

Four-dimensional graphene and chiral fermions

Michael Creutz

Physics Department,

Brookhaven National Laboratory

Upton, NY 11973, USA

(Dated: December 2007)

Abstract

Motivated by the description of the graphene electronic structure in terms of the relativistic Dirac equation, a generalization to four dimensions yields a strictly local fermion action describing two species and possessing an exact chiral symmetry. This is the minimum number of species required by the well known “no-go” theorems.

PACS numbers: 11.15.Ha, 11.30.Rd, 71.10.Fd, 71.20.-b

The structure of graphene, a single layer of graphite consisting of a hexagonal lattice of carbon atoms, has attracted considerable attention recently both from the experimental front and the fact that the low electronic excitations are described by the Dirac equation for massless fermions [1]. From the point of view of a particle physicist, this structure has two particularly striking features. First, the massless structures are robust for topological reasons related to chiral symmetry. Second, it achieves this symmetry in a manner that involves the minimum number of effective massless fermions required by the famous “no-go theorems” for lattice chiral symmetry [2, 3].

Given the importance of chiral symmetry in particle physics and the difficulties with implementing it with a lattice regularization [4], it is natural to ask whether these properties of the graphene electronic structure can be extended to four dimensions. Indeed, this is possible, and provides a remarkable fermion action with an exact chiral symmetry and manifesting two species of massless states, the minimal number consistent with chiral symmetry. This action is strictly local and thus will be vastly faster in simulations than either the overlap operator or domain wall fermions, the only other known ways to have chiral symmetry with only two flavors.

Recently a chiral gauge theory structure on two dimensional graphene has been proposed [5]. Given that we do not yet have a lattice regularization of the standard model, it would be particularly interesting if this construction could be extended to the four dimensional lattices presented here.

Although well known, it is useful to briefly review the standard two-dimensional graphene band structure. We will closely parallel this derivation for the four-dimensional case. Ignored here are all but the pi orbitals in a tight binding approximation. Our electrons hop from neighbor to neighbor around a fixed underlying hexagonal lattice. A fortuitous choice of coordinates makes the problem straightforward to solve. As sketched in Fig. (1), orient a graphene surface with one third of the bonds horizontal, one third sloping up at 60 degrees, and one third sloping down. It is then convenient to collapse the atoms at the opposite ends of each horizontal bond together and call this unit a lattice “site,” as enclosed in ellipses in the figure. For each site, let a^\dagger denote the creation operator for an electron on the left atom, and correspondingly let b^\dagger create an electron on the right atom. The commutation relations are the usual

$$[a_{x_1, x_2}, a_{x'_1, x'_2}^\dagger]_+ = [b_{x_1, x_2}, b_{x'_1, x'_2}^\dagger]_+ = \delta_{x_1, x'_1} \delta_{x_2, x'_2} \quad (1)$$

with the a type operators anti-commuting with the b 's. Finally, it is useful to label the sites using a non-orthogonal coordinate system with axes x_1 sloping up at 30 degrees intersecting the corresponding sites, and similarly x_2 sloping down at 30 degrees. All of this is illustrated in Fig. (1).

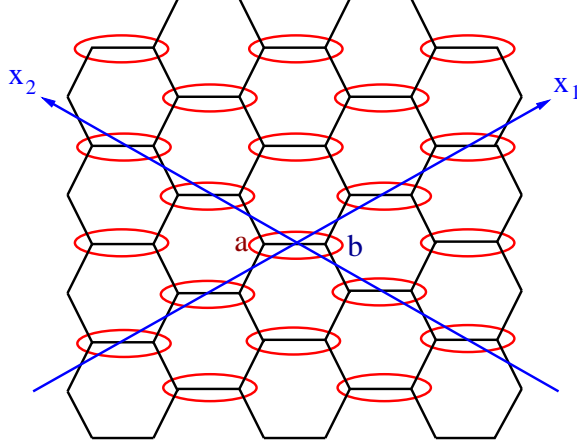


FIG. 1: Organize the graphene structure into two-atom “sites” involving horizontal bonds as shown by ellipses. Call the left hand atom of each site type a and the right hand atom type b . The coordinates of the sites are labeled along the non-orthogonal x_1 and x_2 axes.

