The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/175443

Please be advised that this information was generated on 2019-12-19 and may be subject to change.
Super-Planckian electron cooling in a van der Waals stack

Alessandro Principi,1 Mark B. Lundeberg,2 Niels C.H. Hesp,2
Klaas-Jan Tielrooij,2 Frank H.L. Koppens,2,3 and Marco Polini4

1 Radboud University, Institute for Molecules and Materials, NL-6525 AJ Nijmegen, The Netherlands
2 ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain
3 ICREA-Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain
4 Istituto Italiano di Tecnologia, Graphene Labs, Via Morego 30, I-16163 Genova, Italy

Radiative heat transfer (RHT) between macroscopic bodies at separations that are much smaller than the thermal wavelength is ruled by evanescent electromagnetic modes and can be orders of magnitude more efficient than its far-field counterpart, which is described by the Stefan-Boltzmann law. In this Letter we present a microscopic theory of RHT in van der Waals stacks comprising graphene and a natural hyperbolic material, i.e. hexagonal boron nitride (hBN). We demonstrate that RHT between hot carriers in graphene and hyperbolic phonon-polaritons in hBN is extremely efficient at room temperature, leading to picosecond time scales for the carrier cooling dynamics.

Introduction.—The cooling stages of the temperature dynamics of hot carriers in a crystal typically proceed via energy transfer to phonons [1]. In the case of pristine graphene, ultra-long cooling times, on the order of nanoseconds, have been theoretically predicted [2, 3]. Such slow cooling dynamics is due to energy transfer to graphene acoustic phonons via collisions that conserve momentum. If realized experimentally, this intrinsic relaxation dynamics would imply notable figures of merit for graphene-based photodetectors [4]. Unfortunately, the cooling dynamics in “first-generation” graphene samples [5], i.e. samples deposited on SiO2, is believed to be dominated by far more efficient disorder-assisted momentum-non-conserving collisions between electrons and graphene acoustic phonons [6–9]. According to theory [8, 9], such “supercollisions” are due to short-range (rather than long-range) disorder.

It is therefore not clear yet how to reach the intrinsic regime [2, 3]. In this respect, a natural question arises: What is the fate of the temperature dynamics of hot carriers in “second-generation” samples [10], where graphene is encapsulated between hexagonal boron nitride (hBN) crystals [11–16]? On the one hand, these samples have shown nearly ideal transport characteristics [11–16], whereby momentum-conserving electron-acoustic phonon scattering [17, 18] fully determines dc transport times at room temperature, at least for sufficiently large carrier densities. On the other hand, hBN crystal slabs are known to support low loss standing Fabry-Pérot phonon-polaritons [20–22]. These modes occur because hBN is a uniaxial crystal with intrinsic hyperbolic character [19], i.e. with in-plane ($\epsilon_1$) and out-of-plane ($\epsilon_2$) components of the dielectric tensor $\epsilon$ having opposite signs in the so-called “reststrahlen” frequency bands.

Could radiative heat transfer (RHT) to hyperbolic phonon-polaritons in hBN significantly affect the late stages of the cooling dynamics of hot carriers in graphene? In this Letter we answer this question affirmatively.

RHT between macroscopic bodies has been studied since they early days of 1900, when Planck explained the black-body radiation spectrum. In the regime $d \gg d_T$, where $d$ is the separation between two bodies and $d_T = h c / (k_B T)$ is the thermal wavelength, RHT is due to traveling electromagnetic (EM) waves and is controlled by the Stefan-Boltzmann law. On the contrary, in the limit $d \ll d_T$ RHT is dominated by evanescent modes of the EM field and power transfer can greatly exceed the black-body limit ("super-Planckian" thermal emission). Anomalous RHT between closely spaced bodies was first studied experimentally by Hargreaves [23]. This pioneering work motivated the introduction of a general theoretical formalism [24], which was based on the theory of fluctuating electrodynamics [25–27]. Near-field thermal coupling has been intensively studied in the past both experimentally [28–33] and theoretically [34–41].

