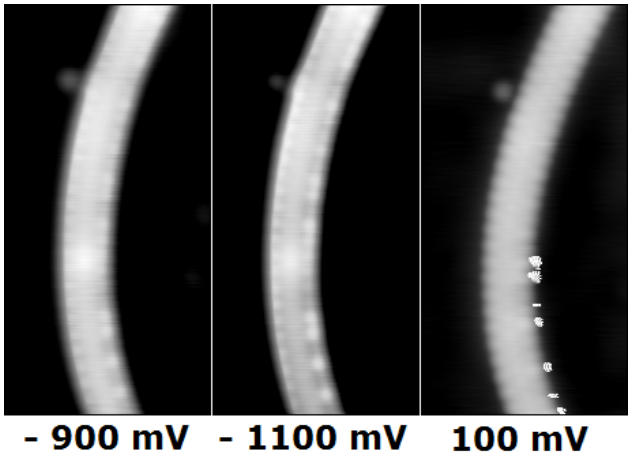


FIG. 14. A ribbon that shows "rippling" on its inside edge. All scans were recorded at a feedback setpoint of 3 nA.

FIG. 15. The "base pinned" ribbon experiment, where a ribbon was bent to a large bending angle while suppressing buckling through pinning.

van der Waals attraction pulls the ripple slightly off the surface, and as a result the tip starts bouncing on the rippled inside.

In one experiment (seen in figure 12), a ribbon was bent and buckled using the anchor point method. A bending angle of 100 degrees was obtained. Then, without pinning, the anchor point was removed. Surprisingly, no elastic whip-back was observed. This suggests that buckling generates a partial relieve of stress. An attempt to bend the ribbon back resulted in straightening of the ribbon, and no signs of damage were observed. Then the ribbon was bent the other way, up to the point that buckling started to appear again. Straightening for a second time showed once more that the structure was not damaged, indicating that buckling is a reversible geometric feature.

We wanted to find a way to bend a nanoribbon to large angles without having it buckle. We therefore devised a procedure where buckling was inhibited by additional pinning at the side of the ribbon. In the experiment of figure 15, a ribbon was found to be relatively free to rotate around its attachment point to the nanoribbon network, at the left side of its base. Initial bending attempts only resulted in the rotation of the complete ribbon around this point. To force bending in stead of rotating, the ribbon was contacted at a single point near the base. Now bending was succesful. When buckling was observed, the ribbon was slightly bent back, and the ribbon was pinned at two more points near the bottom. Subsequently the ribbon was bent to a more extreme angle, pinned down, and the anchor point was removed. This also happened to be one of the first experiments where electronic structure characterizations have been performed. This experiment will from now on be referred to as the "base pinned" ribbon experiment.

In general, when performing STS both on buckling points and on contact points, it was observed that peaks in the spectrum were much sharper and much broader, respectively. This makes sense, since the lifetime broadening is determined by the coupling to the substrate. In attached ribbons, this coupling is much larger, and in protruding points, this coupling is smaller. Furthermore, background tunneling is suppressed for larger values for the tip height.

ELECTRONIC STRUCTURE

Differential conductance mapping

Before analyzing the differential conductance spectra, some work on differential conductance mapping will be described here. Differential conductance maps have been produced in constant-height mode for the base-pinned ribbon (figure 16). Unfortunately, no clear orbital structure is seen, which is ascribed to the tip not begin atomically sharp. The two "bumps" on the left side of the ribbon that are indicated with lines arise from the Au(111) herringbone reconstruction. It is clear from the intensity that at -0.6 V and 1.7 V the tip really comes into resonance with states on the ribbon. As will be shown later, the spectrum showed a characteristic shoulder at -0.6 V in the bent case, as opposed to a sharp onset at -0.7 V in the straight case. As far as the spatial differential conductance is concerned, however, there is no clear distinction between -0.6 V and -0.7 V. On the other hand, going to -0.8 V exposes a resonance with a new orbital pattern. In line with the predictions from the previous chapter, we expect this to be the onset of another band with a new orbital pattern. In fact, a tight binding calculation (using parameters $H_0 = -2.8$ eV,$S_0 = 0.1$ and $\kappa = 3$) on a 27-mer (corresponding to the observed length of 12 nm of the ribbon) yields an energy difference between HOMO and HOMO-1 of 40 meV, and subsequently between HOMO-1 and HOMO-2 of 60 meV. According to the model, the second band should start at 330 meV below the HOMO level. This explains why the bias differences in the experiment were too large to probe the individual particle-in-a-box eigenstates.

We have the best shot at finding the individual states for a short nanoribbon, since the number of states in each band is then low, and the states have a relatively large energy spacing. Therefore, we searched for a very short ribbon, and a differential conductance map series was initiated from -610 mV to -1170 mV in 10 mV intervals, with a oscillation amplitude of 5 mV rms. The tip was not atomically sharp nor CO terminated, but the fingerprints that are looked for are merely the nodal planes in the longitudinal direction of the ribbon. However, as can be seen in figure 17, these states could still not be resolved well. A tight-binding calculation was performed on a model nanoribbon of the same length to find the energy spacings between the eigenstates and calculate the differential conductance maps. To simulate the effect of tip-broadening, the calculated grid of values for $\psi(x, y, z_0)$ was convolved with a Gaussian function of width 5 Å, and $\rho(x, y, z_0)$ was also convolved with this

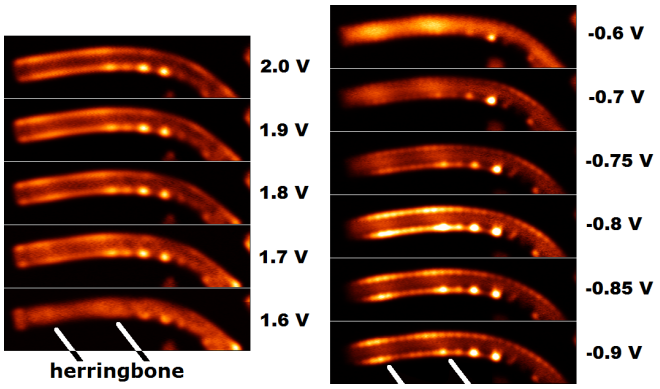


FIG. 16. Differential conductance maps on the base pinned nanoribbon

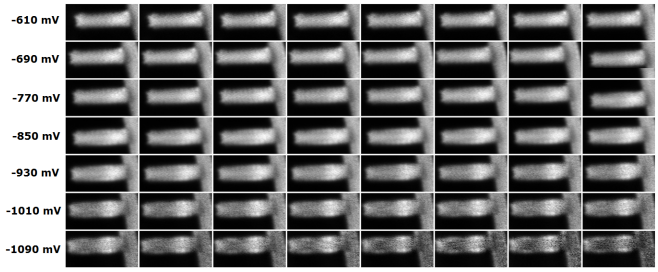


FIG. 17. Differential conductance maps on a short nanoribbon, with 10 meV intervals (left to right, top to bottom; the values denote the first bias values in each row).

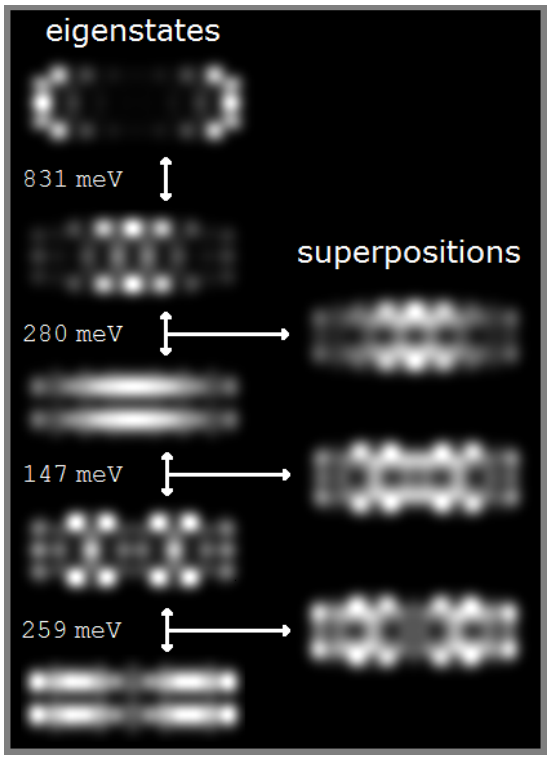


FIG. 18. Calculated wide s-tip differential conductance maps for a 7-mer nanoribbon, with energy spacings between the eigenstates. The maps on the right show (equal) superpositions of successive states.

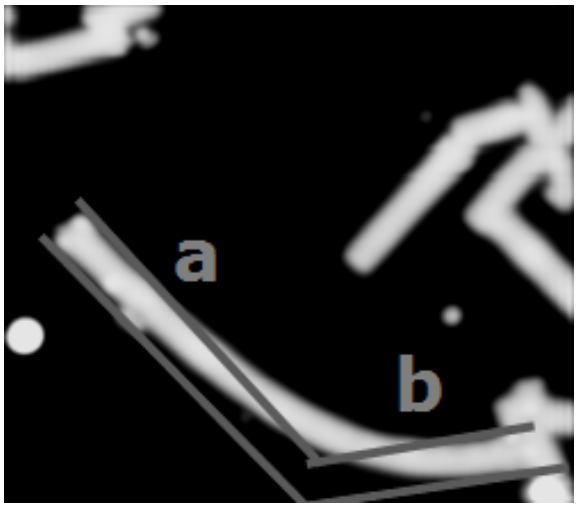


FIG. 19. Determination of the bending angle.

function. The calculated eigenstates with corresponding energy differentials are shown in figure 18. The measured differential conductance maps in figure 17 do not really correspond well with these predictions, although the observed shift from towards a two-lobed orbital structure at higher negative bias has some resemblance to the calculated superposition eigenstates on the right side of figure 18. This leads us to believe that the individual levels are broadened to such an extent that, even with a short ribbon with relatively large energy spacings, the individual levels can not be resolved.

This energy broadening of course has a lower limit by the weak coupling to the gold, and might be increased due to the tip being positioned in very close proximity for the measurement. This last aspect is necessary to create enough signal in the differential conductance for such a small voltage oscillation. Another theory might be that all eigenstates cast their own series of electron-phonon peaks onto the energetic positions of other states, resulting in a convoluted spectrum where a multiple of eigenstates contributes to the total signal. Electron-phonon coupling was indeed briefly considered to be an important factor in the experiments (some aspects of it are described in appendix F), but no effort was made to devise experiments that could really resolve the phononic fine-structure, if any.

Data Analysis

We have also performed many scanning tunneling spectroscopy experiments. Peak-fitting of differential conductance spectra was performed using the peak-o-mat software. It was found quickly that the Gaussian lineshape fit the curves more properly than Lorentzian fits, so onlly Gaussians are used. These curves are related to an ensemble of states, rather than individual states. The fitting procedure was done by optimizing the least-squares fit of a few Gaussian functions on top of a constant background for the negative bias region (-1.5 V to -0.5 V) and for the positive bias region (1 to 2 V).

The bending angle of the GNRs was inferred straightforwardly from image analysis. Tangent lines to the ribbons left and right side were drawn from both termini. The coordinates of the line ends and their intersection were used to calculate the angle with the vector product rule

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{||\mathbf{a}||||\mathbf{b}||}$$

This construction is shown in figure 19. The angles found for the left and right side were averaged. The curvature is determined by dividing the total bending angle through the ribbon length.

Care needs to be taken to keep the tip clean during the experiments. Most manipulation steps are accompanied with changes in the tip, and generally, this means that the tip needs to be conditioned after manipulation before it is clean enough for spectroscopy again. Tip positioning is crucial in the experiments - the distance determines broadening and possibly capacitative coupling. Positioning on a nodal plane of a molecular orbital compromises the ability to observe the corresponding energy level in the recorded density of states. In addition, it was noted that the presence of metal humps (for example, pivot points) increases the number of peaks in the VB region, obscuring fingerprints of the nanoribbon.

Scanning tunneling spectroscopy

Consider the base pinned ribbon experiment of figure 15, the spectra of which are shown in figure 20. Spectra on gold showed that the tip had a flat density of states at the regions where the valence band and conductance band of the ribbon reside, both at the start and at the end of the experiment. One striking feature is that the onset in the bent case contains a small peak, whereas the onset in the straight case is really sharp. The series of peaks have shifted towards the Fermi energy.

The general structure of the peaks - a strong onset and a decay towards higher values - can be understood as the shape of a band of standing wave states. Surprisingly, the bands contain some small regularly-spaced bumps. Fitting with a four-Gaussian model gives a model with uniform widths and a uniform spacing of 90 meV. This led us to believe that the peak-structure contained electron-phonon peaks, corresponding to extra transport channels. Within this framework, the fact that the peak series show an initial increase in the bent case would signify a change towards a phonon-assisted tunneling regime. These considerations are described in appendix F. Although there may indeed be some phononic fine structure

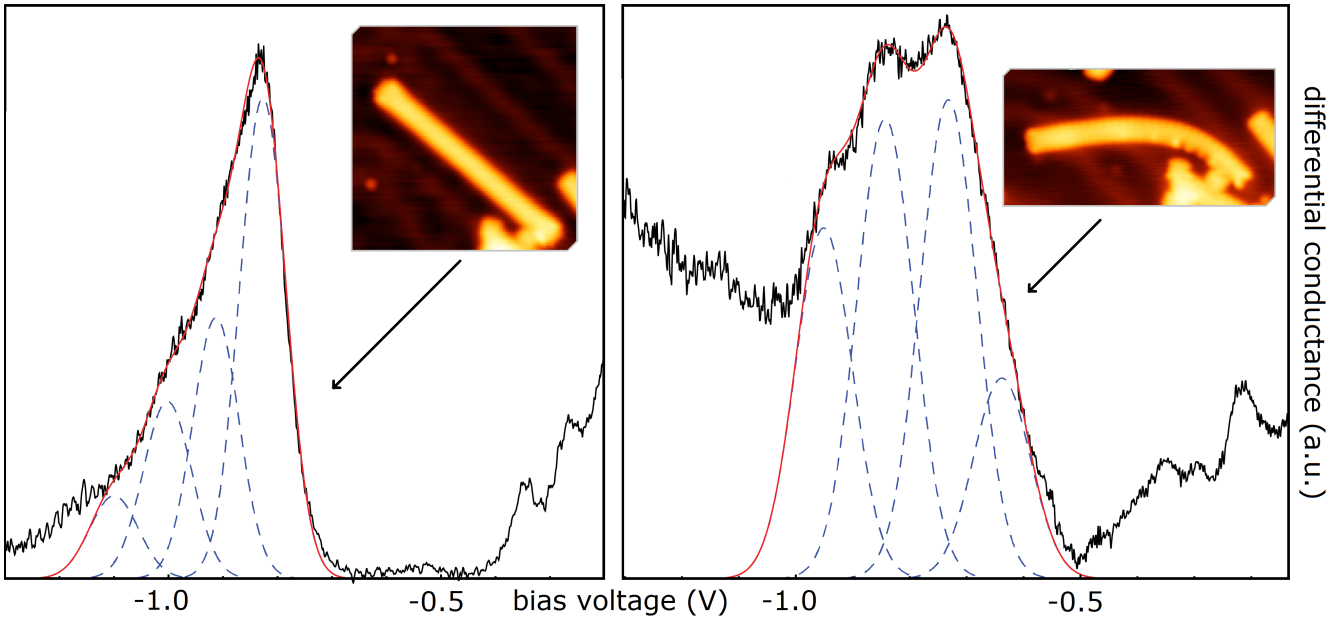


FIG. 20. Four-Gaussian fits to the valence band onsets of the STS spectra of the base pinned ribbon

hidden in the bands, later experiments showed that the general shape of the bands can be described satisfactorily with the tight binding model.