With these conventions, the Hamiltonian of interest involves only nearest neighbor hoppings between a and b type sites

$$\begin{aligned}
H = K \sum_{x_1, x_2} & a_{x_1, x_2}^\dagger b_{x_1, x_2} + b_{x_1, x_2}^\dagger a_{x_1, x_2} \\
& + a_{x_1+1, x_2}^\dagger b_{x_1, x_2} + b_{x_1-1, x_2}^\dagger a_{x_1, x_2} \\
& + a_{x_1, x_2-1}^\dagger b_{x_1, x_2} + b_{x_1, x_2+1}^\dagger a_{x_1, x_2}.
\end{aligned} \tag{2}$$

Here K is the basic hopping parameter. The phase of K is a convention; here I consider positive real hopping. To diagonalize this Hamiltonian go to momentum space

$$a_{x_1, x_2} = \int_{-\pi}^{\pi} \frac{dp_1}{2\pi} \frac{dp_2}{2\pi} e^{ip_1 x_1} e^{ip_2 x_2} \tilde{a}_{p_1, p_2}. \tag{3}$$

This brings the Hamiltonian to the simple form

$$\begin{aligned}
H = K \int_{-\pi}^{\pi} \frac{dp_1}{2\pi} \frac{dp_2}{2\pi} & \tilde{a}_{p_1, p_2}^\dagger \tilde{b}_{p_1, p_2} (1 + e^{-ip_1} + e^{+ip_2}) \\
& + \tilde{b}_{p_1, p_2}^\dagger \tilde{a}_{p_1, p_2} (1 + e^{+ip_1} + e^{-ip_2}).
\end{aligned} \tag{4}$$

The problem reduces to diagonalizing a two by two matrix of form

$$H(p_1, p_2) = K \begin{pmatrix} 0 & z \\ z^* & 0 \end{pmatrix} \tag{5}$$

with

$$z = 1 + e^{-ip_1} + e^{+ip_2}. \tag{6}$$

The energy eigenvalues are

$$E(p_1, p_2) = \pm K |z|. \tag{7}$$

From Eq. (6) it is easy to see that the energy vanishes only at two points, $p_1 = p_2 = \pm 2\pi/3$.

The robustness of these isolated points can be seen by considering contours of constant energy. These are closed curves of constant $|z|$ in p_1, p_2 space. The important point is that for such a contour near one of the zero energy solutions, the phase of z wraps non-trivially around a circle. This non-contractable mapping indicates that on reducing the energy and shrinking the curve to a point, the magnitude of the energy at this point must vanish. This is the mechanism that prevents a band gap from opening in the spectrum.

This robustness is associated with a chiral symmetry. Because the hoppings are always between a and b type sites, we can change the sign of the energy by taking $b \rightarrow -b$. This is equivalent to the statement that σ_3 anti-commutes with the Hamiltonian. For the four-dimensional generalization, this will become the anti-commutation of γ_5 with the Dirac operator.

We wish to extend this formalism to the four-dimensional case. We want an operator D to insert into the Euclidian path integral via the fermion action $\bar{\psi}D\psi$. For low energy excitations this operator should reduce to two massless Dirac fermions, and this reduction should be robust due to a chiral symmetry. At the outset impose γ_5 hermiticity $\gamma_5 D \gamma_5 = D^\dagger$. Using this we can construct a hermitian ‘‘Hamiltonian’’

$$H = \gamma_5 D \tag{8}$$

with which we will parallel the two dimensional discussion. It is important to remember that this is not the Hamiltonian of the three dimensional quantum system, but a convenient operator for leading us back to D . In four-dimensional space, the analog of the curves of constant ‘‘energy’’ are three dimensional manifolds. To maintain a topological argument in analogy to the two dimensional case, we want to consider the situation where these surfaces wrap non-trivially around a three sphere, an S_3 . For this purpose it is quite natural to maintain the form of Eq. (5), but extend z to two by two matrices in a quaternionic space. That is, take

$$z = a_0 + i\vec{a} \cdot \vec{\sigma} \tag{9}$$

with a_μ a real four vector and $\vec{\sigma}$ denotes the traditional Pauli matrices. Then vanishing energy states will correspond to a_μ vanishing as a four vector. The goal is to construct our Hamiltonian so that that constant energy surfaces that wrap around zero energy points within the Brillouin zone will involve a non-trivial mapping in the quaternionic space. Because of the periodicity of the Brillouin zone, these zero energy points must appear in pairs so that the overall wrapping will vanish. Indeed, this is the famous no-go theorem [2, 3].

