Dominant Role of Two-Photon Vertex in Nonlinear Response of Dirac Materials

Habib Rostami1,∗ and Emmanuele Cappelluti2
1 Nordita, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden
2 Istituto di Struttura della Materia (ISM), CNR, 34149 Trieste, Italy

Using a conserving Baym-Kadanoff approach, we present a fully compelling theory of nonlinear dc response of a Dirac system to electric fields in the presence of disorder scattering. We show that the nonlinear terms are strikingly ruled by the appearance of a dominant two-photon vertex which is absent at the bare level and finite even in the weak-coupling limit. Such two-photon vertex self-generation highlights the crucial role of the frequency and field dependence of the scattering rates in the nonlinear regime. Our study reveals a novel many-body mechanism in the nonlinear response of Dirac materials whose effects are predicted to be observable.

Due to their linear dispersion, \( c_k \sim |k| \), and to the underlying chiral structure, Dirac materials show a variety of exotic features that makes them a versatile platform for theoretical investigations of new physics and for application purposes. Despite the complex physics, many properties of these materials are often rationalized using concepts of non-interacting particle or semi-classical model [1–4]. For instance, a standard transport model is conventionally applied for the dc conductivity in highly-doped graphene (Boltzmann regime), where the mobility is evaluated at the non-interacting level, and the interactions enter only through the effective parameter known as transport scattering rate \( \Gamma_{tr} \) [5]. At odds with the above scenario, there is a wide consensus that the quantum regime (low-energy transitions in undoped Dirac model) is much more complex and it might be significantly affected by many-body effects [6].

The nonlinear electromagnetic response of Dirac materials has attracted recently a considerable interest in two [7–14] and three dimensions [15–20]. Widely investigated are the nonlinear optical properties and in particular the appearing of four-wave mixing, nonlinear Kerr effect, second and third-harmonic-generation in single-layer graphene, with remarkable technological interest [21–30]. Peculiar of Dirac material is, due to the linear dispersion, the absence of the bare two-photon-electron coupling, which should give rise to the so-called diamagnetic term. The lack of such term prompts several widely debated issues, as the validity of optical sum rules [31–33]. Most of the theoretical descriptions of nonlinear effects rely at the moment upon non-interacting analyses, or semiclassical approaches [23, 24, 34–36] where, in a similar way as in the Boltzmann theory of linear response, the dominant transition processes resemble the ones of the many-body physics, beyond the semi-classical approaches, can drastically change the above scenario, pointing out that different physical processes can be responsible for the relevant properties of the nonlinear dc transport. Analyzing the case of disorder scattering as a basilar benchmark example, we show how non-conserving phenomenological models of scattering intrinsically fail and high-order vertex processes must properly taken into account. More in particular, we show that, despite the bare diamagnetic two-photon vertex (TPV) being null in Dirac materials, the many-body renormalized TPV is finite and relevant and it can play a dominant role. Our results, besides providing a consistent framework for a proper analysis of nonlinear transport and optical response in realistically interacting Dirac materials, open novel perspectives for understanding and predicting new functional properties of these complex promising systems.

We consider the two-dimensional (2D) Dirac Hamiltonian \( \hat{H}_k = \hbar v \sigma \cdot \mathbf{k} - \mu_0 \mathbf{l} \), where \( \mu_0 \) is the bare chemical potential ruling the charge doping. For realistic purposes we consider the paradigmatic case of graphene [37] where \( \sigma = (\sigma_x, \sigma_y) \), where \( \sigma_i \) stand for the Pauli matrices in the spinor space, and \( \mathbf{r} = \pm \) stands for valley index in the Brillouin zone of graphene. In the dipole approximation the light-matter interaction can be modelled by applying the minimal coupling transformation \( h \mathbf{k} \rightarrow h \mathbf{k} + e \mathbf{A}(t) \) where \( \mathbf{A}(t) \) stands for an external vector potential. The corresponding electric field is given by \( \mathbf{E}(t) = -\partial_t \mathbf{A}(t) \). Due to the linear dispersion, the electron-photon coupling does not present a diamagnetic (two-photon) bare term but only the linear coupling:

\[
\mathcal{H}_{\text{light-matter}} = \hbar e v \int d\mathbf{r} \bar{\psi}(\mathbf{r}) \mathbf{A}(t) \psi(\mathbf{r}) .
\] (1)