In this work we present a theoretical study of RHT between a two-dimensional (2D) gas of hot massless Dirac fermions (MDFs) in graphene [42] and nearby slabs of hBN crystals. We follow an approach that differs from fluctuating electrodynamics [24–27] and relies on a combination of Fermi’s golden rule with an analytic electrostatic calculation of how phonon-polaritons modify the instantaneous free-space photon propagator. We have, however, checked (not shown in this work) that Eqs. (15)-(16) can also be obtained from fluctuating electrodynamics. We demonstrate that the hyperbolic nature of hBN crystals sets an extremely efficient intrinsic pathway for the dissipation of heat stored by graphene carriers. Cooling into non-hyperbolic polar substrates has been studied e.g. in Ref. 43. Below, we set $\hbar = k_B = 1$, unless explicitly stated otherwise.

Electron-photon coupling and the hBN-dressed photon propagator.—RHT occurs because of the coupling between MDFs in graphene and the surrounding three-dimensional (3D) EM field. This coupling is described...
by the usual light-matter interaction Hamiltonian

$$H_{\text{eph}} = \frac{e}{cV} \sum_{q,\nu,\omega} \mathbf{j}_{-\mathbf{q}} \cdot \mathbf{A}_{q,\nu,\omega}(t),$$  

(1)

where $-e$ is the electron charge, $c$ is the speed of light in vacuum, $V$ is the 3D quantization volume, $\nu = \text{TM, TE}$ is the polarization of the EM field, and $\mathbf{j}_q = \sum_{k,\lambda,\lambda'} c_{k,\lambda}^{\dagger} \mathbf{J}_{k-q,\lambda,\lambda'} c_{k',\lambda'}$ is the MDF particle current operator [42]. Here, $c_{k,\lambda}$ ($c_{k,\lambda}$) creates (destroys) an electron with momentum $k$ in band $\lambda = \pm$, and $\mathbf{J}_{k,\lambda}$ are the matrix elements of the current operator in the band representation. The intra-band matrix element is $\mathbf{J}_{k,\lambda}$, $\lambda = \lambda' = \nu$V $\left[\cos(\theta_{k,\lambda}/2), \sin(\theta_{k,\lambda}/2)\right]$, while the inter-band one is $\mathbf{J}_{k,\lambda} \lambda' = i\hat{z} \times \mathbf{J}_{k,\lambda}$ $\lambda'$. Here $\theta_{k,\lambda} \equiv \theta_k + \theta_{\lambda'}$ and $\lambda' = \pm \lambda$. Note that the current operator $\mathbf{j}_q$ (photon field $\mathbf{A}_{q,\nu,\omega}$) in Eq. (1) is represented by a 2D (3D) vector.

We now introduce the matrix element of the light-matter interaction,

$$u_{k,q,\nu} = \frac{2\pi e^2}{\omega_{q,\nu}} e^{i\nu} \cdot \mathbf{J}_{k,\lambda}$$

(2)

where $\omega_{q,\nu}$ are the frequencies of the photonic modes, and the second-quantized expression for the Fourier components of the photon field,

$$\mathbf{A}_{q,\nu,\omega}(t) = \sqrt{\frac{2\pi e^2}{\omega_{q,\nu}}} e^{i\nu} [a_{q,\nu}(t) + a_{-q,-\nu}(t)]$$

(3)

where $a_{q,\nu}(t)$ ($a_{\nu,\omega}(t)$) creates (destroys) a photon with momentum $q$ and polarization $\nu$ at time $t$. The corresponding 3D photon propagator

$$G_{\nu,\alpha,\beta}^{3D}(q, q, t) = -\frac{e^2}{c^2} (TA_{q,\nu,\alpha}(t)A_{q,\nu,\beta}^{\dagger})$$

