Binding energy and lifetime of excitons in metallic nanotubes

L. Shan, M. Agarwal, and E. G. Mishchenko

Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA

The difficulty of describing excitons in semiconducting single-wall nanotubes analytically lies with the fact that excitons can neither be considered strictly one- nor two-dimensional objects. However, the situation changes in the case of metallic nanotubes where, by virtue of screening from gapless metallic subbands, the radius of the exciton becomes much larger than the radius of the nanotube $R_{ex} \gg R$. Taking advantage of this, we develop the theory of excitons in metallic nanotubes, determining that their binding energy is about $0.08v/R$, in agreement with the existing experimental data. Additionally, because of the presence of the gapless subbands, there are processes where bound excitons are scattered into unbound electron-hole pairs belonging to the gapless subbands. Such processes lead to a finite exciton lifetime and the broadening of its spectral function. We calculate the corresponding decay rate of the excitons.

I. INTRODUCTION

An exciton is a bound state of an electron and a hole formed by their Coulomb attraction—the solid-state analogy of a hydrogen atom, but with a larger size and a smaller binding energy. The smaller binding energy is due to the screening of the mutual Coulomb interaction by bound electrons of the medium, described by its dielectric constant. Excitons typically exist in insulators and weakly doped semiconductors. Metallic and strongly doped semiconducting materials disfavor formation of excitons. This is the result of two factors. First, as the conduction band is populated (e.g., by doping), screening by free charges strongly reduces the magnitude of the electron-hole Coulomb interaction and decreases its binding energy. Second, population of the conduction band reduces the number of quantum states available to accommodate the electron after its (virtual) scattering off the hole, further decreasing the binding energy to the point where no meaningful bound state may be formed anymore.

This situation changes in quasi-one-dimensional systems, such as metallic nanotubes [1], where formation of excitons occurs in subbands *different* from the subbands that are responsible for metallic screening. Such separation occurs as a result of quantization of the circumferential momentum. This makes an exciton a well-defined excitation. Excitons in metallic single-walled carbon nanotubes (SWNTs) were experimentally observed in Refs. [2,3]. Their binding energy was found to be about 50 meV, an order of magnitude smaller than the typical band gaps Δ of semiconducting nanotubes. Theoretically, excitons in metallic SWNTs were first studied by Ando under the effective-mass approximation [4]. Later, the binding energy of the exciton was addressed by first-principles calculations [5–7], and also by the density matrix theory [8], with the latter results being in agreement with experimental measurements. A brief review of excitonic effects in metallic SWNTs was given in Ref. [9]. However, to our knowledge, no simple analytical description of excitons in metallic nanotubes has been developed. It is the purpose of the present paper to fill this gap.

Let us illustrate the difficulty of describing excitons analytically in semiconducting SWNTs, stemming from the fact that excitons can neither be considered one-dimensional (1D) nor two-dimensional (2D) excitations. Indeed, consider the lowest energy subbands with the spectrum $\varepsilon(p) = \pm\sqrt{\Delta^2 + v^2 p^2}$, and expand it near the bottom of the subband,

$$\sqrt{\Delta^2 + v^2 p^2} \approx \Delta + \frac{p^2}{2\mu}, \quad \mu = \frac{\Delta}{v^2}. \tag{1}$$

The band gap is typically $\Delta \sim v/R$, where R is the radius of the nanotube. For convenience, we set $\hbar = 1$ throughout the paper. (In particular, in the zone-folding tight-binding approximation both the (8,0) and (10,0) zigzag nanotubes have $\Delta = v/3R$.)

Because the electron-hole interaction energy is $U(r) = -e^2/r$, the exciton binding energy E_b can be estimated by minimizing,

$$-E_b \approx \min\left[\frac{v^2 p^2}{2\Delta} - \frac{e^2}{r}\right], \tag{2}$$

taking into account the uncertainty relation, $r \sim 1/p$. This yields

$$E_b \sim \left(\frac{e^2}{v}\right)^2 \Delta. \tag{3}$$

Since $e^2/v \sim 1$, we obtain that $E_b \sim \Delta$ and, consequently, the exciton radius $R_{ex} \sim v/E_b \sim R$. Because to consider excitons to be 1D one would need to have $R_{ex} \gg R$, and conversely, to view them as 2D one would require $R_{ex} \ll R$, the semiconducting problem falls instead between the two limits where a numerical analysis is necessary. However, as we are going to see below, the situation changes in the case of metallic nanotubes where—by virtue of screening by the metallic subbands—the exciton binding energy decreases significantly (as already evidenced by the experimental data). As a result, the radius of the exciton increases, $R_{ex} \gg R$, and treatment of the exciton as a quasi-one-dimensional object becomes possible. This is what makes the analytic solution viable.

TABLE I. Analytically and numerically calculated (or experimentally measured) binding energies of the excitons in different metallic carbon nanotubes (M-NTs). The values in parenthesis are extrapolated from the binding energy for the (13,1) nanotube obtained in Ref. [8] and reported in it $E_b \propto 1/R$ dependence of the binding energy on the radius of a metallic nanotube.