Without the loss of generality, we assume an electric field along the \( y \) axis. As scattering source we consider long-range impurity centers with standard Born impurity correlations [38–40]. Within this framework we can write the Born impurity self-energy in the complex frequency space: \( \Sigma(z) = \gamma_{\text{imp}} \sum_k \hat{G}(\mathbf{k}, z) \) where the Green’s function follows \( \hat{G}(\mathbf{k}, z) = (z - \hat{H}_k - \Sigma(z))^{-1} \). For isotropic scattering we get a diagonal self-energy in the spinor basis

∗Electronic address: habib.rostami@su.se
as $\hat{\Sigma}(z) = \Sigma(z) \hat{I}$. It is well known that under these conditions the impurity self-energy, as well as the Coulomb and other scattering ones, depends intrinsically on the ultraviolet energy cut-off $W$ representing the range of validity of the Dirac model. In order to provide a conserving approach, this issue needs to be cured by means of a proper regularization [41, 42]. As detailed in the Supplementary Material (SM), we employ standard dimensional regularization [41, 42]. As detailed in the Supplementary Material (SM), we employ standard dimensional regularization.

Conserving approaches, based for instance on a Baym-Kadanoff derivation [43, 44], are fundamental in theoretical physics to ensure that compelling results are obtained. This aim is particularly important in nonlinear response since an arbitrary selection of diagrams can easily lead to spurious conclusions. The choice of the vector-potential gauge, within the paradigmatic Born impurity scattering we consider here, permits us an exact derivation of self-consistent equations (see SM for details [45]) for all the high order processes relevant in the third-order response function which is the leading nonlinear term in centrosymmetric Dirac materials. The diagrammatic expression of the third-order response function is provided in Fig. 1 where, roughly speaking, empty symbols represent $n$-photon vertices ($n = 2, 3$) expressed in terms of the renormalized lower-order vertices (Fig. 1e,f), whereas filled symbols represent the solution of a Bethe-Salpeter-like (BS) self-consistent resummation for a given $n$-photon vertex (Fig. 1b–d). Leaving aside the complexity of the self-consistent set of equations, few relevant things are worth to be underlined here. First of all, we notice that an effective multi-photon coupling is induced by the disorder scattering source even if it is absent in the Hamiltonian at the bare (non-interacting) level (Fig. 1e,f). Second, that the relevance of each $n$-photon vertex is largely governed by the BS many-body resummation as depicted in Fig. 1b–d. This might lead to a reduction (screening) or to an enhancement of different multi-photon scattering depending on the Pauli structure of the corresponding photon vertex, as we discuss more extensively later.

The diagrammatic expressions in Fig. 1 represent in full generality the optical frequency-dependent third-order response function in Dirac materials, including third-harmonic generation, four-wave mixing, etc. For a generic interaction, the effective solution of such coupled equations on the real-frequency axis is a formidable task that does not allow for a practical solution. The focus on the isotropic disorder scattering is on the other hand particularly suitable to investigate many-body effects in nonlinear electromagnetic response since it allows

![Diagram](image-url)
for a set of equations in the Matsubara space which can be generalized in a rigorous way on the real frequency axis, using the well-known procedure of multiple branch cuts in the complex frequency space. The derivation is lengthy and cumbersome but compelling and it is summarized in the SM [45]. We consider first the dc transport limit. Without loss of generality, it is possible to express the linear and the third-order dc conductivity in terms of two dimensionless quantities:

\[
\sigma_{dc}^{(1)} = \sigma_0 f_1 \left( \frac{\mu}{\Gamma(\mu)} ; U \right),
\]

\[
\sigma_{dc}^{(3)} = \sigma_0 \frac{t_0}{E_0} \left[ \Gamma(\mu) \right] f_3 \left( \frac{\mu}{\Gamma(\mu)} ; U \right),
\]

where \( \mu \) is the effective chemical potential \( \mu = \mu_0 - \text{Re}\Sigma(\omega = 0) \) and \( \Gamma(\mu) \) the scattering rate \( \Gamma(\mu) = -\text{Im}\Sigma(\omega = 0) \). \( \sigma_0 \propto e^2/h \) the universal conductivity unit and \( E_0 \propto t_0/ea \) is a characteristic electric field scale determined by inter-atomic hopping energy \( t_0 \) and by the lattice constant \( a \) [37].