(4)

and its Fourier transform

$$G_{\nu,\alpha,\beta}(q, q, \omega) \equiv \int_{-\infty}^{+\infty} dt e^{i\omega t} G_{\nu,\alpha,\beta}^{3D}(q, q, t)$$

(5)

contain all the necessary information. Here, "$T" denotes the time-ordering operator, $\alpha, \beta = x, y, z$ are Cartesian indices, and ($\ldots$) denotes an average over the thermal ensemble. For photons in free space, the propagator (5) reduces to

$$G_{\nu,\alpha,\beta}^{3D}(q, q, \omega) = -\frac{4\pi e^2}{(\omega + i\eta)^2} e^{i\nu} e^{i\nu},$$

(6)

where $e_{\nu,\alpha}$ is the $\alpha$-th component of the polarization vector corresponding to the photonic mode $\nu$.

Because of the phonons in the nearby hBN crystals, however, the bare propagator (6) is substantially altered. This “dressing” can be easily captured analytically in the non-retarded limit, where we can use the following relation between the vector potential and the instantaneous electrostatic potential: $\partial_t \mathbf{A}_{q,\nu,\omega} = icq \phi(q, q, t)$, for $\nu = \text{TM}$. This identity implies that, in the non-retarded limit, the 3D dressed propagator $G_{\nu,\alpha,\beta}^{3D}(q, q, t)$ can be calculated from the knowledge of the 3D instantaneous Coulomb propagator. The latter, in turn, can be calculated by utilizing a straightforward electrostatic approach described in Ref. 22. Finally, the required 2D propagator can be obtained after integrating over $q$. The end result is

$$\text{Im}[G_{\nu,\alpha,\beta}^{2D}(q, \omega)] = \begin{cases} \frac{q \eta q^2}{\omega^2} \lambda_{\nu,\omega}, & \text{if } \nu = \text{TM}, \\ 0, & \text{if } \nu = \text{TE} \end{cases}$$

(7)

where $V_{q,\omega}$ is the instantaneous Coulomb propagator dressed by the presence of the surrounding hBN dielectric slabs. Its frequency dependence stems from the frequency dependence of the in-plane and out-of-plane permittivities $\varepsilon_x(\omega)$ and $\varepsilon_z(\omega)$ of hBN. Explicit expressions for $V_{q,\omega}$ and the permittivities $\varepsilon_x(\omega)$ and $\varepsilon_z(\omega)$ are reported in Eqs. (A1)-(A2) below. As discussed in Ref. 22, poles of the dressed Coulomb interaction $V_{q,\omega}$ yield the dispersion relation of standing phonon-polariton modes in the hBN slabs surrounding graphene. We now proceed to calculate RHT between hot carriers in graphene and the dressed EM field around graphene.

Boltzmann-transport theory of RHT.—The calculation of the 2D dressed propagator (7) allows us to calculate how phonon-polaritons dress the squared matrix element of the light-matter interaction. We start by squaring the bare matrix element in Eq. (2):

$$|u_{k, q, \nu} |^2 = \frac{2\pi e^2}{\omega_{q, \nu}} \sum_{\alpha, \beta = x, y} J_{k, \lambda, k+q, \lambda'} J_{k+q, \lambda', k} e^{i\nu} e^{i\nu},$$

(8)