We want a construction giving precisely one and only one such pair. We also want to involve only local couplings, i.e. with only simple trigonometric functions of the momenta appearing in the dispersion relations. Because of the robustness of the zeros, if such a construction exists, it is clearly not unique. To find one such solution, start with a regular four-dimensional lattice and perform a Fourier transform. Now there will be four momentum variables p_1, p_2, p_3, p_4 , all ranging from $-\pi$ to π . A convenient form to explore is

$$\begin{aligned}
z = -12C &+ e^{ip_1\sigma_x} + e^{ip_2\sigma_x} + e^{-ip_3\sigma_x} + e^{-ip_4\sigma_x} \\
&+ e^{ip_1\sigma_y} + e^{-ip_2\sigma_y} + e^{-ip_3\sigma_y} + e^{ip_4\sigma_y} \\
&+ e^{ip_1\sigma_z} + e^{-ip_2\sigma_z} + e^{ip_3\sigma_z} + e^{-ip_4\sigma_z}.
\end{aligned} \tag{10}$$

Here C is a parameter whose value will be discussed later. The various exponentials can be expanded using $e^{ip\sigma_i} = \cos(p) + i\sigma_i \sin(p)$. For zero energy states we need z to vanish. This gives four equations corresponding to the coefficients of 1 and each σ_i vanishing. From the coefficients of the Pauli matrices we have

$$\begin{aligned}
\sin(p_1) + \sin(p_2) - \sin(p_3) - \sin(p_4) &= 0 \\
\sin(p_1) - \sin(p_2) - \sin(p_3) + \sin(p_4) &= 0 \\
\sin(p_1) - \sin(p_2) + \sin(p_3) - \sin(p_4) &= 0.
\end{aligned} \tag{11}$$

These three relations imply the sines of all the momentum components must be equal. Picking p_1 arbitrarily, each other p_μ must either equal p_1 or $\pi - p_1$. Now turning to the constant part of z , we have

$$\cos(p_1) + \cos(p_2) + \cos(p_3) + \cos(p_4) = 4C. \tag{12}$$

Since the cosine function is bounded by unity, we clearly must take $C < 1$ to have any solutions. To resolve the $p_i \leftrightarrow \pi - p_i$ ambiguity it is convenient to ask that $\cos(p_i)$ be positive. Imposing the constraint $C > 1/2$ ensures this. We will later discuss some interesting specific choices for C .

So with these constraints on the constant C there are exactly two zeros of energy in the Brillouin zone. These occur when all components of p are equal and satisfy $\cos(p) = C$. The two solutions differ in the sign of p . Picking the positive sign for convenience, it is useful to expand about the zero

$$p_\mu = \tilde{p} + q_\mu \tag{13}$$

with $\cos(\tilde{p}) = C$ and $\tilde{p} > 0$. Defining $S = \sin(\tilde{p}) = \sqrt{1 - C^2}$, we have

$$\begin{aligned}
\cos(p_\mu) &= C \cos(q_\mu) - S \sin(q_\mu) = C - Sq_\mu + O(q^2) \\
\sin(p_\mu) &= S \cos(q_\mu) + C \sin(q_\mu) = S + Cq_\mu + O(q^2).
\end{aligned} \tag{14}$$

Inserting all this into our quaternion

$$\begin{aligned}
z = & -3S(q_1 + q_2 + q_3 + q_4) \\
& + iC\sigma_x(q_1 + q_2 - q_3 - q_4) \\
& + iC\sigma_y(q_1 - q_2 - q_3 + q_4) \\
& + iC\sigma_z(q_1 - q_2 + q_3 - q_4) + O(q^2).
\end{aligned} \tag{15}$$

At this point we introduce a convention for the Dirac gamma matrices

$$\begin{aligned}
\vec{\gamma} &= \sigma_x \otimes \vec{\sigma} \\
\gamma_4 &= -\sigma_y \otimes 1 \\
\gamma_5 &= \sigma_z \otimes 1 = \gamma_1 \gamma_2 \gamma_3 \gamma_4
\end{aligned} \tag{16}$$

The direct product notation here is defined so that γ_5 is diagonal with -1 in the last two places.