It is worth to stress again that Eqs. (3)-(4) are tied together since they must descend in a compelling way from a common approximation for the self-energy. Heretofore, although many approaches for the self-energy have been discussed for the linear response, the third-order response has been analyzed only in the simplistic case of a phenomenological constant scattering rate \( \Gamma(\mu) = \Gamma \). Since such phenomenological non-energy does not depend on the applied external field, the third-order response function reduces to the first “square” diagram of Fig. 1a dropping all the vertex renormalization processes, i.e. replacing the filled circles with empty ones (= bare electron-photon coupling). A similar scheme can as well be employed for the linear response. Under these ultrasmallified conditions, one can see that the linear and third-order dc transport depend uniquely on the semiclassical parameter \( x = \mu/\Gamma(\mu) \), i.e. \( f_3(x; y) = f_3(x) \), \( f_3(x; y) = f_3(x) \). An analytical expression for the functions \( f_1(x) \), \( f_3(x) \) is obtained in the SM [45]. In particular, in the Boltzmann regime one gets results \( f_1(\infty) \approx 2\mu/\pi\Gamma \), \( f_3(\infty) \approx -3\pi\Gamma/32\mu \), implying that nonlinear effects lead to a reduction of the dc conductivity in the Boltzmann regime. A similar analysis is performed in the quantum regime, giving \( f_1(0) \approx 8/\pi^2 \), \( f_3(0) \approx 2/5 \), meaning that nonlinear effects should yield an enhancement of the dc conductivity in the quantum regime.

The above predictions, based on the phenomenological model of a constant scattering rate \( \Gamma \), are challenged when many-body effects are computed in a compelling conserving scheme. In Fig. 2 we show the characteristic dc transport [46] function \( f_3 \) as a function of the semiclassical parameter \( x = \mu/\Gamma(\mu) \), from the extreme quantum limit \( x \ll 1 \) to the Boltzmann regime \( x \gg 1 \). From the computational point of view, since the presence of a finite cut-off energy scale \( W \), the quantum limit can be conveniently investigated by fixing \( U \) and varying \( \mu \), whereas the Boltzmann regime is more easily spanned by fixing \( \mu \) and varying \( U \). Let’s discuss first the Boltzmann regime (Fig. 2b). We notice that a compelling many-body analysis recovers qualitatively (but with an increase in the magnitude of factor 10) the predictions of the phenomenological constant-\( \Gamma \) model with \( f_3(x, y) \approx f_3(x) \propto -1/x \) in the Boltzmann regime \( x \gg x^* \approx 10 \). On the other hand, in the quantum regime \( x \ll x^* \) (Fig. 2a), \( f_3 \) shows a significant dependence on \( \mu_0 \), signaling that the nonlinear dc transport properties are no more governed uniquely by the semiclassical parameter \( x \) but that the detailed value of \( \mu_0 \) (or conversely, of \( U \)) starts playing a relevant role. Some striking things are worth being pointed out: (i) counter-intuitively, the third-order contribution to the dc transport appears to be magnified approaching the clean limit \( U \to 0 \); (ii) there are two isosbestic points (i.e. \( x \)-points where \( f_3 \) does not depend on \( U \)) coinciding in a very good approximation with the zeroes of the \( f_3 \) function; (iii) whereas the phenomenological constant-\( \Gamma \) modelling predicts a a well-determined positive sign of the nonlinear dc correction in the quantum regime (implying an increase of the total conductivity), the sign of the nonlinear terms of the full conserving many-body theory in the quantum regime is not univocally determined, presenting a positive region in the crossover range and a negative sign in the extreme quantum limit.