M-NTs (n, m)	Binding energies E_b (meV)			
	This paper	Density matrix	*ab initio*	Exper.
(12,0)	86	(88)	~50 [6]	–
(13,1)	76	78 [8]	–	~50 [3]
(10,10)	60	(61)	~50 [6]	–
(21,21)	28	(29)	–	~50 [2]

In this paper, we determine the binding energy and lifetime of an exciton in metallic SWNTs, taking into account the screening effects within the random phase approximation (RPA). A problematic feature of the 1D Coulomb problem is the r^{-1} singularity in the potential energy, which is not integrable (unlike in the three-dimensional situation). This feature was first addressed by Loudon [10] and later further extensively studied by others [11–15]. Loudon introduced the truncated Coulomb interaction $e^2/(|x| + a_0)$ with a positive constant a_0 to ensure that Coulomb potential is regular at small distances $x \to 0$. In our problem, the nanotube radius R appears naturally and no other cutoff is needed. Using a variational ansatz with a Gaussian trial function, we show below that the binding energy is (the value of Δ corresponds to an armchair metallic nanotube)

$$E_b \approx 0.08\,\Delta, \quad \Delta = \frac{v}{R}. \quad (4)$$

The $1/R$ dependence of the binding energy exhibits a good agreement with the results obtained in Ref. [8] with the density matrix theory, see Table I. In contrast, *ab initio* methods yield only rather crude estimates of the binding energy [2,5,6]. This is because computational limitations do not permit calculations for large nanotubes. In addition, the technique of photoluminescence (PL) spectroscopy cannot be applied to obtain binding energies of metallic nanotubes, which makes it challenging to determine the binding energy of excitons in metallic SWNTs. This is because the exciton is likely to undergo a nonradiative decay to the nearby linear subbands instead of an electron-hole recombination process with a photon released. Nonetheless, the exciton binding energies experimentally observed [2,3] have the same order of magnitude as predicted by our calculations.

Furthermore, we explore the stability of excitons in metallic SWNTs—the question that is of fundamental interest but which remains unexplored in the existing theoretical works. The mechanism of a finite lifetime of the exciton can be illustrated by Fig. 1. The mth subband, corresponding to the integer angular quantum number m, has the energy $\varepsilon_m(p) = \pm v\sqrt{m^2/R^2 + p^2}$. Exciton bound states are formed between gapped, $m \neq 0$ subbands, and the lowest $m = \pm 1$ exciton is indicated by a dashed line in Fig. 1. This energy *overlaps* with the gapless $m = 0$ subbands. Accordingly, an elastic transition of the electron and the hole from gapped to gapless subbands opens up a decay channel for the exciton.
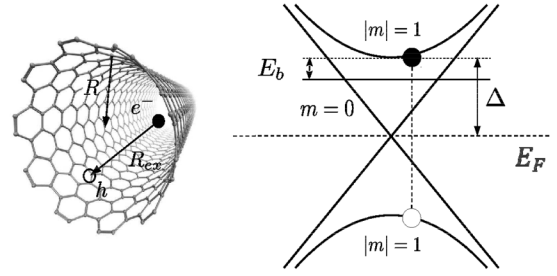

FIG. 1. Exciton state emerges due to the final state interaction of the electron and hole with energy below the parabolic subband.

To better understand the role of the Coulomb interaction in the formation and decay of the exciton, the following picture is helpful. The exciton is produced by multiple virtual scattering events of the electron and the hole within the gapped subbands. Such transitions are controlled by the Coulomb coupling V_0, the subscript indicating that no angular momentum change takes place. These intrasubband transitions determine the binding energy of the exciton (see Fig. 2, left panel). The intersubband transitions (shown on the right panel in Fig. 2) appear as a result of the "dipolar" Coulomb interaction V_1 and occur with a change ± 1 of the angular momentum of the electron (and with the opposite change of the angular momentum of the hole). We obtain that the ratio of the binding energy to the half-width of the exciton spectral function is

$$\Gamma \approx 0.015\,\Delta. \quad (5)$$

Such a ratio indicates that excitons in metallic SWNTs are well-defined excitations. This is consistent with experimental results obtained through ultrafast luminescence [16]. Below we derive our main results, Eqs. (4) and (5).