The rest of the experiments showed similar results. In most cases, the band gap was found to decrease slightly. Furthermore, in a couple of experiments, spectra were recorded before and after pinning, and it appears as if pinning slightly increased the band gap. The most notable case was an increase of exactly 100 mV in the band gap of the dehydrogenated ribbon shown in figure 8. Since the experiments were carried out by combining bending and pinning, this means that opposing effects are induced. Pinning, although a nice effect from the point of view of geometric manipulation, should therefore be avoided when trying to measure the effect of pure bending.

Local density of states variations

To see whether the particle-in-a-box behaviour is present in graphene nanoribbons, we looked at variations in the LDOS as a function of position on a straight nanoribbon. We performed measurements on the 28 nm long ribbon that was emulated for calculations in the previous chapter. The results of these measurements are shown in figure 21. The LDOS was recorded on three different spots - halfway, quarter way and at the end - and at the same time calculated at these positions with tight binding. Although the quantitative values are off (the experimental valence band onset is at higher negative bias), the shape of the valence band onset is nicely reproduced. At the ribbon terminus, the end state shows up and the valence band onset is hardly visible anymore. These variations in the LDOS are in close agreement with the theory and earlier work on particle-in-a-box states on similar π-conjugated systems[10]. Now the second peak in the valence band should be the onset of a new band of states, derived from a different anthracene orbital. This is in agreement with the differential conductance maps shown before (figure 16), where this new band onset marks the point where the contrast changed.

The ribbon was now bent without pinning, anchoring or buckling. Differential conductance spectra halfway the ribbon before and after the manipulation are displayed in figure 27, where also the three-Gaussian fit models are displayed, and the Au(111)-background spectrum to show that the density of states of the tip does not interfere with the observed ribbon DOS. The absolute value

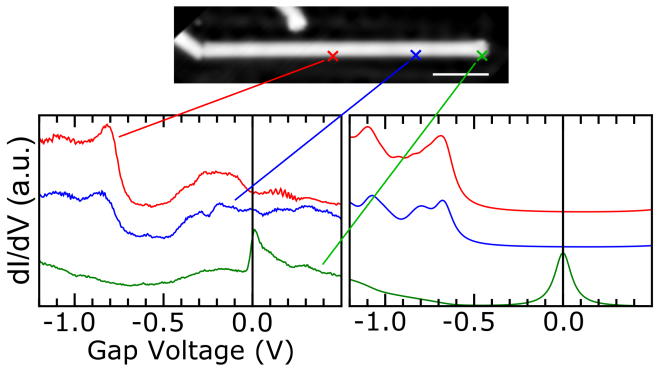


FIG. 21. The local density of states on three positions on a nanoribbon. Left: STS spectra. Right: simulated spectra.

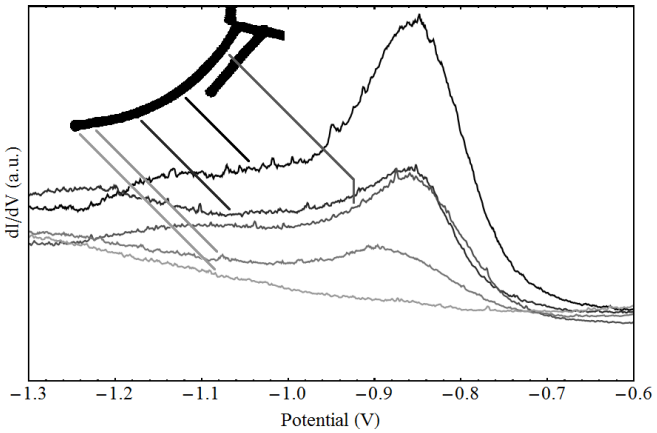


FIG. 22. The shape of the valence band onset in a bent nanoribbon. The sharpest peak, shifted most towards the Fermi level, is seen halfway. Nearer the ends the intensity decreases and shifts, and at the very end, no onset can be identified anymore. Like the straight case, the variations in the LDOS can be ascribed to the probing of one-dimensional particle-in-a-box states with quadratic dispersion.

of the band gap was found to decrease.

Again, spectra were recorded over the length of the ribbon to find the particle-in-a-box behaviour. The variations in valence band onset sharpness and position match perfectly with the standing wave theory, as seen in figure 22. Arguably, this is conclusive evidence that the electronic structure of nanoribbons consists of standing wave-envelope states, which remain unaltered in the bent case. However, we do not know yet what the featureless state at the conductance band onset in the differential conductance at 1.6 V of figure 16. This state gives a strong onset for the conductance band at the end, hiding the slow onset predicted by theory, as seen in figure 23.

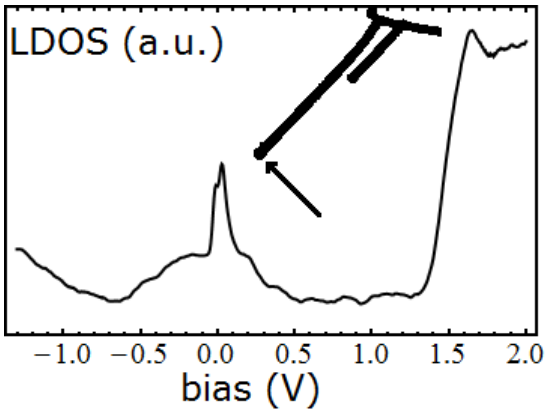


FIG. 23. A differential conductance spectrum on the end of the 28 nm long nanoribbon (position indicated in inset), showing the expected slow increase of the valence band and an unexpected high LDOS at the conductance band side.

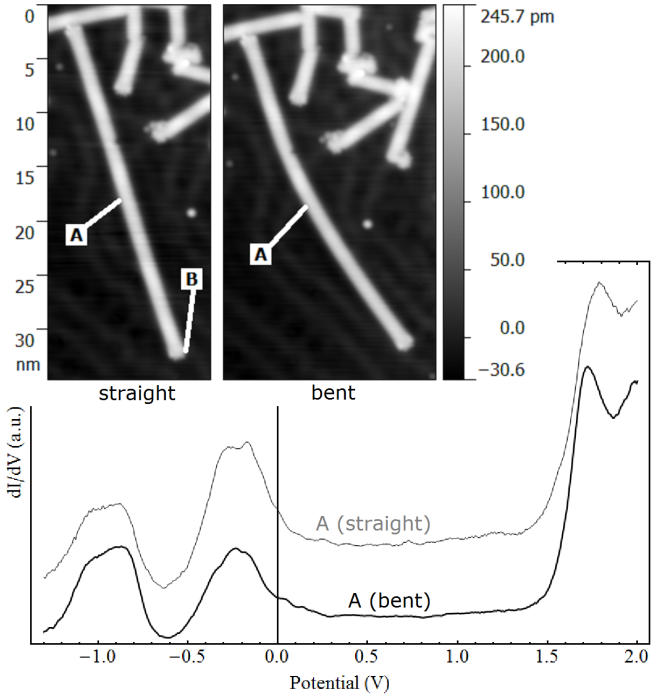


FIG. 24. An experiment, where STS was performed on point A before and after bending.

Band gap variations

For the 28 nm long nanoribbon of figure 21 and 22, we managed to push it into two slightly less curved conformations. Again, STS measurements were performed, and although there was no monotonic trend, the band gap was always lower in the bent conformation than initially in the straight situation.

This ribbon is one of the five ribbons on which good quality spectroscopic measurements have been performed. Figures 26 and 24 show two more experiments. In the last experiment, the change was significant enough to be noticable by eye in the full range spectra (that cover both the valence and conductance band). The results from peak fitting showed a decreasing band gap as a re-

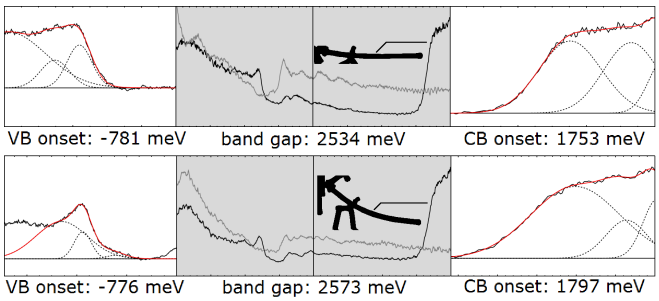


FIG. 25. The experiment, where the change in the band gap upon bending was positive.

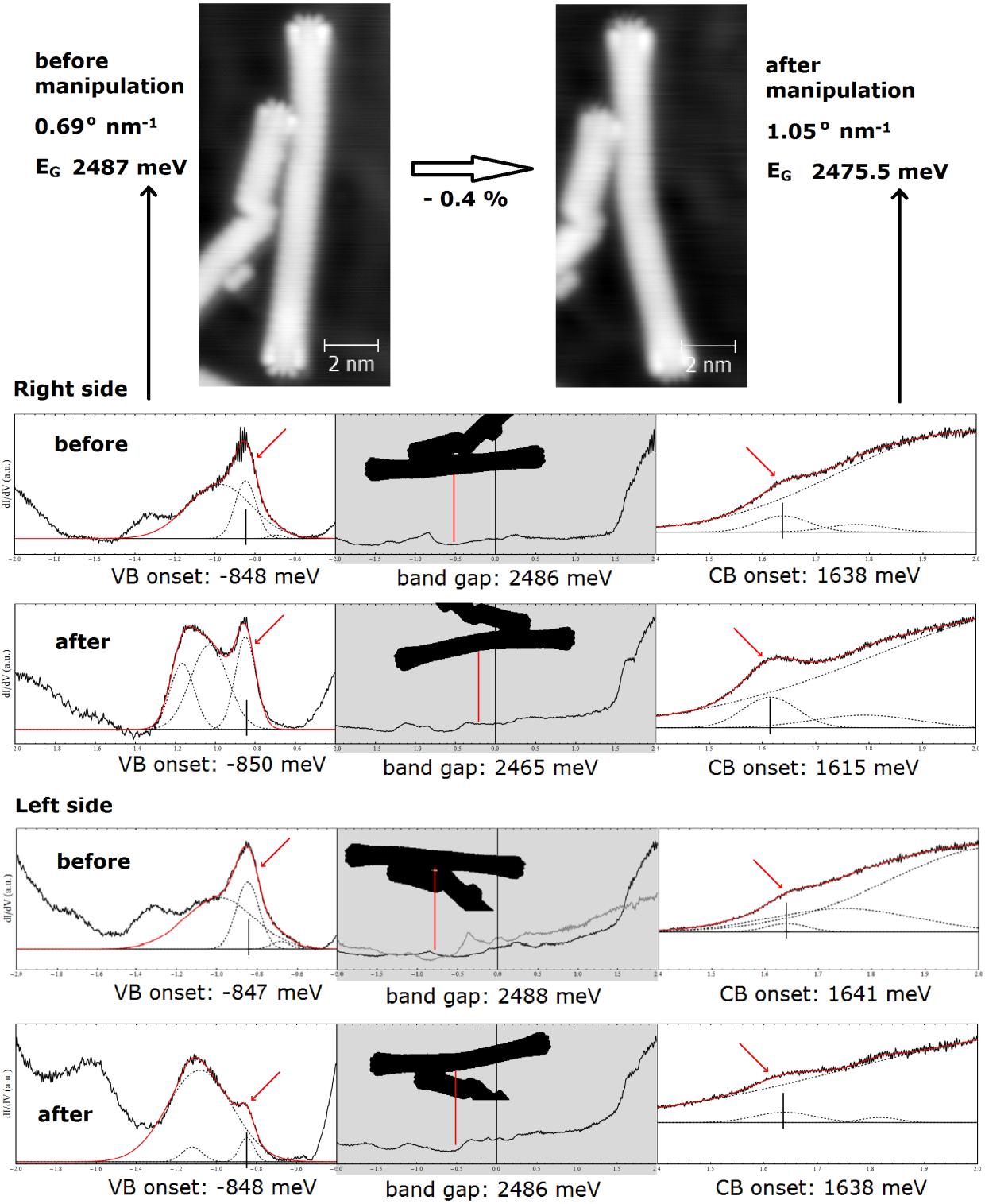


FIG. 26. One total experiment, in which STS was performed before and after manipulation. The curves could all be fitted, and the band gap values for left and right are averaged to obtain the best estimate for the band gap. The error in the peak fitting procedure turns out to be much less than the error left in the STS itself.

