Introduction to the Physical Properties of Graphene

Jean-Noël FUCHS   Mark Oliver GOERBIG

Lecture Notes 2008

Contents

Chapter 1

Introduction to Carbon Materials

The experimental and theoretical study of graphene, two-dimensional (2D) graphite, is an extremely rapidly growing field of today's condensed matter research. A milestone was the experimental evidence of an unusual quantum Hall effect reported in September 2005 by two different groups, the Manchester group led by Andre Geim and a Columbia-Princeton collaboration led by Philip Kim and Horst Stormer [1, 2]. Since this moment and until March 2008, the beginning of the present series of lectures, almost 700 manuscripts with "graphene" in their title have been posted on the preprint server www.arXiv.org (see Fig. 1.1).

The reasons for this enormous scientific interest are manyfold, but one may highlight some major motivations. First, one may emphasise its possible technological potential. One of the first publications on graphene in 2004 by the Geim group proved indeed the possibility of an electric field effect in graphene, i.e. the possibility to control the carrier density in the graphene sheet by simple application of a gate voltage [3]. This effect is a fundamental element for the design of electronic devices. Today's silicon-based electronics reaches its limits in miniaturisation, which is on the order of 50 nm for an electric channel, whereas it has been shown that a narrow graphene strip with a width of only a few nanometers may be used as a transistor [4], i.e. the basic electronics component. One may therefore hope to improve the miniaturisation by one order of magnitude when using graphene-based electronics.

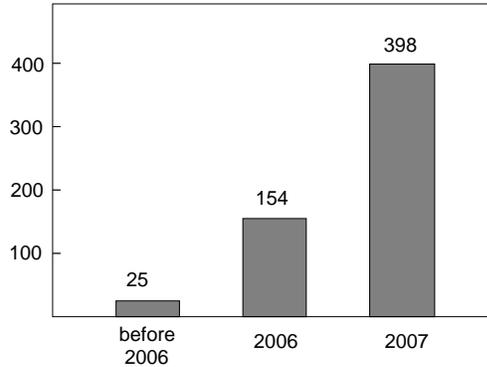Apart from these promising technological applications, two major moti-

Figure 1.1: Number of manuscripts with "graphene" in the title posted on the preprint server. In interpreting these numbers, one must, however, consider that several publications on graphene appeared before 2006, e.g. in the framework of carbon-nanotube or graphite research. At this moment, the name "graphene" was not commonly used.

vations for fundamental research may be emphasised. Graphene is the first truely 2D crystal ever observed in nature. This is remarkable because the existence of 2D crystals has often been doubted in the past, namely due to a theorem (Mermin-Wagner theorem) which states that a 2D crystal looses its long-range order, and thus melts, at any small but non-zero temperature, due to thermal fluctuations. Furthermore, electrons in graphene show relativistic behaviour, and the system is therefore an ideal candidate for the test of quantum-field theoretical models which have been developed in high-energy physics. Most promenently, electrons in graphene may be viewed as massless charged fermions living in 2D space, particles one usually does not encounter in our three-dimensional world. Indeed, all massless elementary particles happen to be electrically neutral, such as photons or neutrinos.[1] Graphene is therefore an exciting bridge between condensed-matter and high-energy physics, and the research on its electronic properties unites scientists with various thematic backgrounds. The discussion of graphene's electronic properties and how such relativistic effects are revealed in electric transport measurements is naturally a prominent part of the present lecture notes.

The interest in graphene is not only limited to the scientific community. An important number of large-audience articles have recently been published. The following list of (more or less serious) citations reveals this broad interest

---

[1]The neutrino example is only partially correct. It has been shown that neutrinos must indeed have a tiny non-zero mass.

> "Electrons travel through it so fast that their behaviour is governed by the theory of relativity rather than classical physics."
>
> (The Economist, 2006)

> "Inside every pencil, there is a neutron star waiting to get out."
>
> (New Scientist, 2006)

> "We'll have to rewrite the theory of metals for this problem."
>
> (Physics Today, 2006)

Notice that the last citation is from a leading scientist in the field and may indeed serve as a guideline for these lectures.[2]

1.1   The Carbon Atom and its Hybridisations

Carbon, the elementary constituent of graphene and graphite, is the 6th element of the periodic table. Its atom is, therefore, built from 6 protons, A neutrons, and 6 electrons, where $A = 6$ and 7 yield the stable isotopes ^{12}C and ^{13}C, respectively, and $A = 8$ characterises the radioactive isotope ^{14}C. The isotope ^{12}C, with a nuclear spin $I = 0$, is the most common one in nature with 99% of all carbon atoms, whereas only 1% are ^{13}C with a nuclear spin $I = 1/2$. There are only traces of ^{14}C (10^{-12} of all carbon atoms) which β-decays into nitrogen ^{14}N. Although ^{14}C only occurs rarely, it is an important isotope used for historical dating (radiocarbon). Due to its half-life of $5\,700$ years, which corresponds to a reasonable time scale in human history, measurement of the ^{14}C concentration of an organic material, mainly wood, allows one to date its biological activity up to a maximum age of roughly $80\,000$ years. In general, carbon is the elementary building block of all organic molecules and, therefore, responsible for life on Earth.

In the atomic ground state, the 6 electrons are in the configuration $1s^2 2s^2 2p^2$, i.e. 2 electrons fill the inner shell 1s, which is close to the nucleus and which is irrelevant for chemical reactions, whereas 4 electrons occupy the outer shell of 2s and 2p orbitals. Because the 2p orbitals ($2p_x$, $2p_y$, and $2p_z$) are roughly 4 eV higher than the 2s orbital, it is energetically favourable to

---

[2]However, our aim is more modest – we do not intend to "rewrite" but rather to "apply" the theory of metals to graphene.
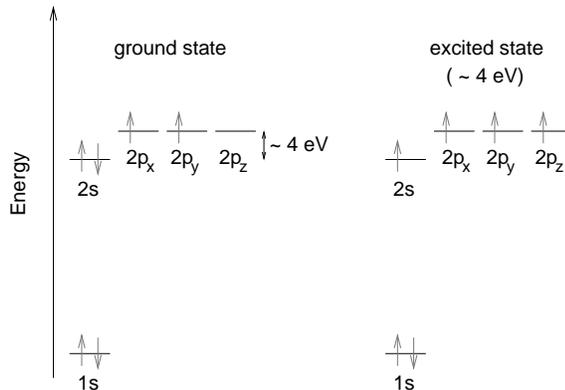
Figure 1.2: Electronic configurations for carbon in the ground state (left) and in the excited state (right).

put 2 electrons in the 2s orbital and only 2 of them in the 2p orbitals (Fig 1.2). It turns out, however, that in the presence of other atoms, such as e.g. H, O, or other C atoms, it is favourable to excite one electron from the 2s to the third 2p orbital, in order to form covalent bonds with the other atoms. The gain in energy from the covalent bond is indeed larger than the 4 eV invested in the electronic excitation.

In the excited state, we therefore have four equivalent quantum-mechanical states, $|2s\rangle$, $|2p_x\rangle$, $|2p_y\rangle$, and $|2p_z\rangle$. A quantum-mechanical superposition of the state $|2s\rangle$ with n $|2p_j\rangle$ states is called sp^n hybridisation, which play an essential role in covalent carbon bonds.

1.1.1 sp^1 hybridisation

In the sp^1 hybridisation,[3] the $|2s\rangle$ state mixes with one of the 2p orbitals. For illustration, we we have chosen the $|2p_x\rangle$ state. A state with equal weight from both original states, is obtained by the symmetric and anti-symmetric combinations

$$|sp_+\rangle = \frac{1}{\sqrt{2}}\left(|2s\rangle + |2p_x\rangle\right), \qquad |sp_-\rangle = \frac{1}{\sqrt{2}}\left(|2s\rangle - |2p_x\rangle\right).$$

The other states, $|2p_y\rangle$ and $|2p_z\rangle$, remain unaffected by this superposition. The electronic density of the hybridised orbitals has the form of a club and

---

[3]The superscript is often omitted, and one may alternatively use "sp hybridisation".

(a)

$|2s\rangle$ $|2p_x\rangle$ $|sp_+\rangle$

$|2s\rangle$ $-|2p_x\rangle$ $|sp_-\rangle$
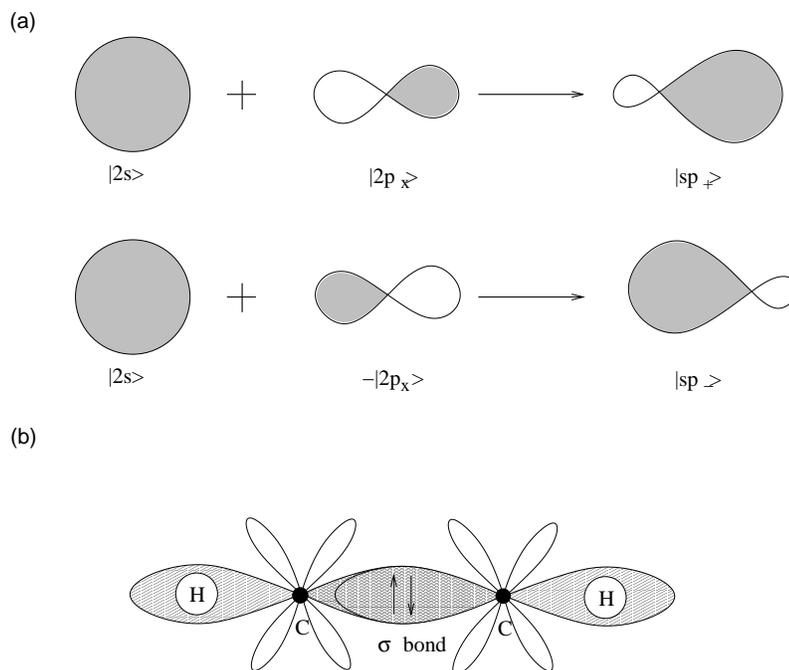
(b)

Figure 1.3: *(a)* Schematic view of the sp^1 hybridisation. The figure shows on the r.h.s. the electronic density of the $|2s\rangle$ and $|2p_x\rangle$ orbitals and on the l.h.s. that of the hybridised ones. *(b)* Acetylene molecule (H–C≡C–H). The propeller-like 2p$_y$ and 2p$_z$ orbitals of the two C atoms strengthen the covalent σ bond by forming two π bonds (not shown).
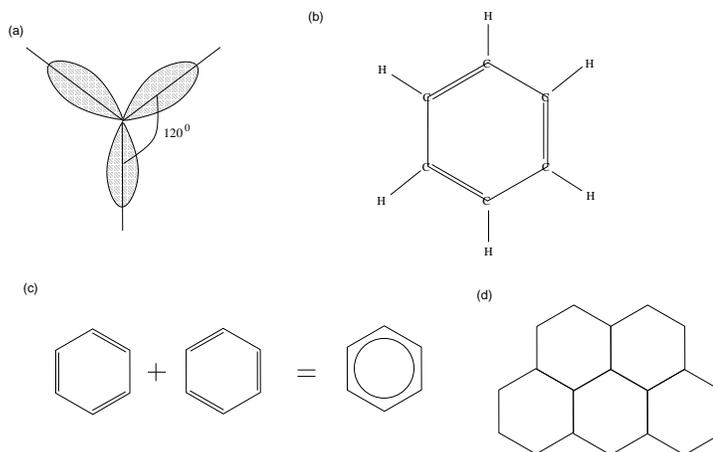
Figure 1.4: *(a)* Schematic view of the sp^2 hybridisation. The orbitals form angles of 120°. *(b)* Benzene molecule (C$_6$H$_6$). The 6 carbon atoms are situated at the corners of a hexagon and form covalent bonds with the H atoms. In addition to the 6 covalent σ bonds between the C atoms, there are three π bonds indicated by the doubled line. *(c)* The quantum-mechanical ground state of the benzene ring is a superposition of the two configurations which differ by the position of the π bonds. The π electrons are, thus, delocalised over the ring. *(d)* Graphene may be viewed as a tiling of benzene hexagons, where the H atoms are replaced by C atoms of neighbouring hexagons and where the π electrons are delocalised over the whole structure.

is elongated in the $+x$ ($-x$) direction for the $|sp_+\rangle$ ($|sp_-\rangle$) states [Fig. 1.3 (a)]. This hybridisation plays a role, e.g., in the formation of the acetylene molecule H–C≡C–H, where overlapping sp^1 orbitals of the two carbon atoms form a strong covalent σ bond [Fig. 1.3 (b)]. The remaining unhybridised 2p orbitals are furthermore involved in the formation of two additional π bonds, which are weaker than the σ bond.

1.1.2   sp^2 hybridisation – graphitic allotopes

In the case of a superposition of the 2s and two 2p orbitals, which we may choose to be the $|2p_x\rangle$ and the $|2p_y\rangle$ states, one obtains the planar sp^2 hy-

bridisation. The three quantum-mechanical states are given by

$$\begin{aligned}
|sp_1^2\rangle &= \frac{1}{\sqrt{3}}|2s\rangle - \sqrt{\frac{2}{3}}|2p_y\rangle, \\
|sp_2^2\rangle &= \frac{1}{\sqrt{3}}|2s\rangle + \sqrt{\frac{2}{3}}\left(\frac{\sqrt{3}}{2}|2p_x\rangle + \frac{1}{2}|2p_y\rangle\right), \\
|sp_3^2\rangle &= -\frac{1}{\sqrt{3}}|2s\rangle + \sqrt{\frac{2}{3}}\left(-\frac{\sqrt{3}}{2}|2p_x\rangle + \frac{1}{2}|2p_y\rangle\right).
\end{aligned} \tag{1.1}$$

These orbitals are oriented in the xy-plane and have mutual 120° angles [Fig. 1.4 (a)]. The remaining unhybridised $2p_z$ orbital is perpendicular to the plane.

A prominent chemical example for such hybridisation is the benzene molecule the chemical structure of which has been analysed by August Kekulé in 1865 [5]. The molecule consists of a hexagon with carbon atoms at the corners linked by σ bonds [Fig. 1.4 (b)]. Each carbon atom has, furthermore, a covalent bond with one of the hydrogen atoms which stick out from the hexagon in a star-like manner. In addition to the 6 σ bonds, the remaining $2p_z$ orbitals form 3 π bonds, and the resulting double bonds alternate with single σ bonds around the hexagon. Because a double bond is stronger than a single σ bond, one may expect that the hexagon is not perfect. A double bond (C=C) yields indeed a carbon-carbon distance of 0.135 nm, whereas it is 0.147 nm for a single σ bond (C–C). However, the measured carbon-carbon distance in benzene is 0.142 nm for all bonds, which is roughly the average length of a single and a double bond. This equivalence of all bonds in benzene was explained by Linus Pauling in 1931 within a quantum-mechanical treatment of the benzene ring [6]. The ground state is indeed a quantum-mechanical superposition of the two possible configurations for the double bonds, as shown schematically in Fig. 1.4 (c).

These chemical considerations indicate the way towards carbon-based condensed matter physics – any graphitic compound has indeed a sheet of graphene as its basic constituent. Such a graphene sheet may be viewed simply as a tiling of benzene hexagons, where one has replaced the hydrogen by carbon atoms to form a neighbouring carbon hexagon [Fig. 1.4 (d)]. However, graphene has remained the basic constituent of graphitic systems during a long time only on the theoretical level. From an experimental point of view, graphene is the youngest allotope and accessible to physical measurements
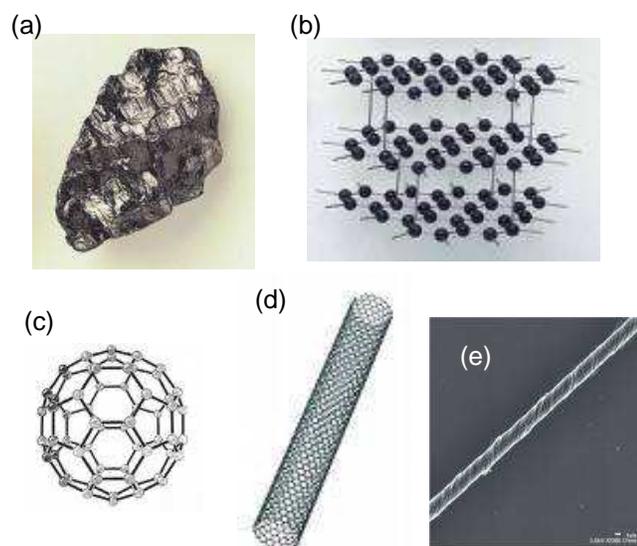
(a)   (b)

(c)   (d)

(e)

Figure 1.5: Graphitic allotopes *(a)* Piece of natural graphite. *(b)* Layered structure of graphite (stacking of graphene layers). *(c)* 0D allotope: C_{60} molecule. *(d)* 1D allotope: single-wall carbon nanotube. *(e)* Optical image of a carbon nanotube.

only since 2004.

Historically, the longest known allotope is 3D graphite [Fig. 1.5 (a)]. Graphite was discovered in a mine near Borrowdale in Cumbria, England in the 16th century, and its use for marking and graphical purposes was almost immediately noticed. Indeed, the nearby farmers used graphite blocks from the mine for marking their sheep. Due to its softness and dark color, graphite was considered during a long time as some particular type of lead. The full name "lead pencil" still witnesses this historical error.[4] That graphite was formed from carbon atoms was discovered by the Swedish-German pharmacist Carl Wilhelm Scheele in the middle of the 18th century. But it was the German chemist Abraham Gottlob Werner in 1789 who coined the material by its current name "graphite", thus emphasising its main use for graphical purposes.[5]

Graphite may be viewed as a stacking of graphene sheets [Fig. 1.5 (b)] that stick together due to the van der Waals interaction, which is much

---

[4]The graphitic core of a pencil is still called "lead", and the German name for "pencil" is "Bleistift", "Blei" being the German name for lead.

[5]The term is derived from the Greek word $\gamma\rho\alpha\varphi\epsilon\iota\nu$ ("graphein", to draw, to write).

weaker than the inplane covalent bonds. This physical property explains the graphic utility of the material: when one writes with a piece of graphite, i.e. when it is scratched over a sufficiently rough surface, such as a piece of paper, thin stacks of graphene sheets are exfoliated from bulk graphite and now stick to the surface. This is possible due to the above-mentioned weak van der Waals interaction between the graphene sheets.

The 0D graphitic allotope (fullerenes) has been discovered in 1985 by Robert Curl, Harold Kroto, and Richard Smalley [7]. Its most prominent representative is the C_{60} molecule which has the form of a football and is also called "buckyball". It consists of a graphene sheet, where some hexagons are replaced by pentagons, which cause a crumbling of the sheet and the final formation of a graphene sphere [Fig. 1.5 (c)]. Its existance had been predicted before, in 1970, by the Japanese theoretician Eiji Ozawa [8].