II. THE HAMILTONIAN OF THE SYSTEM

The Coulomb interaction potential between two electrons located on the surface of a nanotube of radius R at (x_1, ϑ_1) and (x_2, ϑ_2) is given by

$$V(x, \vartheta) = \frac{e^2}{\sqrt{x^2 + 4R^2 \sin^2(\vartheta/2)}}, \quad (6)$$

where $x = x_1 - x_2$ and $\vartheta = \vartheta_1 - \vartheta_2$ are the relative distance along the nanotube axis and the relative angle around the circumference of the nanotube. [In the case of an
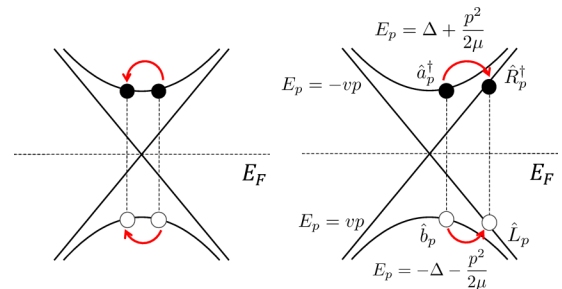

FIG. 2. Scattering processes shown within the massive subband (via U_0 coupling) and between massive and massless subbands (via U_1 coupling).

electron-hole pair, the sign of the interaction (6) is reversed.] If R becomes large compared with a characteristic distance of the electron and the hole trajectories, the interaction approaches the Coulomb potential on a flat two-dimensional plane. On the other hand, if R becomes small, Eq. (6) reduces to the "1D hydrogen" problem studied by Loudon [10]. The Coulomb interaction on the cylinder—the problem interpolating between 1D and 2D—was previously considered by Mahan [17] and Petersen [18].

What is different in our case is the need to take into account screening, introduced by the gapless subbands of metallic SWNTs.

This is most conveniently done by transitioning to momentum space. The Fourier transform of the Coulomb interaction (6) yields the following expression [4,19]:

$$V_m(q) = \int_{-\infty}^{\infty} dx\, e^{-iqx} \int_{-\pi}^{\pi} \frac{d\vartheta}{2\pi}\, e^{-im\vartheta} \frac{e^2}{\sqrt{x^2 + 4R^2 \sin^2(\vartheta/2)}},$$
$$= 2e^2 K_{|m|}(qR) I_{|m|}(qR), \qquad (7)$$

where m is an integer number, and I_m and K_m are the modified Bessel functions of the first kind and second kind, respectively. However, because of the presence of conduction electrons in the gapless subbands, which move around in response to the "bare" Coulomb interaction $V_m(q)$, the actual (screened) interaction is modified: $V_m(q) \rightarrow U_m(\omega, q)$. This screened interaction disturbs the equilibrium of the system and induces density variations, which within the linear response are proportional to the strength of the interaction, $n_m(\omega, q) = \Pi_m(\omega, q) U_m(\omega, q)$, with the coefficient $\Pi_m(\omega, q)$ referred to as the polarization function (i.e., density-density correlation function) associated with the change m in the angular momentum and the change q in the linear momentum. The central tenet of RPA is the assumption of the mean field, which predicts that the variation of the density $n_m(\omega, q)$ induces the additional potential in the system: $e\varphi(\omega, q) = V_m(q) n_m(\omega, q) = V_m(q) \Pi_m(\omega, q) U_m(\omega, q)$. This additional potential together with the bare potential constitutes the total interaction: $e\phi(\omega, q) + V_m(q) = U_m(\omega, q)$. This gives

$$U_m(\omega, q) = \frac{V_m(q)}{1 - V_m(q)\Pi_m(\omega, q)}. \qquad (8)$$

The formation of the exciton is mostly facilitated by the strongest interaction V_0, which at low $qR \ll 1$ becomes logarithmically strong, $V_0(q) \approx -2e^2 \ln(qR)$. In contrast, the higher interaction harmonics, $m \neq 0$, remain constant in this limit, $V_m \rightarrow e^2/|m|$. The screening of the V_0 interaction is determined by the uniform harmonics of the polarization function with no change in the angular momentum, given by

$$\Pi_0(\omega, q) = \frac{N}{\pi v} \frac{q^2 v^2}{\omega^2 - q^2 v^2}, \qquad (9)$$

where $N = 4$ accounts for two spin directions and the presence of the two Dirac points within the Brillouin zone (see Appendix for details).

A large spatial radius of the exciton R_{ex} makes it sufficient to consider only the static limit of Eq. (9), where ω is disregarded compared with qv. Indeed, frequencies involved are of the order of the exciton binding energy, $\omega \sim E_b = 1/(\mu R_{ex}^2)$, where the effective mass $\mu \sim 1/(vR)$, according to Eq. (1).

On the other hand, the involved momenta are $q \sim 1/R_{ex}$. Accordingly, the ratio $\omega^2/q^2 v^2 \sim R^2/R_{ex}^2 \ll 1$. The screened Coulomb interaction in this static limit assumes the form

$$U_0(q) = \frac{2e^2 K_0(qR) I_0(qR)}{1 + \alpha K_0(qR) I_0(qR)}, \quad \alpha = \frac{2Ne^2}{\pi v} \approx 6.9, \quad (10)$$

where we introduced the dimensionless interaction strength $\alpha = 2Ne^2/\pi v$, where v is taken to be the same as the velocity of electrons propagating in graphene, $v = 8 \times 10^5$ m/s.