We can rationalize points (i)-(ii) by assuming that in the quantum regime the nonlinear characteristic dc transport function \( f_3(x, U) \) can be factorized as [45]:

\[
f_3(x; U) \approx \frac{C(x)}{U^{\gamma}},
\]

(with \( \gamma > 0 \)), where the strength \( U \) of the interaction rules governs the intensity of the third-order dc transport, while the semiclassical parameter \( x \) seems to dictate the sign of the third-order correction. To assess the meaningfulness of such description we plot in Fig. 3a in a log-log scale the absolute value of the function \( f_3(x, U) \)
versus $U$ for a representative case $x = 0.03$ in the quantum regime. We find a perfect agreement with a scaling behavior $f_3(x,U) \propto 1/U^\gamma$ with $\gamma$ slightly smaller than 2 ($\gamma \approx 1.82$) signaling that in the clean limit the third-order dc transport is expected to be dominant with respect to the linear one. As detailed in [45], a similar analysis is valid in the whole quantum regime.

In order to gain a full understanding of these novel features, we analyze separately in Fig. 3a,b the relevance of each family of diagrams contributing to the total third-order conductivity as depicted in Fig. 1a. We can thus realize that the contribution of the conventional “square” diagram (which is the only one present in the non-interacting case and for the phenomenological damping model), is essentially marginal, as well as the contribution of the last “bubble” associated with the renormalized three-photon vertex. The dominant role is instead played by the “triangle” diagrams containing the renormalized two-photon vertex. A quantitative analysis shows that each family of diagrams obeys Eq. (5) with an approximately integer exponent (i.e. $\gamma_\star \approx 2$, $\gamma_\triangle \approx 2$, and $\gamma_\square \approx 1$ for the square, triangle and the bubble diagrams). The dominance of the triangle diagrams results thus in an exponent very close to 2 ($\gamma = \gamma_\triangle$). The self-consistent BS renormalization of the TPV (Fig. 1c) is a crucial ingredient in such novel scenario. This can be assessed in Fig. 3a,b where one can see that, once neglected the BS renormalization, the contribution of the triangle diagrams results to be of the same order (or smaller) of that of the conventional square diagram. The dominant role of the TPV renormalization appears even more evident by investigating the scaling of the characteristic third-order dc transport function $f_3(x,U)$ versus $U$. As depicted in Fig. 3a, once replaced the BS renormalized TPV (Fig. 1c) with the “bare” one (Fig. 1e), the third-order dc conductivity scales as $1/U$ ($\gamma_\triangle \approx 1$), with an additional sign change change, as shown in Fig. 3b. This means that the BS renormalization of the TPV gives rise in the quantum regime to an additional dependence $\sim 1/U$ that diverges in the clean limit. The $n$-photon vertex matrix structure, which reads $\Lambda_{\gamma} = (-i\epsilon\sigma_{\gamma})^n \Lambda_{\gamma}^{(0)}$, plays a crucial role in the relevance of the BS renormalization effect. The impressive effect is peculiar of the TPV renormalization $\Lambda_2 = \Lambda_2^{(0)}/[1 - UX_2]$ and does not appear in the BS renormalization of the one-three-photon vertex $A_n = \Lambda_2^{(0)}/[1 - UX_n]$, with $n = 1, 3$ [45]. This different impact can be traced down to the different structure in the Pauli space. As detailed in [45], we get indeed $X_1 = X_3 \propto \text{Tr}[\sigma_y \hat{G}]$, $X_2 \propto \text{Tr}[\hat{G}]$. In the quantum regime, one can thus show that in the dc limit $UX_1 \approx UX_3 \approx U$, whereas $UX_2 \approx 1 + O(U)$, so that resulting in an effective divergence of $\Lambda_2^{(0)}$ in the dc limit at zero temperature and in the clean limit ($U \to 0$).