With these conventions our Euclidean Dirac operator takes the form

$$\begin{aligned}
D = & C(q_1 + q_2 - q_3 - q_4)i\gamma_1 \\
& + C(q_1 - q_2 - q_3 + q_4)i\gamma_2 \\
& + C(q_1 - q_2 + q_3 - q_4)i\gamma_3 \\
& + 3S(q_1 + q_2 + q_3 + q_4)i\gamma_4 + O(q^2).
\end{aligned} \tag{17}$$

This reproduces the desired massless Dirac equation if we identify new momenta

$$\begin{aligned}
k_1 &= C(q_1 + q_2 - q_3 - q_4) \\
k_2 &= C(q_1 - q_2 - q_3 + q_4) \\
k_3 &= C(q_1 - q_2 + q_3 - q_4) \\
k_4 &= 3S(q_1 + q_2 + q_3 + q_4).
\end{aligned} \tag{18}$$

Proper Lorentz invariance requires symmetry between the k 's. This implies that the original lattice, as generated by translations using the q 's, will in general be distorted from simple hyper-cubic.

Inverting the above relation enables us to study the physical angles between the original lattice directions

$$\begin{aligned}
q_1 &= \frac{+k_1+k_2+k_3}{4C} + \frac{k_4}{12S} \\
q_2 &= \frac{+k_1-k_2-k_3}{4C} + \frac{k_4}{12S} \\
q_3 &= \frac{-k_1-k_2+k_3}{4C} + \frac{k_4}{12S} \\
q_4 &= \frac{-k_1+k_2-k_3}{4C} + \frac{k_4}{12S}.
\end{aligned} \tag{19}$$

Since the k coordinates should be orthogonal, we can compute the angle between two of the q axes

$$\frac{q_i \cdot q_j}{|q|^2} = \frac{1 - 10S^2}{1 + 26S^2}. \tag{20}$$

The original axes can be made orthogonal if $S^2 = 1/10$. This corresponds to $C = 3\sqrt{10}/10 = 0.94868\dots$, consistent with the constraints. With this choice of coordinates, gauging the theory is straightforward. Because the starting links are orthogonal, we can use the usual Wilson gauge action with group elements on links interacting with the simple plaquette action.

Another especially interesting choice for C gives a closer analogy with graphene. Imagine the fermion fields on one site to actually be spread along a new bond connecting two atoms in the k_4 direction, similar to the construction for graphene indicated in Fig. (1). This structure would be particularly symmetric if the angles between all five bonds attached to an atom were equal. This occurs when

$$C = \frac{3\sqrt{14}}{14} = 0.80178\dots, \quad (21)$$

which is still consistent with the constraints.

This lattice has an appealing intuitive geometric interpretation in terms of bonds along one direction, analogous to the horizontal bonds in Fig. (1), and then splitting at an atom into four bonds going off symmetrically into four-space to join other horizontal bonds displaced in the various directions. While in two dimensional graphene each carbon is coupled symmetrically to three neighbors, here each atom is directly coupled to 5 others. The entire lattice is then built up of hexagonal chairs with an inter-bond angle of $104.4775\dots$ degrees. The five fold symmetry associated with this particular choice of C should help to reduce lattice artifacts. Note that the diamond lattice in three dimensions represents an intermediate case, where one bond splits into three giving each atom a tetrahedral environment.

This choice of C also clarifies the appropriate gauging procedure. On spreading the k_4 bonds, the plaquettes of the starting theory become hexagonal “chairs.” To keep all directions equivalent, the action should now include terms from all possible six link chairs, including those involving only the four original directions. Without these terms the speed of light associated with gluons or fermions will not be equal. The lattice symmetry requires all such terms to have the same gauge coupling. Although to make the analogy complete one could introduce additional link matrices in the gauge group for the k_4 direction, this is not necessary since these can be thought of as being gauge fixed to unity.

From the standpoint of computational efficiency, it does not appear to matter much what value of C is chosen, within the constraints. Near the limits of $C = 1$ and $C = 1/2$ one should expect lattice artifacts to increase. The hyper-cubic value $C = 3\sqrt{10}/10$ is presumably closest to conventional lattice gauge ideas and is simplest to gauge, with only a traditional plaquette action

required. On the other hand, the value $C = 3\sqrt{14}/14$ may have smaller lattice artifacts due to its high symmetry. Other values of C will require adjusting the strength of six link terms in the action to maintain the same speed of light for the fermions and the gluons.