To similarly calculate the screening of the V_1 interaction, one needs to know the polarization function associated with the ± 1 change of the angular momentum. While the intra-subband value of $\Pi_0(\omega, q)$ at low ω and qv, as seen from Eq. (9), depends on which one of the two quantities tends to zero faster than the other, there is no such ambiguity for the intersubband polarization function $\Pi_1(\omega, q)$: as shown in Appendix, $\Pi_1(0, 0) = -1.16N/(\pi v)$. This gives, for the screened intersubband interaction,

$$U_1 = \frac{e^2}{1 + 1.16\frac{e^2 N}{\pi v}} \approx 0.2\, e^2. \qquad (11)$$

In the limit of large wavelengths, $qR \ll 1$, the two interactions approach each other, $U_0(0) \approx U_1$. It should be pointed out, however, that while the U_1 interaction remains almost constant for finite but small q, the interaction $U_1(q)$ is rather sensitive to q, with the derivative $dU_1(q)/dq$ diverging at $q \rightarrow 0$.

Note that in the absence of spectrum curvature, the polarization function $\Pi_0(\omega, q)$ of one-dimensional subbands is independent of temperature. It is also unmodified by the electron-electron interactions [20]. On the other hand, the polarization constant $\Pi_1(0, 0)$, which involves gapped subbands, has a negligible temperature dependence. Indeed, the population of thermally excited electron-hole pairs is $\sim \exp(-\Delta/k_B T)$, where Δ is much larger than the thermal energy at room temperature. For these reasons it is sufficient to consider the screening of the Coulomb interaction at zero temperature. Similarly, small levels of doping, $\mu \ll \Delta$, do not affect the strength of the Coulomb interaction.

Having determined the magnitude of the electron-electron (and, therefore, electron-hole) interaction, we can proceed to calculate the energy and the lifetime of the exciton. The Hamiltonian of the system has the following form:

$$\hat{H} = \sum_p \left(\frac{p^2}{2\mu} + \Delta \right) (\hat{a}_p^\dagger \hat{a}_p - \hat{b}_p^\dagger \hat{b}_p) + \sum_p vp\, (\hat{R}_p^\dagger \hat{R}_p - \hat{L}_p^\dagger \hat{L}_p)$$
$$+ \frac{1}{L} \sum_{p,k} U_0(p-k) \hat{a}_k^\dagger \hat{b}_p^\dagger \hat{b}_k \hat{a}_p$$
$$+ \frac{U_1}{2L} \sum_{p,k} (\hat{R}_k^\dagger \hat{b}_p^\dagger \hat{L}_k \hat{a}_p + a_p^\dagger \hat{L}_k^\dagger \hat{b}_p \hat{R}_k)$$
$$+ \frac{U_1}{2L} \sum_{p,k} (\hat{L}_k^\dagger \hat{b}_p^\dagger \hat{R}_k \hat{a}_p + \hat{a}_p^\dagger \hat{R}_k^\dagger \hat{b}_p \hat{L}_k). \qquad (12)$$

Here $\hat{a}_p$ and $\hat{b}_p$ are the operators for the particles residing on the upper and lower subbands with $m = 1$ (or $m = -1$); the operators $\hat{R}_p$ and $\hat{L}_p$ correspond to right- and left-moving particles of massless ($m = 0$) subbands; the $U_0(q)$ interaction describes scattering within the massive subbands, whereas

U_1 coupling describes processes where scattering occurs between massive and massless subbands. Note that of all of the Coulomb interaction terms, we have retained only those that are responsible for the formation of the exciton with zero total momentum: for example, the U_0 term describes scattering of the electron with momentum p ($\hat{a}_p$) and the hole with momentum $-p$ ($\hat{b}_p^\dagger$) into a pair of new states with momenta k and $-k$.

The intersubband terms in the Hamiltonian (12) have the extra prefactor $1/2$ compared with the intrasubband transitions. The origin of this difference lies in the pseudospin nature of the underlying Hamiltonian of the two-dimensional graphene crystal that forms the nanotube. (The pseudospin arises from the existence of two atomic sublattices in the graphene honeycomb arrangement of carbon atoms.) As a result, two states of the same energy and opposite momenta are orthogonal to each other. More generally, the amplitude of the transition between the states having momenta **p** and **p** + **q** and the same sign of energy is suppressed by the factor $\cos[(\theta_{\mathbf{p+q}} - \theta_{\mathbf{p}})/2]$, where $\theta_{\mathbf{p}}$ is the angle that the momentum **p** makes with the x axis. When graphene is rolled into a nanotube, the circumferential momenta are quantized, $p_y = m/R$. The gapless states $m = 0$ are those that move along the x axis: **p** $= (p, 0)$ with $\theta_{\mathbf{p}} = 0$ or π. In contrast, in the gapped subbands, **p** $= (p, m/R)$, with the relevant momenta are near the bottom/top of the subbands: $p \ll 1/R$. Accordingly, the relevant states are those that have $\theta_{\mathbf{p}} \approx \pm\pi/2$. Correspondingly, each particle transitioning between a gapless state and a gapped state (close to the bottom of the subband) introduces a factor $\cos(\pi/4) = 1/\sqrt{2}$ into the amplitude of the scattering. For the two-particle Coulomb interaction U_1, the total additional coefficient is, therefore, $1/2$.

