Non-Abelian Gauge Potentials in Graphene Bilayers

P. San-Jose, J. González, and F. Guinea

DOI: 10.1103/PhysRevLett.108.216802
Non-Abelian gauge potentials in graphene bilayers

P. San-Jose\textsuperscript{1}, J. González\textsuperscript{1} and F. Guinea\textsuperscript{2}
\textsuperscript{1}Instituto de Estructura de la Materia (IEM-CSIC), Serrano 123, 28006 Madrid, Spain
\textsuperscript{2}Instituto de Ciencia de Materiales de Madrid (ICMM-CSIC), Cantoblanco, 28049 Madrid, Spain

We study the effect of spatial modulations in the interlayer hopping of graphene bilayers, such as those that arise upon shearing or twisting. We show that their single-particle physics, characterized by charge accumulation and recurrent formation of zero-energy bands as the pattern period $L$ increases, is governed by a non-Abelian gauge potential arising in the low-energy electronic theory due to the coupling between layers. We show that such gauge-type couplings give rise to a potential that, for certain discrete values of $L$, spatially confines states at zero energy in particular regions of the Moiré patterns. We also draw the connection between the recurrence of the flat zero-energy bands and the non-Abelian character of the potential.

Introduction.— The discovery of graphene, the material made of a one-atom-thick carbon layer, has provided the realization of a system where the electrons have conical valence and conduction bands, therefore behaving as massless Dirac fermions [1–3]. A remarkable feature of graphene is that deformations of its honeycomb lattice may produce a similar effect to that of gauge potentials in the low-energy Dirac theory [4]. Recently, it has been shown that the local in-plane deformations induced by strain can be mimicked by an effective vector potential, which may give rise to the analogue of Landau levels in the deformed graphene sheet [5].

In this paper we show that the effect of modulations in the interlayer hopping of graphene bilayers can be represented in general by a non-Abelian background gauge potential in the low-energy electronic theory, and that it is responsible for the zero energy charge density waves and the dispersionless minibands, predicted by theory, and recently measured [6]. The vector components of the potential take values in the space of SU(2) matrices, which correspond to rotations in the Hilbert space of the two layers. This kind of non-Abelian gauge fields [7] is relatively rare in a condensed-matter context [8–10], but it is quite relevant in subatomic physics, being responsible for the interaction between matter fields. The proton and the neutron, for instance, compose an isospin SU(2) doublet. It was proposed long ago that an ideal experiment of scattering of these particles onto a non-Abelian flux line should lead to the transfer of protons into neutrons and vice versa [11]. In general, matter fields pick up a matrix-valued ‘phase’ in their propagation in a non-Abelian gauge field. Interference of such matrix-valued phase along two indistinguishable paths (as opposed to the conventional U(1) phase) leads to an intriguing non-Abelian generalization of the Aharonov-Bohm effect. In our context, this would manifest as coherent layer polarization induced by the interference of two SU(2) phases acquired along the two paths.

Experimental realizations of non-Abelian gauge potentials have been proposed before in the study of ultracold atoms [12, 13]. Investigations have addressed in particular the influence of the non-Abelian gauge potentials in the development of the Landau levels produced by a conventional magnetic field [14, 15]. However, the question of whether pure non-Abelian gauge fields may lead to a phenomenology similar to the magnetic localization of Landau states remains open. Our investigation sheds light on this question, showing that it is possible to develop a zero-energy level of spatially confined states as a consequence of the non-Abelian gauge potential, provided that the fermion fields return to the original internal state around a closed path.

The effective non-Abelian gauge potentials that arise in the bilayers have actually a genuine applied interest, since they induce periodic spatial confinement of electronic states. Indeed, we will see that the one-dimensional (1D) modulation of the interlayer tunneling leads to the confinement of electronic states into narrow 1D channels. We will also extend our approach to the description of twisted bilayers [16–24], where the non-Abelian gauge potential turns out to confine low energy electrons into a triangular array of quantum dots. The problem of confinement of electronic states has particular relevance, given that scalar potential barriers are not effective to constrain the propagation of the electrons in graphene [25], which makes the non-Abelian gauge potentials proposed in this paper an interesting alternative to the confinement and manipulation of electronic states in graphene devices.