and its Fourier transform

$$\text{Im}[G_{\nu,\alpha,\beta}^{2D}(q, \omega)] = \int_{-\infty}^{+\infty} \frac{d\omega}{\pi} \text{Im}[G_{\nu,\alpha,\beta}^{3D}(q, q, \omega)].$$
electron distribution function \( f_{k,\lambda} [2, 3, 8, 9] \). Here,
\[
\mathcal{L}_{k,\lambda} = \sum_{k',\lambda'} \left[ f_{k,\lambda}(1 - f_{k',\lambda'}) W_{k,\lambda \rightarrow k',\lambda'} - f_{k',\lambda'}(1 - f_{k,\lambda}) W_{k',\lambda' \rightarrow k,\lambda} \right],
\]
and the transition probability
\[
W_{k,\lambda \rightarrow k',\lambda'} = 2\pi \sum_{q,\nu} \left| U_{k,\lambda}^{q,\nu} \right|^2 \left[ (n_{q,\nu} + 1) \delta(\Delta q \varepsilon - \omega_{q,\nu}) + \delta(k - k' - q) + n_{q,\nu} \delta(\Delta q \varepsilon + \omega_{q,\nu}) \right] \times \delta(k - k' + q),
\]
where \( n_{q,\nu} \) is the phonon-polariton distribution function and \( \Delta q \varepsilon = \varepsilon_{k,\lambda} - \varepsilon_{k',\lambda'} \) the electronic transition energy.

We are now in the position to calculate the energy transfer rate, which we define to be positive if \( T_e > T_L \). Multiplying both members of the Boltzmann equation by \( \varepsilon_{k,\lambda} - \mu \) and summing over \( k, \lambda \) we find an equation of motion for the energy density \( \varepsilon \):
\[
\partial_t \varepsilon = -\dot{Q},
\]
where the energy transfer rate is given by
\[
\dot{Q} = -\sum_{\nu} \int \frac{d^2 q}{(2\pi)^2} \int_{-\infty}^{\infty} \frac{d\omega}{\omega} \left[ n_B(\omega/T_L) - n_B(\omega/T_e) \right] \times \text{Im}[\chi^{(0)}_{\nu,\alpha,\beta}(q,\omega)] \text{Im}[G^{(0)}_{\nu,\alpha,\beta}(q,\omega)].
\]

Here, \( \chi^{(0)}_{\alpha,\beta}(q,\omega) \) is the current-current response tensor of a 2D system of \textit{non-interacting} MDFs [44]. To obtain Eq. (13) we assumed that both electrons and phonon-polaritons are at equilibrium at the two temperatures \( T_e \) (electron temperature) and \( T_L \) (lattice temperature), respectively. Therefore, \( f_{k,\lambda} \) and \( n_{q,\nu} \) are equilibrium distribution functions: \( f_{k,\lambda} \equiv n_B((\varepsilon_{k,\lambda} - \mu)/T_e) \) and \( n_{q,\nu} \equiv n_B(\omega_{q,\nu}/T_L) \), where \( n_B(x) = (e^x \pm 1)^{-1} \) is the Fermi-Dirac (Bose-Einstein) distribution. The chemical potential \( \mu = \mu(T_e) \) is obtained by requiring the particle density \( n \) to be time-independent.

We finally bootstrap Eq. (13) by introducing dynamical screening at the level of the random phase approximation (RPA) [45]. This is easily accomplished by the formal replacement
\[
\text{Im}[\chi^{(0)}_{\alpha,\beta}(q,\omega)] \rightarrow \frac{\text{Im}[\chi^{(0)}_{\alpha,\beta}(q,\omega)] q_{\alpha} q_{\beta}}{\varepsilon(q,\omega)^2 q^2} + \text{Im}[\chi^{(0)}_{\alpha,\beta}(q,\omega)] \left( \delta_{\alpha\beta} - \frac{q_{\alpha} q_{\beta}}{q^2} \right),
\]
where \( \chi^{(0)}_{\alpha,\beta}(q,\omega) \) and \( \chi^{(0)}_{\alpha,\beta}(q,\omega) \) are the longitudinal and transverse current-current response functions, and \( \varepsilon(q,\omega) = 1 - V_{q,\omega}\chi^{(0)}_{\alpha\beta}(q,\omega) \), with \( \chi^{(0)}_{\alpha\beta}(q,\omega) = \varepsilon(q,\omega)^2 \chi^{(0)}_{\alpha\beta}(q,\omega)/\omega^2 \) the density-density response function [46], is the dynamical RPA screening function [22].

