Excitons in narrow-gap carbon nanotubes

R. R. Hartmann

School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, United Kingdom

I. A. Shelykh

Science Institute, University of Iceland, Dunhagi 3, IS-107, Reykjavik, Iceland and
International Institute of Physics, Av. Odilon Gomes de Lima, 1722, Capim Macio, CEP: 59078-400, Natal - RN, Brazil

M. E. Portnoi[*]

School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, United Kingdom and
International Institute of Physics, Av. Odilon Gomes de Lima, 1722, Capim Macio, CEP: 59078-400, Natal - RN, Brazil

We calculate the exciton binding energy in single-walled carbon nanotubes with narrow band gaps, accounting for the quasirelativistic dispersion of electrons and holes. Exact analytical solutions of the quantum relativistic two-body problem are obtained for several limiting cases. We show that the binding energy scales with the band gap, and conclude on the basis of the data available for semiconductor nanotubes that there is no transition to an excitonic insulator in quasimetallic nanotubes and that their THz applications are feasible.

I. INTRODUCTION

In the 1940s it was predicted by Wallace[1] that due to its honeycomb lattice, graphene, a single monolayer of carbon, should exhibit unusual semimetallic behavior. The gap between the valence and conductance band is exactly zero, and the low-energy excitations are massless chiral Dirac fermions.[2,3] On the other hand, the electronic band structure of other carbon-based materials shows substantial differences from the band structure of graphene. Among them are fullerenes, which can be considered as zero-dimensional carbon molecules with discrete energy spectra, and carbon nanotubes, which are obtained by rolling graphene along a given direction and reconnecting the carbon bonds.[4]

The energy spectrum of a single-walled carbon nanotube is determined by the way it is rolled[5] and is characterized by two integers (n,m) $(0 \leqslant m \leqslant n)$ denoting the relative position, $\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2$, of the pair of atoms on a graphene strip which coincide when the strip is rolled into a tube ($\mathbf{a}_1$ and $\mathbf{a}_2$ are the unit vectors of the hexagonal lattice). For most combinations of n and m the energy spectrum of the nanotube is characterized by the gap, whose value is comparable to those in semiconductor materials. However, for a third of n and m combinations, namely when $n - m = 3p$, $p = 0, 1, 2, \ldots$, the value of the gap is drastically reduced and lies in the terahertz frequency range. Moreover, for $m = n$ the gap vanishes in zero magnetic field, and opens only after the application of a magnetic field parallel to the nanotube axis.[4–6]

Optical properties of carbon nanotubes have been investigated by many authors.[7–11] It was shown that the excitonic effect plays an important role and that the properties of the excitonic resonance can be modulated by applying external fields.[12–14] However, excitons were theoretically studied mostly for semiconductor carbon nanotubes with sufficiently large gaps.[15–18] For metallic nanotubes mostly the excitons associated with high branches of the nanotube spectrum separated by an energy of about 2 eV were considered.[19,20] On the other hand, the analysis of the long-wavelength properties of narrow-band nanotubes is also of high interest, as there is a growing number of proposals using carbon nanotubes of this type for THz applications, including several schemes put forward by the authors of the present work.[21–23] The only work on excitons in narrow-gap nanotubes in which the stability of excitons in metallic carbon nanotubes subjected to Aharonov-Bohm magnetic flux was considered is Ref. 24, where it was claimed that even in this case the exciton binding energy never exceeds the gap. The main conclusions of Ref. 24 are based on numerical calculations in the $\mathbf{k} \cdot \mathbf{p}$ scheme. In our present work we give further consideration to this interesting case, using a noticeably different semianalytical approach based on a Dirac-like matrix equation.

At first glance, narrow-gap carbon nanotubes fit ideally the excitonic insulator picture[25] predicted for the case when the exciton binding energy exceeds the band gap. However, the energy spectrum of a narrow-gap carbon nanotube is quasirelativistic with the effective mass of the charge carriers proportional to the gap leading to a drastic reduction of the exciton binding energy with reducing the gap. In this paper we consider excitons formed by relativistic quasi-one-dimensional electrons and holes interacting via a model yet realistic potential. We show that the binding energy scales with the band gap and conclude on the basis of the data available for semiconductor nanotubes that there is no transition to an excitonic insulator in quasimetallic carbon nanotubes and that their THz applications are feasible.

II. FORMALISM AND INTERACTION POTENTIAL

The Hamiltonian for a single free electron in a narrow-gap carbon nanotube can be written (see Appendix A) as

$$\widehat{H}_0 = \hbar v_{\mathrm{F}} \begin{pmatrix} 0 & b\widehat{q} - i\Delta \\ b\widehat{q} + i\Delta & 0 \end{pmatrix}, \tag{1}$$

where $\hat{q} = -i\frac{\partial}{\partial x}$ is the operator of the wave vector along the nanotube axis and we use the basis $|\psi_A\rangle, |\psi_B\rangle$ with indices A, B corresponding to the carbon atoms of two different sublattices in the honeycomb lattice. Here v_{F} is the Fermi