Model.— The simplest realizations of a non-Abelian gauge potential are found by means of a modulated mismatch in the relative position of the two lattices of a bilayer, obtained either by applying strain or shear in one of the layers or by relative rotation between the two layers. In both instances, the resulting mismatch produces characteristic Moiré patterns, see Fig. 1, which reflect the spatial alternation between AA’-type stacking (perfect alignment of the atoms in the two layers) and AB’-type, BA’-type (Bernal) stacking, where $A'$ and $B'$ correspond to the two sublattices of the lower (upper) lattice.

At energies $\varepsilon \lesssim 1$ eV, the Moiré electron system is described by Dirac fermions on each layer, coupled by a position dependent interlayer hopping amplitude. The
the alternation between confined between the gauge potential energy state concentrated around $V_{AA'}$, together with a typical zero-energy state confined between the $AB'$ and $BA'$ regions, and a finite energy state concentrated around $AA'$.

Hamiltonian takes the form \[ H = v_F \left( \begin{pmatrix} 0 & \Pi_\perp & \hat{V}_{AA'}(r) & \hat{V}_{AB'}(r) \\ \Pi_\perp & 0 & \hat{V}_{BA'}(r) & \hat{V}_{AA'}(r) \\ \hat{V}_{AA'}(r) & \hat{V}_{BA'}(r) & 0 & \Pi_\perp \\ \hat{V}_{AB'}(r) & \hat{V}_{AA'}(r) & \Pi_\perp & 0 \end{pmatrix} \right) \] (1)

where $\Pi_\pm \equiv -i\partial_x + \partial_y \mp (\hat{A}_x + i\hat{A}_y)$. The spatially modulated interlayer coupling functions $\hat{V}$ arise from the Moiré pattern formation, and the intra-layer Abelian gauge field $\pm \hat{A}$ describes the strains in each layer. These strains lead to constant gauge fields in our case, $\hat{A} = \Delta K/2$. Note that, as discussed later, this model also describes twisted bilayers, in which the interlayer Dirac cone shift $\Delta K$ arises due to the relative twist between layers, not strains. Since $\hat{A}$ is anyhow uniform, we can gauge it away by a transformation $U = \exp(i/2)\tau_3 \Delta K \cdot r$, where $\tau_3$ is a Pauli matrix which operates on the layer index. This transforms consequently the interlayer couplings into $\hat{V}_{ij}(r) = V_{ij}(r)e^{-i\Delta K \cdot r}$.

**Low-energy theory of sheared bilayers.** — We consider first the instance in which shear $u_{xy}$ is applied along the $AB$ bonds of a given layer section ($y$ direction). Then a 1D Moiré pattern is produced in the orthogonal $x$ direction, smoothly alternating between $AA'$, $BA'$ and $AB'$ stacking as shown in Fig. 1(a). The corresponding hopping amplitudes are related by $\hat{V}_{AA'}(x) = \hat{V}_{BA'}(x - L/3) = \hat{V}_{AB'}(x + L/3)$, where $\hat{V}_{AA'}(x) \approx (w/v_F) [1 + 2 \cos(2\pi x/L)]$ using a single-harmonic approximation [16] (the interlayer coupling is $w \approx t_{\perp}/3 = 0.11$ eV, where $t_{\perp}$ is the interlayer hopping).

To assist in interpreting the role of the different interlayer couplings, we define the functions $A_x(x) = -(\hat{V}_{AB'}(x) + \hat{V}_{BA'}(x))/2$ and $A_y(x) = (\hat{V}_{AB'}(x) - \hat{V}_{BA'}(x))/2$. Then $\hat{V}_{AB} = -A_x + A_y$, $\hat{V}_{BA'} = -A_x - A_y$, and it becomes clear that $A_x, A_y$ act as off-diagonal vector potentials. Taking Pauli matrices $\sigma$ in the $AB$ pseudospin space and $\tau$ in the space of the two layers, we may recast Eq. (1) into

\[ H = v_F \sigma \cdot (-i \partial - \hat{A}) + v_F \tilde{V}_{AA'} \tau_1 \] (2)

where we have introduced the gauge potential $\hat{A} = (A_x \tau_1, A_y \tau_2)$, which induces a precession of the layer index as an electron moves in real space. This $\hat{A}$ is non-Abelian, since $[\hat{A}(r), \hat{A}(r')] \neq 0$ in general (see also Ref. 28). This formulation highlights the different nature of the $\tilde{V}_{AA'}$ coupling, that acts rather like a scalar potential (proportional to the unit matrix $\sigma_0$).