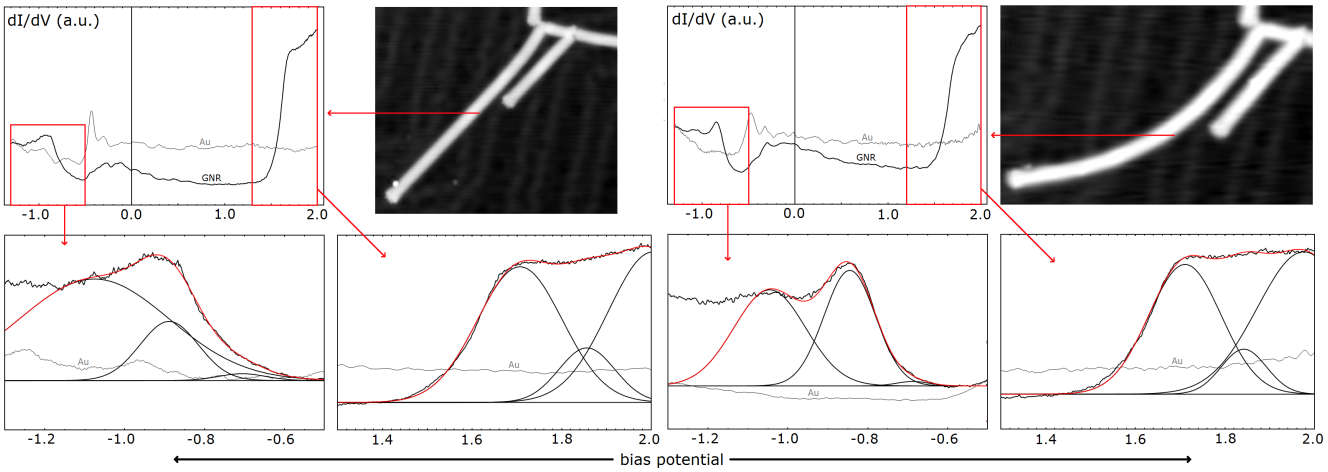


FIG. 27. Fitted spectra to a nanoribbon both in its straight and bent conformation

sult of bending in 4 out of 5 cases. An exception of this is the experiment the spectra in figure 25. This might be due to the coupling to the nanoribbon network near the halfway point. Indeed, in the experiment displayed in figure 26 a similar situation is encountered, where the coupling to an the proximity of parts of the graphene nanoribbon network bring about humps in the differential conductance that are not otherwise expected. In that experiment, fortunately, the change is further down the occupied states, and the valence band onset can still always be identified and fitted properly.

In summary, we were finally able to find a small effect in the band gap. Although the tight binding theory describes the electronic structure of graphene nanoribbons really well, the trend in the band gap as predicted by MM/TB does not correspond well, apart from the fact that the band gap is indeed found to be relatively unresponsive to changes in conformation.

There is a slight variation in valence band position and conductance band position in different experiments. This might be due to different chemical environments due to different attachments to the nanoribbon network. Due to these subtle variations, a trend does not show up nicely in a graph of band gap versus curvature. Therefore, a plot is made where percentage change is shown as a function of curvature. The 0 % value at zero curvature is obtained by extrapolating the data linearly. This representation for the 5 experiments is shown in figure 29. It is seen quickly that the band gap decreases in 4 out of 5 cases, but the

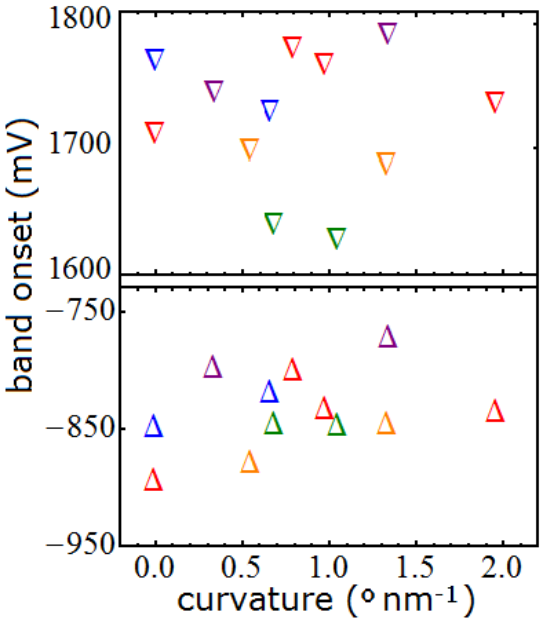


FIG. 28. Measured absolute values of the valence band and conductance band from all experiments that could be fitted well.

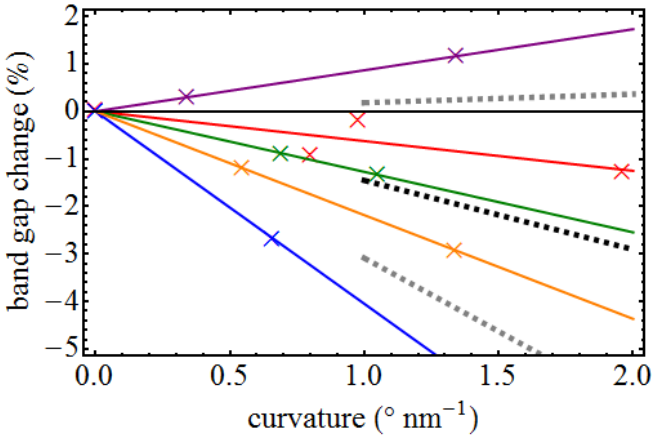


FIG. 29. The percentage change in band gap as a function of curvature for the 5 experiments. The thick, dashed, black line gives the average slope. The gray lines give the average plus and minus the standard deviation.

spread in the data is still quite large. Although the number of experiments is quite low to perform a statistical analysis, we can obtain a value of the average percentage band gap decrease and standard deviation for these 5 experiments. This gives a band gap decrease with increasing curvature of 1.5 ± 1.5 % nm/^{o}. These results are also displayed in figure 29. Even though the standard deviation in these data is as large as the trend, the measurements convinced us that the band gap is still decreasing with increasing curvature, albeit extremely slightly.

[1] J. V. Barth, H. Brune, G. Ertl, and R. J. Behm, Phys. Rev. B **42**, 9307 (1990).

[2] J. Cai, P. Ruffieux, R. Jaafar, M. Bieri, T. Braun, S. Blankenburg, M. Muoth, A. P. Seitsonen, M. Saleh, X. Feng, K. Müllen, and R. Fasel, Nature **466**, 470 (2010).

[3] J. van der Lit, M. P. Boneschanscher, D. Vanmaekelbergh, M. Ijäs, A. Uppstu, M. Ervasti, A. Harju, P. Liljeroth, and I. Swart, Nature communications **4**.

[4] L. Talirz, H. Söde, J. Cai, P. Ruffieux, S. Blankenburg, R. Jafaar, R. Berger, X. Feng, K. Müllen, D. Passerone, R. Fasel, and C. a. Pignedoli, Journal of the American Chemical Society **135**, 2060 (2013).

[5] M. Ijäs, M. Ervasti, A. Uppstu, P. Liljeroth, J. van der Lit, I. Swart, and A. Harju, Phys. Rev. B **88**, 075429 (2013).

[6] P. Ruffieux, J. Cai, N. C. Plumb, L. Patthey, D. Prezzi, A. Ferretti, E. Molinari, X. Feng, K. Mllen, C. A. Pignedoli, and R. Fasel, ACS Nano **6**, 6930 (2012).

[7] J. Repp, G. Meyer, S. Paavilainen, F. E. Olsson, and M. Persson, Science **312**, 1196 (2006).

[8] Q. Yuan, H. Hu, J. Gao, F. Ding, Z. Liu, and B. I. Yakobson, Journal of the American Chemical Society **133**, 16072 (2011).

[9] We did not try to perform tight-binding calculations on buckled nanoribbons, since one of the premises - the σ-π orthogonality - is destroyed once the ribbon comes out of its molecular plane. This complicates the electronic problem tremendously.

[10] J. Repp, P. Liljeroth, and G. Meyer, Nature Physics **6**, 975 (2010).

Conclusion and outlook

We have succesfully bent and characterized nanoribbons with a scanning tunneling microscope. We managed to develop ribbon manipulation schemes that give a large amount of control over the ribbon conformational geometry. The largest curvatures that were obtained by bending alone was approximately $2\ ^o\ nm^{-1}$. We can also pin ribbons, and we can combine these tricks to bring ribbons in extremely strained positions. While doing so, the buckling effect was discovered, which was resolved in detail. First the inside of the ribbon becomes "rippled", then it starts to lift up and stands on its side. Buckling is reversible and was performed in both directions. Lack of "whip back" from the buckled geometry shows that there is a partial strain relieve in having the curvature out-of-plane in stead of in-plane.

We find that the electronic structure of GNR can be described by a (nonperiodic) TB model. This model was calibrated and benchmarked using existing theories and experiments on graphene and graphene nanosystems. General features, like the appearance of GNR in AFM and STM show a good comparison. Upon increasing the simulated ribbon length, the DOS approaches the peaked structure of 1D systems, where every atom introduces a band, and the wave functions within each band can be described by the one-dimensional particle-in-a-box model. The theory predicts a shallower onset of VB and CB near the ribbon ends. These calculated features of the LDOS compare very well with differential conductance spectra, indicating that the standing wave theory is satisfactory for ribbons in practice. By combining the tight binding theory with a molecular mechanics model, some predictions for changes in the electronic structure were formulated. Importantly, the band gap was calculated to remain almost unperturbed, only shifting up with 0.2 % for a curvature of $2\ ^o\ nm^{-1}$.

Upon bending, it is found in scanning tunneling spectroscopy that the band gap of a ribbon decreases by a slight amount of 1.5 % nm/^{o}. Although the spread in the data is large, a significant decrease was observed in 4 out of 5 experiments. The discrepancy in the predicted trend of the band gap appears to be due to the negligence of anharmonicity in the molecular mechanics model, which would result in an average elongation of the ribbon upon bending. Nevertheless, this shift in the band gap is so small that we conclude that nanoribbons seem to be electronically relatively robust against deformations. For electronics this might often be a good thing, but this also means that conformational engineering is not a way to taylor electronic properties.

If the electronic structure of graphene nanoribbons should be engineered, it would be better to apply a larger perturbation to the lattice, for example by doping it or engineering the edges. As conformational engineering appears to be slightly too mild to bring about significant changes, the next step would be to go to structural engineering. Surely, future research in this direction will benefit from a combined theoretical and practical approach like the approach in this research.

But the bending of nanoribbons still has a lot of interesting aspects that require further research. The buckling effect is one of these. An in-plane molecular mechanics model can obviously not account for out-of-plane distortions. Similarly, a tight binding calculation is difficult to implement for buckled nanoribbons because of the fact that the wave function orthogonality is broken. Refined theories might elucidate interesting new physics for these kinds of geometries. And maybe, the scanning probe engineering "toolbox" can be extended so much in the future that ribbons (or other molecules) can be manipulated into knots and all kinds of exotic shapes, some of which might even display more interesting new physics.

Review of the project

I have worked for one and a half year on this research - mostly at a part time basis, to be able to follow classes - and it has been a thoroughly enjoyable time. One of the aspects of this research that I particularly liked was the freedom that I got to define my own project - rather than choosing a project from the shelve. I started with little knowledge on the field of scanning probe microscopy and no knowledge of graphene nanoribbons whatsoever. But I quickly managed to turn this around and put both the microscope and the computer to good work.

Not only my understanding was improved, but also my handling of (expensive) equipment. Subtle improvements in the handling of the scanning probe microscope and related machinery were commonly referred to in our group as *Fingerspitzegefühl*. It is difficult to quantize how much *Fingerspitzegefühl* was used or obtained in my time as a Master student in the basement at CMI, but its importance should not be underestimated.

When I started with reading literature on graphene nanoribbons, I already knew some things about quantum chemistry, but I became fascinated with the extent to which graphene and graphene nanostystems have become a playground for theoreticians. Transport properties, massless Fermions, spin coherence length, half-integer quantum Hall effect, spin-orbit coupling, ferromagnetism, pseudomagnetic field, rehybridization strain, Schottky rectification, gas sensing and spin valves are some of the key concepts in the vast world of physical phenomena that show up in graphene and graphene nanosystems. I thought that I could maybe put some theory in my research as well, and began by performing some DFT calculations.

I grew tired with the "black-boxyness" of quantum chemistry software and started to solve a simple empirical, nearest-neighbour Hamiltonian matrix in Mathematica. Knowing that this was a crude model, I did not expect too much of my first calculation. However, upon finding the band gap, gap states and a near-perfect match between orbitals predicted with this Hückel, or finite tight-binding model, and DFT, I was pleasantly surprised. This was the starting point of a trip into computational chemistry of nanosystems, and upon improving the model (and my understanding of *ab initio* methods) I was able to perform calculations on quantum-confined systems like nanoflakes and ribbons of various widths and relate them to established results. At first I was skeptical about my own model, but slowly I convinced myself to believe it.

At this point I started to make plans on what measurements to do on the STM - while cooperating on projects on quantum dots and epitaxial graphene to get used to operating the STM/AFM. I thought that I would have a chance at predicting something from my own model and backing it up if I could bend nanoribbons in the microscope and model them. There was absolutely no guarantee that this would work and whether there would be any effect. So I propose to do something on deformed ribbons - something that might be achieved with STM. The first step was making physically reliable models of nanoribbons, which I took as an opportunity to get some practice in Molecular Dynamics programming. Then the tight binding model was used and predictions on the band gap and level shifting were proposed.

The measurements were performed in 4 sets, while gradually increasing the *Fingerspitzegefühl* for STM manipulation, tip conditioning and obtaining clear, analysable differential conductance spectra. It was quite rewarding to controllably manipulate nanoribbons and observe the change by rescanning. After significant bending, we found the buckling effect - something I had not even thought about before.

It took some time before differential conductance spectra could be obtained clearly and reproducibly. No clear, consistent trend in band gap was established in the experiments, and level shifting was also not clear. The structure of the best spectra puzzled us at first, and we began brainstorming about what we were actually seeing. Clearly it had nothing to do with the - in retrospect rather naive - models that were made in the beginning. Electron-vibron coupling was a mechanism that was taken for granted in establishing the peak structure of single states, but the suggestion that various peaks in the spectra could

be higher-energy mode phonon replicas was a thought that I never considered and took some time to sink in. It makes a huge difference with what ideas and what mindset data is analysed (or re-analysed). At the same time, I did my best to analyse in an unbiased way.

In the end I can say that the calculations were instrumental in the understanding of what was going on in ribbons. Even though the effect that we eventually found is not really ground breaking, the fact that I was able to actually perform the difficult experiments in the STM ánd relate it to theory was very rewarding.

Acknowledgements

Most thanks go out to Joost van der Lit, Mark Boneschanscher and Ingmar Swart. Joost was my supervisor and learned me about the ins and outs of graphene nanoribbons. Most of anyone, he was the one that shaped and helped me define my project. He taught me how to use the STM, and we did many measurements together. Joost is a truly multi-faceted guy, with steady hands in the fume hood and great competence in cooking, beer brewing, electrical engineering and tinkering with the AFM/STM setup. Mark has taken the role of AFM teacher. In times of despair, he was the one who saved the sample transfer or got the qPlus tip to resonate. I am thankful for the complete course on graphene AFM he has provided, up to the point that I was able to make beautiful scans that we eventually published together. Ingmar has taught me a tremendous amount. His expertise on scanning probe microscopy and quantum modelling were invaluable for my research.