velocity in graphene, connected to the tight-binding matrix element of electron hopping $|t| \approx 3$ eV and the graphene lattice constant a by $\hbar v_F = \frac{\sqrt{3}}{2}|t|a$, where $a = |\mathbf{a}_1| = |\mathbf{a}_2| = 2.46$ Å.[5] For the (n,n) armchair nanotube the value of the band gap $2\hbar v_F|\Delta|$ is determined by the external magnetic field, $\Delta = \frac{2}{a\sqrt{3}}\sin(\frac{\pi\Phi}{n\Phi_0})$, where Φ is the magnetic flux through the nanotube cross section, $\Phi_0 = ch/e$ is the flux quantum, and $b = \sqrt{\frac{4}{3} - \frac{1}{3}\cos^2(\frac{\pi\Phi}{n\Phi_0})}$. Since for experimentally accessible magnetic fields $\Phi/\Phi_0 \ll 1$, in our further consideration we set $b = 1$. A similar Hamiltonian can be written for a narrow-gap carbon nanotube with a gap opened by curvature[26] or for certain types of graphene nanoribbons.[27] The diagonalization of Eq. (1) gives a quasirelativistic dispersion, $E = \pm\hbar v_F\sqrt{\Delta^2 + q^2}$. To go from the case of the electrons to the case of the holes, $|\psi_i^e\rangle \rightarrow |\psi_i^h\rangle$, one should use the substitution $\hat{q} \rightarrow -\hat{q}, \Delta \rightarrow -\Delta$. For a pair of interacting electron and hole the total Hamiltonian can be written in the form of a 4×4 matrix, and the stationary Schrödinger equation for determining the binding energy written in the basis $|\Psi_{ij}\rangle = |\psi_i^e\rangle|\psi_j^h\rangle$ reads

$$
\hbar v_F \begin{pmatrix} 0 & \hat{q}_e - i\Delta & -\hat{q}_h + i\Delta & 0 \\ \hat{q}_e + i\Delta & 0 & 0 & -\hat{q}_h + i\Delta \\ -\hat{q}_h - i\Delta & 0 & 0 & \hat{q}_e - i\Delta \\ 0 & -\hat{q}_h - i\Delta & \hat{q}_e + i\Delta & 0 \end{pmatrix}
$$
$$
\times \begin{pmatrix} \Psi_{AA} \\ \Psi_{BA} \\ \Psi_{AB} \\ \Psi_{BB} \end{pmatrix} = [E - V(x_e - x_h)] \begin{pmatrix} \Psi_{AA} \\ \Psi_{BA} \\ \Psi_{AB} \\ \Psi_{BB} \end{pmatrix}, \quad (2)
$$

where the indices e and h correspond to the electrons and holes, and $\hat{q}_{e,h} = -i\frac{\partial}{\partial x_{e,h}}$. In the absence of interaction and band-filling effects this Hamiltonian yields four energy eigenvalues corresponding to a pair of noninteracting quasiparticles: $E = \hbar v_F(\pm\sqrt{\Delta^2 + q_e^2} \pm \sqrt{\Delta^2 + q_h^2})$; only the solution with positive signs should be taken if we consider a system containing a single electron and a single hole. As the potential of electron-hole interaction $V(x_e - x_h)$ depends only on the distance between the electron and hole, it is convenient to use the center of mass and relative motion coordinates, $X = (x_e + x_h)/2, x = x_e - x_h, \hat{q}_e = \hat{K}/2 + \hat{k}, \hat{q}_h = \hat{K}/2 - \hat{k}$, and represent the exciton wave functions as $\Psi_{ij}(X,x) = e^{iKX}\phi_{ij}(x)$, which permits the substitution of the operator $\hat{K}$ by the number K, having the physical meaning of the wave vector of the exciton as a whole. Considering the case of $K = 0$ corresponding to a static exciton, the equation for the wave function of relative motion reads

$$
\begin{pmatrix} 0 & \hat{k} - i\Delta & \hat{k} + i\Delta & 0 \\ \hat{k} + i\Delta & 0 & 0 & \hat{k} + i\Delta \\ \hat{k} - i\Delta & 0 & 0 & \hat{k} - i\Delta \\ 0 & \hat{k} - i\Delta & \hat{k} + i\Delta & 0 \end{pmatrix} \begin{pmatrix} \phi_{AA} \\ \phi_{BA} \\ \phi_{AB} \\ \phi_{BB} \end{pmatrix}
$$
$$
= [\varepsilon - \tilde{V}(x)] \begin{pmatrix} \phi_{AA} \\ \phi_{BA} \\ \phi_{AB} \\ \phi_{BB} \end{pmatrix}, \quad (3)
$$

where $\varepsilon = E/\hbar v_F, \tilde{V}(x) = V(x)/\hbar v_F$ and $\hat{k} = -i\partial/\partial x$. Equation (3) represents a system of first-order differential equations,

which can be reduced (see Appendix B) to a single second-order equation for ϕ_{AA}:

$$
\frac{d^2\phi_{AA}}{dx^2} + \frac{1}{\varepsilon - \tilde{V}(x)}\frac{d\tilde{V}(x)}{dx}\frac{d\phi_{AA}}{dx}
$$
$$
+ \left[\left(\frac{\varepsilon - \tilde{V}(x)}{2}\right)^2 - \Delta^2\right]\phi_{AA} = 0. \quad (4)
$$

Before solving Eq. (4), we need to specify the interaction potential. It needs to possess the following properties: first, it should remain finite as $x \rightarrow 0$ and for small x scale as $V(x) \approx -e^2/(\epsilon\sqrt{x^2 + d^2})$, where ϵ is the effective dielectric constant and d is the short-range cutoff parameter, which is of the order of the nanotube diameter; second, for large x it should decay exponentially due to the effects of screening necessarily present in nanotubes of the metallic type[19]—it is apparent that the size of the exciton should not exceed the mean separation between quasiparticles. The convenient choice of the potential is

$$
\tilde{V}(x) = -\frac{\alpha}{\cosh(\beta x)}, \quad (5)
$$

where $\alpha = e^2/(\epsilon\hbar v_F d)$ and the effective length $1/\beta$ defines the spatial extent of the interaction. The additional advantage of the potential given by Eq. (5) is that it allows one to obtain some analytical results in the context of graphene physics as was shown in Ref. 28.