III. BINDING ENERGY AND LIFETIME

The wave function of the exciton is sought in the form

$$|\psi\rangle = \sum_q f_q \hat{a}_q^\dagger \hat{b}_q |0\rangle + \sum_{q>0} g_q \hat{R}_q^\dagger \hat{L}_q |0\rangle + \sum_{q<0} g_q \hat{L}_q^\dagger \hat{R}_q |0\rangle,$$

(13)

where $|0\rangle$ is the ground state of the system where all individual electron states with the positive energy are empty and all states with the negative energy are occupied. The function f_p describes the amplitude of the electron-hole pair to be in the gapped states, whereas the function g_q describes the likelihood of the pair to reside in the massless states. Note that in the last two terms we have explicitly taken into account that in the ground state $|0\rangle$ the left-moving states are occupied as long as q is positive whereas the right-moving states are occupied if q is negative. Because the system is symmetric with respect to the symmetry between left- and right-moving states, the function g_q must be symmetric: $g_q = g_{-q}$.

The Schrödinger equation $\hat{H}|0\rangle = E|0\rangle$ separates into two coupled equations for the functions f_q and g_q:

$$(E - 2\Delta - q^2/\mu)f_q = -\frac{1}{L}\sum_p U_0(q-p)f_p - \frac{U_1}{2L}\sum_p g_p,$$

$$(E - 2v|q|)g_q = -\frac{U_1}{2L}\sum_p f_p. \quad (14)$$

Excluding g_q from these equations and replacing the sums by the integrals, $\sum_p \to L\int dp/2\pi$, we arrive at the following integral equation for the function f_q:

$$(E - 2\Delta - q^2/\mu)f_q = \int_{-\infty}^{\infty} \frac{dp}{2\pi} f_p\left[-U_0(p-q), +\frac{U_1^2}{4}\right.$$
$$\left.\times \int_{-\infty}^{\infty} \frac{dp'}{2\pi} \frac{1}{E - 2|p'|v + i\eta}\right]. \quad (15)$$

The singularities at $E = 2|p'|v$ lead to the imaginary part in the energy $E - i\Gamma/2$ of the exciton. Because the resulting imaginary part is small compared with the band gap, $\Gamma \ll \Delta$, it is sufficient to keep the infinitesimal η in the denominator and utilize the Sokhotski identity, $\text{Im}\, 1/(E - 2|p'|v + i\eta) = -i\pi\delta(E - 2|p'|v)$, for the calculation of the integral's imaginary part. In contrast, the real part, which arises from the principal value of the integral, can be ignored. Although the real part appears to diverge logarithmically, such divergence is the artefact of our assumption that U_1 is constant. This approximation, in any case, fails for transferred momenta of the order $1/R$, which should, therefore, be used as the upper cutoff for the logarithmic integral. Finally, because the energy involved is large, $E \sim 2\Delta = 2v/R$ (rather than the small binding energy E_b), the logarithm is of the order 1 and the real part of the last term in the brackets in Eq. (15) merely adds a contribution $\sim U_1^2/8\pi v$. This second-order correction is small compared to the main contribution from the U_0 term and may be ignored. Hence, we obtain from Eq. (15),

$$\left(E - 2\Delta - \frac{q^2}{\mu}\right)f_q = -\int_{-\infty}^{\infty} \frac{dp}{2\pi}\left[U_0(p-q) + i\frac{U_1^2}{8v}\right]f_p.$$

(16)

Below, the binding energy and the lifetime of the exciton,

$$E = 2\Delta - E_b - i\Gamma/2, \quad (17)$$

are determined from the ground-state eigenvalue E of this equation.

A. The shallow potential approximation

The fastest way to *estimate* the exciton binding energy is by utilizing the well-known in quantum mechanics shallow well approximation which in the momentum space amounts to replacing the interaction $U_0(q)$ with its zero-momentum value $U_0 \equiv U_0(0)$. Equation (16) then becomes exactly solvable and yields

$$E_b + i\Gamma/2 = \frac{\mu}{4}\left[U_0 + i\frac{U_1^2}{8v}\right]^2 \approx \frac{\mu}{4}\left[U_0^2 + i\frac{U_0 U_1^2}{4v}\right]. \quad (18)$$

The effective mass, cf. Eq. (1), can now be expressed via the band gap, $\mu = \Delta/v^2$ and the interaction constants $U_0 \approx \pi v/4$ and $U_1 \approx 0.2\, e^2$, given by Eqs. (10) and (11). This gives,

$$E_b = \frac{\pi^2}{64}\Delta, \quad \Gamma/E_b \approx 0.19. \quad (19)$$

It is clear, however, that the value $E_b = 0.15\Delta$ that follows from Eq. (19) significantly *overestimates* the binding energy of the exciton: as evidenced by Eq. (10), the function $U_0(q)$ is irregular at $q \to 0$ where it has an infinite derivative. For

this reason, the integrand in Eq. (16) in fact contributes much less to the integral where p departs from q than predicted by the shallow potential approximation. To make a better approximation, we are now going to utilize the variational approach.