Carbon nanotubes, the 1D allotope, may be viewed as graphene sheets which are rolled up [Fig. 1.5 (d) and (e)], with a diameter of several nanometers. One distinguishes single-wall from multi-wall nanotubes, according to the number of rolled up graphene sheets. The discovery of carbon nanotubes is most often attributed to Sumio Iijima and his 1991 publication in *Nature* [9]. Recently, doubts about this attribution have been evoked because it seems that carbon nanotubes had a longer history in the community of material scientists [10]. Indeed, a publication by the Soviet scientists Radushkevich and Lukyanovich in 1952 contained a transmission electron microscope image showing carbon nanotubes [11]. It is, however, the merit of the 1991 Iijima paper to have attracted the interest of the condensed matter physics community to carbon nanotubes and to have initiated an intense research on this compound, also in the prospect of nanotechnological applications.

1.1.3 sp^3 hybridisation – diamonds

If one superposes the 2s and all three 2p orbitals, one otains the sp^3 hybridisation, which consists of four club-like orbitals that mark a tetrahedron. The orbitals form angles of 109.5^o degrees [Fig. 1.6 (a)]. A chemical example for this hybridisation is methane (CH_4), where the four hybridised orbitals are used to form covalent bonds with the 1s hydrogen atoms. In condensed matter physics, the $2p^3$ hybridisation is at the origin of the formation of diamonds, when liquid carbon condenses under high pressure. The diamond lattice consists of two interpenetrating face-center-cubic (fcc) lattices, with a lattice spacing of 0.357 nm, as shown in Fig. 1.6 (b).
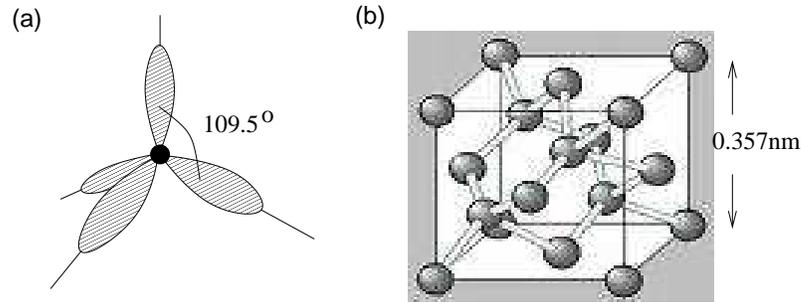
Figure 1.6: *(a)* sp^3 hybridisation with 109.5^o angle between the four orbitals. *(b)* Crystal structure of diamond (two interpenetrating fcc lattices).

Although they consist of the same atomic ingredient, namely carbon, the 3D graphite and diamond crystals are physically extremely different. Graphite, as described above, is a very soft material due to its layered structure, whereas diamond is one of the hardest natural materials because all bonds are covalent σ bonds. The fact that all 4 valence electrons in the outer atomic shell are used in the formation of the σ bondes is also the reason for diamond being an insulator with a large band gap of 5.47 eV. In contrast to insulating diamond, the electrons in the weaker π bonds in graphite are delocalised and, thus, yield good electronic conduction properties.

1.2   Crystal Structure of Graphene and Graphite

1.2.1   Graphene's honeycomb lattice

As already mentioned in the last section, the carbon atoms in graphene condense in a honeycomb lattice due to their sp^2 hybridisation. The honeycomb lattice is not a Bravais lattice because two neighbouring sites are not equivalent. Fig. 1.7 (a) illustrates indeed that a site on the A sublattice has nearest neighbours (*nn*) in the directions north-east, north-west, and south, whereas a site on the B sublattice has *nn*s in the directions north, south-west, and south-east. Both A and B sublattices, however, are triangular[6] Bravais lattices, and one may view the honeycomb lattice as a *triangular Bravais lattice with a two-atom basis* (A and B). The distance between *nn* carbon atoms is 0.142 nm, which is the average of the single (C–C) and double (C=C)

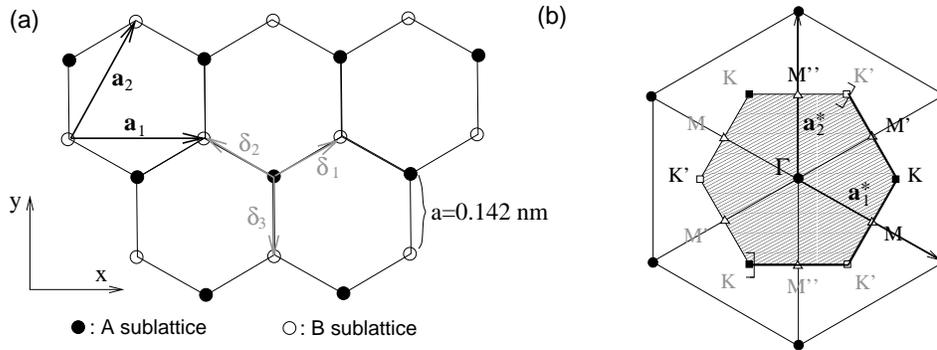[6]The triangular lattice is sometimes also called *hexagonal* lattice.

Figure 1.7: *(a)* Honeycomb lattice. The vectors $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, and $\boldsymbol{\delta}_3$ connect *nn* carbon atoms, separated by a distance $a = 0.142$ nm. The vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are basis vectors of the triangular Bravais lattice. *(b)* Reciprocal lattice of the triangular lattice. Its primitive lattice vectors are $\mathbf{a}_1^*$ and $\mathbf{a}_2^*$. The shaded region represents the first Brillouin zone (BZ), with its centre Γ and the two inequivalent corners K (black squares) and K' (white squares). The thick part of the border of the first BZ represents those points which are counted in the definition such that no points are doubly counted. The first BZ, defined in a strict manner, is, thus, the shaded region plus the thick part of the border. For completeness, we have also shown the three inequivalent cristallographic points M, M', and M'' (white triangles).

covalent σ bonds, as in the case of benzene.

The three vectors which connect a site on the A sublattice with a *nn* on the B sublattice are given by

$$\boldsymbol{\delta}_1 = \frac{a}{2}\left(\sqrt{3}\mathbf{e}_x + \mathbf{e}_y\right), \qquad \boldsymbol{\delta}_2 = \frac{a}{2}\left(-\sqrt{3}\mathbf{e}_x + \mathbf{e}_y\right), \qquad \boldsymbol{\delta}_3 = -a\mathbf{e}_y, \quad (1.2)$$

and the triangular Bravais lattice is spanned by the basis vectors

$$\mathbf{a}_1 = \sqrt{3}a\mathbf{e}_x \qquad \text{and} \qquad \mathbf{a}_2 = \frac{\sqrt{3}a}{2}\left(\mathbf{e}_x + \sqrt{3}\mathbf{e}_y\right). \qquad (1.3)$$

The modulus of the basis vectors yields the lattice spacing, $\tilde{a} = \sqrt{3}a = 0.24$ nm, and the area of the unit cell is $A_{uc} = \sqrt{3}\tilde{a}^2/2 = 0.051$ nm^2. The density of carbon atoms is, therefore, $n_C = 2/A_{uc} = 39$ nm$^{-2} = 3.9 \times 10^{15}$ cm^{-2}. Because there is one π electron per carbon atom that is not involved in a covalent σ bond, there are as many valence electrons than carbon atoms, and their density is, thus, $n_\pi = n_C = 3.9 \times 10^{15}$ cm^{-2}. As is discussed in the following chapter, this density is not equal to the carrier density in graphene, which one measures in electrical transport measurements.

The reciprocal lattice, which is defined with respect to the triangular Bravais lattice, is depicted in Fig. 1.7 (b). It spanned by the vectors

$$\mathbf{a}_1^* = \frac{2\pi}{\sqrt{3}a} \left(\mathbf{e}_x - \frac{\mathbf{e}_y}{\sqrt{3}} \right) \qquad \text{and} \qquad \mathbf{a}_2^* = \frac{4\pi}{3a} \mathbf{e}_y. \qquad (1.4)$$

Physically, all sites of the reciprocal lattice represent equivalent wave vectors. Any wave – be it a vibrational lattice excitation or a quantum-mechanical electronic wave packet – propagating on the lattice with a wave vector differing by a reciprocal lattice vector has indeed the same phase up to a multiple of 2π, due to the relation

$$\mathbf{a}_i \cdot \mathbf{a}_j^* = 2\pi \delta_{ij} \qquad (1.5)$$

(for $i, j = 1, 2$) between direct and reciprocal lattice vectors. The first Brillouin zone [BZ, shaded region and thick part of the border of the hexagon in Fig. 1.7 (b)] represents a set of inequivalent points in the reciprocal space, i.e. of points which may not be connected to one another by a reciprocal lattice vector, or else of physically distinguishable lattice excitations. The long wavelength excitations are situated in the vicinity of the Γ point, in centre of the first BZ. Furthermore, one distinguishes the six corners of the first BZ, which consist of the inequivalent points K and K' represented by the vectors

$$\pm\mathbf{K} = \pm\frac{4\pi}{3\sqrt{3}a} \mathbf{e}_x. \qquad (1.6)$$

The four remaining corners [shown in gray in Fig. 1.7 (b)] may indeed be connected to one of these points via a translation by a reciprocal lattice vector. These cristallographic points play an essential role in the electronic properties of graphene because their low-energy excitations are centered around the two points K and K', as is discussed in detail in the following chapter. We emphasise, because of some confusion in the litterature on this point, that the inequivalence of the two BZ corners, K and K', has nothing to do with the presence of two sublattices, A and B, in the honeycomb lattice. The form of the BZ is an intrinsic property of the *Bravais* lattice, independent of the possible presence of more than one atom in the unit cell. For completeness, we have also shown, in Fig. 1.7 (b), the three crystallographically inequivalent M points in the middle of the BZ edges.
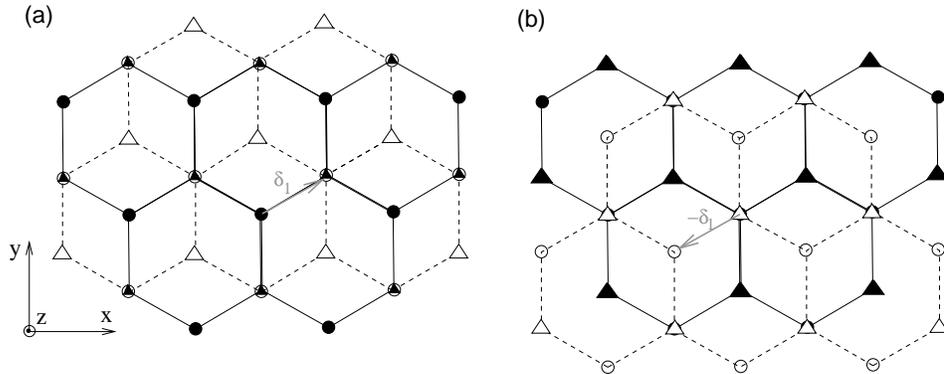
Figure 1.8: Two possibilities for stacking two graphene layers. The symbols correspond to the following lattice sites: black circles are A sites in the lower layer, black triangles B sites in the lower layer, white circles A' sites in the upper layer, and white triangles B' sites in the upper layer. *(a)* The upper layer is translated by $\boldsymbol{\delta}_1$ with respect to the lower one; the A' sites are on top of the B sites. *(b)* The upper layer is translated by $-\boldsymbol{\delta}_1$ with respect to the lower one; the B' sites are on top of the A sites.

1.2.2 Graphene stacking – the different forms of graphite

Graphite consists of stacked graphene layers. One distinguishes crystal (ordered) graphite, with two different basic stacking orders, from turbostratic graphite with a certain amount of disorder in the stacking.

To illustrate the ordering of graphene layers in crystal graphite, we first consider only two graphene layers (bilayer graphene). The distance between layers is roughly $d = 2.4a = 0.34$ nm, and the stacking is such that there are atoms in the upper layer placed at the hexagon centres of the lower layer. The layers are, thus, translated with respect to each other, and one may distinguish two different patterns, as shown in Fig. 1.8. The displacement is given by either $\boldsymbol{\delta}_i$ or $-\boldsymbol{\delta}_i$, where one may choose any of the *nn* vectors (1.2) with $i = 1, 2$, or 3. For the choice $\boldsymbol{\delta}_i$, one obtains a configuration where the A' atoms of the upper layer are on top of the B atoms of the lower one [Fig. 1.8 (a)], and for a translation of $-\boldsymbol{\delta}_i$, the B' atoms in the upper layer are on top of A atoms in the lower one [Fig. 1.8 (b)]. Notice that the two configurations, both of which contain now four atoms (A, B, A', and B') per unit cell, are equivalent if one has a reflexion symmetry in the perpendicular axis, $z \rightarrow -z$.

Although physically not relevant in the bilayer case, this difference in the stackingturns out to be important when considering crystal graphite with an

infinity of stacked layers. Let us consider the situation where the second layer is translated with respect to the first by $\boldsymbol{\delta}_i$. The third layer may be chosen to be translated with respect to the second either by $-\boldsymbol{\delta}_i$, in which case its sites coincide with those of the first one, or again by $\boldsymbol{\delta}_i$, where the atomic configuration of the first layer is not recovered. In the first case, one therefore obtains an ABA stacking, whereas the stacking is ABC in the second case. In general, one distingushes two types of ordering:

1. All layers are translated with respect to their lower neighbour by $\boldsymbol{\delta}_i$. One obtains a rhombohedral stacking, which is also called ABC stacking because one needs three layers to recover the atomic configuration of the original layer. There are, thus, 6 atoms per unit cell which has an extension of $3d$ in the z-direction (β-graphite).

2. If the sign of the translation alternates, i.e. $\boldsymbol{\delta}_i, -\boldsymbol{\delta}_i, \boldsymbol{\delta}_i, -\boldsymbol{\delta}_i, ...$, when stacking the graphene layers, one obtains a hexagonal (or else AB or Bernal) stacking. Here, one has 4 atoms per unit cell which has a height of $2d$ in the z-direction (α-graphite).

In principle, it is possible to have some randomness in the stacking, i.e. ABC parts may randomly substitute AB parts when considering α-graphite. However, a crystalline AB stacking occurs most often in nature, with about 30% of ABC-stacked graphite.

In the case of turbostratic graphite, one may distinguish translational disorder from rotational disorder in the stacking. Generally, the graphene layers are much less bound together in turbostratic than in crystalline graphite, such that turbostratic graphite is a better lubrificant.

As an example of rotational disorder, we consider two graphene layers which are rotated by an angle

$$\phi = \frac{\mathbf{a}_1 \cdot \mathbf{a}_1'}{|\mathbf{a}_1||\mathbf{a}_1'|} = \mathbf{e}_x \cdot \mathbf{e}_x'$$

with respect to each other, where $\mathbf{a}_1'$ denotes the lattice vector in the upper graphene layer corresponding to $\mathbf{a}_1$ in the lower one. If the angle fulfils certain comensurability conditions, one obtains a so-called Moiré pattern, with a larger unit cell as depictecd in Fig. 1.9. The Moiré pattern reproduces a (larger) honeycomb lattice.
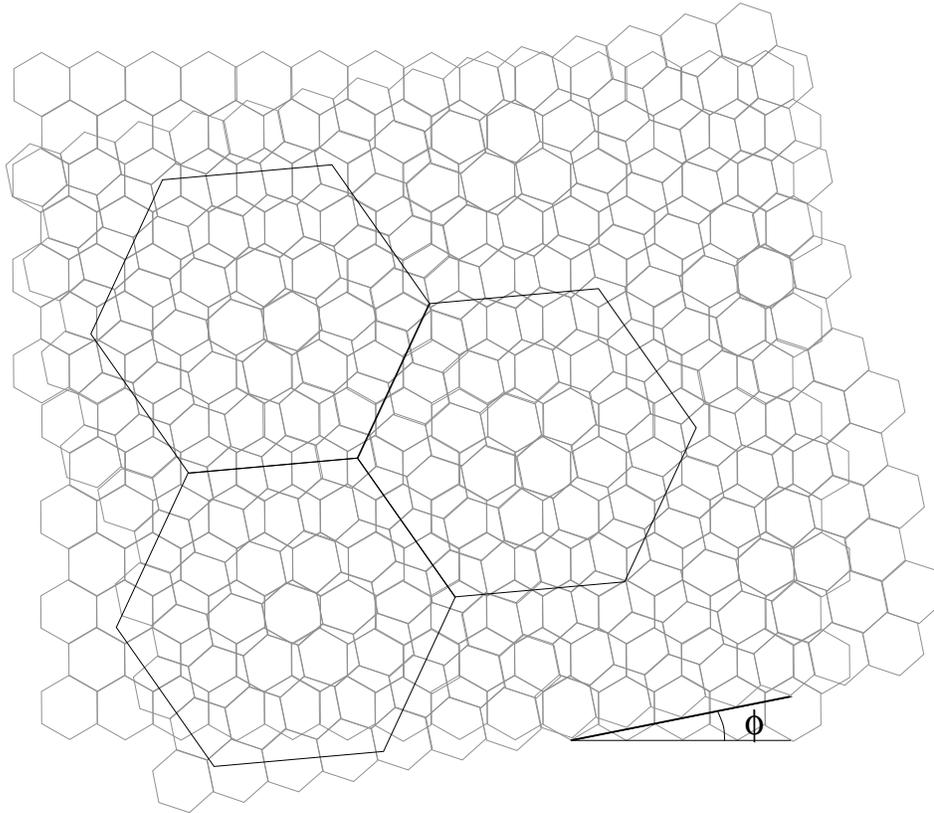
Figure 1.9: Moiré pattern obtained by stacking two honeycomb lattices (gray) with a relative (chiral) angle ϕ. One obtains a hexagonal superstructure indicated by the black hexagons.
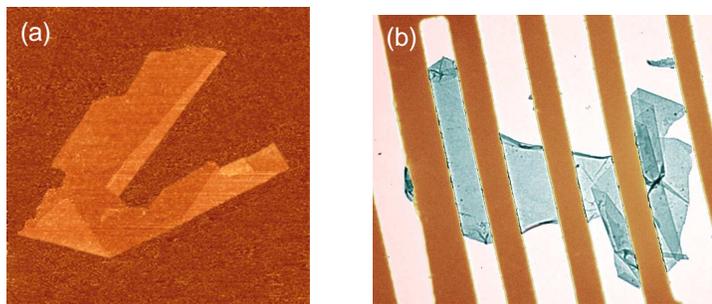
Figure 1.10: *(a)* AFM image of a graphene flake on a SiO_2 substrate. *(b)* Transmission-electron-microscopic image of free-hanging graphene. Both pictures have been taken by the Manchester group.