III. RESULTS AND DISCUSSION

The results of the numerical solution are shown in Figs. 1 and 2. Figures 1(a) and 1(b) show the dependence of the binding energy ε_b calculated as an absolute value of the difference between the eigenenergy of Eq. (4) and the energy of the pair of noninteracting electron and hole, $\varepsilon_b = |\varepsilon - 2\Delta|$, measured in the units of β versus the effective strength of the interaction α/β for two different values of $\Delta/\beta = 1$ and $\Delta/\beta = 0.01$ corresponding to the cases of the semiconductor and narrow-gap quasimetallic nanotube, respectively. In both cases, for small values of α/β there is only one bound s-type state, whose binding energy increases with the increase of the interaction strength, until it reaches the value $\varepsilon_b = \varepsilon_g = 2\Delta$ after which it goes to the continuum of states with negative energies and thus unbinds. On the other hand, the increase of α/β leads to the appearance of higher-order solutions, corresponding to p, d, etc., excitons. This is an adequate picture for a single impuritylike exciton. However, if the interaction between an electron and a hole created across the gap separating the ground and excited states of the system exceeds the band gap in a many-body system, the ground state should be redefined and the many-body effects should govern the value of the gap. To find out whether the gap in a narrow-gap nanotube is mostly governed by single-electron effects such as an external magnetic field and curvature or by many-body effects, the value of the effective interaction strength α/β should be estimated from the known data for excitons in semiconductor nanotubes.

Note also that by reducing Eq. (4) for $\varepsilon = 0$ to the hypergeometric equation (see Appendix B) one can find the
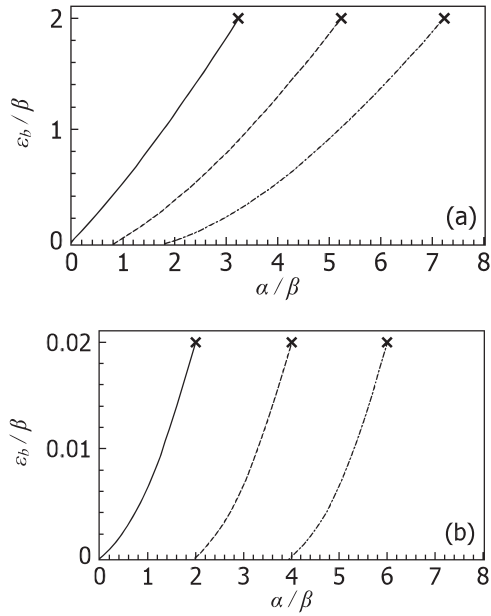
FIG. 1. The dependence of the exciton binding energy ε_b/β on the interaction strength (a) for a semiconductor nanotube with $\Delta/\beta = 1$, (b) for a quasimetallic nanotube with $\Delta/\beta = 0.01$. The different lines correspond to different excitonic states. The crosses indicate the exact analytic results for zero-energy states ($\varepsilon_b = \varepsilon_g$).

analytical expression for the values of the parameter α/β when the exciton binding energy is exactly equal to the band gap:

$$\frac{\alpha}{\beta} = 1 + 2n + \sqrt{1 + 4\frac{\Delta^2}{\beta^2}}, \qquad (6)$$

where $n = 0, 1, 2, \ldots$. These exact values are shown by the crosses in Fig. 1.

It is convenient to express the dimensionless interaction strength α/β as

$$\frac{\alpha}{\beta} = \frac{e^2}{\epsilon d \hbar v_F \beta} = \frac{c}{v_F} \frac{e^2}{\hbar c} \frac{1}{\epsilon \beta d} \approx \frac{300}{137} \frac{1}{\epsilon \beta d}. \qquad (7)$$

When the interaction strength is written in the above form, one can immediately see the direct relation of the strongly bound exciton problem to the long-standing problem of a
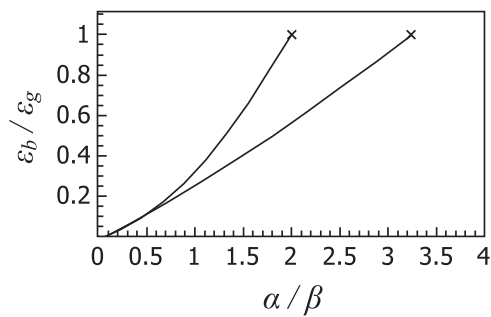

FIG. 2. The dependence of the ratio of the exciton binding energy to the nanotube band gap on the interaction strength for a quasimetallic carbon nanotube with $\Delta/\beta = 0.01$ (upper curve) and a semiconductor nanotube with $\Delta/\beta = 1$. Only s states are shown. The crosses correspond to the exact analytic results when the exciton binding energy is equal to the band gap.

supercritical charge (the nuclear charge with $Z > 137$) and atomic collapse in relativistic quantum mechanics.[29] This problem has recently been revisited and reformulated for graphene[30] where the effective impurity charge is increased by a factor of $c/v_F \approx 300$, which is also present in the numerator of Eq. (7). It follows from Eq. (6) that for an exciton in a narrow-gap carbon nanotube ($\Delta \ll \beta$) the effective supercritical charge corresponding to $\varepsilon_b > \varepsilon_g$ can be achieved for $\alpha/\beta > 2$. In principle, parameters ϵ and β can be controlled by submerging nanotubes in a solution or changing the number of electrons and holes by injection, optical excitation, or varying temperature. However, one should not expect the value of $1/(\epsilon\beta d)$ to be higher for a quasimetallic carbon nanotube than for a semiconductor nanotube of similar diameter.

Figure 2 shows the dependence of the ratio of the binding energy of the $1s$ exciton to the value of the gap on the interaction strength for the cases of the semiconductor and narrow-gap nanotubes. For very small interaction strengths the ratio $\varepsilon_b/\varepsilon_g$ is a universal quadratic function of α/β (see Appendix C). The two curves remain practically indistinguishable from one another with increasing α/β up to $\alpha/\beta \approx 0.6$. A further increase of the interaction strength or range leads to the breakdown of universality for finite range potentials.