B. The variational solution

To apply the variational approach to Eq. (16), we choose the Gaussian trial function,

$$f(x) = \left(\frac{\beta}{\pi}\right)^{1/4} e^{-\beta x^2/2}, \qquad (20)$$

or, equivalently, in the momentum space, $f_p = (4\pi/\beta)^{1/4} e^{-p^2/2\beta}$. According to Eq. (16), the ground-state energy of the exciton is the sum of the kinetic energy,

$$\overline{T} = \int_{-\infty}^{\infty} \frac{dp}{2\pi} \frac{p^2}{\mu} f_p^2 = \frac{\beta}{2\mu}, \qquad (21)$$

and the expectation value of the potential energy. The latter is a complex quantity: $U_0(x) - i\frac{U_1^2}{8v}\delta(x)$. The imaginary part determines the exciton lifetime,

$$\Gamma/2 = \frac{U_1^2}{8v}[f(0)]^2 = \frac{U_1^2}{8v}\sqrt{\frac{\beta}{\pi}}, \qquad (22)$$

while the real part yields, with the help of the momentum representation, the following integral:

$$\overline{U_0} = \int_{-\infty}^{\infty} dx\, U_0(x)[f(x)]^2 = \int_{-\infty}^{\infty} \frac{dp}{2\pi}\, U_0(p)e^{-p^2/4\beta}. \qquad (23)$$

The binding energy should be found by minimizing the sum $\overline{T} - \overline{U_0}$ with respect to β. The integral in $\overline{U_0}$ cannot be calculated exactly but can be approximated rather accurately. First, it is convenient to utilize the dimensionless variables $s = pR$ and $t = \sqrt{\beta}R$ to recast the average potential energy (10) in the form

$$\overline{U_0} = \frac{2e^2}{\pi R} \int_0^{\infty} ds\, e^{-s^2/4t^2} \frac{K_0(s)I_0(s)}{1 + \alpha K_0(s)I_0(s)}. \qquad (24)$$

Next, we anticipate that for the large-radius excitons the small values of $t < 1$ (and hence the small values of s) are relevant. (This expectation is supported by the final result.) The function $I_0(s) \approx 1$, whereas the Macdonald function has a logarithmic singularity, $K_0(s) \approx \ln(2/s) - \gamma$, where $\gamma = 0.577$ is the Euler constant. Finally, we notice that the logarithm depends on its argument s rather weakly, in comparison with the exponential $e^{-s^2/4t^2}$, and hence can be approximated as a constant within the relevant range of the s integration,

$$\int_0^{\infty} ds\, e^{-s^2/4t^2} \frac{K_0(s)I_0(s)}{1 + \alpha K_0(s)I_0(s)} \approx \frac{\sqrt{\pi}\, t \ln(C/t)}{1 + \alpha \ln(C/t)}, \qquad (25)$$

with some fitting parameter C. Numerical calculation demonstrates that the value $C = 3/2$ provides an excellent fit between the exact numerical integration of the left-hand side of Eq. (25) and its right-hand side for the value $\alpha = 6.9$ stated in Eq. (10), which corresponds to $N = 4$ gapless modes.

TABLE II. The binding energies of the excitons in (21,21) armchair nanotubes for different dielectric constants κ.

Substrates	Dielectric constant κ	Binding energy E_b (meV)
SiO$_2$	2.5	22.4
SiC	3.75	18.7
Si/SiO$_2$	4.4	17.9
GaAs	7	13.8

Accordingly, we arrive at the following value of the binding energy, $-E_b = \overline{T} - \overline{U_0}$,

$$\begin{aligned} E_b &= -\frac{t^2}{2\mu R^2} + \frac{2e^2}{\sqrt{\pi}R} \frac{t \ln(C/t)}{1 + \alpha \ln(C/t)} \\ &= \Delta\left[-\frac{t^2}{2} + \frac{\sqrt{\pi}}{N} \frac{\alpha t \ln(C/t)}{1 + \alpha \ln(C/t)}\right], \end{aligned} \qquad (26)$$

where in the last line we used the definition of the coupling constant α from Eq. (10) and also replaced the effective mass $\mu = \Delta/v^2$ in terms of the band gap $\Delta = v/R$.

The binding energy (26) has a maximum at $t = t_0 = 0.36$, where the value E_b stated in Eq. (4) is reached. This binding energy agrees well with the experimental measurements [2]. In turn, the obtained result justifies the approximations of a large exciton radius. Indeed, according to the wave function (20), the radius of the exciton is $R_{ex} \sim \sqrt{2/\beta} = \sqrt{2}R/t_0 \approx 4R$.