Jaco Geuchies used to be my partner in crime in the basement for the first half year, after which he graduated. He was already working on his project on quantum dots when I just arrived. I want to thank him for his help and for his Latin American music, which lightened the atmosphere in the basement at times when the machine was not cooperating.

Hans Ligthart, the technician of the group - apart from being a great guy to have around - is really practised in tinkering with the UHV setup. Stephan Zevenhuizen also deserves acknowledgements, not only for his help with computer problems, but also his dedication to writing scripts and programs facilitating measurements and tip preparation on the AFM/STM.

Fruitful discussions are key in group sessions about recently obtained results. Apart from the continuous input from Ingmar, Joost and Mark, many insights and suggestions were given by prof. Daniel Vanmaekelbergh. It was really nice to see his enthousiasm when my results appeared to conceal non-trivial physical effects, and his enthousiasm was an important motivation to keep working, keep improving on STM and AFM skills and keep thinking. Cristiane des Morais Smith and Simon Stuij deserve an acknowledgement for their involvement with the theoretical part of my research. Simon has already improved on my tight binding model in order to look at other types of strained nanosystems.

I would like to thank all my colleagues from Condensed Matter and Interfaces, for their discussions, viewpoints, support and social company. This includes visiting professor Daniel Gamelin, whose input on my research and career has to be acknowledged. Finally, many thanks go out to two people that, although spending a short time at CMI, have helped me tremendously in my work on the STM and AFM. Many thanks to Kiane de Kleijne (and good luck in Nijmegen) and Marlou Slot.

APPENDIX A. ELECTRON TRANSPORT

Diffusive transport

Charge tansport through metals takes place through all electrons, especially those near the Fermi energy. Charge transport through (intrinsic) semiconductors takes place through electrons promoted to the conductance band and holes remaining in the valence band. This can happen by applying a bias voltage, or through thermalization.

Diffusive transport is described by Drude theory[1]. Electrons and holes travelling through a lattice may be scattered through either phonons (distortions of the lattice itself) or through impurities. Scattering mechanisms act on a transient current by damping it in time

$$I(t) = I(0)e^{-k_{tot}t}$$

Both processes have a characteristic scattering rate and they combine to give a total scattering rate ($k_{tot} = k_{ph} + k_{imp}$). The reciprocal of the scattering rate is the characteristic scattering time τ. Although an electron in an electric field is accelerated according to

$$\dot{\mathbf{v}} = \frac{\mathbf{F}}{m} = \frac{e\mathbf{E}}{m}$$

Drude assumed that scattered electrons start off with a random velocity vector after each scattering event. Effectively, this makes the drift velocity - rather than the acceleration - proportional to the field strength.

$$\mathbf{v}(t) = \mathbf{v}_{rand} + \frac{e\mathbf{E}t}{m} \Rightarrow \langle \mathbf{v} \rangle = \frac{e\mathbf{E}\tau}{m}$$

The mobility μ is defined as $\mu = e\tau/m$, so that it becomes the proportionality constant between the drift velocity and the field. For the general case, in which the electrons and holes have differing mobilities, one can write

$$\mathbf{v}_e = \mu_e \mathbf{E},\ \mathbf{v}_h = \mu_h \mathbf{E}$$

The current is related to the drift velocity through $\mathbf{J} = ne\langle \mathbf{v} \rangle$ so that

$$\mathbf{J} = e(n_e\mu_e + n_h\mu_h)\mathbf{E} = \sigma\mathbf{E}$$

with $\sigma = e(n_e\mu_e + n_h\mu_h) = ne^2\tau/m$ the conductivity. So when a high conductivity is desired, one should increase both the mobilities and carrier concentrations as much as possible. The mobility is strongly dependent on the concentration and nature of impurities in a solid.

The reciprocal of conductivity is the resistivity ρ.

$$\mathbf{J} = \sigma\mathbf{E} \Leftrightarrow \mathbf{E} = \rho\mathbf{J}$$

For electric components with specific dimensions, it is in general not feasible to analyze the electric field and current density. The extrinsic form of the formula above (called Ohm's law) is then used, with conductance G or resistance R.

$$I = GV \Leftrightarrow V = RI$$

Nonlinear (non-Ohmic) behaviour may arise in systems, particularly under high bias conditions. In this case, the conductance is a function of bias voltage

$$I = G(V)V$$

An important material property is the differential conductance dI/dV, which is a key factor in characterization of materials through scanning tunnelling spectroscopy (STS). These parameters are particularly important in ballistic transport, rather than diffusive transport, where the Drude model breaks down.

Both the mobility and charge carrier concentration are material specific. For narrow-gap semiconductors, at non-zero temperatures, charge carriers are always present, as thermal excitations may continuously promote electrons in the conductance band (leaving behind holes in the valence band). Hence, the charge carrier occupation is not a single cut-off at the Fermi level, but follows Fermi-Dirac statistics arising from the steady state thermal energy collection and dissipation by electrons. The Fermi-Dirac (FD) distribution gives the chance to find an electron at a certain energy

$$\mathrm{FD}(E) = \frac{1}{1 + e^{(E-E_F)/kT}}$$

The electron occupation is found by multiplying this probability distribution by the density of states $\mathrm{DOS}(E)$

$$N(E) = \mathrm{DOS}(E)\mathrm{FD}(E)$$

The density of states can be obtained by solving the SE for the system. The total number of free charge carriers in thermal equilibrium is then found by integrating over the total number of holes in the conductance band and the total number of electrons in the valence band

$$n_e = \int_{E_F}^{\infty} \mathrm{DOS}(E)\mathrm{FD}(E)dE$$

$$n_h = \int_{-\infty}^{E_F} \mathrm{DOS}(E)(1 - \mathrm{FD}(E))dE$$

Diffusive conductivity of graphene

The Drude model can be used to write the conductance in terms of the mobility

$$\sigma = n\mu e = (n_h\mu_h + n_e\mu_e)e$$

where the subscripts h and e denote holes and electrons. For intrinsic graphene $n_h = n_e$, and, to a good approximation $\mu_h = \mu_e$, so that $\sigma \approx 2n_e\mu_e$. The density of states is approximately linear within a reasonable thermal region (26 meV at room temperature) from the Fermi level ($DOS(E-E_F) = cE$, with c a proportionality constant), and after setting $E_F = 0$, one can write

$$\sigma = 2\mu ec\int_0^\infty \frac{EdE}{1+e^{-E/(kT)}}$$

$$= 2\mu eck^2T^2\int_0^\infty \frac{xdx}{1+e^{-x}} = 2\frac{c\pi^2k^2\mu e}{12}T^2 \propto \mu eT^2$$

In this way, it is easily seen that the conductance is temperature dependent, and vanishes for zero temperature.

Ballistic transport

By definition, a ballistic conductor is a conductor whose length scale is smaller than the mean free path of an electron, so that there is no such thing as drift velocity. Naïvely one would be tempted to think that an electron could then be accelerated indefinitely, and that resistivity would drop to zero. This is not the case, and interestingly enough, it is established that resistance for the smallest metallic point contacts obtains a mysterious value of around 26 $k\Omega$. This value - the so-called conductance quantum G_0 - happens to be dictated by the Pauli exclusion principle.

A model for ballistic conduction is a twodimensional transport channel between two leads in thermal equilibrium, with chemical potential μ_L and μ_R for the left and right electrode respectively. These potentials are obtained by applying a bias voltage, so that $\mu_L = E_F + eV/2$ and $\mu_L = E_F - eV/2$. The channel has length L and width W, and has a constant potential which is set to the reference value $V = 0$. The outside potential is set to infinity, so that the eigenstates in the channel become Bloch waves of the form

$$\psi(x,y) = \sin\frac{n\pi y}{W}\ e^{ikx},\ n \in N$$

(a detailed explanation why this happens to be the case is given in the theory chapter). The energy of each wave is found by solving the time-independent Schrödinger equation

$$\hat{\mathrm{H}}\psi = \left(V(x,y) - \frac{\hbar^2}{2m}\nabla^2\right)\psi = E\psi$$

$$\Rightarrow E = \frac{n^2h^2}{8mW^2} + \frac{\hbar^2k^2}{2m}$$

This means that, under the assumption of a homogeneous potential, for every transverse wave profile - characterised by n - there is a parabolic band in the dispersion relation $E(k)$.

Now since the chemical potential in the left lead is higher than on the right lead, electrons want to travel from left to right, and in doing so, they can occupy all positive wave numbers k for which $\mu_R < E(k) < \mu_L$. Every electron takes a time $\tau(k) = L/v(k)$ to cross from the left lead to the right lead, and in doing so, it contributes a current $\delta I = e/\tau = ev(k)/L$. The total current is

$$I = 2\sum_k \delta I(k) = \frac{2e}{L\Delta k}\int v(k)dk$$

where the factor 2 accounts for spin and $\Delta k = 2\pi/L$ is the spacing between wave numbers in the conductor in reciprocal space. The integral over k is mapped onto an energy integral, for which we know the boundaries to be μ_l and μ_R

$$I = \frac{2eL}{2\pi L}\int v(k)\left(\frac{dE}{dk}\right)^{-1}dE$$

and using $dE/dk = \hbar v(k)$, the expression evaluates to

$$I = \frac{e}{\pi\hbar}(\mu_R - \mu_L) = \frac{2e^2}{h}V$$

$$G_0 = \frac{dI}{dV} = \frac{2e^2}{h} \approx (26\ k\Omega)^{-1}$$

The crux is that every electron with its own wave vector occupies the conduction channel, and in doing so, blocks the channel for other candidates with the same value of k because of the Pauli principle. This explains how the conductance quantum dictates the maximum achievable current in ballistic conduction *per transport channel*.

APPENDIX B. PRINCIPLES OF QUANTUM PHYSICS

Time-separability

The general electronic SE for a single electron in a potential energy landscape is written as

$$\hat{\mathrm{H}}\psi = \left(V(\mathbf{r},t) - \frac{\hbar^2\nabla^2}{2m}\right)\psi = i\hbar\frac{\partial\psi}{\partial t}$$

If the Hamiltonian is time-independent, it can do no work, and its operation on an eigenstate yields the eigenenergy of that state.

$$\hat{\mathrm{H}}(\mathbf{r})\psi = E_n\psi$$

with n a quantum number or a set of quantum numbers describing the specific state. As long as the Hamiltonian is Hermitian, the set of eigenfunctions of $\hat{\mathrm{H}}$ is orthogonal and its span is defined as the Hilbert space. We will look at time-independent Hamiltonians. In general, a wave function is separable in a space-dependent term and a time-evolving term.

$$\psi = T(t)\psi(\mathbf{r})$$

$$i\hbar\frac{\partial\psi}{\partial t} = i\hbar\psi(\mathbf{r})\frac{\partial T(t)}{\partial t} = T(t)\hat{\mathrm{H}}(\mathbf{r})\psi(\mathbf{r})$$

$$\Rightarrow T(t) = e^{-i\hat{\mathrm{H}}t/\hbar}$$

By absorbing the time-independent Hamiltonian, T becomes an operator. Its operation on $\psi(\mathbf{r})$ yields the full wave function, because if $\psi(\mathbf{r})$ is an eigenfunction of $\hat{\mathrm{H}}$, then

$$\psi = T(t)\psi(\mathbf{r}) = e^{-i\hat{\mathrm{H}}t/\hbar}\psi(\mathbf{r}) = \psi(\mathbf{r})e^{-i\omega t}$$

with $\omega = E/\hbar$. As long as the Hamiltonian is time-independent, we can solve the SE in two steps:

$$\hat{\mathrm{H}}(\mathbf{r})\psi(\mathbf{r}) = E_n\psi(\mathbf{r})$$

$$T(t)\psi(\mathbf{r}) = \psi(\mathbf{r})e^{-i\omega t} = \psi$$

It is noted, that for eigenstates, the operation by T comes down to the multiplication with a time-dependent phase factor $e^{-i\omega t}$.

Plane and standing electron waves

An electron in a large region with constant potential V is described by the following form of the time-independent SE and the following solution

$$\nabla^2\psi(\mathbf{r}) = -\frac{2m_e(E-V)}{\hbar^2}\psi(\mathbf{r}) = -k^2\psi(\mathbf{r})$$

$$\psi(\mathbf{r}) = c_1e^{i\mathbf{k}\cdot\mathbf{r}} + c_2e^{-i\mathbf{k}\cdot\mathbf{r}}$$

with c_1 and c_2 constants and $k^2 = k_x^2+k_y^2+k_z^2$ the squared magnitude of the wavevector $\mathbf{k}$. The eigenenergy of this wave is quadratic in the wave vector

$$E = V + \frac{\hbar^2k^2}{2m}$$

a result known as quadratic dispersion. We will consider the "right-going wave" solution $c_1 = c$ and $c_2 = 0$. c itself is a constant that may contain an initial phase if appropriate

$$c = c'e^{i\phi_0},\ c', \phi_0 \in \Re$$

The time-dependant wavefunction is obtained by simply multiplying this result with the time-evolution factor

$$\psi = ce^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

It will prove useful to point out that this plane wave is an eigenfunction of differential operators with respect to time and space, with the following eigenvalues

$$\nabla\psi = i\mathbf{k}\psi,\ \nabla^2\psi = -k^2\psi$$

$$\frac{\partial}{\partial t}\psi = -i\omega\psi,\ \frac{\partial^2}{\partial t^2}\psi = -\omega^2\psi$$

Also, the electron wave has (directional) momentum $\mathbf{p} = \hbar\mathbf{k}$. The velocity of the electron is the group velocity of the wave.

$$v_x = \frac{\partial\omega}{\partial k_x} = \frac{1}{\hbar}\frac{\partial E}{\partial k_x} \Rightarrow \mathbf{v}(\mathbf{k}) = \frac{1}{\hbar}\nabla_{\mathbf{k}}E$$

The case of an electron whose energy is lower than the background potential is important. From now on, the explicit space dependence is dropped.

$$\nabla^2\psi = -\frac{2m_e(E-V)}{\hbar^2}\psi = k'^2\psi$$

where the negative wavevector constant k^2 is replaced by the positive constant $k'^2 = -k^2$. Solutions are exponentials, rather than waves

$$\psi = c_1e^{\mathbf{k}'\cdot\mathbf{r}} + c_2e^{-\mathbf{k}'\cdot\mathbf{r}}$$