It should be noted that for a zero-range attractive potential between the electron and hole[31] this universality holds for the whole range of interaction strength (see Appendix D). However, for the smooth interaction potential considered here the universality breaks for sufficiently large interaction strength. The difference between the case of semiconductor and quasimetallic nanotubes is not as large as one could expect. Namely, the value of α/β for which the $1s$-exciton binding energy is exactly equal to the band gap is about 3.2 for the case of a semiconductor nanotube with $\Delta/\beta = 1$ and about 2 for a quasimetallic nanotube with $\Delta/\beta = 0.01$. Also, the variational calculations[15] of the binding energy in semiconductor nanotubes supported by the experimental data gives a value for the exciton binding energy of approximately 30% of the band gap, which according to Fig. 2 corresponds to the values of $\alpha/\beta \approx 1$. Using the experimental data for the exciton binding energy for the excited subbands of metallic tubes[13] provides even smaller values for α/β. These estimations ensure that for a quasimetallic nanotube the $1s$ exciton bound state lies within the gap. The hypothetical situation $\alpha/\beta > 2$ corresponds to the case of the so-called "excitonic insulator,"[25] where the account of many-body effects becomes crucially important and goes beyond the scope of our consideration, which is essentially a two-particle approach. Certain many-body aspects of the physics of excitons in narrow-gap nanotubes and their relation to the Luttinger liquid model have been recently studied using a numerical renormalization group.[32]

Figure 3 shows the square modulus of the excitonic wave function in real space for a quasimetallic nanotube. One sees that for the $1s$ state the density in the center of mass has a local minimum, which differs strikingly from the result obtained earlier for semiconductor nanotubes in the effective-mass approximation,[10] in which the probability density in the exciton center of mass has a maximum. This is related to the complex matrix structure of the Hamiltonian (1) resulting in the multicomponent structure of the eigenfunctions. A similar
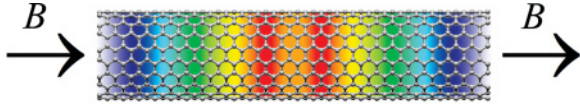
FIG. 3. (Color online) The density of the 1s exciton for a (10,10) carbon nanotube with a magnetic-field-induced gap of 10 meV (2.5 THz) corresponding to a magnetic field of 15 T along the nanotube axis. The density represents the probability of finding the electron and hole comprising the exciton at the indicated relative separation. Red and blue colors correspond to the highest and lowest values of density, respectively.

dip in the ground-state density was previously reported for graphene-based waveguides.[28]

It should be noted that taking into account the valley and spin quantum numbers increases the number of different types of excitons associated with a given carbon nanotube spectrum branch to 16.[18] Their consideration, however, can be done along the lines described above, the only difference being the modified interaction strength α/β for each type of exciton. The exciton with the highest binding energy in a semiconductor nanotube is known to be optically inactive (dark). The difference between the energy of the dark and bright excitons is proportional to the exciton binding energy and exceeds $k_B T$ at room temperature causing a significant suppression in the optical emission from semiconductor nanotubes. As it is shown above, for narrow-gap nanotubes the binding energy is drastically reduced and there should be no noticeable difference in the population of dark and bright excitonic states at any experimentally attainable temperatures. At room temperature, all dark and bright excitons in narrow-gap nanotubes should be fully ionized and the direct interband transitions[21–23] govern the emission in the terahertz range.

IV. CONCLUSIONS

In conclusion, we considered the formation of the exciton in narrow-gap carbon nanotubes characterized by the quasirelativistic spectrum of free particles. We show that the exciton binding energy scales with the band gap and vanishes with decreasing the gap even for strong electron-hole attraction. Therefore excitonic effects including strongly bound dark excitons, which explain the poor electroluminescent properties of semiconducting nanotubes, should not dominate for narrow-gap carbon nanotubes. This opens the possibility of

using quasimetallic carbon nanotubes for various terahertz applications.

APPENDIX A: ARMCHAIR CARBON NANOTUBE IN A MAGNETIC FIELD

Graphene's effective matrix Hamiltonian following the notations of Ref. 5 is given by

$$H = t \begin{pmatrix} 0 & f_k \\ f_k^\star & 0 \end{pmatrix}$$

and the corresponding eigenvalues are

$$E = \pm |t| \sqrt{|f_k|^2},$$

where

$$f_k = \exp\left(i\frac{a}{\sqrt{3}}k_x\right) + 2\exp\left(-i\frac{a}{2\sqrt{3}}k_x\right)\cos\left(k_y\frac{a}{2}\right). \quad \text{(A1)}$$

For an (n,n) armchair carbon nanotube k_x is quantized in the following manner:

$$k_x = \frac{2\pi}{a\sqrt{3}}\frac{l}{n},$$

where l is an integer. Defining k_T as the projection of the wave vector along the nanotube axis Eq. (A1) can be expressed as

$$f_k = \exp\left(i\frac{2\pi}{3}\frac{l}{n}\right) + 2\exp\left(-i\frac{\pi}{3}\frac{l}{n}\right)\cos\left(\frac{k_T a}{2}\right). \quad \text{(A2)}$$

In the presence of a magnetic field, $l \to l + \mathcal{F}$ (here $\mathcal{F} = \Phi/\Phi_0$) and the armchair carbon nanotube energy spectrum becomes

$$E = \pm |t|\sqrt{1 + 4\cos\left(\pi\frac{l+\mathcal{F}}{n}\right)\cos\left(\frac{k_T a}{2}\right) + 4\cos^2\left(\frac{k_T a}{2}\right)}. \quad \text{(A3)}$$

We are interested solely in the lowest conduction and highest valence subbands, which correspond to $l = n$. In this instance the subbands energy spectra are

$$E = \pm |t|\sqrt{1 - 4\cos\left(\pi\frac{\mathcal{F}}{n}\right)\cos\left(\frac{k_T a}{2}\right) + 4\cos^2\left(\frac{k_T a}{2}\right)}, \quad \text{(A4)}$$

where the $+(-)$ sign denotes the lowest conduction (highest valence) subband. The introduction of a magnetic field along the nanotube axis shifts the minimum in the spectrum away from $k_T = 2\pi/3a$, the electrons "acquire mass" and a band gap is opened. The new minimum $k_T^{\min}$ is obtained by differentiating Eq. (A4) and equating it to zero; this yields

$$\frac{1}{2}\cos\left(\pi\frac{\mathcal{F}}{n}\right) = \cos\left(k_T^{\min}\frac{a}{2}\right). \quad \text{(A5)}$$