The exciton radius increases when the Coulomb interaction is further reduced if the nanotube is located on the surface of a dielectric substrate. Correspondingly, the dimensionless interaction strength will be modified as $\alpha = 4Ne^2/\pi v(\kappa + 1)$, where κ is the dielectric constant of the substrate. The binding energy can still be calculated using Eq. (26). Table II shows the calculated binding energies for several substrates.

The decay of the exciton into the linear subbands was first studied numerically in Ref. [7], which concluded that such decay processes lead to a negligible broadening. Our result indicate otherwise and are also in a good agreement with experiments reported in Ref. [16].

The value of the inverse lifetime now follows from Eq. (22):

$$\Gamma = \frac{\Delta t_0}{4\sqrt{\pi}} \frac{U_1^2}{v^2}, \qquad (27)$$

whose numerical value yields Eq. (5). The exciton acquires a significant broadening but nonetheless remains a well-defined excitation. Interestingly, this ratio is numerically very close to the value predicted by the shallow potential approximation.

IV. SUMMARY AND CONCLUSIONS

Because of the presence of the gapless subbands, excitons in metallic carbon nanotubes acquire unique features that distinguish them from excitons in other solid-state systems. First, the quasi-one-dimensional nature of nanotubes makes screening by conduction electrons less effective than in conventional metals. As a result, the electron-hole interaction remains significant enough to ensure the formation of a bound pair. Second, the separation (in the momentum space) of

gapless states ($m = 0$) from the subbands where the exciton is formed ($|m| = 1$) and the fact that the latter subbands are fully gapped allow the electron and the hole to explore fully the gapped subbands, unlike what happens in a conventional doped semiconductor where filling of the conduction band quickly depletes the number of available electron states. Third, the screening by the gapless states is nonetheless significant enough so that the radius of the exciton is greater than the nanotube radius with the binding energy of the order of 0.1Δ. This allows one to treat excitons as quasi-one-dimensional objects, unlike excitons in semiconducting nanotubes which are neither one-dimensional nor two-dimensional objects. Fourth, the presence of the gapless subbands opens up a channel for exciton attenuation where the electron and hole can scatter off each other into the gapless states. The presence of this channel leads to a considerable broadening of the exciton but not so significant as to smear it away completely.

ACKNOWLEDGMENT

APPENDIX: POLARIZATION FUNCTION OF A METALLIC NANOTUBE

The screening of Coulomb interaction in a nanotube is determined by its polarization function $\Pi_m(i\omega, q)$, which in the zone-folding approximation can be obtained from the polarization function of the underlying two-dimensional graphene crystal. The low-energy excitations in graphene are described by the Dirac Hamiltonian, $\hat{H} = v\boldsymbol{\sigma} \cdot \boldsymbol{p}$, where $\boldsymbol{\sigma}$ is the Pauli matrix acting in the pseudospin space of the two triangular sublattices of carbon atoms. The polarization function,

$$\Pi(i\omega, \mathbf{q}) = NT \operatorname{Tr} \sum_{i\epsilon} \int \frac{d^2 p}{(2\pi)^2} \hat{G}^{(0)}(i\epsilon + i\omega, \mathbf{p} + \mathbf{q})$$
$$\times \hat{G}^{(0)}(i\epsilon, \mathbf{p}), \tag{A1}$$

relates to the product of two Green's functions of π electrons consisting of the contributions from both the upper cone ($\beta = 1$) and lower cone ($\beta = -1$),

$$G^{(0)}(i\epsilon, \mathbf{p}) = \frac{1}{i\epsilon - v\mathbf{p} \cdot \hat{\boldsymbol{\sigma}}} = \frac{1}{2} \sum_{\beta = \pm 1} \frac{1 + \beta\hat{\sigma}_p}{i\epsilon - \beta v p}, \tag{A2}$$

where $\hat{\sigma}_p = \hat{\boldsymbol{\sigma}} \cdot \mathbf{p}/p$ is the projection of the pseudospin Pauli matrix onto the direction of electron momentum. At zero temperature the $T = 0$ $\beta' = -\beta$ terms contribute,

$$\Pi(i\omega, \mathbf{q}) = -\frac{N}{4} \sum_{\beta} \int \frac{d^2 p}{(2\pi)^2} \frac{\beta}{i\omega + v\beta p + v\beta|\mathbf{p} + \mathbf{q}|}$$
$$\times \operatorname{Tr}[(1 - \beta\sigma_{\mathbf{p}})(1 + \beta\sigma_{\mathbf{p+q}})]. \tag{A3}$$

The retarded counterpart of the polarization operator can be obtained through analytical continuation $i\omega \to \omega + i\eta$. However, since virtual transitions avoid all singular poles, a small imaginary constant $i\eta$ can be disregarded for our purposes,

$$\Pi(\omega, \mathbf{q}) = \frac{N}{8\pi^2} \sum_{\beta} \int_{-\infty}^{\infty} d^2 p \frac{\beta[\cos(\theta_{\mathbf{p+q}} - \theta_{\mathbf{p}}) - 1]}{\omega + v\beta p + v\beta|\mathbf{p} + \mathbf{q}|}. \tag{A4}$$