The nearly-free electron model

In the nearly-free electron model, the constant potential background V of the free-electron model is replaced by a perturbed potential $V(\mathbf{r})$ that varies across the unit cell. It is set up by the nuclei, which are periodic in the lattice, so V must also respect this periodicity, so that

$$V(\mathbf{r}+\mathbf{T}) = V(\mathbf{r})\ \forall\ \mathbf{T} \in \{\mathbf{T}\}$$

$$\Rightarrow V(\mathbf{r}) = \sum_{\mathbf{Q}} V_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\mathbf{r}}$$

The Bloch criterion states that plane electron waves in a non-constant potential also have a localized spatial part, which is periodic with the lattice

$$\psi = u(\mathbf{r}) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)},\ u(\mathbf{r}+\mathbf{T}) = u(\mathbf{r})\ \forall\ \mathbf{T} \in \{\mathbf{T}\}$$

$$\Rightarrow u(\mathbf{r}) = \sum_{\mathbf{Q}} u_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\mathbf{r}}$$

The Bloch waves themselves, as defined by the wave vector $\mathbf{k}$ within the Brillouin zone, are mixed linearly to arrive at energy minimization. This method is known as the variation principle, and it will come back more often. So every electron is described by a Fourier expansion

$$\psi(\mathbf{r}) = \int \psi(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{Q}} u_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\mathbf{r}} d\mathbf{r}$$

The time-independent SE becomes

$$\left[\sum_{\mathbf{Q}} V_{\mathbf{Q}} e^{-i\mathbf{Q}\cdot\mathbf{r}} - \frac{\hbar^2\nabla^2}{2m}\right] \int e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{Q}'} u'_{\mathbf{Q}} e^{-i\mathbf{Q}'\cdot\mathbf{r}} d\mathbf{r}$$

$$= \int \sum_{\mathbf{Q}} V_{\mathbf{Q}} e^{-i(\mathbf{k}-\mathbf{Q})\cdot\mathbf{r}} d\mathbf{r}$$

$$+ \int \frac{\hbar^2(\mathbf{k}-\mathbf{Q}')^2}{2m} \sum_{\mathbf{Q}'} u'_{\mathbf{Q}} e^{i(\mathbf{k}-\mathbf{Q}')\cdot\mathbf{r}} d\mathbf{r}$$

$$= E \int \sum_{\mathbf{Q}'} u_{\mathbf{Q}'} e^{i(\mathbf{k}-\mathbf{Q}')\cdot\mathbf{r}} d\mathbf{r}$$

$$\Rightarrow$$

$$E - \frac{\hbar^2(\mathbf{k}-\mathbf{Q}')^2}{2m} \sum_{\mathbf{Q}'} u'_{\mathbf{Q}} e^{i(\mathbf{k}-\mathbf{Q}')\cdot\mathbf{r}}$$

After plugging all expansions into the SE, it is effectively solved in $\mathbf{k}$-space to give a set of coupled equations.

$$\left(E_i - \frac{\hbar^2 k^2}{2m_e}\right)\psi(\mathbf{k}) = \sum_{\mathbf{Q}} V_{\mathbf{Q}} \psi(\mathbf{k}-\mathbf{Q})$$

These equations are the central equations. When a discrete set of wave vectors $\{\mathbf{k}\}$ in the first Brillouin zone is chosen, one arrives at a set of equations, that can be cast in matrix form. They become then the periodic variant of the secular equations.

The relativistic Klein-Gordon free-electron model

The extrapolation of the SE to make it relativistically correct is quite straightforward, but might not always be easy in its use. The kinetic energy operator in the SE

$$T = \frac{p^2}{2m_e} \Rightarrow \hat{\mathrm{T}}\psi = -\frac{\hbar^2}{2m_e}\nabla^2\psi$$

is replaced by the relativistic kinetic energy operator

$$T = \sqrt{(pc)^2 + (m_e c^2)^2}$$

$$\Rightarrow \hat{\mathrm{T}}\psi = \sqrt{-\hbar^2 c^2\nabla^2 + (m_e c^2)^2}\psi$$

Upon ignoring the potential energy, the kinetic energy becomes the total energy, so that the time-dependent SE becomes

$$\sqrt{-\hbar^2 c^2\nabla^2 + (m_e c^2)^2}\psi = i\hbar\frac{\partial}{\partial t}\psi$$

Now the procedure is to take the square of both operators

$$(-\hbar^2 c^2\nabla^2 + (m_e c^2)^2)\psi = -\hbar^2\frac{\partial^2\psi}{\partial t^2}$$

Upon rearranging terms, this is an inhomogeneous wave equation

$$(\Box + \mu^2)\psi = 0$$

with

$$\Box = \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \nabla^2,\ \mu = \frac{m_e c}{\hbar}$$

Substituting $\psi = ce^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ gives

$$(\Box + \mu^2)\psi = \left(-\frac{\omega^2}{c^2} + k^2 + \mu^2\right)\psi = 0$$

which is solved for

$$\frac{\omega}{c} = \pm\sqrt{k^2+\mu^2}$$

$$E(\mathbf{k}) = \pm\mu\hbar c\sqrt{1+\frac{k^2}{\mu^2}}$$

where the negative frequency solution is discarded. Now when the square root is expanded ($\sqrt{1+x} \approx 1+\frac{x}{2}$), we retrieve the relativistic rest mass and the free electron dispersion.

$$E(\mathbf{k}) = \mu\hbar c\left(1+\frac{k^2}{2\mu^2}\right) = mc^2 + \frac{\hbar^2 k^2}{2m}$$

However, in the limit that the mass goes to zero, we have

$$E(\mathbf{k}) = \lim_{\mu\to 0}\hbar c\sqrt{k^2+\mu^2} = \pm\hbar c|\mathbf{k}|$$

So once the problem of graphene is treated as a relativistic free electron gas, it is indeed possible to arrive at linear dispersion - in addition to parabolic dispersion.

APPENDIX C. LINEARITY OF THE GRAPHENE DISPERSION RELATION

Within the tight-binding model, the dispersion relation of graphene is

$$E_{\pm}(\mathbf{k}) = \mp\frac{\beta\sqrt{f(\mathbf{k})}}{1 \mp s\sqrt{f(\mathbf{k})}}$$

with

$$f(\mathbf{k}) = 3 + 2\cos\mathbf{k}\cdot\mathbf{T}_1 + 2\cos\mathbf{k}\cdot\mathbf{T}_2$$

$$+2\cos\mathbf{k}\cdot(\mathbf{T}_2 - \mathbf{T}_1)$$

The points K in reciprocal space are defined as the six points where the two bands touch (at the Fermi level). In terms of the reciprocal lattice vectors defined before, one of these points is at $\mathbf{k}_K = (\mathbf{Q}_1 - \mathbf{Q}_2)/3$. The inner product of this K-point vector with the lattice vectors is

$$\mathbf{k}_K\cdot\mathbf{T}_1 = \frac{\mathbf{Q}_1\cdot\mathbf{T}_1}{3} - \frac{\mathbf{Q}_2\cdot\mathbf{T}_1}{3} = \frac{2\pi}{3} - 0$$

$$\mathbf{k}_K\cdot\mathbf{T}_2 = 0 - \frac{2\pi}{3}$$

$$\mathbf{k}_K\cdot(\mathbf{T}_2 - \mathbf{T}_1) = \mathbf{k}_K\cdot\mathbf{T}_2 - \mathbf{k}_K\cdot\mathbf{T}_1 = -\frac{4\pi}{3}$$

We want to expand around the K-point. For small values of E (close to the Fermi level), so small values of $f(\mathbf{k})$, $1 + s\sqrt{f(\mathbf{k})} \approx 1$ so that $E \approx \pm\beta\sqrt{f(\mathbf{k})}$. We choose $\mathbf{k}$ close to $\mathbf{k}_K$: $\mathbf{k} = \mathbf{k}_K + \delta\mathbf{k}$, with $\delta\mathbf{k}$ close to zero. The vector products become

$$\mathbf{k}\cdot\mathbf{T}_1 = \delta\mathbf{k}\cdot\mathbf{T}_1 + \frac{2\pi}{3} = \theta_1 + \frac{2\pi}{3}$$

$$\mathbf{k}\cdot\mathbf{T}_2 = \theta_2 - \frac{2\pi}{3}$$

$$\mathbf{k}\cdot(\mathbf{T}_2 - \mathbf{T}_1) = \theta_2 - \theta_1 - \frac{4\pi}{3}$$

with

$$\theta_i = \delta\mathbf{k}\cdot\mathbf{T}_i = \frac{\sqrt{3}}{2}r_0\left(\sqrt{3}\delta k_x \pm \delta k_y\right)$$

Now

$$f(\mathbf{k}) = 3 + 2\cos\left(\theta_1 + \frac{2\pi}{3}\right) + 2\cos\left(\theta_2 - \frac{2\pi}{3}\right)$$

$$+2\cos\left(\theta_2 - \theta_1 - \frac{4\pi}{3}\right)$$

For small values of θ, the harmonic functions can be expanded

$$2\cos\left(\theta - \frac{2\pi}{3}\right) = \sqrt{3}\sin\theta - \cos\theta \approx \frac{\theta^2}{2} + \sqrt{3}\,\theta - 1$$

$$2\cos\left(\theta + \frac{2\pi}{3}\right) = -\sqrt{3}\sin\theta - \cos\theta \approx \frac{\theta^2}{2} - \sqrt{3}\,\theta - 1$$

After substituting these results into f, together with the identity $\cos(\theta - 4\pi/3) = \cos(\theta + 2\pi/3)$, all constant and linear terms vanish

$$f(\mathbf{k}) = 3 + \left(\frac{\theta_1^2}{2} + \sqrt{3}\,\theta_1 - 1\right) + \left(\frac{\theta_2^2}{2} - \sqrt{3}\,\theta_2 - 1\right)$$

$$+\left(\frac{(\theta_2 - \theta_1)^2}{2} + \sqrt{3}\,(\theta_2 - \theta_1) - 1\right)$$

$$= \theta_1^2 + \theta_2^2 - \theta_1\theta_2$$

Calculating the squares gives

$$f(\mathbf{k}) = \frac{9}{4}\left(\delta k_x^2 + \delta k_x^2\right)$$

so

$$E(\mathbf{k}) \approx \pm\beta\sqrt{f(\mathbf{k})} = \pm\frac{3}{2}\beta\sqrt{\delta k_x^2 + \delta k_x^2}$$

f has the formula of a circular paraboloid. The square root of f is a circular cone, so E is conical around the Fermi wave vector.

APPENDIX D. THE DRIVEN DAMPED HARMONIC OSCILLATOR

In a quadratic potential, a mass m on a spring with stiffness k will experience a restoring force $\mathbf{F} = m\ddot{\mathbf{x}} = -k\mathbf{x}$ as a result of a deflection, causing the mass to oscillate. Aditionally, there will always be some damping present, which is proportional to the velocity of the mass (with constant c). The equation of motion is

$$\ddot{\mathbf{x}} + \frac{c}{m}\dot{\mathbf{x}} + \omega_0^2\mathbf{x} = 0 = \ddot{\mathbf{x}} + 2\zeta\omega_0\dot{\mathbf{x}} + \omega_0^2\mathbf{x}$$

with $\omega_0 = \sqrt{k_0/m}$ and $\zeta = c/(2\sqrt{mk})$. The solutions are found by substituting the exponential *Ansatz* $\mathbf{x} = \mathbf{x}_0 e^{rt}$

$$(r^2 + 2\zeta\omega_0 r + \omega_0^2)\mathbf{x}_0 e^{rt} = 0$$

$$\Rightarrow r = \omega_0(-\zeta \pm i\sqrt{1-\zeta^2}) = -\omega_0\zeta \pm i\omega_1$$

$\omega_0\zeta$ represents the amplitude decay rate and $\omega_1 = \omega_0\sqrt{1-\zeta^2}$ the new eigenfrequency, which is slightly diminished from the undamped case. $1/\zeta\omega_0$ is the time after which the oscillation amplitude decays to e^{-1} of its original value and $2\pi/\zeta$ is the associated number of oscillations. The functional solution to the damped oscillator problem is

$$\mathbf{x}_T(t) = \mathbf{x}_0 e^{(\pm i\omega_1 - \zeta\omega_0)t}$$

The subscript T denotes that this solution is transient. The damping rate is related to the quality factor as

$$Q = \frac{1}{2\zeta}$$

The Q-factor is proportional to the number of oscillations before a certain amplitude decay is achieved.

An AFM-tip is dynamically driven. Consider an external harmonic force $\mathbf{F}_d e^{i\omega t}$ parallel to $\mathbf{x}$ with a driving frequency ω (without subscript). The new equation of motion becomes

$$\ddot{\mathbf{x}} + 2\zeta\omega_0\dot{\mathbf{x}} + \omega_0^2\mathbf{x} = \frac{\mathbf{F}_d}{m}e^{i\omega t}$$

The transient term is again the solution to the homogeneous equation, as derived above, but the steady state response of the oscillator has the same frequency as the driving force and is undamped. This is the term that eventually remains.

$$\mathbf{x} = \mathbf{x}_0 e^{i\omega t}$$

$$\Rightarrow (-\omega^2 + 2\zeta\omega_0\omega i + \omega_0^2)\mathbf{x}_0 e^{i\omega t} = \frac{\mathbf{F}_d}{m}e^{i\omega t}$$

This is harmonic oscillation with a complex amplitude of

$$\mathbf{x}_0 = \frac{\mathbf{F}_d}{m(2\zeta\omega_0\omega i + \omega_0^2 - \omega^2)}$$

This variable is readily decomposed into a amplitude magnitude and a phase

$$|\mathbf{x}_0| = \frac{|\mathbf{F}_d|}{m\sqrt{(2\omega\omega_0\zeta)^2 + (\omega^2 - \omega_0^2)^2}}$$

$$\arg\mathbf{x}_0 = \arctan\frac{2\omega\omega_0\zeta}{\omega^2 - \omega_0^2}$$

The maximum amplitude response is achieved by driving the oscillator at its eigenfrequency ($\omega = \omega_0$). When this condition is satisfied, the oscillator position lags the driving force by a phase shift of $\pi/2$.

$$|\mathbf{x}_0|_{max} = |\mathbf{x}_0|_{\omega=\omega_1} = \frac{|\mathbf{F}_d|}{2m\omega_0^2\zeta}$$

$$\arg\mathbf{x}_0(\omega = \omega_0) = \frac{\pi}{2}$$

APPENDIX E. MECHANISM OF GRAPHENE NANORIBBON SYNTHESIS

It is interesting to look a little bit into the chemistry of nanoribbons as they are prepared in ultrahigh vacuum following the method of Cai and coworkers[2]. UHV surface synthesis is really different kind of chemistry than the more usual "wet" synthetic chemistry, but also obeys some simple kinetic and thermodynamic principles that can be understood very well. Grasping the important points in the formation is important for the optimization of the growth process and for the design of new synthetic routes towards similar compounds.