Since we are interested in particle behavior in the vicinity of $k_T^{\min}$, it is natural to re-express the electron energy spectrum in terms of q_T, defined as the momentum measured relative to $k_T^{\min}$, i.e., $k_T = k_T^{\min} + q_T$, thus Eq. (A2) can be expressed by

$$f_k = \exp\left[-i\frac{\pi}{3}\left(1+\frac{\mathcal{F}}{n}\right)\right]\left\{-\exp\left(i\pi\frac{\mathcal{F}}{n}\right) + 2\cos\left[\left(k_T^{\min} + q_T\right)\frac{a}{2}\right]\right\}.$$

Using the identity Eq. (A5) f_k becomes

$$f_k = \exp\left[-i\frac{\pi}{3}\left(1+\frac{\mathcal{F}}{n}\right)\right]\left[\cos\left(\pi\frac{\mathcal{F}}{n}\right)\cos\left(\frac{aq_T}{2}\right) - 2\sqrt{1-\frac{1}{4}\cos^2\left(\pi\frac{\mathcal{F}}{n}\right)}\sin\left(\frac{aq_T}{2}\right) - \exp\left(i\pi\frac{\mathcal{F}}{n}\right)\right]. \quad \text{(A6)}$$

Expanding Eq. (A6) in terms of q_T and retaining first-order terms only yields

$$f_k = \exp\left[-i\frac{\pi}{3}\left(1+\frac{\mathcal{F}}{n}\right)\right]\left[-i\sin\left(\pi\frac{\mathcal{F}}{n}\right) - \sqrt{1-\frac{1}{4}\cos^2\left(\pi\frac{\mathcal{F}}{n}\right)}aq_T\right]. \quad \text{(A7)}$$

The effective matrix Hamiltonian can therefore be written as

$$\hat{H}_0 = t\begin{pmatrix} 0 & e^{-i\theta}\left[-i\sin\left(\pi\frac{\mathcal{F}}{n}\right) - \sqrt{1-\frac{1}{4}\cos^2\left(\pi\frac{\mathcal{F}}{n}\right)}a\hat{q}\right] \\ e^{i\theta}\left[i\sin\left(\pi\frac{\mathcal{F}}{n}\right) - \sqrt{1-\frac{1}{4}\cos^2\left(\pi\frac{\mathcal{F}}{n}\right)}a\hat{q}\right] & 0 \end{pmatrix}, \quad \text{(A8)}$$

where $\theta = \frac{\pi}{3}(1+\frac{\mathcal{F}}{n})$ and $\hat{q} = -i\frac{\partial}{\partial y}$. This 2×2 Hamiltonian acts on a two-component Dirac wave function with components χ_1 and χ_2 associated with the A and B sublattices of graphene, respectively. By changing the basis wave functions from χ_1 and χ_2 to $\psi_A = -\chi_1$ and $\psi_B = -\exp(-i\theta)\chi_2$, and changing the variable $y \to -x$, so that the x coordinate is now along the nanotube axis, the effective matrix Hamiltonian can be expressed as

$$\hat{H}_0 = |t|\begin{pmatrix} 0 & \sqrt{1-\frac{1}{4}\cos^2\left(\pi\frac{\mathcal{F}}{n}\right)}a\hat{q} - i\sin\left(\pi\frac{\mathcal{F}}{n}\right) \\ \sqrt{1-\frac{1}{4}\cos^2\left(\pi\frac{\mathcal{F}}{n}\right)}a\hat{q} + i\sin\left(\pi\frac{\mathcal{F}}{n}\right) & 0 \end{pmatrix}, \quad \text{(A9)}$$

which acts on the two-component Dirac wave function $(\psi_A, \psi_B)^{\mathrm{T}}$. For brevity let $b = \sqrt{\frac{4}{3} - \frac{1}{3}\cos^2(\frac{\pi}{n}\mathcal{F})}$ and $\Delta = \frac{2}{a\sqrt{3}}\sin(\frac{\pi}{n}\mathcal{F})$. Thus

$$\hat{H}_0 = \hbar v_{\mathrm{F}}\begin{pmatrix} 0 & b\hat{q} - i\Delta \\ b\hat{q} + i\Delta & 0 \end{pmatrix}, \quad \text{(A10)}$$

where now $\hat{q} = -i\frac{\partial}{\partial x}$.

APPENDIX B: DERIVATION OF THE ANALYTIC SOLUTION FOR $E = 0$

For the case of total momentum $K = 0$, corresponding to the static exciton, the multicomponent wave function of relative motion satisfies the matrix equation given by Eq. (3) of the main text. Since $\phi_{AA} = \phi_{BB}$, the system of equations (3) reduces to

$$\phi_{BA} = \frac{2}{[\varepsilon - \tilde{V}(x)]}(\hat{k} + i\Delta)\phi_{AA}; \quad \text{(B1)}$$

$$\phi_{AB} = \frac{2}{[\varepsilon - \tilde{V}(x)]}(\hat{k} - i\Delta)\phi_{AA}; \quad \text{(B2)}$$

$$(\hat{k} - i\Delta)\phi_{BA} + (\hat{k} + i\Delta)\phi_{AB} = [\varepsilon - \tilde{V}(x)]\phi_{AA}. \quad \text{(B3)}$$

Substituting Eqs. (B1) and (B2) into Eq. (B3) yields the following second-order differential equation:

$$\frac{d^2\phi_{AA}}{dx^2} + \frac{1}{\varepsilon - \tilde{V}(x)}\frac{d\tilde{V}(x)}{dx}\frac{d\phi_{AA}}{dx}$$
$$+ \left[\left(\frac{\varepsilon - \tilde{V}(x)}{2}\right)^2 - \Delta^2\right]\phi_{AA} = 0, \quad \text{(B4)}$$

which coincides with Eq. (4) of the main text.