Using Eq. (14) in Eq. (13) and restoring \( \hbar \) and \( \hbar_B \), we finally find the desired expression for the energy transfer rate:
\[
Q = \frac{\hbar}{4} \int \frac{d^2 q}{(2\pi)^2} \int_{-\infty}^{\infty} \frac{d\omega}{\omega} [n_B(\omega_e) - n_B(\omega_L)] Z(q,\omega),
\]
where \( Z(q,\omega) \) is the dynamical screening function [47].

\[\text{FIG. 1. (Color online) Color plots of the dimensionless function } Z(q,\omega) \text{ in Eq. (16) for an electron temperature } T_e = 300 \text{ K and a top (bottom) hBN thickness } d' = 9 \text{ nm (d = } 27 \text{ nm). All other heterostructure and hBN parameters (phonon frequencies and lifetimes) are reported in the Appendix below. Panel (a) } n = 5.0 \times 10^{12} \text{ cm}^{-2}. \text{ Panel (b) } n = 1.0 \times 10^{12} \text{ cm}^{-2}. \text{ Dashed lines indicate the edges of the graphene intra-band and inter-band particle-hole continua [42]. In panel (a), the bottom edge of the inter-band continuum is not present since it occurs well above the hBN reststrahlen bands for } n = 5 \times 10^{12} \text{ cm}^{-2}. \text{ In both panels, we clearly see that } Z(q,\omega) \text{ is maximum at the location of the poles of the dressed Coulomb interaction } V_{q,\omega} \text{ (standing hBN phonon-polariton modes [20, 21]) and at the zeroes of the dynamical dielectric function } \varepsilon(q,\omega) \text{ (plasmon-phonon polariton branches [15, 22]).} \]
where $\omega_{c,L} \equiv \hbar \omega/(k_B T_{c,L})$ and

$$
\mathcal{Z}(q, \omega) \equiv 4 \sum_{n} \left| V_{q,\omega} \right|^2 \text{Im} \left[ \chi_{n,0}(0, q, \omega) \right].
$$

A color plot of the real function $\mathcal{Z}(q, \omega)$ for typical values of microscopic parameters is reported in Fig. 1. Eqs. (15)-(16) are the most important results of this work. The transverse part of the current–current response function in Eq. (7) drops out of the problem since the non-retarded 2D propagator (7) is purely longitudinal. We note that the quantity $\mathcal{Z}(q, \omega)$ in Eq. (16) is dimensionless and bounded, $0 \leq \mathcal{Z}(q, \omega) \leq 1$. The super-Planckian nature of the energy transfer rate (15) stems from contributions to the integral coming from phonon-polariton modes with $q \gg \omega/c$, the only natural short-wavelength cut-off ($\sim k_F$) for the integral being provided by the graphene response function $\text{Im}[\chi_{n=0}^{(0)}(q, \omega)]$.

Cooling time and temperature dynamics.—The temperature dynamics $T_e(t)$ can be calculated from the differential equation (12), provided that we introduce the heat capacity. Indeed, using that $\mathcal{E} = \sum_{k,\lambda} (\varepsilon_{k,\lambda} - \mu) f_{k,\lambda}$, we can explicitly calculate $\partial_t \mathcal{E}$. We find $\partial_t \mathcal{E} = C_n \partial_t T_e$, where

$$
C_n = \sum_{k,\lambda} \left( \frac{\partial f_{k,\lambda}}{\partial \varepsilon_{k,\lambda}} \right)^2 \left[ \varepsilon_{k,\lambda} - \mu(T_e) \right]^2 \frac{k_B T_e^2}{k_B T_e^2} + \frac{\partial \mu(T_e)}{\partial T_e} \sum_{k,\lambda} \left( \frac{\partial f_{k,\lambda}}{\partial \varepsilon_{k,\lambda}} \right) \frac{\varepsilon_{k,\lambda} - \mu(T_e)}{k_B T_e}
$$