1.3 Fabrication of Graphene

We briefly review the two techniques used in fabricating graphene, mechanical exfoliation of graphene from bulk graphite and graphitisation of epitaxially grown SiC crystals. The resulting graphene probes have similar, but not always equal, physical properties. It is therefore important to describe both systems separately and to clearly indicate on what type of graphene, exfoliated or epitaxial, a particular measurement is performed. A systematic comparison of both graphene types is still an open research field.

1.3.1 Exfoliated graphene

The mechanical exfoliation technique, elaborated mainly by the Manchester group, consistsy of peeling a layered material [13]. In the case of graphite, it uses its primary graphical capacity, which we have already alluded to above: if one scratches a piece of graphite on a substrate, thin graphite stacks are exfoliated from the bulk and left behind on the substrate. Most of these thin stacks comprise several (tens or hundreds) of graphene sheets, but a few of them only consist of a single graphene layer. The fabrication of graphene is, thus, extremely simple, and we produce single layers of graphene whenever we write with a pencil of sufficiently high graphitic quality.[7]

In practice, one does not use a pencil to fabricate graphene layers, but one prepares very thin samples by peeling a small graphite crystallite in a piece

---

[7]Not all pencils are of the required quality. The graphite lead of most pencils contains normally a certain amount of clay impurities.

of folded scotchtape. Each time one peels away the tape, the graphite flake is cleaved into two parts and thus becomes thinner. After several cycles, the scotchtape with the graphene sheets stuck to it is glued to the SiO_2 substrate, prepared by a mix of hydrochloric acid and hydrogen peroxide to accept better the graphene sheets from the scothtape. When the tape is carefully peeled away, the graphene sheets remain glued to the substrate.

The main problem consists of the identification of the few single graphene layers, which are randomly distributed on the substrate. A definitve identification may be achieved by atomic-force microscopy (AFM), the scanning capacity of which is unfortunately limited to a very small surface during a reasonable time. One therefore needs to have a hint of where to search for mono-layer graphene. This hint comes from a first (optical) glance at the substrate. The 300 nm thick SiO_2 substrate, which was originally used by the Manchester group and which is now most common, turns out to yield an optimal contrast such that one may, by optical meanse, identify mono-layer graphene sheets with a high probability. AFM imaging is then used to confirm this first glance. An example of such AFM image is shown in Fig. 1.10 (a). Notice that it has been shown that one may also depose an exfoliated graphene sheet on a scaffold such that it is free-hanging over a rather large surface of several μm^2 [Fig. 1.10 (b)].

The exfoliation technique is not limited to graphite, but may also be used in the fabrication of ultra-thin samples from other layered crystals. The Manchester group has indeed shown that one may, by this technique, fabricate e.g. single layer BN, $NbSe_2$, $Bi_2Sr_2CaCu_2O_x$, and MoS_2 crystallites of a typical size of several μm^2 [13]. The exfoliation technique is, therefore, extremely promising in the study of truly 2D crystals.

One may, furthermore, control the carrier density in the metallic 2D crystals, such as graphene, by the electric field effect; the 300 nm thick insulating SiO_2 layer is indeed on top of a positively doped metallic Si substrate, which serves as a backgate. The combined system graphene-SiO_2-backgate may, thus, be viewed as a capacitor (see Fig. 1.11) the capacity of which is

$$C = \frac{Q}{V_g} = \frac{\epsilon_0 \epsilon A}{d}, \tag{1.7}$$

where $Q = en_{2D}A$ is the capacitor charge, in terms of the total surface A, V_g is the gate voltage, $d = 300$ nm is the thickness of the SiO_2 layer with the dielectric constant $\epsilon = 3.7$. The field-effect induced 2D carrier density is
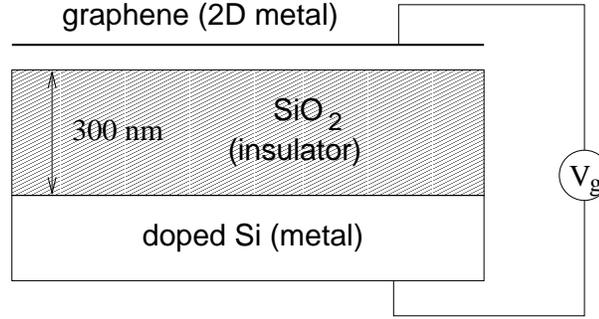
Figure 1.11: Schematic view of graphene on a SiO_2 substrate with a doped Si (metallic) backgate. The system graphene-SiO_2-backgate may be viewed as a capacitor the charge density of which is controled by a gate voltage V_g.

thus given by

$$n_{2D} = \alpha V_g \qquad \text{with} \qquad \alpha \equiv \frac{\epsilon_0 \epsilon}{ed} \simeq 7.2 \times 10^{10} \frac{\text{cm}^{-2}}{\text{V}}. \qquad (1.8)$$

The gate voltage may vary roughly between -100 and 100 V, such that one may induce maximal carrier densities on the order of 10^{12} cm^{-2}, on top of the intrinsic carrier density which turns out to be zero in graphene, as will be discussed in the next chapter. At gate voltages above ± 100 V, the capacitor breaks down (electrical breakdown).

1.3.2 Epitaxial graphene

An alternative method to fabricate graphene has been developed by the Atlanta group, led by Walt de Heer and Claire Berger [14]. It consists of exposing an epitaxially grown hexagonal (4H or 6H-) SiC crystal[8] to temperatures of about 1300° C in order to evaporate the less tightly bound Si atoms from the surface. The remaining carbon atoms on the surface form a graphitic layer (graphitsation). The physical properties of this graphitic layer depend on the chosen SiC surface. In the case of the Si-terminated (0001) surface, the graphitisation process is slow, and one may thus control the number of formed graphene layers (usually one or two). The resulting electron mobility,

---

[8]SiC exists also in several other crystal structures, such as the zinc blende structure (3C-SiC), i.e. a diamond structure with different atom types on the two distinct fcc sublattices.
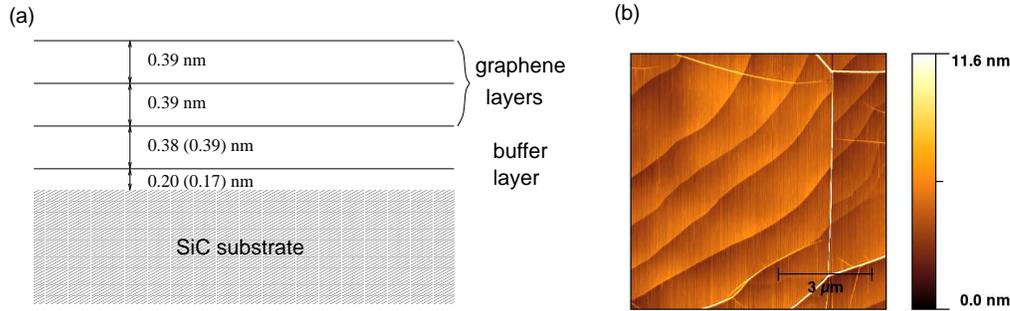
Figure 1.12: Epitaxial Graphene. *(a)* Schematic view on epitaxial graphene. On top of the SiC substrate, a graphitic layer. It consists of several one-atom thick caron layers. The first one is a buffer layer, which is tightly bound to the substrate, at a distance of 0.2 (0.17) nm for a Si- (C-)terminated surface. The graphene layers are formed on top of this buffer layer at a distance of 0.38 (0.39) nm, and they are equally spaced by 0.39 nm. *(b)* AFM image of epitaxial graphene on C-terminated SiC substrate. The steps those of the SiC substrate. The 5-10 graphene layers lie on the substrate similar to a carpet which has folds visible as white lines on the image.

however, turns out to be rather low, such that the Si-terminated surface is less chosen for the fabrication of samples used in transport measurements. For a C-terminated $(000\bar{1})$ surface, the graphitisation process is very fast, and a large number of graphene layers are formed (up to 100). In contrast to epitaxial graphene on the Si-terminated surface, the electron mobility is, here, rather high.

In contrast to exfoliated graphene, the SiC substrate must be considered as an integral part of the whole system of epitaxial graphene. It is indeed the mother compound, and the first graphitic layer formed during the graphitisation process is tightly bound to the SiC substrate (Fig. 1.12). The distance between this layer and the substrate has been estimated numerically to be 0.20 nm for a Si-terminated surface and 0.17 nm for a C-terminated surface [15]. These distances are much smaller than the distance between graphene sheets in crystalline graphite ($2.4a = 0.34$ nm). 0.38 (0.39) nm above this buffer layer for Si- (C-)terminated SiC, several graphene layers are formed, which are separated by a distance of 0.39 nm. These characteristic distances are roughly 20% larger than the layer distances in crystalline graphite, and one may therefore expect them to be less tightly bound. The above-mentioned numerical values for the distances have been confirmed by X-ray measurements [15].

Furthermore, X-ray diffraction measurements have revealed a certain amount of rotational disorder in the stacking of the graphene layers [16]. This, together with the rather large spacing between the graphene sheets on top of the buffer layer, corroborates the view that the graphitic layer on the SiC substrate consists indeed of almost independent graphene layers rather than of a thin graphite flake. However, the charge is not homogeneously distributed between the different graphene layers – it turns out that the graphene layer,[9] which is closest to the substrate, is electron-doped due to a charge transfer from the bulk SiC, whereas the subsequent layers are hardly charged. As will be discussed in the following chapters, this inhomogeneous charge distribution in epitaxial graphene may be at the origin of different results in transport measurements when compared with exfoliated graphene and theoretical predictions.

---

[9]We refer to the layers on top of the buffer layer when using the term "graphene". The buffer layer may not be counted due to its tight bonding to the SiC substrate.

Chapter 2

Electronic Band Structure of Graphene

As we have discussed in the introduction, three electrons per carbon atom in graphene are involved in the formation of strong covalent σ bonds, and one electron per atom yields the π bonds. The π electrons happen to be those responsible for the electronic properties at low energies, whereas the σ electrons form energy bands far away from the Fermi energy. This chapter is, thus, devoted to a discussion of the energy bands of π electrons within the tight-binding approximation, which was originally calculated for the honeycomb lattice by P. R. Wallace in 1947 [17]. We do not consider the σ electrons, here, and refer the interested reader to the book by Saito, Dresselhaus, and Dresselhaus for a tight-binding calculation of the energy bands formed by the σ electrons [12], which are far away from the Fermi level.

The chapter consists of two sections. The first one is devoted to the calculation of the π energy bands in graphene, where one is confronted with the problem of two atoms per unit cell. After a brief discussion of Bloch's theorem, and a formal solution of the tight-binding model, we calculate the energy dispersion of π electrons in graphene, taking into account nearest-neighbour (*nn*) and next-nearest-neighbour (*nnn*) hopping and *nn* overlap corrections. The second section consists of the continuum limit, which describes the low-energy properties of electrons in graphene.

2.1 Tight-Binding Model for Electrons on the Honeycomb Lattice

The general idea of the tight-binding model is to write down a trial wave-function constructed from the orbital wavefunctions, $\phi^{(a)}(\mathbf{r} - \mathbf{R}_j)$, of the atoms forming a particular lattice described by the (Bravais) lattice vectors $\mathbf{R}_j = m_j\mathbf{a}_1 + n_j\mathbf{a}_2$, where m_j and n_j are integers.[1] In addition, the trial wave-function must reflect the symmetry of the underlying lattice, i.e. it must be invariant under a translation by any arbitrary lattice vector $\mathbf{R}_i$. We consider, for simplicity in a first step, the case of a Bravais lattice with one atom per unit cell and one electron per atom. The Hamiltonian for an arbitrary electron, labelled by the integer l, is given by

$$H_l = -\frac{\hbar^2}{2m}\Delta_l + \sum_{j=1}^{N} V(\mathbf{r}_l - \mathbf{R}_j), \tag{2.1}$$

where $\Delta_l = \nabla_l^2$ is the 2D Laplacian operator, in terms of the 2D gradient $\nabla_l = \partial/\partial x_l + \partial/\partial y_l$ with respect to the electron's position $\mathbf{r}_l = (x_l, y_l)$, and m is the electron mass. Each ion on site $\mathbf{R}_j$ yields an electrostatic potential felt by the electron, and its overall potential energy $\sum_j^N V(\mathbf{r}_l - \mathbf{R}_j)$, where N is the number of lattice sites, is, therefore, a periodic function with respect to an arbitrary translation by a lattice vector $\mathbf{R}_i$ in the thermodynamic limit $N \to \infty$. The total Hamiltonian is the sum over all electrons,

$$H = \sum_{l}^{N} H_l, \tag{2.2}$$

if we suppose one electron per lattice site, as mentioned above.

The tight-binding approach is based on the assumption that the electron l is originally bound to a particular ion at the lattice site $\mathbf{R}_l$, i.e. it is described to great accuracy by a bound state of the (atomic) Hamiltonian

$$H_l^a = -\frac{\hbar^2}{2m}\Delta_l + V(\mathbf{r}_l - \mathbf{R}_l),$$

whereas the contributions to the potential energy $\Delta V = \sum_{j \neq l}^{N} V(\mathbf{r}_l - \mathbf{R}_j)$ from the other ions at the sites $\mathbf{R}_j$, $j \neq l$, may be treated perturbatively. The

---

[1]Because we are interested in 2D lattices, we limit the discussion to two dimensions, the generalisation to arbitrary dimensions being straight-forward.

bound state of H_l^a is described by the above-mentioned atomic wavefunction $\phi^{(a)}(\mathbf{r} - \mathbf{R}_j)$.

2.1.1 Bloch's theorem

Another ingredient, apart from the atomic wavefunction, of the trial wavefunction is a symmetry consideration – the trial wavefunction must respect the discrete translation symmetry of the lattice. This is the essence of Bloch's theorem. In quantum mechanics, a translation by a lattice vector $\mathbf{R}_i$ may be described by the operator

$$\mathcal{T}_{\mathbf{R}_i} = e^{\frac{i}{\hbar}\hat{\mathbf{p}}\cdot\mathbf{R}_i}, \tag{2.3}$$

in terms of the momentum operator $\hat{\mathbf{p}}$, which needs to be appropriately defined for the crystal and which we may call quasi-momentum operator, as will become explicit below. The symmetry operator, because it describes a symmetry operation under which the physical problem is left invariant, commutes with the full Hamiltonian (2.2), $[\mathcal{T}_{\mathbf{R}_i}, H] = 0$. The eigenstates of H are, therefore, necessarily also eigenstates of $\mathcal{T}_{\mathbf{R}_i}$, for any lattice vector $\mathbf{R}_i$, and the momentum $\mathbf{p}$, which is the eigenvalue of the momentum operator $\hat{\mathbf{p}}$, is a good quantum number. Because of the relation (1.5) between the basis vectors of the direct and the reciprocal lattices, this momentum is only defined modulo a reciprocal lattice vector $\mathbf{G}_j = m_j^* \mathbf{a}_1^* + n_j^* \mathbf{a}_2^*$, where m_j^* and n_j^* are arbitrary integers. Indeed, if we had chosen the momentum operator $\hat{\mathbf{p}}' = \hat{\mathbf{p}} + \hbar \mathbf{G}_j$ instead of $\hat{\mathbf{p}}$ in the definition (2.3) of the discrete translation operator, we would have simply multiplied it with a factor of $\exp(i\mathbf{G}_j \cdot \mathbf{R}_i) = \exp(i2\pi n) = 1$ because of the integer value $n = m_j^* m_i + n_j^* n_i$. We, thus, need to identify, as pointed out in the last chapter, identify all momenta which differ by a reciprocal lattice vector, and it is more convenient to speak of a quasi-momentum $\mathbf{p} = \hbar \mathbf{k}$, which is restricted to the first BZ.

The trial wavefunction, constructed from the atomic orbital wavefunctions $\phi^{(a)}(\mathbf{r} - \mathbf{R}_j)$,

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}_j} e^{i\mathbf{k}\cdot\mathbf{R}_j} \phi^{(a)}(\mathbf{r} - \mathbf{R}_j) \tag{2.4}$$

fulfils the above-mentioned requirements, i.e. it is an eigenstate of the translation operator (2.3).[2] That the wavefunction (2.4) is indeed an eigenstate

---

[2]The sum $\sum_{\mathbf{R}_j}$ is a short notation for $\sum_{m_j,n_j}$ and runs over all 2D lattice vectors $\mathbf{R}_j = m_j \mathbf{a}_1 + n_j \mathbf{a}_2$, in the thermodynamic limit of an infinite lattice.

of $\mathcal{T}_{\mathbf{R}_i}$ may be seen from

$$
\begin{aligned}
\mathcal{T}_{\mathbf{R}_i}\psi_{\mathbf{k}}(\mathbf{r}) &= \psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}_i) \\
&= \sum_{\mathbf{R}_j} e^{i\mathbf{k}\cdot\mathbf{R}_j}\phi^{(a)}\left[\mathbf{r} - (\mathbf{R}_j - \mathbf{R}_i)\right] \\
&= e^{i\mathbf{k}\cdot\mathbf{R}_i}\sum_{\mathbf{R}_m} e^{i\mathbf{k}\cdot\mathbf{R}_m}\phi^{(a)}(\mathbf{r} - \mathbf{R}_m) = e^{i\mathbf{k}\cdot\mathbf{R}_i}\psi_{\mathbf{k}}(\mathbf{r}),
\end{aligned}
$$

where the first line reflects the translation by the lattice vector $\mathbf{R}_i$ in the argument of the wavefunction, and we have resummed the lattice vectors in the last line with the redefinition $\mathbf{R}_m = \mathbf{R}_j - \mathbf{R}_i$.

2.1.2 Lattice with several atoms per unit cell

If we have several atoms per unit cell, such as in the case of the honeycomb lattice, the reasoning described above must be modified. Notice first that a translation by any vector $\boldsymbol{\delta}_j$ that relates a site on one sublattice to that on a second sublattice is not a symmetry operation, i.e. $[\mathcal{T}_{\delta_j}, H] \neq 0$, if we define $\mathcal{T}_{\delta_j} \equiv \exp(i\hat{\mathbf{p}} \cdot \delta_j/\hbar)$ in the same manner as the translation operator (2.3). One must, therefore, treat the different sublattices apart.