Let us now consider the case of

$$\tilde{V}(x) = -\frac{\alpha}{\cosh(\beta x)}$$

and $\varepsilon = 0$. Making the change of variable $z = \beta x$ transforms Eq. (B4) to

$$\frac{d^2\phi_{AA}}{dz^2} + \tanh(z)\frac{d\phi_{AA}}{dz}$$
$$+ \left[-\frac{1}{4}\omega^2\tanh^2(z) + \frac{1}{4}\omega^2 - \tilde{\Delta}^2\right]\phi_{AA} = 0, \quad \text{(B5)}$$

where $\omega = \alpha/\beta$ and $\tilde{\Delta} = \Delta/\beta$. The change of variable $\chi = \tanh(z)$ allows Eq. (B5) to be expressed as

$$(\chi^2 - 1)^2\frac{\partial^2\phi_{AA}}{\partial\chi^2} + \chi(\chi^2 - 1)\frac{\partial\phi_{AA}}{\partial\chi} + (c\chi^2 + e)\phi_{AA} = 0, \quad \text{(B6)}$$

where $c = -\frac{1}{4}\omega^2$ and $e = \frac{1}{4}\omega^2 - \Delta^2$. Equation (B6) is of a known form and has the solutions[33]

$$\phi_{AA} = A_1(\chi + 1)^p(\chi - 1)^q\eta\left[\frac{1}{2}(\chi + 1)\right], \quad \text{(B7)}$$

where A_1 is a constant, the function $\eta[\frac{1}{2}(\chi + 1)]$ is to be found, and p and q are found from the following conditions:

$$4q(q-1) + 2q + c + e = 0;$$

$$(p-q)[2(p+q)-1] = 0. \tag{B8}$$

Equations (B6)–(B8) yield

$$(\chi^2 - 1)\frac{d^2\eta}{d\chi^2} + [(2p + 2q + 1)\chi - 2(p - q)]\frac{d\eta}{d\chi}$$
$$+ [(p+q)^2 + c]\eta = 0. \tag{B9}$$

Performing a change of variable $(\chi + 1)/2 \to \kappa$ reduces Eq. (B9) to the hypergeometric equation

$$\kappa(\kappa - 1)\frac{d^2\eta}{d\kappa^2} + [(a_1 + a_2 + 1)\kappa - a_3]\frac{d\eta}{d\kappa} + a_1 a_2 \eta = 0,$$

where $a_1 = p + q \pm \sqrt{-c}$, $a_2 = p + q \mp \sqrt{-c}$, and $a_3 = 2p + \frac{1}{2}$. Thus the form of η is

$$\eta = {}_2F_1\left(p + q - \sqrt{-c}, p + q + \sqrt{-c}; 2p + \frac{1}{2}; \kappa\right). \tag{B10}$$

Hence the wave function ϕ_{AA} is given by

$$\phi_{AA} = A_1(1 + \chi)^p(1 - \chi)^q$$
$$\times {}_2F_1\left(p + q - \frac{\omega}{2}, p + q + \frac{\omega}{2}; 2p + \frac{1}{2}; \frac{1+\chi}{2}\right). \tag{B11}$$

From Eq. (B8) q is found to be

$$q = \frac{1 \pm \sqrt{1 + 4\tilde{\Delta}^2}}{4}, \tag{B12}$$

and p can take the values

$$p = q \tag{B13}$$

or

$$p = \frac{1}{2} - q. \tag{B14}$$

Let us first consider the case of $p = q$. In this instance Eq. (B11) becomes

$$\phi_{AA} = A_1[1 - \chi^2]^q$$
$$\times {}_2F_1\left(2q - \frac{\omega}{2}, 2q + \frac{\omega}{2}; 2q + \frac{1}{2}; \frac{1+\chi}{2}\right). \tag{B15}$$

For the function ϕ_{AA} to vanish as $z \to \infty$ we require

$$q = \frac{1 \pm \sqrt{1 + 4\tilde{\Delta}^2}}{4} > 0 \tag{B16}$$

and that the hypergeometric series is terminated. This can be satisfied if we take the positive root of q and restrict ω such that $2q - \frac{\omega}{2} = -n$, where n is a positive integer, thus we arrive at the condition that

$$\omega = 1 + 2n + \sqrt{1 + 4\tilde{\Delta}^2}. \tag{B17}$$

The functions ϕ_{BA} and ϕ_{AB}, which can be found from Eqs. (B1) and (B2), must also vanish as $z \to \infty$. For simplicity let us analyze the linear combinations $\Psi_+ = \phi_{BA} + \phi_{AB}$ and $\Psi_- = \phi_{BA} - \phi_{AB}$:

$$\Psi_+ = i\frac{4\beta}{V(z)}\frac{\partial\phi_{AA}}{\partial z} = -i\frac{4}{\omega}\cosh(z)\frac{\partial\phi_{AA}}{\partial z}; \tag{B18}$$

$$\Psi_- = -i\frac{4\Delta}{V(z)}\phi_{AA} = i\frac{4\tilde{\Delta}}{\omega}\cosh(z)\phi_{AA}. \tag{B19}$$

For the functions satisfying Eqs. (B18) and (B19) to vanish in the limit of $z \to \infty$, we require

$$q = \frac{1 + \sqrt{1 + 4\tilde{\Delta}^2}}{4} > \frac{1}{2}. \tag{B20}$$

This is automatically satisfied for any finite $\tilde{\Delta}$. Therefore providing ϕ_{AA} vanishes as $z \to \infty$, ϕ_{BA} and ϕ_{AB} will also vanish.