is the heat capacity at a constant density $n$. Using Eq. (17), we can rewrite Eqs. (12)-(15) as $C_n \partial_t T_e = -Q$. We now note that we can formally rewrite the latter equation as

$$
\partial_t T_e = -\frac{T_e - T_L}{\tau(T_e, T_L)},
$$

provided that we introduce $\tau(T_e, T_L) \equiv C_n (T_e - T_L)/Q$. Eq. (18) can be integrated numerically, as we will discuss below. However, in the limit $\Delta T = T_e - T_L \to 0$, one can expand the energy transfer rate $Q$ for small values of $\Delta T$ and the ratio $(T_e - T_L)/Q$ does not depend on $T_e$, i.e.

$$
\tau^* \equiv \lim_{\Delta T \to 0} \tau(T_e, T_L) = \left( \frac{1}{C_n \partial Q} \right)^{-1}.
$$

In this case, Eq. (18) implies a simple exponential decay, $T_e(t) = T_e(0) \exp(-t/\tau^*)$ and $\tau^*$ acquires the physical meaning of a cooling time scale.

Numerical results for $\tau^*$ are shown in Figs. 2 and 3. In particular, in Fig. 2(a) we plot $\tau^*$ as a function of carrier density $n$, for different values of the electron temperature $T_e = T_L$, while in Fig. 2(b) we plot $\tau^*$ as a function of $T_e$, for different values of $n$. Note that, in the limit of $n \to 0$, $\tau^*$ saturates to a constant, since, in the same limit, the heat capacity converges to a function that depends only on the electron temperature ($C_n \propto T_e^2$) and

FIG. 2. (Color online) Panel (a) The cooling time $\tau^*$ (19) as a function of the electron density $n$, for different values of the electron temperature $T_e$. Panel (b) The cooling time $\tau^*$ as a function of $T_e$ for different values of $n$.
so does $\partial Q/\partial T_e$. For most values of the electron density away from the $n = 0$ charge-neutrality point (CNP), $\tau^*$ shows a weak dependence on $n$, because of a cancellation that we now proceed to discuss. Because the integrand in Eq. (15) is proportional to $\text{Im}\{\chi^{(0)}_{\text{nn}}(q, \omega)\}$, we can separate out contributions to $Q$ that are due to intra-band (i.e. $\omega < \nu_{Fq}$) and inter-band excitations (i.e. $\omega > \text{max} (\nu_{Fq}, 2\varepsilon_F / h - \nu_{Fq})$). There is also a contribution due to plasmon-phonon polaritons (zeros of $\varepsilon(q, \omega)$, Ref. 22), which we define by considering contributions to the frequency integral in Eq. (15) coming from the $T = 0$ Pauli-blocking gap, i.e. $\nu_{Fq} < \omega < 2\varepsilon_F / h - \nu_{Fq}$. These three contributions to the cooling rate $1/\tau^*$ are shown in Fig. 3(a). We clearly see that the intra-band contribution is dominant for most values of the carrier density, with the exclusion of the low-density regime, where intra- and inter-band contributions become comparable in magnitude. Note also that the increase with $n$ of the intra-band contribution is nearly exactly cancelled by a decrease with $n$ of the inter-band contribution. This explains the weak dependence of $\tau^*$ on $n$ away from the CNP displayed in Figs. 2(a). The contribution due to the plasmon-phonon polariton branch is negligible. In Fig. 3(b), we show the cooling efficiency of the process investigated in this work by comparing RHT into hyperbolic phonon-polaritons (solid line) to RHT into non-hyperbolic phonon-polaritons (grey-shaded area). The latter is calculated by using Eqs. (15), (16), and (19) one time with $\epsilon_x(\omega) \rightarrow \epsilon_z(\omega)$ in the equation for $V_{q,\omega}$, and one time with $\epsilon_z(\omega) \rightarrow \epsilon_x(\omega)$. These replacements make sure that the crystal slabs surrounding graphene are non-hyperbolic. We clearly see that RHT into standing hyperbolic phonon-polariton modes is far more efficient.