In the case of two atoms per unit cell, we may write down the trial wavefunction as

$$
\psi_{\mathbf{k}}(\mathbf{r}) = a_{\mathbf{k}}\psi_{\mathbf{k}}^{(A)}(\mathbf{r}) + b_{\mathbf{k}}\psi_{\mathbf{k}}^{(B)}(\mathbf{r}), \tag{2.5}
$$

where $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ are complex functions of the quasi-momentum $\mathbf{k}$. Both $\psi_{\mathbf{k}}^{(A)}(\mathbf{r})$ and $\psi_{\mathbf{k}}^{(B)}(\mathbf{r})$ are Bloch functions with

$$
\psi_{\mathbf{k}}^{(j)}(\mathbf{r}) = \sum_{\mathbf{R}_l} e^{i\mathbf{k}\cdot\mathbf{R}_l}\phi^{(j)}(\mathbf{r} + \boldsymbol{\delta}_j - \mathbf{R}_l), \tag{2.6}
$$

where $j = A/B$ labels the atoms on the two sublattices A and B, and $\boldsymbol{\delta}_j$ is the vector which connects the sites of the underlying Bravais lattice with the site of the j atom within the unit cell. Typically one chooses the sites of one of the sublattices, e.g. the A sublattice, to coincide with the sites of the Bravais lattice. Notice furthermore that there is some arbitrariness in the choice of the phase in Eq. (2.6) – instead of choosing $\exp(i\mathbf{k} \cdot \mathbf{R}_l)$, one may also have chosen $\exp[i\mathbf{k} \cdot (\mathbf{R}_l - \boldsymbol{\delta}_j)$, as for the arguments of the atomic wavefunctions. The choice, however, does not affect the physical properties

of the system because it simply leads to a redefinition of the weights $a_\mathbf{k}$ and $b_\mathbf{k}$ which aquire a different relative phase [18].

With the help of these wavefunctions, we may now search the solutions of the Schrödinger equation

$$H\psi_\mathbf{k} = \epsilon_\mathbf{k}\psi_\mathbf{k}.$$

Here, we have chosen an arbitrary representation, which is not necessarily that in real space.[3] Multiplication of the Schrödinger equation by $\psi_\mathbf{k}^*$ from the left yields the equation $\psi_\mathbf{k}^* H\psi_\mathbf{k} = \epsilon_\mathbf{k}\psi_\mathbf{k}^*\psi_\mathbf{k}$, which may be rewritten in matrix form with the help of Eqs. (2.5) and (2.6)

$$(a_\mathbf{k}^*, b_\mathbf{k}^*)\,\mathcal{H}_\mathbf{k}\begin{pmatrix} a_\mathbf{k} \\ b_\mathbf{k} \end{pmatrix} = \epsilon_\mathbf{k}\,(a_\mathbf{k}^*, b_\mathbf{k}^*)\,\mathcal{S}_\mathbf{k}\begin{pmatrix} a_\mathbf{k} \\ b_\mathbf{k} \end{pmatrix}. \tag{2.7}$$

Here, the Hamiltonian matrix is defined as

$$\mathcal{H}_\mathbf{k} \equiv \begin{pmatrix} \psi_\mathbf{k}^{(A)*} H\psi_\mathbf{k}^{(A)} & \psi_\mathbf{k}^{(A)*} H\psi_\mathbf{k}^{(B)} \\ \psi_\mathbf{k}^{(B)*} H\psi_\mathbf{k}^{(A)} & \psi_\mathbf{k}^{(B)*} H\psi_\mathbf{k}^{(B)} \end{pmatrix} = \mathcal{H}_\mathbf{k}^\dagger, \tag{2.8}$$

and the overlap matrix

$$\mathcal{S}_\mathbf{k} \equiv \begin{pmatrix} \psi_\mathbf{k}^{(A)*}\psi_\mathbf{k}^{(A)} & \psi_\mathbf{k}^{(A)*}\psi_\mathbf{k}^{(B)} \\ \psi_\mathbf{k}^{(B)*}\psi_\mathbf{k}^{(A)} & \psi_\mathbf{k}^{(B)*}\psi_\mathbf{k}^{(B)} \end{pmatrix} = \mathcal{S}_\mathbf{k}^\dagger \tag{2.9}$$

accounts for the non-orthogonality of the trial wavefunctions. The eigenvalues $\epsilon_\mathbf{k}$ of the Schrödinger equation are the energy dispersions or energy bands, and they may be obtained from the secular equation

$$\det\left[\mathcal{H}_\mathbf{k} - \epsilon_\mathbf{k}^\lambda \mathcal{S}_\mathbf{k}\right] = 0, \tag{2.10}$$

which needs to be satisfied for a non-zero solution of the wavefunctions, i.e. for $a_\mathbf{k} \neq 0$ and $b_\mathbf{k} \neq 0$. The label λ denotes the energy bands, and it is clear that there are as many energy bands as solutions of the secular equation (2.10), i.e. two bands for the case of two atoms per unit cell. Notice that the generalisation to n atoms per unit cell is straight-forwards – in this case, the wavefunction (2.5) needs to be written as

$$\psi_\mathbf{k} = \sum_{j=1}^{n} a_\mathbf{k}^{(j)}\psi_\mathbf{k}^{(j)},$$

---

[3]The wavefunction $\psi_\mathbf{k}(\mathbf{r})$ is, thus, the real space representation of the Hilbert vector $\psi_\mathbf{k}$.

where the superscript j denotes the different atoms per unit cell. The secular equation (2.10) remains valid, in terms of the Hermitian $n \times n$ matrices

$$\mathcal{H}_{\mathbf{k}}^{ij} \equiv \psi_{\mathbf{k}}^{(i)*} H \psi_{\mathbf{k}}^{(j)} \qquad \text{and} \qquad \mathcal{S}_{\mathbf{k}}^{ij} \equiv \psi_{\mathbf{k}}^{(i)*} \psi_{\mathbf{k}}^{(j)}, \qquad (2.11)$$

and it is now an equation of degree n. This means that there are n energy bands, i.e. as many energy bands as atoms per unit cell.

In a first step, one often neglects the overlap corrections, i.e. one assumes a quasi-orthogonality of the wavefunctions, $\mathcal{S}_{\mathbf{k}}^{ij} = N\delta_{ij}$. It turns out, however, to keep track of these overlap corrections in the case of graphene. As is discussed in the following section, they yield a contribution which is on the same order of magnitude as the *nnn* hopping corrections.

Formal solution

Before turning to the specific case of graphene and its energy bands, we solve formally the secular equation for an arbitrary lattice with several atoms per unit cell. The Hamiltonian matrix (2.11) may be written, with the help of Eq. (2.6), as

$$\begin{aligned}
\mathcal{H}_{\mathbf{k}}^{ij} &= \sum_{\mathbf{R}_l, \mathbf{R}_m} e^{i\mathbf{k}\cdot(\mathbf{R}_l - \mathbf{R}_m)} \int d^2r \, \phi^{(i)*}(\mathbf{r} + \boldsymbol{\delta}_i - \mathbf{R}_k) H \phi^{(j)}(\mathbf{r} + \boldsymbol{\delta}_j - \mathbf{R}_m) \\
&= N \sum_{\mathbf{k}_l} e^{i\mathbf{k}\cdot\mathbf{R}_l} \int d^2r \, \phi^{(i)*}(\mathbf{r}) \left[H^a + \Delta V \right] \phi^{(j)}(\mathbf{r} + \boldsymbol{\delta}_{ij} - \mathbf{R}_l) \\
&= N \left(\epsilon^{(i)} s_{\mathbf{k}}^{ij} + t_{\mathbf{k}}^{ij} \right) \qquad\qquad\qquad\qquad (2.12)
\end{aligned}$$

where $\delta_{ij} \equiv \delta_j - \delta_i$,

$$s_{\mathbf{k}}^{ij} \equiv \sum_{\mathbf{R}_l} e^{i\mathbf{k}\cdot\mathbf{R}_l} \int d^2r \, \phi^{(i)*}(\mathbf{r} + \boldsymbol{\delta}_i - \mathbf{R}_k) \phi^{(j)}(\mathbf{r} + \boldsymbol{\delta}_j - \mathbf{R}_m) = \frac{\mathcal{S}_{\mathbf{k}}^{ij}}{N} \qquad (2.13)$$

and we have defined the *hopping matrix*

$$t_{\mathbf{k}}^{ij} \equiv \sum_{\mathbf{R}_l} e^{i\mathbf{k}\cdot\mathbf{R}_l} \int d^2r \, \phi^{(i)*}(\mathbf{r} + \boldsymbol{\delta}_i - \mathbf{R}_k) \Delta V \phi^{(j)}(\mathbf{r} + \boldsymbol{\delta}_j - \mathbf{R}_m) . \qquad (2.14)$$

The last line in Eq. (2.12) has been obtained from the fact that the atomic wavefunctions $\phi^{(i)}(\mathbf{r})$ are eigenstates of the atomic Hamiltonian H^a with the
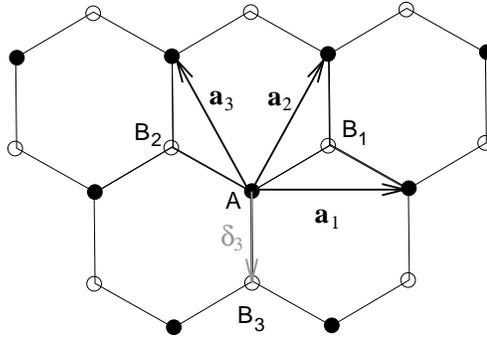
Figure 2.1: Tight-binding model for the honeycomb lattice.

atomic energy $\epsilon^{(i)}$ for an orbital of type i. This atomic energy plays the role of an onsite energy. The secular equation now reads

$$\det\left[t_{\mathbf{k}}^{ij} - \left(\epsilon_{\mathbf{k}}^{\lambda} - \epsilon^{(i)}\right)\right] = 0. \tag{2.15}$$

Notice that, if the the atoms on the different sublattices are all of the same electronic configuration, one has $\epsilon^{(i)} = \epsilon_0$ for all i, and one may omit this onsite energy, which yields only a constant physically irrelevant shift of the energy bands.

2.1.3 Solution for graphene with nearest-neighbour and next-nearest-neighour hopping

After these formal considerations, we now study the particular case of the tight-binding model on the honeycomb lattice, which yields, to great accuracy, the π energy bands of graphene. Because all atomic orbitals are p_z orbitals of carbon atoms, we may omit the onsite energy ϵ_0, as discussed in the last paragraph. We choose the Bravais lattice vectors to be those of the A sublattice, i.e. $\boldsymbol{\delta}_A = 0$, and the equivalent site on the B sublattice is obtained by the displacement $\boldsymbol{\delta}_B = \delta_{AB} = \boldsymbol{\delta}_3$ (see Fig. 2.1). The hopping amplitude between *nn* is given by the expression

$$t \equiv \int d^2r\, \phi^{A*}(\mathbf{r})\Delta V \phi^B(\mathbf{r} + \boldsymbol{\delta}_3), \tag{2.16}$$

and we also take into account *nnn* hopping which connects *nn* sites on the *same* sublattice

$$t_{nnn} \equiv \int d^2r \, \phi^{A*}(\mathbf{r})\Delta V \phi^A(\mathbf{r} + \mathbf{a}_1) = \int d^2r \, \phi^{B*}(\mathbf{r})\Delta V \phi^B(\mathbf{r} + \mathbf{a}_1). \quad (2.17)$$

Notice that one may have chosen any other vector $\boldsymbol{\delta}_j$ or $\mathbf{a}_2$, respectively, in the calculation of the hopping amplitudes. Because of the normalisation of the atomic wavefunctions, we have $\int d^2r \phi^{(j)*}(\mathbf{r})\phi^{(j)}(\mathbf{r}) = 1$, and we consider furthermore the overlap correction between orbitals on *nn* sites,

$$s \equiv \int d^2r \, \phi^{A*}(\mathbf{r})\phi^B(\mathbf{r} + \boldsymbol{\delta}_3). \quad (2.18)$$

We neglect overlap corrections between all other orbitals which are not *nn*, as well as hopping amplitudes for larger distances than *nnn*.

If we now consider an arbitrary site A on the A sublattice (Fig. 2.1), we may see that the off-diagonal terms of the hopping matrix (2.14) consist of three terms corresponding to the *nn* B_1, B_2, and B_3, all of which have the same hopping amplitude t. However, only the site B_3 is described by the same lattice vector (shifted by $\boldsymbol{\delta}_3$) as the site A and thus yields a zero phase to the hopping matrix. The sites B_1 and B_2 correspond to lattice vectors shifted by $\mathbf{a}_1$ and $\mathbf{a}_3 \equiv \mathbf{a}_2 - \mathbf{a}_1$, respectively. Therefore, they contribute a phase factor $\exp(i\mathbf{k} \cdot \mathbf{a}_1)$ and $\exp(i\mathbf{k} \cdot \mathbf{a}_3)$, respectively. The off-diagonal elements of the hopping matrix may therefore be written as[4]

$$t_{\mathbf{k}}^{AB} = t\gamma_{\mathbf{k}}^* = \left(t_{\mathbf{k}}^{BA}\right)^*,$$

as well as those of the overlap matrix

$$s_{\mathbf{k}}^{AB} = s\gamma_{\mathbf{k}}^* = \left(s_{\mathbf{k}}^{BA}\right)^*,$$

($s_{\mathbf{k}}^{AA} = s_{\mathbf{k}}^{BB} = 1$, due to the above-mentioned normalisation of the atomic wavefunctions), where we have defined the sum of the *nn* phase factors

$$\gamma_{\mathbf{k}} \equiv 1 + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_3}. \quad (2.19)$$

---

[4]The hopping matrix element $t_{\mathbf{k}}^{AB}$ corresponds to a hopping from the B to the A sublattice.

The *nnn* hopping amplitudes yield the diagonal elements of the hopping matrix,

$$
\begin{aligned}
t_{\mathbf{k}}^{AA} = t_{\mathbf{k}}^{BB} &= t_{nnn} \left(e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2} + e^{-i\mathbf{k}\cdot\mathbf{a}_2} + e^{i\mathbf{k}\cdot\mathbf{a}_3} + e^{-i\mathbf{k}\cdot\mathbf{a}_3} \right) \\
&= 2t_{nnn} \sum_{i=1}^{3} \cos(\mathbf{k}\cdot\mathbf{a}_i) = t_{nnn} \left(|\gamma_{\mathbf{k}}|^2 - 3 \right),
\end{aligned}
$$

and one obtains, thus, the secular equation

$$
\det \begin{bmatrix} t_{\mathbf{k}}^{AA} - \epsilon_{\mathbf{k}} & (t - s\epsilon_{\mathbf{k}})\gamma_{\mathbf{k}}^* \\ (t - s\epsilon_{\mathbf{k}})\gamma_{\mathbf{k}} & t_{\mathbf{k}}^{AA} - \epsilon_{\mathbf{k}} \end{bmatrix} = 0 \tag{2.20}
$$

with the two solutions ($\lambda = \pm$)

$$
\epsilon_{\mathbf{k}}^{\lambda} = \frac{t_{\mathbf{k}}^{AA} + \lambda t |\gamma_{\mathbf{k}}|}{1 + \lambda s |\gamma_{\mathbf{k}}|} \ . \tag{2.21}
$$

This expression may be expanded under the reasonable assumptions $s \ll 1$ and $t_{nnn} \ll t$, which we further justify at the end of the paragraph,

$$
\begin{aligned}
\epsilon_{\mathbf{k}}^{\lambda} &= t_{\mathbf{k}}^{AA} + \lambda t |\gamma_{\mathbf{k}}| - st |\gamma_{\mathbf{k}}|^2 = t_{nnn}' |\gamma_{\mathbf{k}}|^2 + \lambda t |\gamma_{\mathbf{k}}| \\
&= 2t_{nnn}' \sum_{i=1}^{3} \cos(\mathbf{k}\cdot\mathbf{a}_i) + \lambda t \sqrt{3 + 2\sum_{i=1}^{3} \cos(\mathbf{k}\cdot\mathbf{a}_i)} \tag{2.22}
\end{aligned}
$$

where we have defined the effective *nnn* hopping amplitude

$$
t_{nnn}' \equiv t_{nnn} - st \ , \tag{2.23}
$$

and we have omitted the unimportant constant $-3t_{nnn}$ in the last equation. One, therefore, notices that the overlap corrections simply yield a renormalisation of the *nnn* hopping amplitudes. The hopping amplitudes may be determined by fitting the energy dispersion (2.22) obtained within the tight-binding approximation to those calculated numerically in more sophisticated band-structure calculations. These yield a value of $t \simeq -3$ eV for the *nn* hopping amplitude and $t_{nnn}' \simeq 0.1t$, which justifies the above-mentioned expansion for $t_{nnn}'/t \ll 1$. Notice that this fitting procedure does not allow for a distinction between the "true" *nnn* hopping amplitude t_{nnn} and the contribution from the overlap correction $-st$. We, therefore, omit this distinction in the following discussion and omit the prime at the effective *nnn* hopping amplitude, but one should keep in mind that it is an effective parameter with a contribution from *nn* overlap corrections.
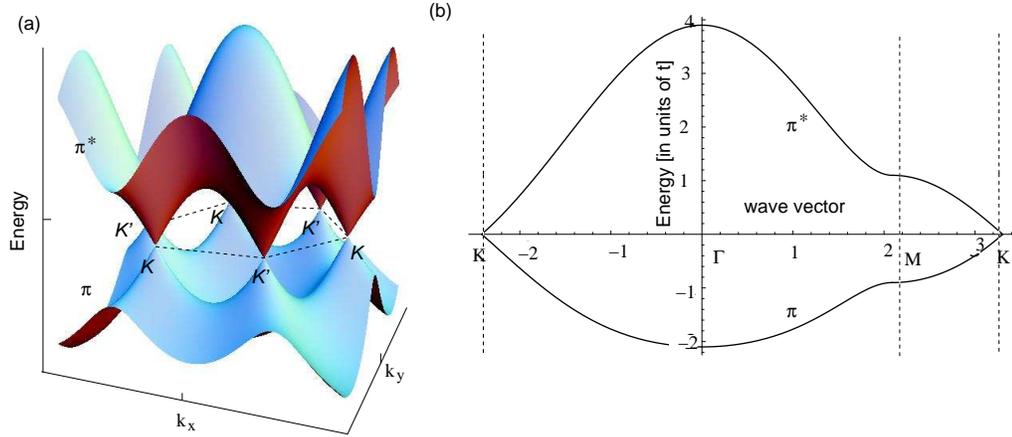
Figure 2.2: Energy dispersion obtained within the tight-binding approximation, for $t_{nnn}/t = 0.1$. One distinguishes the valence (π) band from the conduction (π^*) band. The Fermi level is situated at the points where the π band touches the π^* band. *(a)* Energy dispersion as a function of the wave-vector components k_x and k_y. *(b)* Cut throught the energy dispersion along characteristic lines (connecting the points $K \rightarrow \Gamma \rightarrow M \rightarrow K$. The energy is measured in units of t and the wave vectors in units of $1/a$.