The impact of the two-photon renormalization can be understood in more details by investigating the two-photon renormalization factor, [45]

$$Q_2(z_1,z_2) = \frac{1}{1 - UX_2(z_1,z_2)} = \frac{S(z_1) - S(z_2)}{z_1 - z_2},$$

where $X_2(z_1,z_2) \propto \sum_k \text{Tr}[\hat{G}(k,z_1)\hat{G}(k,z_2)]$ and where $z_1$ and $z_2$ are the electronic frequencies in the complex plane. Particularly enlightening is the analysis of the retarded-retarded (RR) channel. In the dc limit ($\omega \to 0$) $Q_2^{RR}(\epsilon,\epsilon + \omega)$ is simply given by

$$\lim_{\omega \to 0} Q_2^{RR}(\epsilon,\epsilon + \omega) = \frac{dS(\epsilon)}{d\epsilon} = \frac{S(\epsilon)}{2US(\epsilon) + \mu_0 + \epsilon},$$

where $\omega$ is the photon energy and $\epsilon$ is the electronic energy from the Fermi surface. For low-energy excitations $\epsilon = 0$ we have thus $Q_2^{RR} = S(0)/[2US(0) + \mu_0]$. The Boltzmann regime is achieved as $\mu_0 \gg 2US(0)$. In the clean limit $U \to 0$ we get thus $S(0) = \mu_0$ and $Q_2^{RR} = 1$.

---

**FIG. 3:** Nonlinear conductivity of graphene at zero temperature. (a) Log-log scale plot for the absolute value of the characteristic third-order transport function $f_3(\mu,U)$ versus $U$ at $x = 0.03$. Different lines correspond to the individual contribution of diagrams in Fig. 1a as mentioned the plot-legend. (b) $f_3(\mu,U)$ as a function of the chemical potential for $U = 0.11$. Similar to the panel (a), different curves correspond to different diagram’s contribution as mentioned the plot-legend. (c) Colormap plot for $g = |\sigma_{\text{dc}}| \cdot E^2/|\sigma_{\text{dc}}|$ factor with $E = 1\text{mV/nm}$ versus chemical potential $\mu$ and relaxation rate $\Gamma(\mu)$ in the full conserving models. The sign of $\sigma_{\text{dc}}$ is written on the plot where two sign-switch borders are highlighted by dashed red lines. Green and blue dashed lines stand for the contour lines with $g = 1$ and $g = 0.1$, respectively. Similar colormap plot for the constant-Γ model is given in the SM [45].
The quantum regime is on the other hand characterized by $\mu_0 \ll 2U S(0)$, and we get $Q_2 = 1/2U$, leading thus to a huge enhancement (divergence) in the clean limit. A similar behavior $Q_2 \approx 1/U$ appears also in the retarded-advance (RA) channel, although is a more delicate way. In the dc limit, we find indeed the leading term in $Q_2 \delta^{(2)}(\epsilon, \omega) \approx -i2\Gamma(\epsilon)/\omega$ which shows a divergence as a function of the photon energy $\omega \to 0$. Once plugged this behavior in the response function associated with $\chi$ of the same order than the linear dc term $g$, a net result, the divergence $1/Q$ way. In the dc limit, we find indeed the leading term $\sigma^{(2)}_{dc}(0) = -1/[4U^2\Gamma(0)]$ and higher orders. As a net result, the divergence $1/\omega$ in $Q_2^{RA}(\epsilon, \epsilon + \omega)$ is reflected in a consequent divergence $-1/\omega$ in the response function, similarly as for the RR channel, but with a negative sign. The balance between RR and RA terms determines the change of sign of the third-order dc conductivity as a function of $x$ in the quantum regime. We must stress that the huge enhancement of the third-order dc transport is governed by the dominant role of the two-phonon vertex renormalization. Since $\chi^{(0)}_{\text{triangles}}$ scales as $U$ and $1/[1 - U X_d]$ scales as $1/U$, such enhancement can be regarded as TPV self-generation ($\Delta_2 \neq 0$) which survives in the weak-coupling (clean) limit $U \to 0$.