Before concluding, we would like to discuss temperature dynamics in the overheating $T_e \gg T_L$ regime. As we have seen above, for $T_e \approx T_L$ the function $T_e(t)$ is an exponential with time scale $\tau^*$. This exponentially fast equilibration does not occur, however, for $T_e \gg T_L$. In this case, the temperature dynamics $T_e(t)$ can be found by solving Eq. (18) with an initial condition, $T_e(0)$. In Fig. 4(a) we show that the solution of this equation for $T_e(0) = 1000 \, \text{K} \gg T_L = 1 \, \text{K}$ (solid line) displays a slow decay. Note that, even after 30 ps, the electrons are not equilibrated with the lattice. On the contrary, for $T_L = 300 \, \text{K}$, the dynamics is exponential (dashed line).

In summary, we have presented a theory of near-field thermal radiation transfer between hot carriers in graphene and hyperbolic phonon-polaritons in hBN slabs. Our theory is relevant for understanding the cooling dynamics in ultra-clean encapsulated samples [10–16], where extrinsic mechanisms [6, 7] due to disorder are expected not to be at play. We have discovered that hyperbolic-phonon polaritons in hBN are extremely efficient heat sinks for hot carriers in graphene at room temperature, leading to picosecond time scales for the carrier cooling dynamics in graphene. Within the realm of high-quality samples, this understanding offers a pathway to tuning cooling times by the hBN thickness, which controls the standing phonon-polaritons shown in Fig. 1. Thinner hBN slabs tend to lengthen the cooling time, which is a relevant direction for greatly improving the sensitivity of photodetectors [4].

This work was supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 696656 “GrapheneCore1”, Fondazione Istituto Italiano di Tecnologia, the Spanish Ministry of Economy and Competitiveness through the “Severo Ochoa” Programme for Centres of Excellence in R&D (SEV-2015-0522), Fundacio Cellex Barcelona, the Mineco grants Ramón y Cajal (RYC-2012-12281), Plan Nacional (FIS2013-47161-P), a “Young Investigator Grant” (FIS2014-59639-JIN), the Government of Catalonia trough the SGR grant (2014-SGR-1535), the ERC StG “CarbonLight” (307806), and the ERC AdG “FEMTO/NANO” (338957). M.B.L. wishes to thank M. Jablan for useful conversations. M.P. is extremely grateful for the financial support granted by ICFO during a visit in August 2016.

**Electrostatics of the hBN/Graphene/hBN vertical stack**

We consider a vertical heterostructure composed of: 1) a graphene sheet located at $z = 0$; 2) a homogeneous but uniaxial insulator of thickness $d'$ with dielectric tensor $\tilde{\varepsilon} = \text{diag}(\varepsilon_x, \varepsilon_x, \varepsilon_z)$ placed above graphene; 3) a homogeneous but uniaxial insulator of thickness $d$ with dielectric tensor $\varepsilon = \text{diag}(\varepsilon_x, \varepsilon_x, \varepsilon_z)$ placed below graphene; 4) homogeneous and isotropic insulators with dielectric constants $\varepsilon_a$ and $\varepsilon_b$ filling the two half-spaces $z > d'$ and $z < -d$, respectively.
We calculate the electrical potential created by an electron sitting at the origin in graphene by following the approach of Ref. 22. The three-dimensional displacement field $D(r, z)$ in cylindrical coordinates must satisfy the condition $\nabla \cdot D(r, z) = 0$ everywhere in space.