Energy dispersion of π electrons in graphene

The energy dispersion (2.22) is plotted in Fig. 2.2 for $t_{nnn}/t = 0.1$. It consists of two bands, labeled by the index $\lambda = \pm$, each of which contains the same number of states. Because each carbon atom contributes one π electron and each electron may occupy either a spin-up or a spin-down state, the lower band with $\lambda = -$ (the π or valence band) is completely filled and that with $\lambda = +$ (the π^* or conduction band) completely empty. The Fermi level is, therefore, situated at the points where the π band touches the π^* band. Notice that, if $t_{nnn} = 0$, the energy dispersion (2.22) is electron-hole symmetric, i.e. $\epsilon_{\mathbf{k}}^{\lambda} = -\epsilon_{\mathbf{k}}^{-\lambda}$. This means that *nnn* hopping and *nn* overlap corrections break the electron-hole symmetry. The points, where the π band touches the π^* band, are called *Dirac points*, for reasons that are explained in the following chapter. They situated at the points $\mathbf{k}^D$ where the energy dispersion (2.22) is zero,

$$\epsilon_{\mathbf{k}^D}^{\lambda} = 0. \tag{2.24}$$

Eq. (2.24) is satisfied when $\gamma_{\mathbf{k}^D} = 0$, i.e. when

$$
\begin{aligned}
\mathrm{Re}\gamma_{\mathbf{k}^D} &= 1 + \cos(\mathbf{k}^D \cdot \mathbf{a}_2) + \cos(\mathbf{k}^D \cdot \mathbf{a}_3) \quad\quad\quad\quad (2.25) \\
&= 1 + \cos\left[\frac{\sqrt{3}a}{2}(k_x^D + \sqrt{3}k_y^D)\right] + \cos\left[\frac{\sqrt{3}a}{2}(-k_x^D + \sqrt{3}k_y^D)\right] = 0
\end{aligned}
$$

and, equally,

$$
\mathrm{Im}\gamma_{\mathbf{k}^D} = \sin\left[\frac{\sqrt{3}a}{2}(k_x^D + \sqrt{3}k_y^D)\right] + \sin\left[\frac{\sqrt{3}a}{2}(-k_x^D + \sqrt{3}k_y^D)\right] = 0. \quad (2.26)
$$

The last equation may be satisfied by the choice $k_y^D = 0$, and Eq. (2.25), thus, when

$$
1 + 2\cos\left(\frac{\sqrt{3}a}{2}k_x^D = 0\right) \quad\quad \Rightarrow \quad\quad k_x^D = \pm\frac{4\pi}{3\sqrt{3}a}.
$$

Comparison with Eq. (1.6) shows that there are, thus, two[5] inequivalent Dirac points D and D', which are situated at the points K and K', respectively,

$$
\mathbf{k}^D = \pm\mathbf{K} = \pm\frac{4\pi}{3\sqrt{3}a}\mathbf{e}_x. \quad\quad\quad\quad (2.27)
$$

Although situated at the same position in the first BZ, it is useful to make a clear conceptual distinction between the Dirac points, which are defined as the points where the two bands π and π^* touch each other, and the purely crystallographic points K and K', which are defined as the corners of the first BZ. There are, indeed, situations where the Dirac points move away from the points K and K', e.g. when the *nn* hopping amplitudes are no longer the same in the directions $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, and $\boldsymbol{\delta}_3$ [19]. In the following chapters, we will, however, consider the natural situation for graphene where the Dirac points are situated at the BZ corners and use the notation K and K' for the Dirac points.

Notice that because of the symmetry $\epsilon_{-\mathbf{k}} = \epsilon_{\mathbf{k}}$, which is a consequence of time-reversal symmetry, Dirac points occur necessarily in pairs – if $\mathbf{k}^D$ is a

---

[5]We remind the reader that there are only two inequivalent points, and not six as Fig. 2.2 (b) might suggest. As pointed out in Sec. 1.2.1, there are pairs of three points that may be connected to one another by a reciprocal lattice vector and that are, thus, crystallographically equivalent.

solution of $\epsilon_{\mathbf{k}} = 0$, so is $-\mathbf{k}^D$. In graphene, there is one pair of Dirac points, and the zero-energy states are, therefore, doubly degenerate. One speaks of a twofold *valley degeneracy*, which survives when we consider low-energy electronic excitations that are restricted to the vicinity of the Dirac points, as is discussed in section 2.2.

Effective tight-binding Hamiltonian

Before considering the low-energy excitations and the continuum limit, it is useful to define an effective tight-binding Hamiltonian,

$$\mathcal{H}_{\mathbf{k}} \equiv t_{nnn}|\gamma_{\mathbf{k}}|^2 \mathbb{1} + t \begin{pmatrix} 0 & \gamma_{\mathbf{k}}^* \\ \gamma_{\mathbf{k}} & 0 \end{pmatrix} . \tag{2.28}$$

Here, $\mathbb{1}$ represents the 2×2 one-matrix

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

This Hamiltonian effectively omits the problem of non-orthogonality of the wavefunctions by a simple renormalisation of the *nnn* hopping amplitude, as alluded to above. The eigenstates of the effective Hamiltonian (2.28) are the spinors

$$\Psi_{\mathbf{k}}^{\lambda} = \begin{pmatrix} a_{\mathbf{k}}^{\lambda} \\ b_{\mathbf{k}}^{\lambda} \end{pmatrix} , \tag{2.29}$$

the components of which are the probability amplitudes of the Bloch wave-function (2.5) on the two different sublattices A and B. They may be determined by considering the eigenvalue equation $\mathcal{H}_{\mathbf{k}}(t_{nnn} = 0)\Psi_{\mathbf{k}}^{\lambda} = \lambda t|\gamma_{\mathbf{k}}|\Psi_{\mathbf{k}}^{\lambda}$, which does not take into account the *nnn* hopping correction. Indeed, these eigenstates are also those of the Hamiltonian with $t_{nnn} \neq 0$ because the *nnn* term is proportional to the one-matrix $\mathbb{1}$. The solution of the eigenvalue equation (2.29) yields

$$a_{\mathbf{k}}^{\lambda} = \lambda \frac{\gamma_{\mathbf{k}}^*}{|\gamma_{\mathbf{k}}|} b_{\mathbf{k}}^{\lambda} = \lambda e^{-i\varphi_{\mathbf{k}}} b_{\mathbf{k}}^{\lambda}$$

and, thus, the eigenstates[6]

$$\Psi_{\mathbf{k}}^{\lambda} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \lambda e^{i\varphi_{\mathbf{k}}} \end{pmatrix},$$ (2.30)

where we have defined the angle

$$\varphi_{\mathbf{k}} = \arctan\left(\frac{\mathrm{Im}\gamma_{\mathbf{k}}}{\mathrm{Re}\gamma_{\mathbf{k}}}\right).$$ (2.31)

As one may have expected, the spinor represents an equal probability to find an electron in the state $\Psi_{\mathbf{k}}^{\lambda}$ on the A as on the B sublattice because both sublattices are built from carbon atoms with the same onsite energy $\epsilon^{(i)}$.[7]

2.2 Continuum Limit

In order to describe the low-energy excitations, i.e. electronic excitations the characteristic energy of which is much smaller than the band width $\sim |t|$, one may concentrate on excitations at the Fermi level. This amounts to restricting the excitations to quantum states in the vicinity of the Dirac points, and one may expand the energy dispersion around $\pm\mathbf{K}$. The wave

---

[6]The eigenstates are defined up to a global (but $\mathbf{k}$-dependent) phase, and one may also choose

$$\Psi_{\mathbf{k}}^{\lambda} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\varphi_{\mathbf{k}}} \\ \lambda \end{pmatrix},$$

which is also found in the litterature.

[7]This is not the case for other lattices such as 2D boron nitride (BN), which also forms a hexagonal lattice. However, in the case of BN, one sublattice consists of boron atoms with an atomic (onsite) energy ϵ^A and the other one of nitrogen atoms with an onsite energy $\epsilon^B \neq \epsilon^A$. The difference in the onsite energy $\mu = \epsilon^A - \epsilon^B$ may be accounted for in the effective Hamiltonian (2.28) if one adds a term

$$\frac{\mu}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This term opens a gap in the energy dispersion at the points K and K',

$$\epsilon_{\mathbf{k}}^{\lambda} = t_{nnn}|\gamma_{\mathbf{k}}|^2 + \lambda\sqrt{t^2|\gamma_{\mathbf{k}}|^2 + \frac{\mu^2}{4}},$$

and it is energetically favourable to fill preferentially the sublattice with lower onsite energy.

vector is, thus, decomposed as $\mathbf{k} = \pm\mathbf{K} + \mathbf{q}$, where $|\mathbf{q}| \ll |\mathbf{K}| \sim 1/a$. The small parameter, which governs the expansion of the energy dispersion, is, therefore, $|\mathbf{q}|a \ll 1$.

It is evident from the form of the energy dispersion (2.22) and the effective Hamiltonian that the basic entity to be expanded is the sum of the phase factors $\gamma_{\mathbf{k}}$. We need to distinguish the sum at the K point from that at the K' point,

$$
\begin{aligned}
\gamma_{\mathbf{q}}^{\pm} \equiv \gamma_{\mathbf{k}=\pm\mathbf{K}+\mathbf{q}} &= 1 + e^{\pm i\mathbf{K}\cdot\mathbf{a}_2}e^{i\mathbf{q}\cdot\mathbf{a}_2} + e^{\pm i\mathbf{K}\cdot\mathbf{a}_3}e^{i\mathbf{q}\cdot\mathbf{a}_3} \\
&\simeq 1 + e^{\pm i2\pi/3}\left[1 + i\mathbf{q}\cdot\mathbf{a}_2 - \frac{1}{2}(\mathbf{q}\cdot\mathbf{a}_2)^2\right] \\
&\quad + e^{\mp i2\pi/3}\left[1 + i\mathbf{q}\cdot\mathbf{a}_3 - \frac{1}{2}(\mathbf{q}\cdot\mathbf{a}_3)^2\right] \\
&= \gamma_{\mathbf{q}}^{\pm(0)} + \gamma_{\mathbf{q}}^{\pm(1)} + \gamma_{\mathbf{q}}^{\pm(2)}
\end{aligned}
$$

By definition of the Dirac points and their position at the BZ corners K and K', we have $\gamma_{\mathbf{q}}^{\pm(0)} = \gamma_{\pm\mathbf{K}} = 0$. We limit the expansion to second order in $|\mathbf{q}|a$.

First order in $|\mathbf{q}|a$

The first order term is given by

$$
\begin{aligned}
\gamma_{\mathbf{q}}^{\pm(1)} &= i\frac{\sqrt{3}a}{2}\left[(q_x + \sqrt{3}q_y)e^{\pm i2\pi/3} + (-q_x + \sqrt{3}q_y)e^{\mp i2\pi/3}\right] \\
&= \mp\frac{3a}{2}(q_x \pm iq_y),
\end{aligned}
\tag{2.32}
$$

which is obtained with the help of $\sin(\pm 2\pi/3) = \pm\sqrt{3}/2$ and $\cos(\pm 2\pi/3) = -1/2$. This yields the effective low-energy Hamiltonian

$$
\mathcal{H}_{\mathbf{q}}^{\mathrm{eff},\xi} = \xi\hbar v_F(q_x\sigma^x + \xi q_y\sigma^y),
\tag{2.33}
$$

where we have defined the Fermi velocity[8]

$$
v_F \equiv -\frac{3ta}{2\hbar} = \frac{3|t|a}{2\hbar}
\tag{2.34}
$$

---

[8]The minus sign in the definition is added to render the Fermi velocity positive because the hopping parameter $t \simeq -3\text{eV}$ happens to be negative, as mentioned in the last section.

and the Pauli matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \text{and} \qquad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Furthermore, we have introduced the *valley isospin* $\xi = \pm$, where $\xi = +$ denotes the K point at $+\mathbf{K}$ and $\xi = -$ the K' point at $-\mathbf{K}$. The low-energy Hamiltonian (2.33) does not take into account *nnn* hopping corrections, which are proportional to $|\gamma_\mathbf{k}|^2$ and, thus, occur only in the second-order expansion of the energy dispersion [at order $(|\mathbf{q}|a)^2$]. The energy dispersion (2.22), therefore, reads

$$\epsilon^\lambda_{\mathbf{q},\xi=\pm} = \lambda\hbar v_F|\mathbf{q}|, \tag{2.35}$$

independent of the the valley isospin ξ. We have already alluded to this twofold valley degeneracy in Sec. 2.1.3, in the framework of the discussion of the zero-energy states at the BZ corners. The twofold valley degeneracy, thus, survives when considering the low-energy excitations in the vicinity of the Dirac points. From Eq. (2.35) it is apparent that the continuum limit $|\mathbf{q}|a \ll 1$ coincides with the limit $|\epsilon| \ll |t|$, as described above, because $|\epsilon_\mathbf{q}| = 3ta|\mathbf{q}|/2 \ll |t|$.

It is convenient to invert the spinor components at the K' point (for $\xi = -$),

$$\Psi_{\mathbf{k},\xi=+} = \begin{pmatrix} \psi^A_{\mathbf{k},+} \\ \psi^B_{\mathbf{k},+} \end{pmatrix}, \qquad \Psi_{\mathbf{k},\xi=-} = \begin{pmatrix} \psi^B_{\mathbf{k},-} \\ \psi^A_{\mathbf{k},-} \end{pmatrix},$$

i.e. to invert the role of the two sublattices. In this case, the effective low-energy Hamiltonian may be represented as

$$\mathcal{H}^{\text{eff},\xi}_\mathbf{q} = \xi\hbar v_F(q_x\sigma^x + q_y\sigma^y) = \hbar v_F \tau^z \otimes \mathbf{q} \cdot \boldsymbol{\sigma}, \tag{2.36}$$

where we have introduced the four-spinor representation

$$\Psi_\mathbf{k} = \begin{pmatrix} \psi^A_{\mathbf{k},+} \\ \psi^B_{\mathbf{k},+} \\ \psi^B_{\mathbf{k},-} \\ \psi^A_{\mathbf{k},-} \end{pmatrix}$$

in the last line via the 4×4 matrices

$$\tau^z \otimes \boldsymbol{\sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & -\boldsymbol{\sigma} \end{pmatrix},$$

and $\boldsymbol{\sigma} \equiv (\sigma^x, \sigma^y)$. In this four-spinor representation, the first two components represent the lattice components at the K point and the last two components those at the K' point. We emphasise that one must clearly distinguish both types of isospin: (a) the *sublattice isospin* is represented by the Pauli matrices σ^j, where "spin up" corresponds to the component on one sublattice and "spin down" to that on the other one. A rotation within the SU(2) sublattice isospin space yields the band indices $\lambda = \pm$, and the band index (or *band isospin*) is, thus, intimitely related to the sublattice isospin. (b) The *valley isospin*, which is described by a second set of Pauli matrices τ^j, the z-component of which appears in the Hamiltonian (2.36), is due to the twofold valley degeneracy and is only indirectly related to the presence of two sublattices.

The eigenstates of the Hamiltonian (2.36) are the four-spinors

$$
\Psi_{\mathbf{k},\lambda}^{\xi=+} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \lambda e^{i\varphi_\mathbf{q}} \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \Psi_{\mathbf{k},\lambda}^{\xi=-} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -\lambda e^{i\varphi_\mathbf{q}} \end{pmatrix}, \quad (2.37)
$$

where we have, now,

$$
\varphi_\mathbf{q} = \arctan\left(\frac{q_y}{q_x}\right) . \quad (2.38)
$$

Density of states

The electronic density of states $\rho(\epsilon)$ counts the number of quantum states in the vicinity of a fixed energy ϵ. It turns out to be a very useful quantity when discussing the electric transport in the diffusive regime. It may be obtained from the total number of states below the energy ϵ,

$$
N^{\lambda=+} = g \sum_{\mathbf{k}/\epsilon\mathbf{k}^+ \leq \epsilon} = A \int_0^\epsilon d\epsilon' \rho(\epsilon'), \quad (2.39)
$$

where we have chosen the electronic π^* band, by derivation with respect to the energy

$$
\rho(\epsilon) = \frac{1}{A} \frac{\partial N^+}{\partial \epsilon}, \quad (2.40)
$$

in terms of the total surface A.[9] The factor g takes into account the degeneracy due to internal degrees of freedom – $g = 4$ in the case of graphene because of the twofold valley degeneracy and the electron spin.

The calculation of the density of states may be rather involved because one needs to calculate a 2D integral over the wave vector as a function of energy, by inverting the energy dispersion. In the vicinity of the Dirac points, however, it may be calculated quite easily because of the isotropy of the energy dispersion (2.35). One may identify

$$\sum_{\mathbf{k}/\epsilon_{\mathbf{k}}^+ \leq \epsilon} \simeq \frac{Ag}{2\pi} \int_0^{q(\epsilon)} dq \, q \ ,$$

where we have profited from the isotropy of the energy dispersion in the vicinity of the Dirac point, with the right hand side of Eq. (2.39), and one obtains

$$\rho(\epsilon) = \frac{g}{2\pi} \frac{q(\epsilon)}{\partial \epsilon / \partial q} \ . \tag{2.41}$$

By inverting the energy dispersion (2.35), one obtains

$$\rho(\epsilon \geq 0) = \frac{g\epsilon}{2\pi \hbar^2 v_F^2} \ ,$$

and for the full (positive and negative) energy range[10]

$$\rho(\epsilon) = \frac{g|\epsilon|}{2\pi \hbar^2 v_F^2} \ , \tag{2.42}$$

as long as $|\epsilon| \ll |t|$.

The density of states, therefore, vanishes linearly at the Dirac points, at zero energy. This is a direct consequence of the linearity of the energy dispersion in the vicinity of the Dirac points. This particular situation needs to be contrasted to the conventional case of electrons in 2D metals, with an energy dispersion of $\epsilon = \hbar^2 q^2 / 2m^*$, in terms of the band mass m^*, where one obtains a constant density of states, $\rho^{2D}(\epsilon) = gm^*/2\pi\hbar^2$.

The full density of states, in the absence of *nnn* hopping corrections, has been calculated by Hobson and Nierenberger in 1953 [20]. It is schematically

---

[9]The density of states defined in this manner is indeed the density of states per unit surface, or unit volume in 3D.