Intuitively, one can single out the following steps in the reaction mechanism:

- 10,10'-dibromo-9,9'-bianthryl (DBBA) molecules are evaporated onto the surface. As long as the temperature is high enough, they will constantly diffuse over the surface.
- The molecules remain unreactive, until the carbon-bromine bond is thermally broken. This heterolytic reaction results in a free precursor radical and a bromine radical. When bromine radicals encounter precursor radicals, they are likely to recombine. This would mean that there is an equilibrium between radicals and intact molecules.
- The precursors will diffuse over the surface, the dynamics of which are governed by the energetic barriers of moving from one site to the next[3].
- When bromine radical encounter each other, they will form Br_2. When precursor radicals encounter each other, they will terminate by dimerization. This dimerization is an Ullmann coupling reaction[4]. Initiation can again proceed thermally in the dimer to allow the polymerization to propagate.
- Br_2 will very easily detach from the surface at the synthesis temperature, effectively driving the equilibrium mentioned before towards radical formation. Its desorption temperature was found to be about 240 ^{o} C[5].
- Of course, the remaining bromine atoms can also detach from the dimer, and the same steps as mentioned above will result in the formation of a trimer, and higher polymers.
- After the polymerization, the temperature is increased and the cyclodehydrogenation step can commence. Cyclodehydrogenation discards of hydrogen atoms in the longitudinal direction of the ribbon while forming carbon-carbon bonds between adjacent monomer units. The resulting ribbon is planar, as opposed to the eclipsed DBBA monomer. It is pulled down to the surface by its interaction with the gold, and in doing so, facilitates the dehydrogenation at the next site[6].
- The hydrogen atoms that are released on the surface in the cyclodehydrogenation can recombine at the nanoribbon radical end. Such a recombination event effectively terminates nanoribbon growth altogether.

The different mechanistic steps have been treated experimentally[5] and in computer models, and much insight has come from calculations[3][6]. The dimer formation can be seen as nucleation, and one can claim that key to the formation of long polymers is to suppress the rate of this step with respect to the polymerization rate. But the optimization of nanoribbon formation - finding the best synthetic parameters to grow long, free nanoribbons - is still research that needs to be carried out.

APPENDIX F. ELECTRON-VIBRON COUPLING

In general, electron-phonon coupling or electron-vibron coupling plays an important role in transport of electrons through molecules. If electrons can couple to a single molecular normal mode upon being transported through the double-barrier junction, then the conductivity can increase with a certain factor each time the bias window is increased by an amount $\hbar\omega$, where ω is the normal mode frequency.

Electron-phonon and electron-vibron coupling arise in electronic transistions as well as charging transitions, since the geometric ground state of the initial and final electronic situation is different. The addition of an electron to a nanoribbon will on average cause the carbon-carbon equilibrium bond distances to change due to the new electronic potential that the nuclei reside in. The probability of a so-called vibronic (vibrational electronic) transition is given by the Franck-Condon factor: the overlap of the nuclear ground state wave function in the electronic ground state with the nuclear nth excited state wave function of the charged state. So if the bias window is such that the electron can only enter the nuclear ground state of the charged state, it can do so with the Franck-Condon factor F_{00}. If the bias window is opened further by $\hbar\omega$, there is an additional tunneling pathway with probability proportional to F_{01}. In the differential conductance spectrum, therefore, all these vibronic pathways give rise to vibronic peaks, spaced by $\hbar\omega$. Since the spacing between the last peaks presented in the previous experiment set was so uniform, phonon replicas seems intuitively a fair explanation for the observed effects. However, it is difficult to rationalize which modes play a role in graphene nanoribbons. Also, as molecular systems become larger, multiple vibrons that may contribute to the differential conductance condense together into a continuous phononic density of states.

Classically, the forces in the nanoribbon framework can be approximated with a molecular mechanics method, where a spring exists between all atoms. Consider an atom with mass m, connected by stretched springs with constant k to its neighbours. The force on and resultant acceleration of the mass is

$$\mathbf{F} = k\sum_i(\mathbf{r}_i - \mathbf{r}),\ \ \ddot{\mathbf{r}} = \omega^2\sum_i(\mathbf{r}_i - \mathbf{r})$$

with $\omega^2 = k/m$ and where i runs over all nearest neighbours. Eigenmodes are vibrations for which all atoms undergo harmonic oscillation with the same eigenfrequency ω_0, so that

$$\ddot{\mathbf{r}} = \mathbf{r}_0\frac{\partial^2}{\partial t}e^{-i\omega_0 t} = -\omega_0^2\ \mathbf{r}$$

Therefore

$$\omega_0^2\ \mathbf{r} + \omega^2\sum_i(\mathbf{r}_i - \mathbf{r}) = 0$$

This last equation can be written in matrix form for all atoms as

$$\omega_0^2\begin{pmatrix}\vdots\\ \mathbf{r}_{nn1}\\ \mathbf{r}\\ \mathbf{r}_{nn2}\\ \vdots\end{pmatrix} + \mathrm{H}\begin{pmatrix}\vdots\\ \mathbf{r}_{nn1}\\ \mathbf{r}\\ \mathbf{r}_{nn2}\\ \vdots\end{pmatrix} = 0$$

with H the off-diagonal matrix coupling the displacements of the nearest neighbours.

$$\mathrm{H} = \begin{pmatrix}\ddots & \vdots & \vdots & \vdots & \\ \cdots & 0 & \omega^2 & 0 & \cdots\\ \cdots & \omega^2 & 0 & \omega^2 & \cdots\\ \cdots & 0 & \omega^2 & 0 & \cdots\\ & \vdots & \vdots & \vdots & \ddots\end{pmatrix}$$

This set of equations can be solved by writing

$$\left(\mathrm{H} + \omega_0^2\ \mathrm{I}\right)\begin{pmatrix}\vdots\\ \mathbf{r}_{nn1}\\ \mathbf{r}\\ \mathbf{r}_{nn2}\\ \vdots\end{pmatrix} = \mathbf{0}$$

which is exactly the same matrix eigendecomposition as in the tight-binding problems. Since H only contains nearest neighbour elements, the set of squared eigenfrequencies ω_0^2 corresponds very well to the eigenenergies of the tight-binding problem, which will be calculated in the next chapter. Although this is a really crude model of the phononic structure of GNR, it can be expected that the phononic DOS has some characteristics of the electronic DOS. Indeed, this is seen to be the case in graphene[7]. A harmonic nearest-neighbour model was used for graphene nanoribbons in work by Sanders *et al.*, showing that the lowest eigenfrequency modes in the 7-AGNR phonon DOS are radial breathing-like modes[8][9]. He went on to show that electrons can couple to this oscillation. Therefore, it can be expected that the electronic structure of nanoribbons, as measured with scanning tunneling spectroscopy, may contain characteristic phonon peaks.

APPENDIX G. ATOMIC FORCE MICROSCOPY

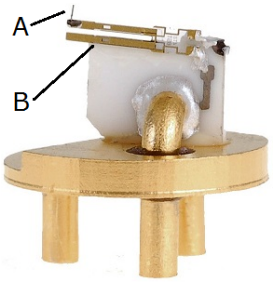


FIG. 1. A qPlus AFM/STM cantilever. A: tip. B: cantilever.

Atomic force microscopy (AFM) - like STM - uses a very small, sharp tip to image surfaces with possible atomic resolution. The difference is that in AFM - rather than the tunneling current - the mechanical response of the tip cantilever due to interaction forces with the substrate is measured. In the earliest (contact mode) AFM experiments, a tip was physically dragged across a surface, and the cantilever deflection could be recorded with great precision[10]. Later, non-contact techniques were developed.

The particular AFM method used in this research is non-contact frequency modulated AFM (nc FM-AFM)[11]. This technique involves setting up an oscillation in the tip at its resonance frequency. The oscillation is initially harmonic, meaning that the restoring force is linear with the cantilever displacement. However, in proximity with the surface, interactions give additional forces on the AFM tip, either adding to the restoring force or diminishing it. This results in a shift in resonance frequency (Δf).

The most important forces on the AFM tip are

- van der Waals attraction: This is an attractive force close to the surface, which lowers the resonance frequency of the tip
- Pauli repulsion: This is a repulsive force at much shorter range than the vdW force. The combined effect of vdW attraction and Pauli repulsion is the Lennard-Jones potential
- In some cases, electrostatic or magnetostatic effects may come into play[12][13].

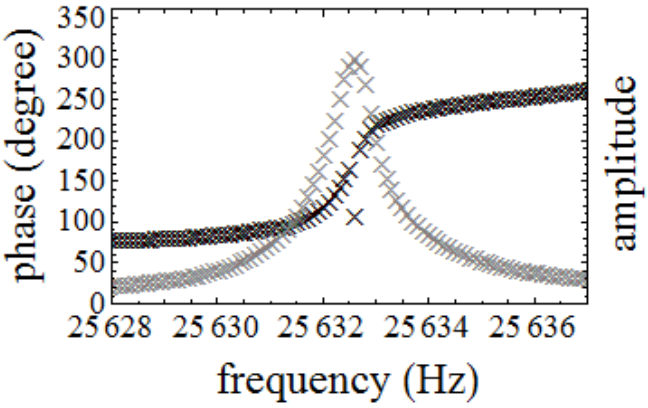


FIG. 2. Measured response curves (amplitude and phase shift) for a qPlus tip in a frequency sweep.

The AFM tip can be thought of as an driven harmonic oscillator, the dynamics of which are shown in appendix D and figure 2. The van der Waals and Pauli forces can be added as additional spring constants in the oscillator model, as long as they are approximately linear with displacement.

$$k = k_0 + k_{vdW} + k_{Pauli}$$

$$\omega = \omega_{eff}(1-\zeta^2) = \sqrt{\frac{k_0 + k_{vdW} + k_P}{m}(1-\zeta^2)}$$

The qPlus sensor is a specific AFM cantilever design (figure 1) allowing for an oscillation with extremely small amplitude - of the order of tens of pm. This makes the qPlus sensor suited for AFM measurements on organic and aromatic systems with atomic precision[14][**?**][15]. The quality factor - a measure for the sharpness of the resonance peak - can be as high as 200000, so that the damping factor ζ is negligible[16]. The fractional change in resonance frequency is

$$\frac{\Delta\omega}{\omega_0} = (\omega_{eff} - \omega_0)\omega_0^{-1}$$

$$\approx \left(\sqrt{\frac{k_0 + k_{vdW} + k_P}{m}} - \sqrt{\frac{k_0}{m}}\right)\sqrt{\frac{m}{k_0}}$$

$$= \sqrt{\frac{k_0 + k_{vdW} + k_P}{k_0}} - 1 = \sqrt{1 + \frac{k_{vdW} + k_P}{k_0}} - 1$$

Expanding the square root up to first order then gives

$$\frac{\Delta f}{f_0} = \frac{\Delta\omega}{\omega_0} \approx \frac{k_{vdW} + k_P}{2k_0}$$

The effective spring constant model is justified for the qPlus, but in general this model breaks down as a result of significant nonlinearity in the forces over the oscillation amplitude[17]. For larger oscillation amplitudes, the resonance frequency shift is found by integrating the interaction forces over the complete distance that the tip travels in an oscillation[18]. Since for large oscillation amplitudes the cycle averaged force constant only contains small contributions from the repulsive and attractive regimes, the imaging capability is compromised.

For atomic resolution the use of the qPlus force sensor is preferred[17][16].

The main difference between AFM and STM is that STM only probes states within the bias window - close to the Fermi energy - whereas the contrast in AFM is implicitly dependent - through the van der Waals and Pauli contributions - on the entire electron density. This means that, in a sense, STM really looks at the electronic structure whereas AFM looks at the geometric structure of a molecule.

Tip passivation

Although STM is sensitive to the electronic structure of the molecule, it might sometimes be difficult to get a good quality atomically sharp tip. For AFM, the proximity of the tip to the sample may induce large attractive interactions, and imaging in the repulsive regime might be difficult due to the reactivity of the metal atom on the end. The spatial resolution of STM and AFM images can be increased by passivating the tip, by picking up small molecules, for example carbon monoxide (CO)[19][20]. CO molecules may adsorb very well on metal surfaces - especially at extremely low temperatures - and they are relatively difficult to pump away in UHV conditions. The setup in the AFM-STM used in the group of Condensed Matter and Interfaces also allows for the gentle administering of a small amount of CO-gas to the cold sample. The contrast in AFM now contains mostly the repulsive Pauli repulsion, and for STM, sharp images are almost guaranteed as the tip terminates at a single point - the molecule. Now when one looks at the tunnel equation, there is a contribution of the electronic structure of the tip, and this significantly alters the appearance of features in STM.

Feedback and control

The setup used in this research (Omicron LT-AFM/STM with qPlus) allows AFM and STM to be performed at the same time. In microscopy mode, the three key parameters are current I, height z and frequency shift Δf, and feedback or control can be exerted on any of the three while measuring the other two. It is thus possible to perform crossover experiments like STM feedback Δf mapping, or AFM feedback I mapping.

The ability to perform AFM and STM calls for various feedback mechanisms. To be able to measure a frequency shift in AFM, it is necessary to have a stable oscillation in the first place. To achieve this, the qPlus sensor is driven by a harmonic oscillation voltage on the tip cantilever at the correct frequency. Here the driving voltage is converted to physical oscillations of the tip, which may drive the cantilever into resonance. Before measuring, the resonance frequency needs to be determined, which is done by applying a frequency sweep on the z piezo and reading out the response of the qPlus in terms of its phase and amplitude response. The response curves are displayed in figure 2. The eigenfrequency contains contributions of the interaction forces, and approaches the damped resonance value ω_1 when the tip is not in contact.