Let us now consider the case of $p + q = \frac{1}{2}$, in this instance Eq. (B11) becomes

$$\phi_{AA} = \tilde{A}_1(1 - \chi^2)^q(1 + \chi)^{1/2-2q}$$
$$\times {}_2F_1\left(\frac{1}{2} - \frac{\omega}{2}, \frac{1}{2} + \frac{\omega}{2}; \frac{3}{2} - 2q; \frac{1+\chi}{2}\right). \tag{B21}$$

Using the identity[34]

$$z^{1-c}{}_2F_1(a + 1 - c, b + 1 - c; 2 - c; z) = {}_2F_1(a, b; c; z), \tag{B22}$$

where $a = 2q - \frac{\omega}{2}$, $b = 2q + \frac{\omega}{2}$, $c = 2q + \frac{1}{2}$, and $z = (1 + \chi)/2$, Eq. (B21) becomes

$$\phi_{AA} = \tilde{A}_2[1 - \chi^2]^q$$
$$\times {}_2F_1\left(2q - \frac{\omega}{2}, 2q + \frac{\omega}{2}; 2q + \frac{1}{2}; \frac{1+\chi}{2}\right), \tag{B23}$$

where $\tilde{A}_2$ is a constant. This is of the same form as Eq. (B15).

APPENDIX C: BINDING ENERGY IN THE LIMIT OF WEAK ELECTRON-HOLE ATTRACTION

Let us consider the limit of very weak electron-hole attraction: $|\tilde{V}(x)| \ll \Delta$ for any inter-particle separation. Substituting $\varepsilon = 2\Delta - \varepsilon_b$ into Eq. (4) and retaining only the first-order terms in $\tilde{V}/\Delta$ and ε_b/Δ transforms Eq. (4) into

$$-\frac{1}{\Delta}\frac{d^2\phi_{AA}}{dx^2} + \tilde{V}(x)\phi_{AA} = -\varepsilon_b\phi_{AA}. \tag{C1}$$

Equation (C1) is of the form of the nonrelativistic one-dimensional Schrödinger equation, where the parameter Δ plays the role of the effective mass. For a particle of mass m in a weak one-dimensional potential $U(x)$, the textbook[35] treatment of the potential as a perturbation yields the following result for the binding energy: $|E| = (m/2\hbar^2)[\int_{-\infty}^{\infty} U(x)dx]^2$, which can be reformulated for Eq. (C1) as

$$\varepsilon_b = \frac{\Delta}{4}\left[\int_{-\infty}^{\infty} \tilde{V}(x)dx\right]^2. \tag{C2}$$

For $\tilde{V}(x) = -\alpha/\cosh(\beta x)$, Eq. (C2) yields

$$\varepsilon_b = \frac{\pi^2}{4}\Delta(\alpha/\beta)^2. \tag{C3}$$

Thus we have shown that for a weak hyperbolic secant potential, when $\alpha \ll \Delta$, the binding energy scales with the band gap and has the universal quadratic dependence on α/β. Notably, Eq. (C3) is also true for a weak interaction potential given by a Lorentzian, $\tilde{V}(x) = \alpha/(1 + \beta^2 x^2)$.

APPENDIX D: EXCITON WITH ATTRACTIVE δ-FUNCTION POTENTIAL

The problem of calculating the energy of an exciton in a carbon nanotube can be solved exactly in the case when interaction potential between the electron and hole is taken to be a δ function, $\tilde{V}(x) = -U_0\delta(x)$, where the strength of the potential U_0 is positive and can be estimated as a product of the strength of the realistic potential and its width.

Let us consider an electron-hole pair in a narrow-gap one-dimensional (1D) carbon nanotube with a band-gap energy of $4|\Delta|$. Measuring all quantities in units of $\hbar v_F$ as above, one can write the Hamiltonian as

$$\widehat{H} = \sqrt{\Delta^2 + \widehat{q}_e^2} + \sqrt{\Delta^2 + \widehat{q}_h^2} + \tilde{V}(|x_e - x_h|), \quad \text{(D1)}$$

where

$$\widehat{q}_{e,h} = -i\frac{\partial}{\partial x_{e,h}}. \quad \text{(D2)}$$

Here we have retained only the eigenstates of the general graphene-type Hamiltonian with positive energies. The Schrodinger equation for the problem we consider thus reads

$$\left[\sqrt{\Delta^2 + \widehat{q}_e^2} + \sqrt{\Delta^2 + \widehat{q}_h^2} + \tilde{V}(|x_e - x_h|)\right]$$
$$\times \Psi(x_e, x_h) = \varepsilon \Psi(x_e, x_h). \quad \text{(D3)}$$

Introducing new variables corresponding to the center of mass and relative motion in the manner explained in the text of the paper and putting the wave vector of the center-of-mass motion equal to zero, one gets the following expression:

$$2\sqrt{\Delta^2 - \frac{d^2}{dx^2}}\psi(x) = [\varepsilon - \tilde{V}(x)]\psi(x), \quad \text{(D4)}$$

where $\psi(x)$ is a wave function of relative motion. The above expression is a 1D Schrödinger equation for a particle with a complicated dispersion placed in an external potential $\tilde{V}(x)$, which can be solved for the case when $\tilde{V}(x) = -U_0\delta(x)$. Indeed, in this case the solution of the Schrödinger equation

corresponding to a bound state reads

$$\psi(x) = Ae^{\kappa(\epsilon)x}, \quad x < 0, \quad \text{(D5)}$$
$$\psi(x) = Ae^{-\kappa(\epsilon)x}, \quad x > 0, \quad \text{(D6)}$$

where $\varepsilon = 2\sqrt{\Delta^2 + \kappa^2}$. It is well known that the zero-range δ potential is equivalent to the introduction of the specific boundary condition for the derivative of the solution at $x = 0$ [the function itself should be continuous, $\psi(+0) = \psi(-0)$]. To obtain a condition for the derivative, let us integrate the Schrödinger equation

$$\left[\widehat{H}_0\left(\frac{d}{dx}\right) - \varepsilon\right]\psi(x) = U_0\delta(x)\psi(x), \quad \text{(D7)}$$

where $\widehat{H}_0(\frac{d}{dx}) = 2\sqrt{\Delta^2 + \frac{d^2}{dx^2}}$ across the interval $[-0; +0]$; this yields

$$\int \widehat{H}_0\left(\frac{d}{dx}\right)\psi(x)dx|_{x=+0}$$
$$- \int \widehat{H}_0\left(\frac{d}{dx}\right)\psi(x)dx|_{x=-0} = U_0\psi(0). \quad \text{(D8)}$$