Within the zone-folding approximation, which ignores any curvature effects on the electronic spectrum arising from the rolling of the graphene sheet, the polarization function for a nanotube is obtained by quantizing the circumferential momenta, $q_y = m/R$, $p_y = n/R$, and replacing the integral with the sum, $R \int dp_y \to \sum_m$. Additionally, to relate the resulting polarization function to the one-dimensional density, the quantity (A4) should be multiplied by the factor $2\pi R$. Expressing the cosine function in terms of the momentum component, we obtain

$$\Pi_m(\omega, q_x) = \frac{N}{2\pi} \sum_n \int_{-\infty}^{\infty} dp_x \left[1 - \frac{(p_x + q_x)p_x + (m+n)n/R^2}{\sqrt{(p_x^2 + n^2/R^2)}\sqrt{(p_x + q_x)^2 + (m+n)^2/R^2}} \right]$$
$$\times \frac{v\sqrt{(p_x^2 + m^2/R^2)} + v\sqrt{(p_x + q_x)^2 + (m+n)^2/R^2}}{\omega^2 - [v\sqrt{(p_x^2 + m^2/R^2)} + v\sqrt{(p_x + q_x)^2 + (m+n)^2/R^2}]^2}. \tag{A5}$$

Of interest to us here are the polarization function for $m = 0$ and $m = \pm 1$. For $m = 0$ and $q_x \ll 1/R$, only the $n = 0$ terms should be retained. For example, for $q_x > 0$, the integral only extends over the interval $-q_x < p_x < 1$ (where the integrand does not depend on p_x) and the expression (9) is recovered, the subscript in q_x being omitted. (The same result follows for $q_x < 0$.)

For $m = 1$, because only the static and homogeneous limit is important for our purposes, one can set $\omega = 0$ and $q_x = 0$ in Eq. (A5). The remaining p_x integral is independent of R and can be (together with the summation over m) calculated numerically. This yields

$$\Pi_1(0, 0) = -1.16 \frac{N}{\pi v}, \tag{A6}$$

which yields the value of the screened U_1 interaction as in Eq. (11).

Note that the dominant contribution into the polarization function (A6) comes from the lowest-order virtual transitions, $n = -1$ and $n = 0$. Retaining only these contributions, one would obtain the estimate $\Pi_1(0, 0) = -\frac{N}{\pi v}$.

Another good order-of-magnitude estimate could be obtained from the polarization function for graphene, $\Pi(\omega, q) = -Nq^2/(16\sqrt{q^2 v^2 - \omega^2})$. In the static limit, $\omega = 0$, replacing $q \to 1/R$, and multiplying the result by $2\pi R$, as explained above, one would obtain, $\Pi_1(0, 0) = -\frac{\pi N}{8v}$, which overestimates the exact value (A6) by only 6%.

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

[2] F. Wang, D. J. Cho, B. Kessler, J. Deslippe, P. J. Schuck, S. G. Louie, A. Zettl, T. F. Heinz, and Y. R. Shen, Phys. Rev. Lett. **99**, 227401 (2007).

[3] P. May, H. Telg, G. Zhong, J. Robertson, C. Thomsen, and J. Maultzsch, Phys. Rev. B **82**, 195412 (2010).

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

[5] C. D. Spataru, S. Ismail-Beigi, L. X. Benedict, and S. G. Louie, Phys. Rev. Lett. **92**, 077402 (2004).

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

[7] S. Uryu and T. Ando, Phys. Rev. B **77**, 205407 (2008).

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

[9] T. Ando and S. Uryu, Phys. Status Solidi C **6**, 173 (2009).

[10] R. Loudon, Am. J. Phys. **27**, 649 (1959).

[11] M. Andrew, Am. J. Phys. **34**, 1194 (1966).

[12] L. K. Haines and D. H. Roberts, Am. J. Phys. **37**, 1145 (1969).

[13] M. Andrew, Am. J. Phys. **44**, 1064 (1976).

[14] J. F. Gomes and A. H. Zimerman, Am. J. Phys. **48**, 579 (1980).

[15] R. E. Moss, Am. J. Phys. **55**, 397 (1987).

[16] T. Koyama, S. Shimizu, T. Saito, Y. Miyata, H. Shinohara, and A. Nakamura, Phys. Rev. B **85**, 045428 (2012).

[17] M. K. Kostov, M. W. Cole, and G. D. Mahan, Phys. Rev. B **66**, 075407 (2002).

[18] T. G. Pedersen, Phys. Rev. B **67**, 073401 (2003).

[19] E. G. Mishchenko, A. V. Andreev, and L. I. Glazman, Phys. Rev. Lett. **87**, 246801 (2001).

[20] I. E. Dzyaloshinskii and A. I. Larkin, Sov. Phys. JETP **38**, 202 (1974).