When approaching the surface, the eigenfrequency changes from ω_1 to ω_2, where now the van der Waals and Pauli (and electrostatic) contributions add to the effective spring constant. Would the tip still be driven at ω_1, then the resonance condition is not fulfilled anymore, and the oscillation decays quickly. Therefore, the AFM feedback circuit works by reading out the tip signal which contains the actual phase. The phase signal is amplified and shifted by 90 degrees, and is subsequently fed back to the tip, so that the oscillation remains stable at its dynamic eigenfrequency. The amplification factor in between is very precarious - slightly too low and the oscillation will decay, slightly too high and the oscillation will get out of hand. Also, the surface may induce additional damping, which needs to be compensated for. Therefore, the amplification factor itself - the gain - is determined by an amplitude feedback loop, that takes the magnitude of the actual oscillation as an input.

These feedback circuits ensure that a stable oscillation can occur at dynamically changing frequencies. The next thing is to either record this frequency - in constant height mode or STM feedback - or to use the frequency shift again as a feedback parameter to regulate the height of the tip (called frequency-modulated AFM feedback). A drawback is that a small change in eigenfrequency takes some time - of the order of ms to be determined correctly, so it is even better to look at the phase. The derivative of the phase response is maximal

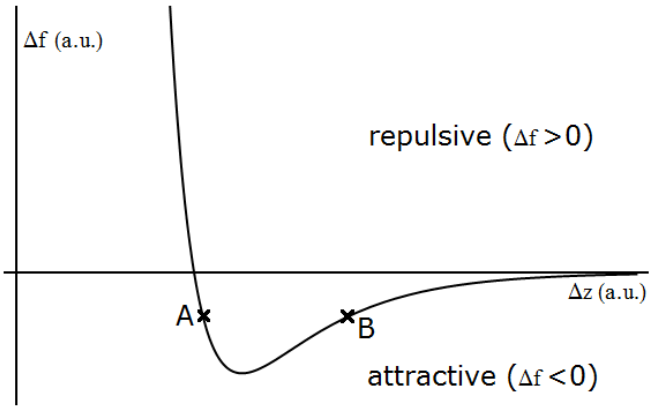


FIG. 3. The resonance frequency shift as a function of the tip height. Point A is a feedback setpoint on the repulsive branch, and point B is a feedback setpoint corresponding to the same frequency shift on the attractive branch.

at the resonance frequency, so that any change in the eigenfrequency shows up even more clearly in the phase shift, and hence feedback can be enhanced in the phase-locked loop (PLL) technique.

So what value of the frequency shift should be used when performing AFM in feedback? This is a slightly more difficult procedure than in STM, since the Lennards-Jones or Morse potential (the interaction curve that is the sum of the attractive van der Waals energy and the repulsive Pauli energy) is neither uniformly increasing nor decreasing. The resonance frequecy shift is proportional to the effective spring constant, which is the second derivative of the interaction curve an has again a similar shape, as shown in figure 3. Now when the feedback setpoint is in the well, the lock may in some cases jump from the repulsive branch (point A in figure 3) to the attractive branch (point B), where the feedback set point for the resonance frequency shift is identical for the two points. This should obviously be avoided. But since the well depth and distance are generally specific of the molecule or metal, scanning across both may induce flipping between attractive contrast to repulsive contrast. Of course, when going further away in the attractive regime, this is not a big issue, but at the same time one loses all atomic contrast[15]. Therefore, AFM in feedback is almost never used in our lab. Nevertheless, when preparing for AFM in feedback, one starts with a spectroscopy experiment. Δf can be measured by changing the tip-sample distance z. Having performed a $\Delta f(z)$ spectroscopy experiment, a frequency setpoint in the Lennard-Jones curve can be selected to allow for AFM feedback microscopy.

Using the qPlus sensor, AFM and STM can be combined in various experiments. An AFM excitation can be set up without significantly disturbing STM experiments, and both current and frequency shift can be monitored. Knowing that both variables can be used as feedback parameters, this allows the following crossover imaging experiments to be performed

- Simultaneous mapping of the current and the frequency shift at a constant height
- Performing AFM frequency-shift measurements on the height profile defined by STM feedback
- Performing STM current measurements on the height profile defined by AFM feedback
- Mapping the force (by integrating the frequency shift at different heights) as a function of bias

A cooperation where I have used AFM is the experimental work on the graphene moiré corrugation as described in appendix H.

APPENDIX H. COOPERATIONS

During my research, I have also been involved in projects with colleagues, which are related to my main research. These have provided me with more experience with scanning probe and simulations.

Analysis of the Moiré corrugation

We have synthesized graphene by chemical vapor deposition (CVD). In CVD, ethylene gas is passed over a metal at elevated temperature, allowing it to decompose. Hydrogen atoms will then dissociate from the surface after combining to dihydrogen. Upon gentle cooling, the carbon atoms will form flakes of monolayer graphene[21]. In this research, the Ir(111) surface was used for graphene CVD.

The hexagonal Ir(111) lattice has a slightly larger lattice constant than graphene, causing a lattice mismatch. The lattices are said to be incommensurate. The position of the honeycombs on the iridium is periodic with 10 iridium lattice constants and 10 graphene lattice constants, so that the mismatch can be written as Ir(111)-(9×9)-graphene-(10×10)[22]. This causes the attraction between the graphene sheet and the metal to depend on the position in this mismatch pattern. Like a blanket, the graphene sheet bulges between these positions. This is called the Moiré corrugation. Both the Ir(111) lattice and any graphene sublattice are hexagonal, causing the Moiré pattern to be hexagonal as well. We have looked at this Moiré corrugation in detail using non-contact AFM[23], some results of which are shown in figure 4.

A nice way to look at graphene and its corrugation is by means of image Fourier analysis. The twodimensional fast Fourier transform (FFT) basically calculates the amplitudes of discrete waves that make up the image in two dimensions. The apparent height of the graphene is a convolution of the atomic graphene lattice and the Moiré corrugation. As both effects feature a well-defined periodicity they are easily identified in the 2D image FFT. The Moiré pattern has a much larger wavelength than the atomic lattice, so it will be seen in the Fourier transform as a series of peaks at much lower wavevector. Applying a high-pass filter basically filters out the components of the Moiré corrugation. The periodic component that is maintained is simply the uncorrugated honeycomb lattice of graphene. Applying a low pass filter does the exact opposite. The honeycomb lattice is now disregarded, and the remaining image is the Moiré corrugation only. This Fourier filtering is shown in figure 4.

Periodic density functional theory calculations and LEED-I(V) measurements have been performed on the iridium-graphene Moiré system. We have supplemented these data by performing atomic force microscopy measurements (AFM) to determine the height of the Moiré corrugation. As explained in the theory, AFM measurements are capable of giving an atomically precise topological images, yielding an apparent height as a function of tip position. Noise in AFM measurements can be as low as a few pm[23]. From AFM image analysis it was concluded that the corrugation height was between 47 ± 5pm. Furthermore, the local corrugation height was found to be uncorrelated to the graphene island size and orientation. The "second order Moiré corrugation", that would arise because of a slight remaining lattice mismatch in the first Moiré pattern, was not found to be preserved accurately over its range. The local corrugation was measured over a relatively large area (see figure 4). Apart from most peaks fallign in the range of 47 ± 5 pm, a few outliers were observed in a relatively narrow range between 58 and 62 pm. These might exist because of defects in the iridium, reflecting local strain relief rather than a second order Moiré. These findings have been the subject of a publication[23].

A striking feature that clearly shows up in the graphene atomic lattice after applying a high-pass filter (figure 4) is the apparent lateral distortion of the lattice. This distortion is not predicted by periodic density functional theory calculations, so it was thought in our group that the origin is a tip relaxation effect. So the next question was: What are the bond lengths throughout the Moiré supercell, and how do tip relaxation effects affect the apparent bond length? The tip relaxation effects throughout the graphene were nicely quantized in the subsequent paper by Mark Boneschanscher and Ingmar Swart[24].

Artificial graphene and silicene

Quantum dots are tiny pieces or clusters of metal or semiconductor containing a number of electrons. The accessible region for electrons can be defined electrostatically, or by simply making the material into nanometer-sized particles. Rather than residing in continuous energy bands, the electrons in a quantum dot can really be observed to reside at discrete energy levels due to their quantum confinement (which is explained in chapter 2). In many respects, quantum dots should be viewed as "artificial atoms". Since the properties of graphene arise mainly from its honeycomb lattice, substitution of quantum dots as artificial atoms in stead of carbon atoms can open up a new world of tunable "designer Dirac materials"[25].

Effort has been done in the group of Condensed Matter and Interfaces to build artificial graphene lattices from CdSe quantum dots. Tight-binding calculations show that the band structure of artificial graphene are indeed similar to those of normal, carbon-based graphene. However, because of the different (quasi)atomic system, the Fermi level is not at the Dirac point. After synthesis,

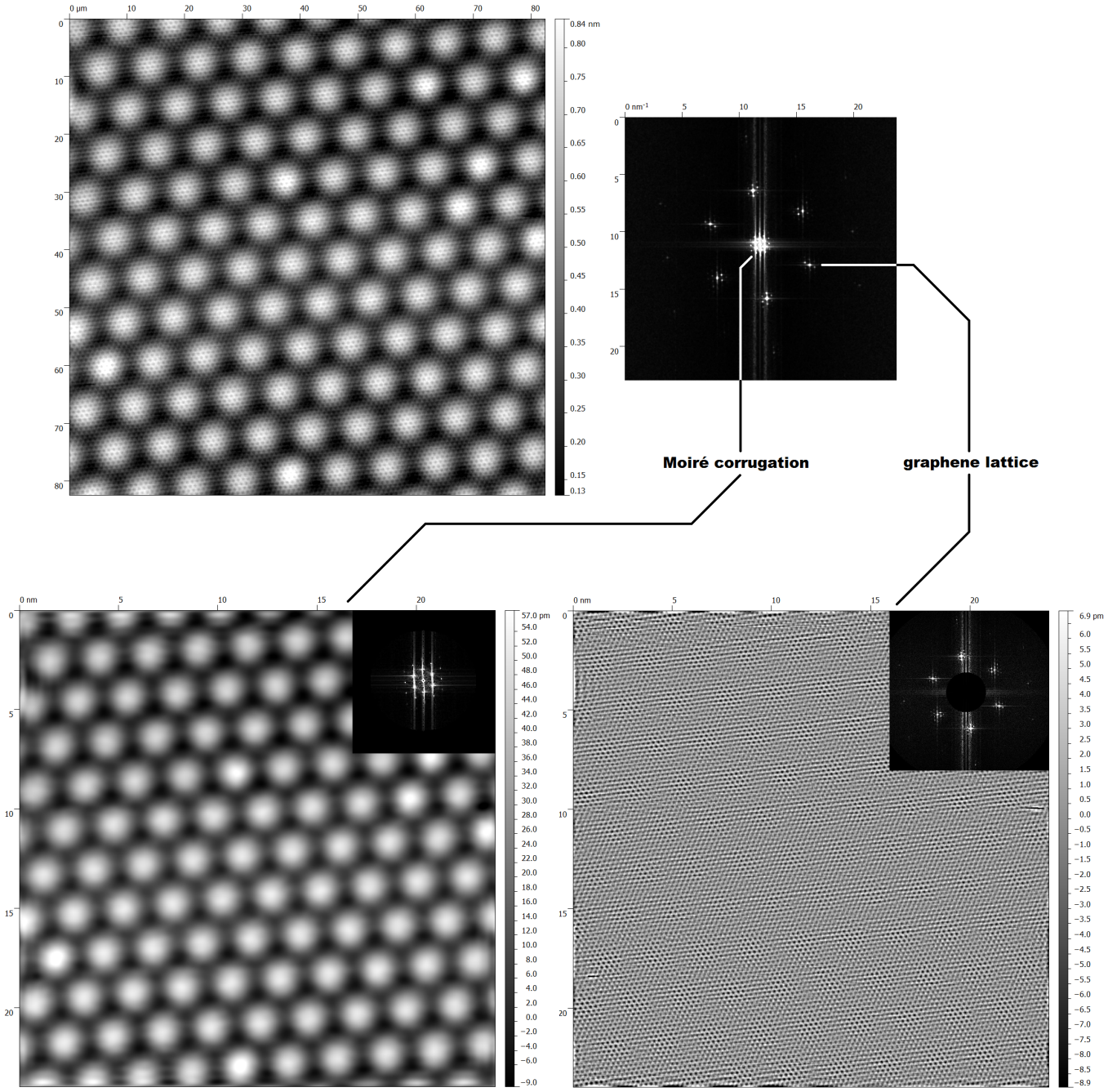


FIG. 4. Constant frequeny shift atomic force microscopy image on graphene, grown epitaxially on Ir(111). Both the hexagonal Moiré pattern and the graphene lattice are clearly visible. The 2D image Fourier transform features characteristic Moiré peaks at small wavevector and graphene lattice peaks at higher wavevector. Low-pass FFT filtering removes the atomic lattice, yielding the pure Moiré pattern (left). High pass filtering removes the Moiré corrugation and gives an image of the pure lattice (right).

characterisation of the macrostructure with various techniques - TEM and STM - has shown the presence of a honeycomb lattice, which turned ou to be of silicene type. Silicene is the silicon equivalent of graphene, where the two sublattices are at a slightly different height.

The author - despite not generating the data that was used for the publications - has been involved in the STM measurements on the quantum dot superlattices. Nevertheless, he was glad to see this project culminate to a publication in Science and see it grow to become a large interuniversity program[26].

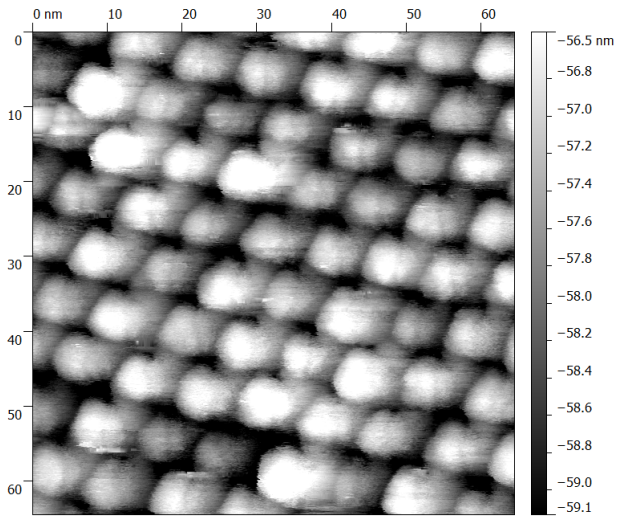


FIG. 5. Quantum dots, self-assembled into a silicene-type lattice

Self-assembly of organic networks

Another promising road ahead for artificial graphenes is the construction of twodimensional organic and metal-organic frameworks. The idea is similar to quantum dot graphene. The organic moieties should have delocalized orbitals, and upon interconnecting them on a surface into a honeycomb geometry, these molecular orbitals should again serve as a basis of the graphene bands. The ultimate goal is again to find the relativistic conduction properties - the Dirac cones. Like quantum dot graphene, the choice of multiple precursors and reaction conditions make these kinds of artificial graphenes much more flexible than the "boring" carbon graphene.