[10]The negative branch of the density of states may be calculated in the same manner.
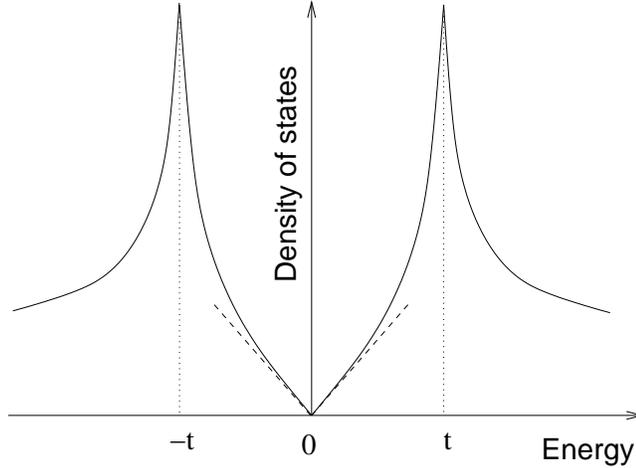
Figure 2.3: Schematic plot of the density of states for electrons in graphene in the absence of *nnn* hopping. The dashed line indicates the density of states obtained at linear order in the continuum limit.

shown in Fig 2.3. In agreement with the result (2.42) obtained at linear order in the continuum limit, it vanishes linearly at zero energy. The divergencies at $\pm t$, called *van-Hove singularities*, are due to the saddle points of the energy dispersion at the M points at the borders of the first BZ (see Fig. 2.2). This divergency may be understood qualitatively from Eq. (2.42) – roughly speaking, the derivative of the energy dispersion occurs in the denominator of the expression for the density of states and, thus, causes a divergency whenever it becomes zero, such as at the extrema or saddle points of the energy dispersion.

Second order in $|\mathbf{q}|a$

Although most of the fundamental properties of graphene are captured within the effective model obtained at first order in the expansion of the energy dispersion, it is useful to take into account second-order terms. These corrections include *nnn* hopping corrections and off-diagonal second-order contributions from the expansion of $\gamma_{\mathbf{k}}$. The latter yield the so-called *trigonal warping*, which consist of an anisotropy of the energy dispersion around the Dirac points.

The diagonal second-order correction, which stems from the *nnn* hopping,

is readily obtained from Eq. (2.32),

$$\mathcal{H}_{nnn}^{\xi} = t_{nnn}|\gamma_{\mathbf{q}}^{\xi}|^2 \mathbb{1} \simeq t_{nnn}|\gamma_{\mathbf{q}}^{\xi(1)}|^2 \mathbb{1} = \frac{9a^2}{4} t_{nnn}|\mathbf{q}|^2 \mathbb{1}. \qquad (2.43)$$

The off-diagonal second-order corrections are somewhat more involved because we need to calculate the modulus of $\gamma_{\mathbf{q}}^{\xi}$,

$$|\gamma_{\mathbf{q}}^{\xi}| \simeq \left[|\gamma_{\mathbf{q}}^{\xi(1)}|^2 + 2\mathrm{Re}\gamma_{\mathbf{q}}^{\xi(1)*}\gamma_{\mathbf{q}}^{\xi(2)}\right]^{1/2} \simeq |\gamma_{\mathbf{q}}^{\xi(1)}| \left[1 + \frac{\mathrm{Re}\gamma_{\mathbf{q}}^{\xi(1)*}\gamma_{\mathbf{q}}^{\xi(2)}}{|\gamma_{\mathbf{q}}^{\xi(1)}|^2}\right] \quad .$$

Because of

$$\mathrm{Re}\gamma_{\mathbf{q}}^{\xi(1)*}\gamma_{\mathbf{q}}^{\xi(2)} = -\xi\frac{9a^2}{16}q_x(q_x^2 - 3q_y^2) ,$$

one obtains at second order

$$|\gamma_{\mathbf{q}}^{\xi}| \simeq \frac{3a}{2}|\mathbf{q}| \left[1 - \xi\frac{|\mathbf{q}|a}{4}\cos(3\varphi_{\mathbf{q}})\right],$$

where we have used the parametrisation $q_x = |\mathbf{q}|\cos\varphi_{\mathbf{q}}$ and $q_y = |\mathbf{q}|\sin\varphi_{\mathbf{q}}$. Finally, the energy dispersion (2.22) expanded to second order in $|\mathbf{q}|a$ reads

$$\epsilon_{\mathbf{q},\xi}^{\lambda} = \frac{9a^2}{4} t_{nnn}|\mathbf{q}|^2 + \lambda\hbar v_F|\mathbf{q}| \left[1 - \xi\frac{|\mathbf{q}|a}{4}\cos(3\varphi_{\mathbf{q}})\right] \quad . \qquad (2.44)$$

As we have already mentioned in Sec. 2.1.3, it is apparent from Eq. (2.44) that the *nnn* correction (which also accounts for *nn* overlap corrections) breaks the electron-hole symmetry $\epsilon_{\mathbf{q},\xi}^{-\lambda} = -\epsilon_{\mathbf{q},\xi}^{\lambda}$. This correction is, however, a rather small correction, of order $|\mathbf{q}|at_{nnn}/t$, to the first-order effective Hamiltonian (2.36). The second-order expansion of the phase factor sum $\gamma_{\mathbf{q}}$ yields a more relevant correction – the third term in Eq. (2.44), of order $|\mathbf{q}|a \gg |\mathbf{q}|at_{nnn}/t$ – to the linear theory. It depends explicitely on the valley isospin ξ and renders the energy dispersion anisotropic in $\mathbf{q}$ around the K and K' point. The tripling of the period, due to the term $\cos(3\varphi_{\mathbf{q}})$, is a consequence of the symmetry of the underlying lattice and is also called *trigonal warping*, as mentioned above.

The trigonal warping of the dispersion relation is visualised in Fig. 2.4, where we have plotted the contours of constant (positive) energy in Fourier space. The closed energy contours around the K and K' points at low energy are sparated by the high-energy contours around the Γ point by the dashed
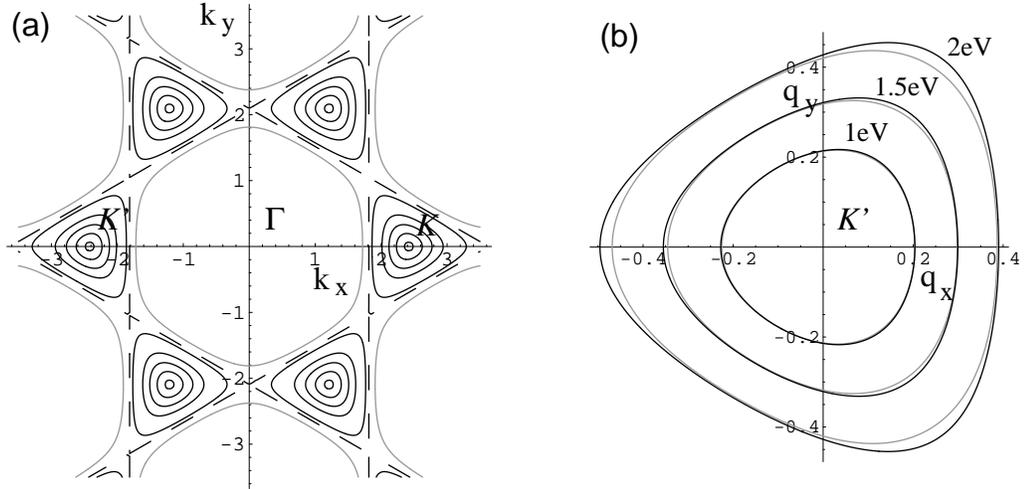
Figure 2.4: Contours of constant (positive) energy in the wave-vector space. *(a)* Contours obtained from the full dispersion relation (2.22). The dashed line corresponds to the energy $t+t_{nnn}$, which separates closed orbits around the K and K' points (black lines, with energy $\epsilon < t + t_{nnn}$) from those around the Γ point (gray line, with energy $\epsilon < t + t_{nnn}$). *(b)* Comparison of the contours at energy $\epsilon = 1$ eV, 1.5 eV, and 2 eV around the K' point. The black lines correspond to the energies calculated from the full dispersion relation (2.22) and the gray ones to those calculated to second order within the continuum limit (2.44).
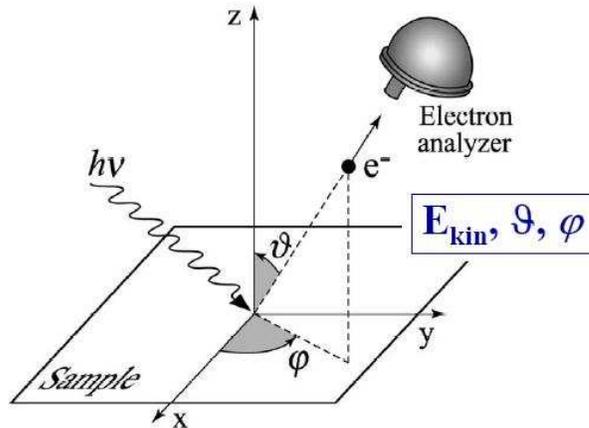
Figure 2.5: Schematic view of an ARPES measurement (from Ref. [21]) The analyser detects the energy E_f of the photoemitted electron as a function of the angles ϑ and φ, which are related to the momentum of the electronic state.

lines in Fig. 2.4 (a) at energy $|t+t_{nnn}|$ the crossing points of which correspond to the M points. As mentioned above, the dispersion relation has saddle points at these points at the border of the first BZ, which yield the van-Hove singularities in the density of states (see Fig. 2.3). In Fig. 2.4 (b), we compare constant-energy contours of the full dispersion relation to those obtained from Eq. (2.44) calculated within a second-order expansion. The contours are indistinguishable for an energy of $\epsilon = |t|/3 \simeq 1$ eV, and the continuum limit yields rather accoucate results up to energies as large as 2 eV. Notice that, in today's exfoliated graphene samples on a SiO_2 substrates, one may probe, by field-effect doping of the graphene sheet, energies which are on the order of 100 meV. Above these energies the condenser breaks down, and Fig. 2.4 (a) indicates that the continuum limit (2.44) yields extremely accurate results at these energies.

2.3 Experimental Characterisation of the Electronic Band Structure

The energy dispersion relation of solids may be determined by ARPES (angle resolved photoemission spectroscopy). A typical ARPES measurement consists of the detection of the direction and energy of a photoemitted elec-
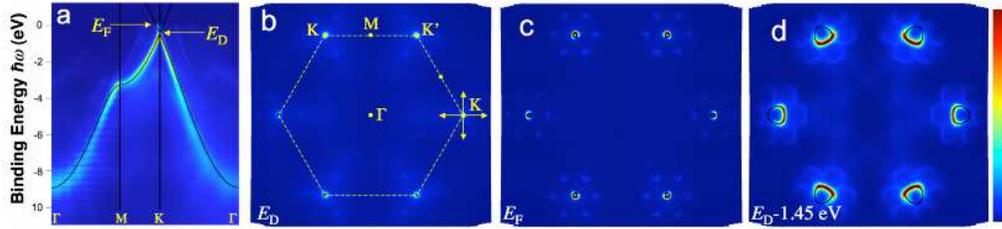
Figure 2.6: ARPES measurement of the dispersion relation of epitaxial graphene (from Ref. [22]).

tron, i.e. an electron emitted from the solid to be studied after prior optical excitation by a photon of well-defined energy and momentum, $\hbar\nu$ and $\hbar\mathbf{k}$, respectively (see Fig. 2.5). The technique, which measures the spectral function, the imaginary part of the electronic Green's function $1/(\omega - \epsilon\mathbf{k}^\lambda + i\eta)$, is based on energy and momentum conservation in a electron-photon scattering process,

$$\hbar\nu = E_f - E_i, \qquad \mathbf{k} = \mathbf{k}_f - \mathbf{k}_i,$$

where E_i and $\mathbf{k}_i$ represent the energy and momentum, respectively, of the initial electronic state and E_f and $\mathbf{k}_f$ those of the final electronic state, i.e. the state of the photoemitted and detected electron. The direction of the latter, in terms of the angles ϑ and φ shown in Fig. 2.5, yields precisely the momentum-dependence of the energy and thus the dispersion relation.

Fig. 2.6 shows ARPES results obtained by Bostwick *et al.* on epitaxial graphene [22]. Similar results have been obtained on thin graphitic samples [23]. The studied sample has been obtained by graphitisation of the Si-terminated (0001) surface. Fig. 2.6 (a) shows the energy dispersion of the π band and yields a good agreement with that calculated in Sec. 2.1.3 [see Fig. 2.2 (b)]. In order to obtain an ARPES signal, the initial state must be occupied, and one may see from Fig. 2.6 (a) that the Fermi level is in the π^* band at roughly $E_F \simeq 0.45$ eV. This is in agreement with the statement in Sec. 1.3.2 that the graphene layer on SiC is naturally electron-doped due to charge transfer from the SiC substrate. Fig. 2.6 (b) shows the form of the first BZ, which is obtained by an energy cut through the dispersion relation at zero energy (at the Dirac points). The energy cut at the Fermi level, $E_F \simeq 0.45$ eV, is shown in Fig. 2.6 (b) and reveals the circular Fermi surface around the K and K' points. The faint satellites are due to the underlying (hexagonal) structure of the SiC substrate.

Finally, Fig. 2.6 (d) corresponds to a cut at -1.0 eV, i.e. in the π band at an energy rather remote from the Dirac points. The abovementioned trigonal warping of the Fermi surface is clearly visible at this energy.

Chapter 3

The Dirac Equation for Relativistic Fermions

As we have discussed in the last chapter, the low-energy properties of electrons in graphene are described by the Hamiltonian (2.36),

$$\mathcal{H}_{\mathbf{q}}^{\xi} = \xi \hbar v_F \mathbf{q} \cdot \boldsymbol{\sigma}$$

which happens to be of the form of the Dirac equation describing relativistic massless fermions. We, therefore, devote this chapter to a brief introduction to relativistic quantum mechanics with a particular emphasis on two spatial dimensions, which are of interest in the context of graphene. This chapter is not meant to be a complete review of relativistic quantum mechanics of quantum field theory, and we refer the interested reader to more specialised books on this fascinating field of physics [24, 25, 26]. We have, furthermore, tried to be as concrete as possible and to avoid formal aspects related to the Lorentz invariance of relativistic space-time, which may be useful for high-energy physicists but which are less relevant for condensed-matter physicists.

After retracing the historical path, which led to the elaboration of relativistic quantum mechanics and quantum field theory in the late 20ies of the last century, we concentrate on the particular form of the Dirac equation in 2D (i.e. two space and one time dimensions) and its amazing physical consequences such as the localisation problem, the resulting jittery motion ("Zitterbewegung"), and Klein tunneling.

3.1   Relativistic Wave Equations

In order to obtain a relativistic wave equation, one may proceed in the same manner as in the derivation of the Schrödinger equation for non-relativistic particles from classical Hamiltonian mechanics. Remember that the classical Hamiltonian of a particle in a potential $V(\mathbf{r})$ is a function

$$H(\mathbf{p},\mathbf{r}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}), \qquad (3.1)$$

defined in phase space, the space spanned by the particle's position $\mathbf{r}$ and its canonical momentum $\mathbf{p}$. In quantum mechanics, the position and the momentum no longer commute, and one reinterprets Eq. (3.1) in terms of operators acting on vectors of a Hilbert space, when one replaces

$$H \to i\hbar\partial_t \qquad \text{and} \qquad \mathbf{p} \to -i\hbar\nabla \ , \qquad (3.2)$$

where ∇ is the gradient with respect to the position $\mathbf{r}$, and ∂_t is the time derivative. This procedure yields the well-known Schrödinger equation

$$i\hbar\partial_t\psi(\mathbf{r},t) = \left[-\frac{\hbar}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r},t) \qquad (3.3)$$

of basic quantum mechanics, where the wave function $\psi(\mathbf{r},t)$ is the spatial representation of the Hilbert vector.

One may indeed apply this procedure to a free relativistic particle of mass m the energy dispersion of which is given by

$$E = \sqrt{m^2c^4 + \mathbf{p}^2c^2} \ , \qquad (3.4)$$

in terms of the velocity of light c. Whereas the taylor expansion for $|\mathbf{p}| \ll mc$ yields the non-relativistic dispersion $mc^2 + \mathbf{p}^2/2m$, apart from the mass energy mc^2, the ultra-relativistic limit of vanishing mass yields $E = c|\mathbf{p}|$. This indicates that we need to treat, in relativistic (quantum) mechanics the energy E and the momentum $\mathbf{p}$ on the same footing, in contrast to the non-relativistic limit, where the energy is proportional to the *square* of the momentum. In relativistic mechanics, one treats indeed the energy, due to the Lorentz invariance, as the time-component of a $D + 1$ dimensional momentum vector, $p^\mu \equiv (E/c, \mathbf{p})$, or its "covariant" vector $p_\mu = (E/c, -\mathbf{p})$, to account for the particular metric of the $D+1$-dimensional space-time. The
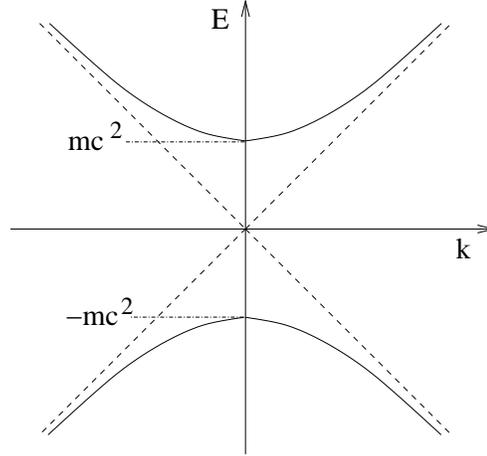
Figure 3.1: Dispersion relation of a relativistic particle. The dispersion consists of two energy (positive and negative) branches, which are separated by twice the particle's mass energy mc^2.

energy dispersion is obtained from the Lorentz-invariant norm of this vector, $\sum_\mu p^\mu p_\mu \equiv E^2/c^2 - \mathbf{p}^2 = m^2c^2$, given in terms of the mass of the particle, which is an invariant scalar with respect to the Lorentz transformations in space-time. A slightly more detailed discussion of these aspects, concerning the covariance of special relativity, may be found in Sec. 3.2.2.