This electron system has the characteristic property of developing flat bands of spatially confined states at large $L$, whose formation is fully controlled by the effect of the gauge potential $\hat{A}$. Computing the energy levels of the Hamiltonian (2), one observes that at large $L$ the system develops two increasingly narrow subbands around zero energy of states confined between $AB'$ and $BA'$ regions, see Fig. 2. Their energy, for any given momentum $k_x$ and $|k_y| < 3w/v_F$, oscillates towards zero, crossing it periodically as $L$ increases (e.g. whenever $wL/2\pi v_F$ is an integer if $k_y = 0$, see inset on the right panel of Fig. 2). Additionally, a second pair of flat bands appear at a finite energy, corresponding to states confined around $AA'$. All these bands become $AA'$-confined and linearly dispersive in $k_y$ for $|k_y| < 3w/v_F$, although they remain non-dispersive in the $x$ direction. These features are strongly reminiscent of the Landau-level to snake-state transition in carbon nanotubes of large radius in a real perpendicular magnetic field [29], which have also an effectively modulated magnetic flux.

This confinement phenomenon may be understood from the effect of a confining potential created purely by the gauge field $\hat{A}$. The equation for the eigenstates $\Psi$ of $H$ can be expressed after squaring the Hamiltonian (and disregarding for simplicity the scalar potential at this point) as

\[ (-\partial^2 + i\partial \cdot \hat{A} + 2i\hat{A} \cdot \partial + A_x^2 + A_y^2 - \sigma_z \tilde{F}_{xy}) \Psi = (\epsilon/v_F)^2 \Psi \] (3)

where the field strength is conventionally defined in terms of the matrix-valued potential $\hat{A}_\mu \sim \tilde{F}_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu - i[\hat{A}_\mu, \hat{A}_\nu]$. Given the invariance of $H$ under the combined operation of charge conjugation and parity, the eigenstates can be chosen in the form $\Psi(r) = [\phi_A(-r), \phi_A(r), \phi_B^*(-r), \phi_B(r)]$, for some $\phi_{1,2}$. In the limit of zero transverse momentum $k_y$ the combinations $\phi_{\pm}(r) \equiv \phi_A(r) \pm \phi_B^*(-r)$ decouple, and the above equation translates, at large $L$, into

\[ -v_F^2 \phi_{\pm}''(x) = -V_{eff}(x) \phi_{\pm}(x) + O\left(\frac{v_F}{wL}\right) \] (4)

with $V_{eff}(x) \equiv -(\pm \varepsilon + A_x + A_y)(\pm \varepsilon + A_x - A_y)$ [30]. This is the wave equation of a scalar mode with eigenvalue $E = 0$.
under the influence of an \(\varepsilon\)-dependent confining potential \(V_{\text{eff}}^B(x)\), sketched in Fig. 1. \(\varepsilon = 0\) eigenstates centered around \(AB\) and \(BA\) regions will arise whenever a level of such potential crosses \(E = 0\). Such states will be peaked exactly at \(AB\) and \(BA\), since the well has \(E = 0\) turning points at said regions. Moreover, a discrete set of \(E = 0\) eigenstates centered around the \(AA\) local minimum will arise at energy \(\varepsilon \sim 1/\sqrt{L}\). These two types of states are apparent in the numerical bandstructure plotted on the right panel of Fig. 2.

The above analysis in terms of \(V_{\text{eff}}\) relies crucially on the non-Abelian character of the gauge potential, \([A_x, A_y] \neq 0\). Without this property, the recurrence of zero-energy states as \(L\) increases would not appear. This may be appreciated from an alternative point of view. In order for a (normalizable) zero-energy state to exist, the operator \(W_{\varepsilon=0}\) relating the wavefunction at \(x = 0\) and \(x = L\), \([\phi_1(L), \phi_2(L)] = W_{\varepsilon=0}[\phi_1(0), \phi_2(0)]\), must have at least one eigenvalue of modulus one. Since at zero energy Eq. (2) leads to

\[-i\partial_x \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = (ik_y + A_x\tau_1 - iA_y\tau_2) \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix},\]

we have for \(k_y = 0\)

\[W_{\varepsilon=0} = \text{Pexp} \left\{ i \int_0^L dx \left[ A_x(x)\tau_1 - iA_y(x)\tau_2 \right] \right\},\]

where “Pexp” denotes the path-ordered product of exponentials of differential line elements.[31] One can check that this operator becomes unitary when \(wL/v_F = 2\pi n\) for integer \(n\). This is the condition for the existence of normalizable zero-energy modes, in agreement with the numerical results.