Apart from the usual considerations in electronic structure, one should look into the mechanism of oriented attachment to be able to understand and predict how organic superlattices, and what kind of metal-organic superlattices are formed. Experimental research was conducted in the group of Condensed Matter and Interfaces into evaporating molecular precursors onto metal surfaces, allowing them to self-assemble. The resulting superstructures were probed with STM and AFM, both with and without metal atoms. The molecules that have been investigated so far were bi-parapyridine acetylene (BPPA), bi-paracyano phenyl acetylene (BPCA) and bi-parabenzoic acid acetylene (BPBA) (see figure 6) .

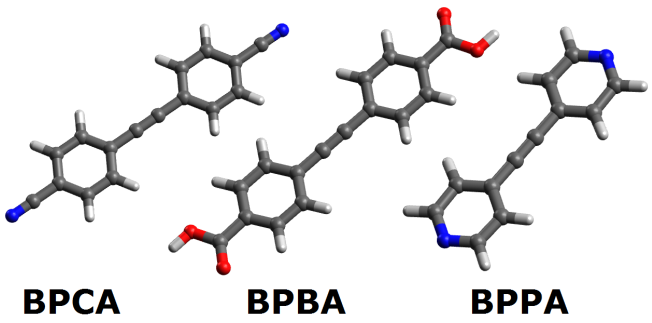


FIG. 6. The three molecules on which the Monte Carlo simulations were performed.

The binding interaction between these molecules is - to first order - van der Waals attraction. Van der Waals attraction can roughly be explained as a correlation effect, in which the momentary dipole of one molecule induces a dipole in the other, causing intermolecular attraction. Empirical models exist in which the van der Waals attraction strength is parametrized in terms of an atomic van der Waals radius R. The intermolecular interaction is found by summing over 6:12 potentials of the form

$$E = 4\epsilon_{ij}\left(\frac{\Sigma R}{\Delta r}\right)^6\left[\left(\frac{\Sigma R}{\Delta r}\right)^6 - 1\right]$$

with $\Sigma R = R_i + R_j$. However, these dispersive forces alone cannot account for all the structures that are seen. For example, pyridines are sometimes observed to form ribbons, and acids can bind with their carbonyl groups to form linear, square, or triangular intersections. Obviously, there is some electrostatic origin to these effects. Indeed, it was found that the mutual interactions between molecules can be described by an electrostatic model based on the atomic charges and the average positions of electrons in their $p_\perp$-orbitals, above and below the molecular plane of these flat molecules. This model, put forward by Hunter and Sanders, treats every atom in the molecule as a quadrupole, containing a positive nucleus monopole and two negative valence electron monopoles on either side of the molecular plane[27]. The Coulomb energy is calculated simply by summing over all point charges

$$E = \sum_{ij}\frac{q_i q_j}{4\pi\epsilon_0 \Delta r_{ij}}$$

where i is in molecule 1 and j is in molecule 2.

To first order, a carbon atom is treated as a $+1$ nuclear charge with two -0.5 electrons above and below the plane. Hydrogen simply doesn't have any out-of-plane charge. Nitrogen has a higher nuclear charge, and it is assigned $+1.5$ and -0.75, for the in-plane and out-of-plane contributions respectively. Oxygen is then $+2$ and -1 respectively. The polarization on heteroatoms and hydrogens can be accounted for by manipulating the charge distribution of the electron cloud. This charge distribution was adjusted according to first-principles calculations.

Also, since the molecules are evaporated on a conducting metal surface, the electric field lines from the molecules should be perpendicular to the metal surface.

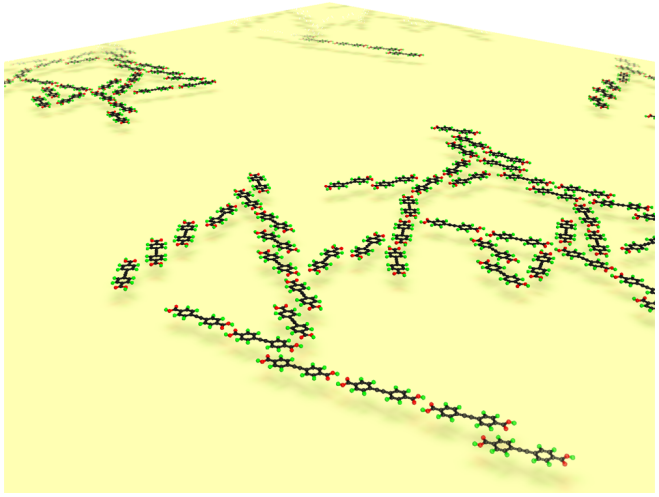


FIG. 7. An artistically rendered still from the molecular monte Carlo simulations on biacid molecules on gold

This is accounted for by generating a mirror image of the molecule, with inverse point charges. The effect introduces an interaction between one molecule with the mirror image of the other molecule, and vice versa. Using this collection of point charges, it is possible to find energetic minima as a function of the coordinates (x, y, θ) of each molecule in an ensemble of molecules. Multiple methods were used to find the energetic minima. The first approach - devised by Joost van der Lit - uses a Newton minimization method to displace a set of molecules, placed at random, towards their energetic minima. This method is very costly for multiple molecules, as this is basically a $3N$-dimensional problem. Therefore, I have implemented a Monte Carlo simulation, which works as follows

- Start with an ensemble of molecules, either randomized or in a crystal structure.
- Pick a molecule i. Calculate the electrostatic energy E_i with respect to its environment by summing over all electrostatic and van der Waals contributions of all atoms of all other molecules. The van der Waals contribution is found using a slightly different Lennard-Jones formula, used in molecular mechanics force fields. The van der Waals parameters were taken from the AMBER force field. The resulting energy is called E_{old}.

$$E_{old} = \sum_{ij} \sqrt{\epsilon_i \epsilon_j} \left(\frac{\Sigma R*}{\Delta r} \right)^6 \left[\left(\frac{\Sigma R*}{\Delta r} \right)^6 - 2 \right]$$

$$+ \sum_{ij} \sum_{kl} \frac{q_i k q_j l}{4\pi\epsilon_0 \Delta r_{ik,jl}}$$

The last term not only sums over atom pairs ij, but over all monopole charges in the quadrupole expansion kl and their mirror image contributions.

- Propose a displacement $(dx, dy, d\theta)$ and update the molecular coordinates (x, y, θ). Displace the molecule across the surface to this new position. The implementation initially performs a rotation matrix operation on the coordinates of the molecule in the origin, and then places its center at (x, y).
- Calculate the energy of molecule i again. This energy is called E_{new}.
- Calculate the normalized Boltzmann weight of this displacement. This is

$$W = \exp \frac{E_{new} - E_{old}}{kT}$$

- Pick a random real number x from the interval $[0, 1]$. If $W < x$: reject the step by moving the molecule back, using a displacement $-(dx, dy, d\theta)$. If $W > x$: accept the step.
- Pick a new molecule and repeat.

The molecules are read out from *.xyz* input files. The Monte Carlo model was created using periodic boundary conditions, using the method of nearest images.

It took us a while to get to the point where the simulation runs gave good results - the self-assembly of the

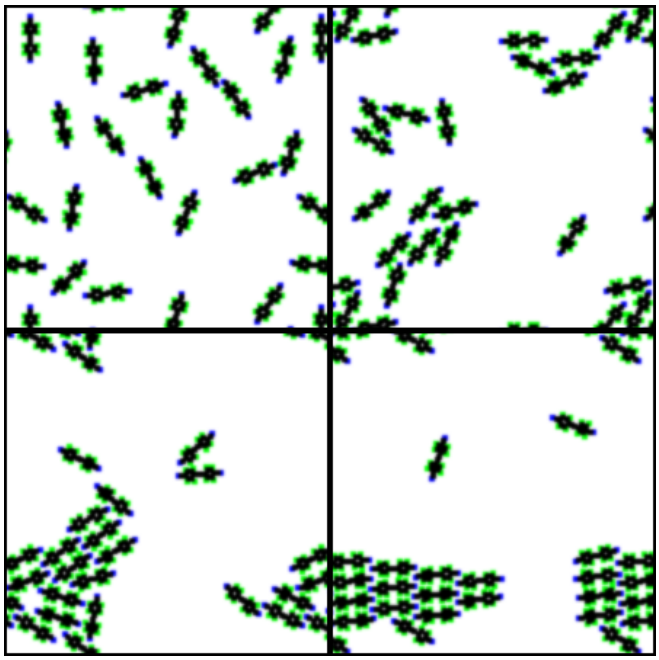


FIG. 8. Self-assembly of bicyano molecules into a hexagonal packed structure, as found with Monte Carlo

surface structures towards the experimental results. Examples are shown in figure 7 for the acids, which primarily form chains and later stick together side-by-side, and in figure 8, where the cyanide molecules form a hexagonal close-packed structure. The first implementation used an average van der Waals well depth. Mixing up calories and Joules, it turned out to be a factor 4.18 too deep. But after correcting this, strange behaviour was seen - arising from a bug in the cut-off radius. Then still, the behaviour of the molecules was not as we hoped for, and searches through parameter space were performed - varying δ, the van der Waals depth and the absorption height. It turned out to be good practice to energetically compare multiple observed crystal structures - also the wrong ones - while doing so. Eventually, we found that the absorption height had only a minor effect over the permissible range between roughly 2.5 to 4 Å. I eventually decided to look up the AMBER parameters to end the simulations where the "effective average van der Waals depth" had to be refined. Still, this left δ to be varied.

Every time we had some promising parameters, a melting simulation was carried out for each molecule to determine the temperature at which the molecules have a twodimensional fluid or gas phase in equilibrium with the solid phase. After this, we performed constant temperature experiments at this melting transition - sometimes combined with slow simulated annealing through the melting point. It is important to give these simulations the time, as carrying out an annealing step too fast results in a quenched amorphic structure. This results from molecules not having enough time to find their global minimum positions and becoming trapped in a local minimum.

We have a good hope in using this model in the future to improve our understanding of the self-assembly mechanism of organic molecules on surfaces and use it to predict or design twodimensional frameworks that host advantageous electronic properties.

[1] P. Drude, Annalen der Physik **306**, 566 (1900).

[2] J. Cai, P. Ruffieux, R. Jaafar, M. Bieri, T. Braun, S. Blankenburg, M. Muoth, A. P. Seitsonen, M. Saleh, X. Feng, K. Müllen, and R. Fasel, Nature **466**, 470 (2010).

[3] J. Björk, F. Hanke, and S. Stafström, Journal of the American Chemical Society **135**, 5768 (2013).

[4] M. Xi and B. E. Bent, Journal of the American Chemical Society **115**, 7426 (1993).

[5] A. Batra, D. Cvetko, G. Kladnik, O. Adak, C. Cardoso, A. Ferretti, D. Prezzi, E. Molinari, A. Morgante, and L. Venkataraman, Chemical Science **00**, 1 (2014).

[6] J. Björk, S. Stafström, and F. Hanke, Journal of the American Chemical Society **133**, 14884 (2011).

[7] L. Wirtz and A. Rubio, Solid State Communications **131**, 141 (2004).

[8] G. Sanders, A. Nugraha, R. Saito, and C. Stanton, Physical Review B , 1 (2012), arXiv:1201.5339v3.

[9] G. D. Sanders, A. R. T. Nugraha, K. Sato, J.-H. Kim, J. Kono, R. Saito, and C. J. Stanton, Journal of physics **25**, 144201 (2013).

[10] G. Binnig, "Atomic force microscope and method for imaging surfaces with atomic resolution," (1988), uS Patent 4,724,318.

[11] T. R. Albrecht, P. Grütter, D. Horne, and D. Rugar, Journal of Applied Physics **69**, 668 (1991).

[12] Y. Martin and H. K. Wickramasinghe, Applied Physics Letters **50** (1987).

[13] M. Nonnenmacher, M. P. OBoyle, and H. K. Wickramasinghe, Applied Physics Letters **58** (1991).

[14] L. Gross, F. Mohn, N. Moll, P. Liljeroth, and G. Meyer, Science (New York, N.Y.) **325**, 1110 (2009).

[15] L. Gross, Nature chemistry **3**, 273 (2011).

[16] F. J. Giessibl, F. Pielmeier, T. Eguchi, T. An, and Y. Hasegawa, Phys. Rev. B **84**, 125409 (2011).

[17] F. J. Giessibl, Rev. Mod. Phys. **75**, 949 (2003).

[18] I. Swart, L. Gross, and P. Liljeroth, Chem. Commun. **47**, 9011 (2011).

[19] L. Gross, F. Mohn, N. Moll, P. Liljeroth, and G. Meyer, Science **325**, 1110 (2009).

[20] Z. Sun, M. P. Boneschanscher, I. Swart, D. Vanmaekelbergh, and P. Liljeroth, Phys. Rev. Lett. **106**, 046104 (2011).

[21] T. A. Land, T. Michely, R. J. Behm, and J. C. Hemminger, Surface Science **264**, 261 (1992).

[22] A. T. N'Diaye, J. Coraux, T. N. Plasa, C. Busse, and T. Michely, New Journal of Physics **10**, 043033 (2008).

[23] S. K. Hämäläinen, M. P. Boneschanscher, P. H. Jacobse, I. Swart, K. Pussi, W. Moritz, J. Lahtinen, P. Liljeroth, and J. Sainio, Phys. Rev. B **88**, 201406 (2013).

[24] M. P. Boneschanscher, S. K. Hmlinen, P. Liljeroth, and I. Swart, ACS Nano **8**, 3006 (2014).

[25] M. Polini, F. Guinea, M. Lewenstein, H. C. Manoharan, and V. Pellegrini, Nature nanotechnology **8**, 625 (2013).

[26] M. P. Boneschanscher, W. H. Evers, J. J. Geuchies, T. Altantzis, B. Goris, F. T. Rabouw, S. A. P. van Rossum, H. S. J. van der Zant, L. D. A. Siebbeles, G. Van Tendeloo, I. Swart, J. Hilhorst, A. V. Petukhov, S. Bals, and D. Vanmaekelbergh, Science **344**, 1377 (2014).

[27] C. A. Hunter and J. K. M. Sanders, Journal of the American Chemical Society **112**, 5525 (1990).