3.1.1 Relativistic Schrödinger/Klein-Gordon equation

Evidently, it is more useful, in order to avoid complicated functions of differential operators, to introduce the "quantum" substitution (3.2) not directly in Eq. (3.4), but in its square. This substitution yields the equation

$$-\hbar^2\partial_t^2\psi(\mathbf{r},t) = \left(m^2c^4 - \hbar^2c^2\nabla^2\right)\psi(\mathbf{r},t), \qquad (3.5)$$

which has been derived independently by Schrödinger, Klein and Gordon in 1926/27 and which is called relativistic Schrödinger or else Klein-Gordon equation. The latter name is more common, and we follow this convention here.

The Klein-Gordon equation raises, however, more puzzling questions for the physicists of that time than it was supposed to solve. A first consequence of the Klein-Gordon equation is that it possesses solutions of *negative* energy,

as may be seen from the plane-wave ansatz

$$\psi(\mathbf{r}, t) = u\, e^{-i(Et/\hbar - \mathbf{k}\cdot\mathbf{r})},$$

which, substituted in Eq. (3.5), yields $E^2 u = (m^2 c^4 + \hbar^2 c^2 \mathbf{k}^2)u$ and, thus, the *two* solutions

$$E^\pm = \pm\sqrt{m^2 c^4 + \hbar^2 c^2 \mathbf{k}^2}, \qquad (3.6)$$

instead of a single positive one. Both solutions are plotted in Fig. 3.1. Notice that negative-energy states are not problematic in a classical theory because of the gap $2mc^2$ between the two energy branches – negative and positive energy particles would simply not communicate. However, such states are a severe problem for a quantum theory, where the ground state, alias the lowest energy state, may not be correctly defined for an energy spectrum, which has no lower bound. A particle may, in principle, "fall" from a positive-energy state to one of negative energy by emitting, e.g., a photon the energy of which corresponds to the energy difference between the initial and the final states. We will discuss this problem in more detail in Sec. 3.3.

Another problem of the Klein-Gordon equation (3.5) is related to its second-order time derivative. One would, in principle, need to specify not only $\psi(\mathbf{r}, t = t_0)$, i.e. the wave function at some reference time t_0, but also its time derivative. As a consequence, the probability density $\rho(\mathbf{r}, t)$, which appears in the continuity equation

$$\partial_t \rho(\mathbf{r}, t) + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0, \qquad (3.7)$$

is no longer necessarily positive, in contradiction to a physicists intuition of a probability density. In order to illustrate this point, we consider first the non-relativistic Schrödinger equation (3.3). If we multiply Eq. (3.3) by $\psi(\mathbf{r}, t)^*$ and substract the result from the complex conjugate of Eq. (3.3) times $\psi(\mathbf{r}, t)$, we obtain

$$i\hbar \left[\psi(\mathbf{r}, t)^* \partial_t \psi(\mathbf{r}, t) \;+\; \psi(\mathbf{r}, t) \partial_t \psi(\mathbf{r}, t)^* \right]$$
$$= -\frac{\hbar^2}{2m} \left[\psi(\mathbf{r}, t)^* \nabla^2 \psi(\mathbf{r}, t) - \psi(\mathbf{r}, t) \nabla^2 \psi(\mathbf{r}, t)^* \right]$$

or else

$$\partial_t \left[\psi(\mathbf{r}, t)^* \psi(\mathbf{r}, t) \right] + \nabla \cdot \left[\psi(\mathbf{r}, t)^* \frac{\hbar}{2im} \nabla \psi(\mathbf{r}, t) - \psi(\mathbf{r}, t) \frac{\hbar}{2im} \nabla \psi(\mathbf{r}, t)^* \right] = 0.$$
$$(3.8)$$

Comparison with the continuity equation (3.7) yields the probability density

$$\rho(\mathbf{r}, t) = \psi(\mathbf{r}, t)^* \psi(\mathbf{r}, t), \tag{3.9}$$

which corresponds to the probability of finding the quantum particle at the position $\mathbf{r}$ at time t, and the probability current density

$$\mathbf{j}(\mathbf{r}, t) = \frac{\hbar}{2im} \left[\psi(\mathbf{r}, t)^* \nabla \psi(\mathbf{r}, t) - \psi(\mathbf{r}, t) \nabla \psi(\mathbf{r}, t)^* \right] . \tag{3.10}$$

If we proceed in the same manner for a quantum particle described in terms of the Klein-Gordon equation (3.5), one obtains the equation

$$\begin{aligned}
\left[\psi(\mathbf{r}, t)^* \partial_t^2 \psi(\mathbf{r}, t) \right. &\quad - \quad \psi(\mathbf{r}, t) \partial_t^2 \psi(\mathbf{r}, t)^* \right] \\
&\quad - \quad c^2 \left[\psi(\mathbf{r}, t)^* \nabla^2 \psi(\mathbf{r}, t) - \psi(\mathbf{r}, t) \nabla^2 \psi(\mathbf{r}, t)^* \right] = 0,
\end{aligned}$$

and the corresponding probability and probability current densities are, thus, after multiplication of the above equation by $i\hbar/2mc^2$,

$$\rho(\mathbf{r}, t) = \frac{i\hbar}{2mc^2} \left[\psi(\mathbf{r}, t)^* \partial_t \psi(\mathbf{r}, t) - \psi(\mathbf{r}, t) \partial_t \psi(\mathbf{r}, t)^* \right] \tag{3.11}$$

and

$$\mathbf{j}(\mathbf{r}, t) = \frac{\hbar}{2im} \left[\psi(\mathbf{r}, t)^* \nabla \psi(\mathbf{r}, t) - \psi(\mathbf{r}, t) \nabla \psi(\mathbf{r}, t)^* \right] , \tag{3.12}$$

respectively. Whereas Eq. (3.12) coincides with the probability current density (3.10) of the Schrödinger case, the probability density (3.11) is clearly different from Eq. (3.9). In contrast to our intuition and non-relativistic quantum mechanics, this probability density may become negative, and it is, therefore, no longer possible to interpret it as the probability of finding a (relativistic) quantum particle at the position $\mathbf{r}$ at time t.

3.1.2 Dirac equation

The above-mentioned problems of the Klein-Gordon equation led Dirac, in 1928, to propose an alternative relativistic wave equation [27]. Whereas its solutions yield positive probability densities, which are more straight-forward to interpret physically, it did not solve the problem of negative energy states. The latter have finally found a compelling interpretation by Dirac in terms of *anti-particles*, as we will discuss later in this chapter.

In order to avoid negative probability densities, Dirac understood that one had to search for an equation that is linear in E, i.e. of first order in the time derivative ∂_t. For dimensional reasons, the only possible ansatz for such an equation is

$$H = \beta mc^2 + c\boldsymbol{\alpha} \cdot \mathbf{p} \qquad \rightarrow \qquad i\hbar\left[\partial_t + c\boldsymbol{\alpha} \cdot \nabla\right]\psi(\mathbf{r}, t) = \beta mc^2\psi(\mathbf{r}, t), \quad (3.13)$$

where β and[1] $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_D)$ are dimensionless objects the mathematical properties of which are fixed by the requirement that the square of the operator equation (3.13) must satisfy the relativistic dispersion relation

$$E^2 = m^2c^4 + c^2\mathbf{p}^2 = \beta^2 m^2c^4 + c^2(\boldsymbol{\alpha} \cdot \mathbf{p})^2 + mc^3(\beta\boldsymbol{\alpha} \cdot \mathbf{p} + \boldsymbol{\alpha} \cdot \mathbf{p}\beta) \ .$$

In order to satisfy this equation, the introduced mathematical objects must *anti-commute*,

$$\beta^2 = 1, \qquad \{\alpha_i, \beta\} = 0, \qquad \text{and} \qquad \{\alpha_i, \alpha_j\} = 2\delta_{ij}, \qquad (3.14)$$

in terms of the anti-commutator $\{A, B\} \equiv AB + BA$. Eq. (3.14) defines the so-called *Clifford algebra*, and one immediately realises that the objects β and α_i must be $N \times N$ matrices, with $N > 1$. We summarise their main properties.

(i) Because the Hamiltonian (3.13) is Hermitian, the matrices must be Hermitian,
$$\alpha_i = \alpha_i^\dagger \qquad \text{and} \qquad \beta = \beta^\dagger.$$

(ii) Furthermore, because of $\alpha_i^2 = \beta^2 = 1$ [see Eq. (3.14)], the matrices must be unitary
$$\alpha_i = \alpha_i^\dagger = \alpha_i^{-1}, \qquad \text{and} \qquad \beta = \beta^\dagger = \beta^{-1},$$
and their eigenvalues are, therefore, ± 1.

(iii) If one applies the determinant to the anti-commutator, $\alpha_i\alpha_j = -\alpha_j\alpha_i$, one obtains with the help of the elementary properties of the determinant,[2]
$$\det(\alpha_i\alpha_j) = \det(-\alpha_j\alpha_i) = (-1)^N \det(\alpha_j\alpha_i) = (-1)^N \det(\alpha_i\alpha_j).$$

In order to satisfy this equation, one needs to choose N to be *even*. The lowest possible value is, therefore, $N = 2$.

---

[1] One needs to introduce one α_i for each space dimension.

[2] We use the fact that if one multiplies an $N \times N$ matrix A by a scalar λ, one has $\det(\lambda A) = \lambda^N \det(A)$, and that the determinant of a product of two matrices A and B is the product of their determinants, $\det(AB) = \det(A)\det(B)$.

Influence of the space dimension on the matrix dimension N

Notice that the dimension N of the matrix must not be confunded with the space dimension D. The latter influences, however, the possible choices for N.

In the case of two space dimensions ($D = 2$), which is of interest in graphene, we have three mutually anti-commuting objects, β, α_1, and α_2. It happens that there exist, with the Pauli matrices,

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.15}$$

a set of three mutually 2×2 matrices, which satisfy the Clifford algebra (3.14). One may identify[3] $\beta = \sigma^z$, $\alpha_1 = \sigma^x$, and $\alpha_2 = \sigma^y$, which yields the 2D Dirac Hamiltonian

$$H^{2D} = c\boldsymbol{\sigma} \cdot \mathbf{p} + mc^2 \sigma^z, \tag{3.16}$$

where $\boldsymbol{\sigma} = (\sigma^x, \sigma^y)$, as in the previous chapter. If we identify c with the Fermi velocity v_F, one immediately sees that this Hamiltonian, for $m = 0$, has the same form as the effective Hamiltonian (2.36) derived in the last chapter for the description of the low-energy properties of electrons in graphene, at the K' point ($\xi = 1$). The effective Hamiltonian at the K point ($\xi = -1$) may be identified with a representation of the Dirac Hamiltonian in terms of the matrices $\beta = -\sigma^z$, $\alpha_1 = -\sigma^x$, and $\alpha_2 = -\sigma^y$, which are an equally valid choice for representing the Clifford algebra (3.14).

For three space dimensions ($D = 3$), the representation of the matrices β and α_i is slightly more involved. Whereas one may choose again the 2×2 Pauli matrices for zero-mass particles, i.e. when one may omit the matrix β,

$$H^{3D}_{m=0} = c\boldsymbol{\alpha} \cdot \mathbf{p}, \tag{3.17}$$

with e.g. $\alpha_1 = \sigma^x$, $\alpha_2 = \sigma^y$, and $\alpha_3 = \sigma^z$, this is no longer possible for massive 3D particles, as may be seen from the following argument. Let us assume that the Pauli matrices represent α_1, α_2, and β, respectively. One must, therefore, find a 2×2 matrix, α_3, which anti-commutes with the Pauli matrices. However, the latter also form, together with the 2×2 unit matrix $\mathbb{1}$, a basis of all 2×2 matrices. One may, thus, decompose

---

[3]Notice that all other permutations are equally possible.

$\alpha_3 = a_0 \mathbb{1} + a_1 \sigma^x + a_2 \sigma^y + a_3 \sigma^z$, in terms of the real numbers a_0, a_1, a_2, and a_3. The anti-commutation requirement $\{\alpha_3, \sigma^i\} = 2(a_0 + a_i) = 0$ is satisfied only if we choose $a_i = -a_0$, for all values of i. However, this is at odds with the condition $\alpha_3^2 = 1$ because one would have $\alpha_3^2 = 2a_0(2 - \sigma^x - \sigma^y - \sigma^z)$. It is, therefore, impossible to represent the matrices β and α_i by 2×2 matrices.

Due to the condition that N must be even, the lowest possible value for representing massive 3D particles is $N = 4$, and indeed one may satisfy the Clifford algebra (3.14) if one chooses, e.g., the so-called standard representation

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \qquad \text{and} \qquad \alpha_i = \begin{pmatrix} 0 & \sigma^i \\ \sigma^i & 0 \end{pmatrix}, \tag{3.18}$$

where each component is a 2×2 matrix, as may be verified in a straight-forward manner. As a consequence, the Dirac Hamiltonian has also a 4×4 matrix structure,

$$H_{m \neq 0}^{3D} = \begin{pmatrix} mc^2 \mathbb{1} & c\mathbf{p} \cdot \boldsymbol{\sigma} \\ c\mathbf{p} \cdot \boldsymbol{\sigma} & -mc^2 \mathbb{1} \end{pmatrix}, \tag{3.19}$$

and the quantum states must be represented by 4-spinors,

$$\Psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \tag{3.20}$$

in terms of the 2-spinors

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \qquad \text{and} \qquad \chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix}.$$

In order to obtain a better understanding of this spinor form, we may investigate the most simple case of a particle in its rest frame, i.e. with $\mathbf{p} = 0$, in which case the Hamiltonian (3.19) becomes diagonal. The 2-spinors ϕ and χ are, therefore, decoupled,

$$mc^2 \phi = E\phi \qquad \text{and} \qquad -mc^2 \chi = E\chi.$$

One immediately notices that the components ϕ correspond to components of positive energy (particle components) and χ to those of negative energy (anti-particle components). One may furthermore show that the components ϕ_1, ϕ_2 and χ_1, χ_2 of the 2-spinors ϕ and χ, respectively, correspond to the different possible spin states of the particle [24, 25].

Continuity equation

We finally notice that the Dirac equation yields positive probability densities, which were Dirac's main motivation to propose an alternative to the Klein-Gordon equation, as mentioned above. In order to obtain the continuity equation (3.7), we proceed in the same manner as for the Schrödinger or the Klein-Gordon equation, i.e. we multiply the differential equation (3.13) from the left by the spinor $\psi^\dagger(\mathbf{r}, t)$ and subtract the Hermitian conjugate of Eq. (3.13), multiplied from the right by $\psi(\mathbf{r}, t)$. This yields the continuity equation

$$\partial_t \left[\psi^\dagger(\mathbf{r}, t)\psi(\mathbf{r}, t) \right] + c\nabla \cdot \left[\psi^\dagger(\mathbf{r}, t)\boldsymbol{\alpha}\psi(\mathbf{r}, t) \right] \ ,$$

and one thus obtains the positive probability density

$$\rho(\mathbf{r}, t) = \psi^\dagger(\mathbf{r}, t)\psi(\mathbf{r}, t)$$

and for the probability current density

$$\mathbf{j}(\mathbf{r}, t) = c\psi^\dagger(\mathbf{r}, t)\boldsymbol{\alpha}\psi(\mathbf{r}, t).$$

3.2 The 2D Dirac Equation

After this short excursion to the 3D Dirac equation, we now concentrate on the 2D equation (3.16),

$$H^{2D} = \mathbf{p} \cdot \boldsymbol{\sigma} + m\sigma^z,$$

which interests us in the case of graphene, where we choose a unit system with $c = \hbar = 1$, for the remainder of this section. We may rewrite the Hamiltonian as

$$H^{2D} = \epsilon \begin{pmatrix} \cos\beta & \sin\beta e^{-i\varphi_\mathbf{p}} \\ \sin\beta e^{i\varphi_\mathbf{p}} & -\cos\beta \end{pmatrix}, \tag{3.21}$$

where we have defined $\epsilon = \sqrt{m^2 + \mathbf{p}^2}$, $\cos\beta = m/\epsilon$, $\sin\beta = |\mathbf{p}|/\epsilon$, and $\varphi_\mathbf{p} = \arctan(p_y/p_x)$, as in Eq. (2.38) for the wave vector $\mathbf{q}$, which is replaced by the momentum $\mathbf{p}$. The eigenvalues of the unitary and Hermitian matrix on the r.h.s. of Eq. (3.21) are $\lambda = \pm$ and correspond to the positve and negative energy states, respectively, and we have

$$E^\lambda = \lambda\epsilon = \lambda\sqrt{m^2 + \mathbf{p}^2}, \tag{3.22}$$

in agreement with Eq. (3.6).

3.2.1 Eigenstates of the 2D Dirac Hamiltonian

In order to obtain the eigenstates, we make the plane-wave ansatz for a 2-spinor,

$$\Psi_{\lambda,\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{2}} \begin{pmatrix} u_\lambda \\ v_\lambda \end{pmatrix} e^{i\mathbf{p}\cdot\mathbf{r}}, \tag{3.23}$$

and one obtains from its substitution in Eq. (3.21)

$$(\cos\beta - \lambda)u_\lambda + \sin\beta e^{-i\varphi_\mathbf{p}}v_\lambda = 0.$$

A possible choice for the 2-spinor components, which respects the normalisation, is

$$u_\lambda = \sqrt{1 + \lambda\cos\beta} \qquad \text{and} \qquad v_\lambda = \lambda\sqrt{1 - \lambda\cos\beta}e^{i\varphi_\mathbf{p}},$$

which yields for the eigenstates

$$\Psi_{\lambda,\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1 + \frac{\lambda m}{\sqrt{m^2+\mathbf{p}^2}}} \\ \lambda\sqrt{1 - \frac{\lambda m}{\sqrt{m^2+\mathbf{p}^2}}}\, e^{i\varphi_\mathbf{p}} \end{pmatrix} e^{i\mathbf{p}\cdot\mathbf{r}}. \tag{3.24}$$

One notices that the ultra-relativistic limit $(m \to 0)$ yields the same states,

$$\Psi_{\lambda,\mathbf{p}}(\mathbf{r}) \to \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \lambda e^{i\varphi_\mathbf{p}} \end{pmatrix} e^{i\mathbf{p}\cdot\mathbf{r}} \tag{3.25}$$

as those of low-energy electrons in graphene (2.37) if we consider one Dirac point K or K'. These low-energy electrons may, therefore, indeed be viewed as massless 2D relativistic particles. In the opposite limit of large mass $(m \gg |\mathbf{p}|)$, one obtains, as in the 3D case, two decoupled equations for the components, with

$$\Psi_{\lambda=+,\mathbf{p}}(\mathbf{r}) \;\rightarrow\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{i\mathbf{p}\cdot\mathbf{r}},$$

$$\Psi_{\lambda=-,\mathbf{p}}(\mathbf{r}) \;\rightarrow\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i\mathbf{p}\cdot\mathbf{r}},$$

apart from an unimportant global phase factor in the second line. In the case of small, but non-zero, values of the momentum, one obtains a small admixture by the "small" components $v_{\lambda=+}$ and $u_{\lambda=-}$. In relativistic quantum mechanics, the components $u_{\lambda=+}$ and $v_{\lambda=-}$ are sometimes also called "large" components.
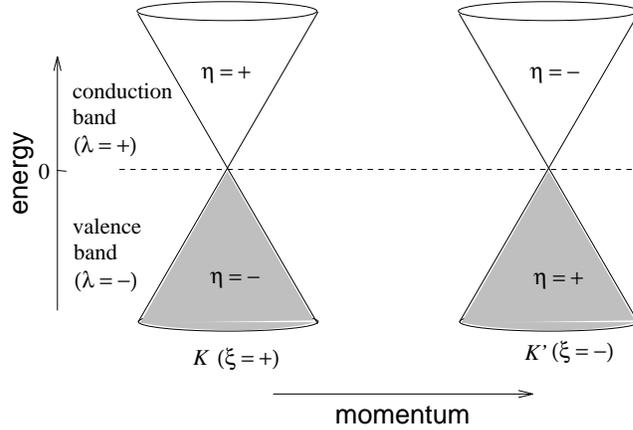
Figure 3.2: Relation between band index λ, valley isospin ξ, and chirality η in graphene.