Low energy description of twisted bilayers.— At energies below 1 eV, a twisted bilayer may be accurately modeled by Hamiltonian (1), where the shift \(\Delta K\) in the relative position of the Dirac points in each layer comes as a consequence of the rotation by the twist angle \(\theta\). If we take the original position of the \(K\) points as \(K = (4\pi/3a_0, 0)\), the shift in each layer is given by \(\pm\Delta K/2 = (0, \pm K\sin(\theta/2))\). On the other hand, \(\theta\) also fixes the size of the Moiré pattern unit cell, which grows as \(\theta\) decreases. More precisely, the Bravais superlattice formed by the Moiré pattern has primitive vectors \(L_\pm = L(\sqrt{3}/2, \pm 1/2)\), where \(L = a_0/2\sin(\theta/2)\).

The interlayer coupling may be written in terms of a single periodic profile \(V(r) = V(r + L_+) = V(r + L_-)\), in such a way that if we fix \(V_{AA}(r) = V(r)\), then \(V_{AB}(r) = V(r + (L_+ + L_-)/3)\) and \(V_{BA}(r) = V(r - (L_+ + L_-)/3)\). A common procedure is to assume that the interlayer hopping is dominated by processes with momentum-transfer \(Q_0 = 0\) or equal to the reciprocal vectors \(Q_{1,2} = (\pm 2\pi/\sqrt{3}, 2\pi)\)/\(L\) [16, 20], so that \(V(r) = (w/v_F)\sum_j \text{exp}(i\mathbf{q}_j \cdot r)\). Coupling \(V\) is complex in this case, however, but we can still carry out the procedure of the preceding section by defining \(A_{1z} = -\text{Re}(V_{AB} - V_{BA})/2, A_{2z} = \text{Im}(V_{AB} - V_{BA})/2, A_{1y} = \text{Im}(V_{AB} - V_{BA})/2\) and \(A_{2y} = \text{Re}(V_{AB} - V_{BA})/2\). We can then write the Hamiltonian for the twisted bilayer as

\[H = v_F\mathbf{\sigma} \cdot (-i\partial - \tau_3\Delta K/2 - \mathbf{A}) + v_F\mathbf{\Phi},\]

with non-Abelian potentials \(\mathbf{A} = (A_{1z}\tau_1 + A_{2z}\tau_2, A_{1y}\tau_1 + A_{2y}\tau_2)\) and \(\mathbf{\Phi} = \text{Re}(V_{AB}/\tau_1 - \text{Im}(V_{AA})\tau_2)\) [32].

The mismatch \(\Delta K\) of the Fermi points may be removed by carrying out a gauge transformation on the spinors, \(\Psi = \text{exp}(i/2\tau_3\Delta K\mathbf{\cdot r})\mathbf{\Phi}\), at the expense of introducing new potentials \(\tilde{V}_{ij}(r) = V_{ij}(r)e^{-i\mathbf{A}\mathbf{\cdot r}}\). We finally get a modified expansion \(\tilde{V}(r) = (w/v_F)\sum_j \text{exp}(i\mathbf{q}_j \cdot r)\), with a star of three vectors \(\mathbf{q}_i\). Note that \(|\tilde{V}(r)| = |V(r)|\) has conical singularities at the center of \(AB'/BA'\) regions.