Using this relation and Eqs. (D5), (D6), and (D4) one gets the following expression for determining the parameter κ:

$$H_0(-\kappa) + H_0(\kappa) - 2H_0(0) = -\kappa U_0. \quad \text{(D9)}$$

Thus

$$\sqrt{\Delta^2 - \kappa^2} = \Delta - \frac{\kappa U_0}{4}, \quad \text{(D10)}$$

which yields

$$\kappa = \frac{8U_0\Delta}{16 + U_0^2}, \quad \text{(D11)}$$

where the energy is given by

$$\varepsilon = 2\sqrt{\Delta^2 - \kappa^2}. \quad \text{(D12)}$$

Note that this result is valid only for the case of $U_0 < 4$, as for $U_0 > 4$, the right-hand side of Eq. (D10) is negative. The binding energy of the exciton can be determined as

$$\varepsilon_b = 2\Delta - \epsilon = 4\Delta U_0^2/(16 + U_0^2), \quad \text{(D13)}$$

which tends to the nonrelativistic result of $\varepsilon_b = U_0^2\Delta/4$ for $U_0 \ll 1$; cf. this result with Eq. (C2). Note that in both relativistic and nonrelativistic cases the binding energy is proportional to the gap.

---

*m.e.portnoi@exeter.ac.uk

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

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

[3] See, e.g., a review article A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).

[4] J.-C. Charlier, X. Blase, and S. Roche, Rev. Mod. Phys. **79**, 677 (2007).

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

[6] H. Ajiki and T. Ando, J. Phys. Soc. Jpn. **62**, 1255 (1993).

[7] S. M. Bachilo, M. S. Strano, C. Kittrell, R. H. Hauge, R. E. Smalley, and R. B. Weisman, Science **298**, 2361 (2002).

[8] A. Hartschuh, H. N. Pedrosa, L. Novotny, and T. D. Krauss, Science **301**, 1354 (2003).

[9] F. Wang, G. Dukovic, L. E. Brus, and T. F. Heinz, Phys. Rev. Lett. **92**, 177401 (2004).

[10]F. Wang, G. Dukovic, L. E. Brus, and T. F. Heinz, Science **308**, 838 (2005).

[11]S. Berciaud, L. Cognet, and B. Lounis, Phys. Rev. Lett. **101**, 077402 (2008).

[12]J. Shaver, J. Kono, O. Portugall, V. Krstic, G. L. J. A. Rikken, Y. Miyauchi, S. Maruyama, and V. Perebeinos, Nano Lett. **7**, 1851 (2007).

[13]J. Shaver and J. Kono, Laser & Photon. Rev. **1**, 260 (2007), and references therein.

[14]A. D. Mohite, P. Gopinath, H. M. Shah, and B. W. Alphenaar, Nano Lett. **8**, 142 (2008).

[15]T. G. Pedersen, Carbon **42**, 1007 (2004).

[16]V. Perebeinos, J. Tersoff, and P. Avouris, Phys. Rev. Lett. **92**, 257402 (2004).

[17]S. Uryu and T. Ando, Phys. Rev. B **76**, 115420 (2007).

[18]For a review, see T. Ando and S. Uryu, Phys. Status Solidi C **6**, 173 (2009).

[19]J. Deslippe, C. D. Spataru, D. Prendergast, and S. G. Louie, Nano Lett. **7**, 1626 (2007).

[20]E. Malic, J. Maultzsch, S. Reich, and A. Knorr, Phys. Rev. B **82**, 035433 (2010).

[21]O. V. Kibis, M. Rosenau da Costa, and M. E. Portnoi, Nano Lett. **7**, 3414 (2007).

[22]M. E. Portnoi, O. V. Kibis, and M. Rosenau da Costa, Superlattices Microstruct. **43**, 399 (2008).

[23]M. E. Portnoi, M. Rosenau da Costa, O. V. Kibis, and I. A. Shelykh, Int. J. Mod. Phys. B **23**, 2846 (2009).

[24]T. Ando, J. Phys. Soc. Jpn. **66**, 1066 (1997).

[25]D. Jerome, T. M. Rice, and W. Kohn, Phys. Rev. **158**, 462 (1967).

[26]C. L. Kane and E. J. Mele, Phys. Rev. Lett. **78**, 1932 (1997).

[27]L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).

[28]R. R. Hartmann, N. J. Robinson, and M. E. Portnoi, Phys. Rev. B **81**, 245431 (2010).

[29]Ya. B. Zeldovich and V. S. Popov, Usp. Fiz. Nauk **105**, 403 (1971) [Sov. Phys. Usp. **14**, 673 (1972)].

[30]A. Shytov, M. Rudner, N. Gu, M. Katsnelson, and L. Levitov, Solid State Commun. **149**, 1087 (2009).

[31]A. M. Fischer, V. L. Campo Jr., M. E. Portnoi, and R. A. Römer, Phys. Rev. Lett. **102**, 096405 (2009).

[32]R. M. Konik, Phys. Rev. Lett. **106**, 136805 (2011).

[33]E. Kamke, *Differentialgleichungen—Lösungsmethoden und Lösungen* (Chelsea, New York, 1974).

[34]M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1972).

[35]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Pergamon, Oxford, 1991), p. 163.