3.2.2 Symmetries and Lorentz transformations

Chirality for zero-mass particles

In high-energy physics, one defines the helicity of a particle as the projection of its spin onto the direction of propagation,

$$h_{\mathbf{p}} = \frac{\mathbf{p} \cdot \boldsymbol{\sigma}}{|\mathbf{p}|}, \tag{3.26}$$

which is a Hermitian and unitary operator with the eigenvalues $\eta = \pm$,

$$h_{\mathbf{p}}|\eta = \pm\rangle = \pm|\eta = \pm\rangle. \tag{3.27}$$

Notice that $\boldsymbol{\sigma}$ describes, in this case, the true physical spin of the particle. In the absence of a mass term, the helicity operator commutes with the Dirac Hamiltonian, and the helicity is, therefore, a good quantum number, e.g. in the description of neutrinos, which have approximately zero mass. One finds indeed, in nature, that all neutrinos are "left-handed" ($\eta = -$), i.e. their spin is antiparallel to their momentum, whereas all anti-neutrinos are "right-handed" ($\eta = +$).

For the case of graphene, described in terms of the 2D Dirac equation (3.13), one may use the same definition (3.26) for a Hermitian and unitary operator $h_{\mathbf{p}}$. Here, the Pauli matrices no longer define the true spin, but the sublattice isospin, as described in Sec. 2.2. Because of this difference, one also calls $h_{\mathbf{p}}$ the *chirality* operator. It clearly commutes with the massless

2D Dirac Hamiltonian (3.16), and one may even express it in terms of the chirality operator

$$H^{2D} = |\mathbf{p}|h_{\mathbf{p}}, \qquad \text{and} \qquad H_{\mathbf{p}}^{\xi} = \xi H^{2D} = \xi|\mathbf{p}|h_{\mathbf{p}} \ ,$$

where the second expression takes into account the two-fold valley degeneracy, in terms of the valley isospin $\xi = \pm$. The band index λ, which describes the valence and the conduction band, is, therefore, entirely determined by the chirality and the valley isospin, and one finds

$$\lambda = \xi\eta \ , \tag{3.28}$$

which is depicted in Fig. 3.2.

Lorentz transformations of scalars and vectors

Although the covariance of the Dirac equation and Lorentz transformations in general are less relevant for the condensed-matter physicist working on graphene, we intend to give a very brief overview on these slightly more formal aspects. They allow us to define a parity and a time-reversal operator, which are the special forms of Lorentz transformations useful in the context of graphene.

In special relativity, all physical laws must be invariant under the Lorentz transformations (e.g. rotations and Lorentz "boosts") of the $D + 1$ dimensional space time. This the meaning of the statement that all physical equations must be covariant. As for the more common D-dimensional space, one may distinguish different mathematical objects in this space with respect to their behaviour under a space symmetry transformation[4] $\mathcal{R}$: a scalar remains invariant under this transformation,

$$\mathcal{R}: \qquad s \to s' = s \ ,$$

whereas a vector is transformed by the particular law

$$\mathcal{R}: \qquad \mathbf{v} \to \mathbf{v}' = R\mathbf{v},$$

where R is the rotation matrix associated with $\mathcal{R}$. One may, furthermore define a tensor (of rank r), all the r components of which are transformed by the rotation matrix

$$\mathcal{R}: \qquad T^{i_1,i_2,...,i_r} \to T'^{i'_1,i'_2...,i'_r} = R^{i'_1,i_1}R^{i'_2,i_2}...R^{i'_r,i_r}T^{i_1,i_2,...,i_r}.$$

---

[4]We consider, for simplicity a rotation in $D = 3$, which is a symmetry operation for an isotropic space.

We may proceed in the same manner in the case of the $D+1$ dimensional space-time, where a scalar is invariant under a Lorentz transformation $\mathcal{L}$,

$$\mathcal{L}: \qquad s \to s' = s,$$

whereas a $D+1$ vector transforms as[5]

$$\mathcal{L}: \qquad v^\mu \to v'^{\mu'} = \Lambda^{\mu'}{}_\mu v^\mu, \tag{3.29}$$

in terms of the $(D+1) \times (D+1)$ matrix Λ. The generalisation to a tensor is again straight-forward. In contrast to the above-mentioned D-dimensional space, one must clearly distinguish between vectors with an upper index and those with a lower one. This is due to the special metric of space-time

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

and the vectors with upper and lower index are related by the metric $v^\mu = g^{\mu\nu} v_\nu$. Thus, if $v^\mu = (v_0, \mathbf{v})$, $v_\mu = (v_0, -\mathbf{v})$, where $\mathbf{v}$ regroups the D space components and v_0 is the time component of the $D+1$ vector. One may, furthermore, define the scalar product between two $D+1$ vectors as

$$(v, w) \equiv v^\mu w_\mu \ ,$$

which must, as the name "scalar" indicates, be invariant under a Lorentz transformation. As an example to illustrate these rather formal aspects, we consider the position vector $x^\mu = (ct, \mathbf{r})$ in space-time or the momentum vector $p^\mu = (E/c, \mathbf{p})$, which we have already alluded to at the beginning of this chapter. The scalar products yield the mass term,

$$p^\mu p_\mu = E^2/c^2 - \mathbf{p}^2 = m^2 c^2,$$

which is indeed a scalar, and the scalar

$$x^\mu x_\mu = c^2 t^2 - \mathbf{r}^2 \ . \tag{3.30}$$

---

[5]We use the Einstein convention, where one sums over repeating (upper and lower) indices omitting the sum sign,

$$v^\mu v_\mu \equiv \sum_\mu v^\mu v_\mu \ .$$

Instead of a classification of all types of Lorentz transformations, which may be found in standard text books on relativity and quantum field theory, e.g. [24, 25], we simply discuss some explicit examples. The simplest example is a rotation in the space part, which does not mix time and space components. One may indeed use the D-dimensional rotation matrix R, and embed it into a $(D + 1) \times (D + 1)$ matrix,

$$(R^{D+1})^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} .$$

A less trivial example is the so-called Lorentz boost, a transformation into a frame of reference that moves at the speed v with respect to the original one, in a direction that we may choose to be the x-direction. In this case, we need to account for the invariance of the scalar product $x^{\mu} x_{\mu}$ in Eq. (3.30), and we need $c^2 t^2 - x^2 = c^2 t'^2 - x'^2$, because the y- and z-components are not affected by the transformation. The transformation, which respects this metric, may be expressed in terms of hyperbolic functions,[6]

$$\begin{aligned} ct' &= ct \cosh \zeta - x \sinh \zeta \\ x' &= -ct \sinh \zeta + x \cosh \zeta , \end{aligned}$$

which respects the metric because of $\cosh^2 \zeta - \sinh^2 \zeta = 1$. These functions of the "angle" ζ, which is also called *rapidity*, may be related to the velocity v of the Lorentz boost, by

$$\cosh \zeta = \sqrt{1 - \frac{v^2}{c^2}} \qquad \text{and} \qquad \tanh \zeta = \frac{v}{c},$$

which is a manifestation of the *time dilatation* and the *Lorentz contraction* in a Lorentz boost. In matrix form, we may thus describe the Lorentz boost in the x-direction by

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \cosh \zeta & -\sinh \zeta & 0 & 0 \\ -\sinh \zeta & \cosh \zeta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{3.31}$$

---

[6]Remember that for a metric which leaves $x^2 + y^2$ invariant, we would have chosen trigonometric functions, which define indeed the rotation.

However, Lorentz boosts play a subordinate role in the discussion of electrons in graphene, as compared for example to the parity transformation,

$$P^{\mu}_{\ \nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad (3.32)$$

which transforms the space components, $\mathbf{r} \to -\mathbf{r}$, but leaves the time component invariant, and the time-reversal transformation

$$T^{\mu}_{\ \nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (3.33)$$

which inverts the direction of time flow, $t \to -t$, but leaves invariant the space components.

Lorentz transformations of spinors

In the last paragraph, we have briefly reviewed the covariance of classical relativistic mechanics, i.e. discussed how the laws of physics, in terms of scalars, vectors, and tensors, behave under Lorentz transformations. If we consider the quantum aspect of the laws of physics, we must also analyse how the state vectors (spinors) in the Hilbert space of quantum mechanics transform in a Lorentz transformation. The spinor in a transformed frame of reference must be related to that of the original one,

$$\psi'(\mathbf{r}', t') = S(\mathcal{L})\psi(\mathbf{r}, t) \qquad \Leftrightarrow \qquad \psi'(\mathbf{r}', t') = S(\mathcal{L})\psi\left[(\Lambda^{-1})^{\mu}_{\ \nu}x^{\nu}\right], \quad (3.34)$$

where $(\Lambda^{-1})^{\mu}_{\ \nu}$ is the inverse of the matrix, which corresponds to the Lorentz transformation $\mathcal{L}$, and $S(\mathcal{L})$ is an $N \times N$ matrix that acts on the spinor.

It is, furthermore, convenient to rewrite the Dirac equation (3.13) in a covariant manner,

$$(p_{\mu}\gamma^{\mu} - mc)\,\psi(\mathbf{r}, t) = 0, \qquad (3.35)$$

where we have defined the γ matrices $\gamma^0 = \beta$ and $\gamma^i = \beta\alpha^i$. In order for this equation to be Lorentz invariant, i.e. valid in any frame of reference, we must have

$$\left(p'_{\mu}\gamma^{\mu} - mc\right)\psi'(\mathbf{r}', t') = 0,$$

which yields, with the above definitions,

$$\left[S^{-1}(\mathcal{L}) \Lambda_\mu^{\ \nu} p_\nu \gamma^\mu S(\mathcal{L}) - mc \right] \psi(\mathbf{r}, t) = 0. \tag{3.36}$$

Comparison with Eq. (3.35), thus, imposes

$$S^{-1}(\mathcal{L}) \gamma^\mu S(\mathcal{L}) = \Lambda^\mu_{\ \nu} \gamma^\nu. \tag{3.37}$$

Parity and time reversal

Eq. (3.37) allows us to construct the matrices $S(\mathcal{L})$ for the different Lorentz transformations. As already mentioned above, we are not interested, in the case of graphene, in those corresponding to Lorentz boosts. Instead, we concentrate on the parity and time-reversal operations for the 2D Dirac equation. Notice that, in this case, the γ matrices read

$$\gamma^0 = \sigma^z \qquad \text{and} \qquad \gamma^i = \sigma^z \sigma^i, \tag{3.38}$$

for $i = x, y$.

For the parity operator, Eqs. (3.33) and (3.37) impose

$$S^{-1}(P) \sigma^z S(P) = \sigma^z \qquad \text{and} \qquad S^{-1}(P) \sigma^z \sigma^i S(P) = -\sigma^z \sigma^i.$$

One notices that the choice[7]

$$S(P) = \sigma^z \mathcal{P} \tag{3.39}$$

satisfies this equation, where $\mathcal{P}$ replaces $\mathbf{r}$ by $-\mathbf{r}$ in the argument of the spinor. Notice that the eigenstates of the parity operator are only eigenstates of the 2D Dirac Hamiltonian in the particle's rest frame for $m \neq 0$ and $\mathbf{p} = 0$. In their rest frame, particles and anti-particles have, thus, opposite parity.

The definition of the time-reversal operator is slightly more involved. Eqs. (3.34) and (3.37) impose

$$S^{-1}(PT) \sigma^z S(T) = -\sigma^z \qquad \text{and} \qquad S^{-1}(T) \sigma^z \sigma^i S(T) = \sigma^z \sigma^i.$$

In this case, one needs to accompany the 2×2 transformation matrix σ^y by the conjugation operation $C : \ z \to z^*$, and one finds

$$S(T) = \sigma^y C \mathcal{T}, \tag{3.40}$$

---

[7]An equally valid choice is σ^z multiplied by a phase factor $\exp(i\phi)$.

where $\mathcal{T}$ replaces t by $-t$ in the spinor argument, which only plays a role in non-stationary spinors. We may investigate the effect of $S(T)$ on the state (3.25) $\Psi_{\lambda,\mathbf{k}}^{K}(\mathbf{r})$ at the K point in graphene,

$$S(T)\Psi_{\lambda,\mathbf{k}}^{K}(\mathbf{r}) = \frac{-i}{\sqrt{2}} \begin{pmatrix} \lambda e^{i\varphi_{-\mathbf{k}}} \\ 1 \end{pmatrix} e^{-i\mathbf{k}\cdot\mathbf{r}} \propto \Psi_{\lambda,-\mathbf{k}}^{K'}(\mathbf{r}),$$

where we have used $\exp(i\varphi_{-\mathbf{k}}) = \exp[i(\pi - \varphi_{\mathbf{k}})] = -\exp(i\varphi_{\mathbf{k}})$. One, thus, notices that the time reversal operation maps a state at the K point at wave vector $\mathbf{k}$ onto a state at K' at wave vector $-\mathbf{k}$.[8] It is indeed a general statement that time reversal yields $\mathbf{k} \to -\mathbf{k}$ and that time-reversal symmetry in a crystal, therefore, leads to a band dispersion with $\varepsilon_{-\mathbf{k}} = \varepsilon_{\mathbf{k}}$.

3.3 Physical Consequences of the Dirac Equation

Limits of the validity of the Dirac equation if interpreted as a one-particle quantum theory

3.3.1 Minimal length for the localisation of a relativistic particle

3.3.2 Velocity operator and "Zitterbewegung"

3.3.3 Klein tunneling and the absence of backscattering

---

[8]Notice that comparison with Eq. (2.37) requires the interchange of the sublattice components of the spinor at the K' point, as mentioned in Sec. 2.2.

Bibliography

[1] K. S. Novoselov, A. K. Geim, S. V. Morosov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature **438**, 197 (2005).

[2] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, Nature **438**, 201 (2005).

[3] K. S. Novoselov, A. K. Geim, S. V. Morosov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306**, 666 (2004).

[4] nano transistor IBM

[5] A. Kekulé, Bulletin de la Societe Chimique de Paris **3** (2), 98 (1865); Annalen der Chemie und Pharmazie **137** (2), 129 (1866).

[6] L. Pauling, *The Nature of Chemical Bonds*, Cornell UP (1960).

[7] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, Nature **318**, 162 (1985).

[8] Original paper in Japanese, translated in: E. Ozawa, H. W. Kroto, P. W. Fowler, E. Wassermann, Phil. Trans. R. Soc. (London) **A 343**, 1 (1993).

[9] S. Iijima, Nature **354**, 56 (1991).

[10] M. Monthioux and V. L. Kuznetsov, Carbon **44**, 1621 (2006).

[11] L. V. Radushkevich and V. M. Lukyanovich, Zurn. Fisic. Chim. **26**, 88 (1952).

[12] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes*, Imperial College Press (1998).

[13] K. S. Novoselov, D. Jiang, T. Booth, V. V. Khotkevich, S. M. Morozov, and A. K. Geim, PNAS **102**, 10451 (2005).

[14] C. Berger, Z. M. Zong, T. B. Li, X. B. Li, A. Y. Ogbazghi, R. Feng, Z. T. Dai, A. N. Marchenkov, E. H. Conrad, P. N. First, and W. A. de Heer, J. Phys. Chem. B, **108**, 19912 (2004); *for a review on the fabrication technique see:* W. A. de Heer, C. Berger, X. Wu, P. N. First, E. H. Conrad, X. B. Li, T. B. Li, M. Sprinkle, J. Hass, M. Sadowski, M. Potemski, and G. Martinez, Solid State Comm. **143**, 92 (2007).

[15] F. Varchon, R. Feng, J. Hass, X. Li, B. Ngoc Nguyen, C. Naud, P. Mallet, J.-Y. Veuillen, C. Berger, E. H. Conrad, and L. Magaud, Phys. Rev. Lett. **99**, 126805 (2007)

[16] J. Hass, F. Varchon, J. E. Millan-Otoya, M. Sprinkle, W. A. de Heer, C. Berger, P. N. First, L. Magaud, and E. H. Conrad, arXiv:0706.2134.

[17] P. R. Wallace, Phys. Rev. **71**, 622 (1947).

[18] C. Bena and G. Montambaux, arXiv:0712.0765.

[19] P. Dietl, F. Piéchon, and G. Montambaux, arXiv:0707.0219; M.O. Goerbig, J.-N. Fuchs, G. Montambaux, and F. Piéchon, arXiv:0803.0912.

[20] J. P. Hobson and W. A. Nierenberger, Phys. Rev. **89**, 662 (1953).

[21] A. Damascelli, Physica Scripta **T109**, 61 (2004).

[22] A. Bostwick, T. Ohta, Th. Seyller, K. Horn, and E. Rotenberg, Nature Physics **3**, 36 (2007).

[23] S. Y. Zhou, G.-H. Gweon, J. Graf, A. V. Fedorov, C. D. Spataru, R. D. Diehl, K. Kopelevich, D.-H. Lee, S. G. Louie, A. Lanzara, Nature Phys. **2**, 595-599 (2006).

[24] C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, McGraw-Hill (1985).

[25] S. Weinberg, *The Quantum Theory of Fields I*, Cambridge UP (1995).

[26] B. Thaller, *The Dirac Equation*, Springer-Verlag (1992).

[27] P. A. M. Dirac, Proc. Roy. Soc. (London) **A 117**, 610 (1928); *ibid.* **A 118**, 351 (1928).