Two representative bandstructures obtained numerically from the Hamiltonian (5) for different values of \(\theta\) are plotted in Fig. 3. The first corresponds to an index \(n = 20\) and exhibits a lowest subband with vanishing energy at the two Dirac points originating from the graphene layers. As the angle \(\theta\) is decreased, the energy scale of the lowest subband is significantly lowered, until it becomes remarkably flat for values of \(n\) around \(n = 31\) \((\theta \approx 1^\circ)\), exhibiting zero Fermi velocity at the \(K\) point and a bandwidth that is more than 100 times smaller than the scale of the next subband. (Note however that this is not a topological zero mode in the sense of a standard zero Landau level [33], since the Atiyah-Singer index
states show a strong confinement in the regions with stacking, as shown in Fig. 3, which is confirmed by atomistic tight binding calculations [18]. This confinement of the Moiré pattern [20]. These will not only give rise to Abelian gauge potentials in the SU(2) flavor space. Unlike for that first instance, higher values of \( n \) giving rise to a flat lowest-energy subband do depend on the strength of the \( V_{AA'} \) coupling [35]. However, the essential spatial confinement properties of the corresponding lowest-energy eigenstates do not. They remain confined around \( AA' \) stacking. They also acquire higher angular momentum components and become increasingly ring-shaped for higher values of \( n \) (see Fig. 3), as expected for the excited states of a 2D potential well centered around \( AA' \) stacking.

Experimental measures of the low-energy electronic properties of twisted bilayers have been reported in particular in Ref. 6. It has been found that, at a certain value \( \theta \sim 1^\circ \), the renormalized Fermi velocity near the \( K \) point of the twisted bilayer becomes so small that the picture based on Dirac quasiparticles breaks down. This comes together with the observation of a clear pattern of spatial confinement in the local density of states, which adopts the form of a triangular charge density wave following the modulation of the Moiré pattern. These features are fully consistent with the confinement of the low-energy eigenstates in the regions of \( AA' \) stacking due to the action of the gauge potential, which provides a strong confinement mechanism according to the preceding discussion. This single-particle mechanism will cooperate with the additional many-body effects that may also contribute to the modulation of the charge in the system.

**Conclusion.**—We have shown that the Moiré-like modulation of the interlayer hopping in graphene bilayers leads to a very rich phenomenology, which can be described in terms of effective non-Abelian gauge potentials in the low-energy electronic theory. We have shown that any additional terms arising from the stacking modulation, such as non-Abelian scalar potentials, do not qualitatively modify the low energy electronic structure. In the case of sheared bilayers with quasi-1D Moiré patterns, the gauge potential is equivalent to a confining potential that leads to low-energy charge accumulation along 1D strips. The effect of the non-Abelian gauge potential in rotationally faulted bilayers is also to develop a characteristic spatial pattern of confinement, and the formation of dispersionless bands for discrete value of the Moiré periods. We conclude these two effects are the characteristic signature of Moiré-induced non-Abelian gauge potentials in graphene bilayers.

The emergence of these types of gauge fields is generic to systems of coupled Dirac equations, and the analysis presented here can be extended to multilayered systems with SU(N) gauge groups. One may also furthermore envision the possibility of tuning the non-Abelian fields caused by stacking by applying generic strain fields to Moiré bilayers. These will not only give rise to Abelian fields as in monolayers, but also to small modifications
of the stacking non-Abelian fields [36], whose interplay is known to produce a rich phenomenology [15].

We acknowledge financial support from MICINN (Spain) through grants FIS2008-00124 and CON- SOLIDER CSD2007-00010. This research was supported in part by the National Science Foundation under Grant No. PHY11-25915.

[8] G. Moore and N. Read, Nuclear Physics B 360, 362 (1991), note that the non-Abelian fields in this reference differ from those in our work in that they are fluctuating quantum fields.
[27] The absence of a mass term in the diagonal is justified by the symmetry between the two sub-lattices within each layer, which is preserved for smooth Moiré patterns.
[30] In Eq. (4), we have neglected corrections of the order $-i\partial_x \phi'(-x) = (\varepsilon - W_{BA}(x))\phi(x)$.
[31] Incidentally, the transfer operator $W$ takes the form of an open Wilson loop, due to the first order character of the Dirac equation.
[32] This procedure shows that it is always possible to trade the complex couplings $V_{AA}, V_{AB}, V_{BA}$ by the potentials $A, B, \Phi$, which encode 6 independent real functions taking into account the different $\tau_1, \tau_2$ projections.
[35] To be precise, while a finite value of $V_{AA}$ does not destroy subsequent instances of flat-band formation, it does lower their corresponding $L$ as compared to the case $V_{AA} = 0$, for which they satisfy the simple relation $wL/v_F \approx 2\pi (j + \frac{1}{2})$, for integer $j$. Moreover, a finite $V_{AA}$ renders the lowest subband with a small residual bandwidth that is non-existent in the pure magnetic case.