The net result on the dc transport is summarized in Fig. 3c where we plot the sign and the magnitude of the third-order conductivity for a given electric field $E \approx 1$ mV/nm normalized to the linear order conductivity, $g = [\sigma^{(3)}_{dc} E^2/\sigma^{(1)}_{dc}]$, in the physical space of the effective chemical potential $\mu$ and scattering rate $\Gamma(\mu)$, as they can be obtained directly in an experimental way. The Boltzmann regime corresponds thus to the right-lower corner whereas the extreme quantum regime ($x \to 0$) is recovered in the left-upper corner. As noticed before, at odds with the predictions of the phenomenological model, we find that the third-order conductivity $\sigma^{(3)}_{dc}$ is negative not only in the Boltzmann regime, but also in the quantum regime. Note that the zeroes of the third-order dc conductivity, see $x_1$ and $x_2$ in Fig. 2, appear in this plot as straight dashed lines. This is a consequence of the factorizable expression for the characteristic third-order dc transport function as shown in Eq. (5). We mark with tiny dotted in this plot the regions where the third-order terms start to be relevant $g \approx 0.1$ and where they become of the same order than the linear dc term $g \approx 1$ [47].

Note that in the Boltzmann regime, by increasing $\mu \to \infty$, one should correspondingly need infinitesimally small $\Gamma \to 0$ in order to detect third-order corrections, while in the quantum regime $\mu \to 0$ third-order corrections appear to be relevant up to large values of $\Gamma$ dictated only by the applied electric field (and by the ultra-violet cut-off of the Dirac model). An alternative and maybe more direct way to assess the relevance of the nonlinear conductivity is to evaluate in graphene [37] as a paradigmatic 2D Dirac material the critical electric field $E_{\text{max}}$ above which third-order corrections to the dc transport become of the same order of the linear term, $|\sigma^{(2)}_{dc}| E_{\text{max}}^2 - |\sigma^{(1)}_{dc}|$. In the phenomenological constant-$\Gamma$ model we obtain $E_{\text{max}} = \alpha E_0$ with $\alpha \approx 1.47 \times \mu \Gamma/\Gamma_0$ in the Boltzmann regime and $\alpha \approx 1.42 \times \Gamma^2/\Gamma_0^2$ in the quantum limit. These values can be compared with the estimates for the full conserving theory that gives $\alpha \approx 0.32 \times \mu \Gamma(\mu)/\Gamma_0$ in the Boltzmann regime and $\alpha \approx 1.14 \mu U \times \Gamma(0)^2/\Gamma_0^2$ in the quantum limit with $U = 1/\ln[\Gamma^2/\Gamma_0^2(0)]$. In Ref. [48] a roughly constant value $\Gamma \approx 15$ meV was estimated in the wide range $\mu \approx 0$–200 meV. With these values the phenomenological model would estimate a critical field $E_{\text{max}} \approx 10.8$ mV/nm for $\mu \approx 200$ meV and $E_{\text{max}} \approx 74.7$ mV/nm for $\mu = 0$ with a quantum-Boltzmann crossover at $\mu^* \approx 150$ meV, whereas the full conserving theory predicts $E_{\text{max}} \approx 2.35$ mV/nm for $\mu \approx 200$ meV and $E_{\text{max}} \approx 0.051$ mV/nm for $\mu = 0$, in a more observable range.

In conclusion, in this Letter, we have presented a fully conserving theory of nonlinear transport response in 2D Dirac materials. Our results show that the previous analyses in literature, based on phenomenological models, can be qualitatively (but not quantitatively) reliable in the Boltzmann regime but they completely fail in the quantum regime. We have shown that, in a wide region of the phase diagram, close to the neutral point, the nonlinear dc transport response is dominated by novel physical processes where the two-photon vertex, absent in the bare Dirac Hamiltonian, plays a relevant role. It should be furthermore stressed that our results, focused on the dc limit, implies that the current knowledge about the nonlinear optical response in the terahertz regime should be deeply revised. Our work opens new scenarios for a deep understanding of the electromagnetic response of Dirac systems, whose relevance ranges from condensed matter to high-energy physics.

Acknowledgements—H.R. acknowledges the support from the Swedish Research Council (VR 2018-04252).