\[ V_{q,\omega} = v_q \left\{ \frac{\sqrt{\epsilon_x \epsilon_z}}{\sqrt{\epsilon_x \epsilon_z}} + (\epsilon_a + \epsilon_b) \tanh \left[ q \frac{\sqrt{\epsilon_x}}{\epsilon_z} (d + d') \right] + (\epsilon_b - \epsilon_a) \frac{\sinh \left[ q \frac{\sqrt{\epsilon_x}}{\epsilon_z} (d - d') \right]}{\cosh \left[ q \frac{\sqrt{\epsilon_x}}{\epsilon_z} (d + d') \right]} \right\} \]

where $v_q = 4\pi e^2/[q(\epsilon_a + \epsilon_b)]$ and $\bar{\epsilon} = (\epsilon_a \epsilon_b + \epsilon_x \epsilon_z)/(\epsilon_a + \epsilon_b)$. In the limit $d' \to 0$, Eq. (A1) reduces to a recently derived result [22].

The frequency dependence of the dressed Coulomb interaction $V_{q,\omega}$ is due to optical phonons in the hBN slabs [22]. Indeed, in the case of hBN, the components of the uniaxial dielectric tensor have an important dependence on frequency in the mid infrared, which is usually parametrized in the following form

\[ \epsilon_{\ell}(\omega) = \epsilon_{\ell,\infty} + \frac{\epsilon_{\ell,0} - \epsilon_{\ell,\infty}}{1 - (\omega / \omega_{T,\ell})^2 + i \gamma_{\ell} / \omega / (\omega_{T,\ell}^2) \omega^2}, \quad (A2) \]

with $\ell = x$ or $z$. Here $\epsilon_{\ell,0}$ and $\epsilon_{\ell,\infty}$ are the static and high-frequency dielectric constants, respectively, while $\omega_{T,\ell}^2$ is the transverse optical phonon frequency in the direction $\ell$. The longitudinal optical phonon frequency $\omega_{L,\ell}^T$ satisfies the Lyddane-Sachs-Teller relation $\omega_{L,\ell}^T = \omega_{T,\ell}^T \sqrt{\epsilon_{\ell,0}/\epsilon_{\ell,\infty}}$. The parameter $\gamma_{\ell}$ represents hBN phonon losses (in meV).

All the calculations reported in the main text have been done with the dielectric parameters reported in Table 1. Also, we have taken $\epsilon_a = 1$, $\epsilon_b = 3.9$, $d' = 9$ nm, and $d = 27$ nm.

However, the presence of an electron with charge density $-e\delta^2(r)\delta(z)$ at $z = 0$ implies a discontinuity of the normal component $D_z$ of the displacement field across $z = 0$, while the tangential components $E_x, E_y$ of the electric field $E(r, z)$ must be continuous.

Since the electric field $E(r, z)$ is irrotational everywhere in space, we can introduce the electric potential $\phi(r, z)$ in the four regions of space $z > d'$, $d' < z < 0$, $-d < z < 0$, and $z < -d$. The Laplace equation $-\epsilon_z \partial^2\phi(r, z) = -\epsilon_y \partial^2\phi(r, z) - \epsilon_x \partial^2\phi(r, z) = 0$ in the uniaxial dielectrics (i.e. for $d' < z < 0$ and $-d < z < 0$) can be reduced [47] to an ordinary Laplace equation by scaling $x \to x/\sqrt{\epsilon_x}$, $y \to y/\sqrt{\epsilon_y}$, and $z \to z/\sqrt{\epsilon_z}$.

Imposing the aforementioned boundary conditions and carrying out tedious but elementary algebraic steps, we find the following expression for the dressed 2D Coulomb interaction on the real-frequency axis:

\[ \epsilon_{\ell}(\omega) = \epsilon_{\ell,\infty} + \frac{\epsilon_{\ell,0} - \epsilon_{\ell,\infty}}{1 - (\omega / \omega_{T,\ell})^2 + i \gamma_{\ell} / \omega / (\omega_{T,\ell}^2) \omega^2}, \quad (A2) \]