Returning from reciprocal to position space, the fermionic action involves several terms. First from the collapsed k_4 bond there is a site diagonal term $12iC\bar{\psi}\gamma_4\psi$. Then for a forward step in the various directions we pick up a factor of the hopping parameter K multiplied by different combinations of gamma matrices, as listed here:

$$\begin{aligned}
\text{for a hop in direction 1: } & +\gamma_1 + \gamma_2 + \gamma_3 - 3i\gamma_4 \\
\text{for a hop in direction 2: } & +\gamma_1 - \gamma_2 - \gamma_3 - 3i\gamma_4 \\
\text{for a hop in direction 3: } & -\gamma_1 - \gamma_2 + \gamma_3 - 3i\gamma_4 \\
\text{for a hop in direction 4: } & -\gamma_1 + \gamma_2 - \gamma_3 - 3i\gamma_4
\end{aligned} \tag{22}$$

Keeping the operator D antihermitean, the reverse hops involve minus the conjugate of these factors. Note the factor of i in front of γ_4 which is absent for the γ_{1-3} terms. This twisting of the phase gives rise to the required factors of sin or cos in Eqs. (11) and Eqs. (12). This action is only marginally more complicated than that of naive fermions; so, it should be easy to insert into simulations.

Chiral symmetry is manifested in the exact anti-commutation of γ_5 with D . This is actually a flavored chiral symmetry since the expansion about the negative solution for p flips the sign of the gamma matrix associated with k_0 . Note that as with naive fermions, D is purely anti-hermitean. The chiral symmetry can easily be broken with the addition of a term proportional to γ_5 to H that splits the degeneracy of the type a and type b sites. This gives each physical fermion a common mass. Because the chiral symmetry remains exact on gauging, an additive renormalization for the physical fermion mass is forbidden, unlike with Wilson fermions. The further addition of a Wilson type mass term would enable splitting the degeneracy of the two species. Note that if we wish to extend this formalism to more flavors, the no-go theorem restricts us to an even number of species. The best we can do for three flavors is to start with four and, using a chiral symmetry breaking operator, make one of the flavors heavier. Such a term will unfortunately introduce the possibility of an additive mass renormalization for the third flavor.

It is perhaps interesting to compare this formalism with staggered fermions [6–8]. The latter also have an exact chiral symmetry and the appearance of several degenerate flavors (sometimes called tastes to distinguish them from independent lattice fields). In both formalisms the chiral symmetry is flavored, with the different species rotating differently under a chiral rotation. Also,

in both cases mixing between the flavors will result in the zero modes associated with topologically non-trivial gauge fields to no longer be exact at finite lattice spacing [9, 10]. However the usual staggered approach has the doublers appearing in multiples of four, and thus is further from the situation desired for the physical case with only the light up and down quarks. Staggered fermions do have only one component of the field per site, yielding potentially faster simulations at the expense of an intricate non-local formalism in constructing composite fields.

Acknowledgements

This manuscript has been authored under contract number DE-AC02-98CH10886 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

-
- [1] For a recent review see A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, arXiv:0709.1163v1 [cond-mat.other].
 - [2] H. B. Nielsen and M. Ninomiya, Nucl. Phys. B **185**, 20 (1981) [Erratum-ibid. B **195**, 541 (1982)].
 - [3] F. Wilczek, Phys. Rev. Lett. **59**, 2397 (1987).
 - [4] M. Creutz, Rev. Mod. Phys. **73**, 119 (2001) [arXiv:hep-lat/0007032].
 - [5] R. Jackiw and S. Y. Pi, Phys. Rev. Lett. **98**, 266402 (2007) [arXiv:cond-mat/0701760].
 - [6] J. B. Kogut and L. Susskind, Phys. Rev. D **11**, 395 (1975).
 - [7] L. Susskind, Phys. Rev. D **16**, 3031 (1977).
 - [8] H. S. Sharatchandra, H. J. Thun and P. Weisz, Nucl. Phys. B **192**, 205 (1981).
 - [9] L. H. Karsten and J. Smit, Nucl. Phys. B **183**, 103 (1981).
 - [10] J. Smit and J. C. Vink, Nucl. Phys. B **286**, 485 (1987).