[1] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. 81, 109 (2009), URL https://link.aps.org/doi/10.1103/RevModPhys.81.109.
[2] T. Wehling, A. Black-Schaffer, and A. Balatsky, Advances in Physics 63, 1 (2014), https://doi.org/10.1080/00018732.2014.927109.
[3] N. P. Armitage, E. J. Mele, and A. Vishwanath, Rev. Mod. Phys. 90, 015001 (2018), URL https://link.aps.org/doi/10.1103/RevModPhys.90.015001.
[4] M. Katsnelson, Graphene: Carbon in Two Dimensions (Cambridge University Press, 2012), ISBN 9780521195409.
[5] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi, Rev. Mod. Phys. 83, 407 (2011), URL https://link.aps.org/doi/10.1103/RevModPhys.83.407.

[6] N. M. R. Peres, Rev. Mod. Phys. 82, 2673 (2010), URL https://link.aps.org/doi/10.1103/RevModPhys.82.2673.

[7] E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, and S. A. Mikhailov, Phys. Rev. Lett. 105, 097401 (2010), URL https://link.aps.org/doi/10.1103/PhysRevLett.105.097401.

[8] H. Zhang, S. Viralq, B. Qiao, L. Kian Ping, S. Masrur, N. Godbout, and P. Kockert, Optics Letters 37, 1856 (2012). URL http://ol.osa.org/abstract.cfm?URI=ol-37-11-1856.

[9] N. Kumar, J. Kumar, C. Gerstenkorn, R. Wang, H.-Y. Chiu, A. L. Smirl, and H. Zhao, Phys. Rev. B 87, 121406 (2013). URL https://link.aps.org/doi/10.1103/PhysRevB.87.121406.

[10] R. I. Woodward, R. T. Murray, C. F. Phelan, R. E. P. de Oliveira, T. H. Runcorn, E. J. R. Kelleher, S. Li, E. C. de Oliveira, G. J. M. Fechine, G. Eda, et al., 2D Materials 4, 011006 (2016). URL https://doi.org/10.1088%2F2053-1583%2F4%2F1%2F011006.

[11] G. Soavi, G. Wang, H. Rostami, D. G. Purdie, D. De Fazio, T. Ma, B. Luo, J. Wang, A. K. Ott, D. Yoon, et al., Nature Nanotechnology 13, 583 (2018). URL https://doi.org/10.1038/s41565-018-0145-8.

[12] J.-H. Chen, J. Feng, and D. Sun, Nature Materials 17, 15995 (2017). URL https://doi.org/10.1038/ncomms15995.

[13] J. Sabio, J. Nilsson, and A. H. Castro Neto, Phys. Rev. B 78, 075410 (2019). URL https://link.aps.org/doi/10.1103/PhysRevB.78.075410.

[14] R. Cenni, Nuclear Physics A 696, 605 (2001). ISSN 0375-9474. URL http://www.sciencedirect.com/science/article/pii/S0375947401011514.

[15] D. E. Parker, T. Morimoto, J. Orenstein, and J. E. Moore, Phys. Rev. B 95, 045121 (2019), URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.013069.

[16] X. Bao, T. Palacios, H. Lin, S. Jia, et al., Nature 565, 507 (2018). URL https://doi.org/10.1038/s41586-018-0508-1.

[17] X. Bao, T. Palacios, H. Rostami, A. Tomadin, O. Balci, I. Paradissos, E. A. A. Pogna, G. Cullero, E. Lidorikis, M. Polini, et al., ACS Photonics 6, 2841 (2019). URL https://doi.org/10.1021/acsphotonics.9b00928.

[18] J. Sabio, J. Nilsson, and A. H. Castro Neto, Phys. Rev. B 78, 075410 (2008). URL https://link.aps.org/doi/10.1103/PhysRevB.78.075410.

[19] X. Bao, T. Palacios, H. Rostami, A. Tomadin, O. Balci, I. Paradissos, E. A. A. Pogna, G. Cullero, E. Lidorikis, M. Polini, et al., ACS Photonics 6, 2841 (2019). URL https://doi.org/10.1021/acsphotonics.9b00928.

[20] J. Sabio, J. Nilsson, and A. H. Castro Neto, Phys. Rev. B 78, 075410 (2008). URL https://link.aps.org/doi/10.1103/PhysRevB.78.075410.

[21] X. Bao, T. Palacios, H. Rostami, A. Tomadin, O. Balci, I. Paradissos, E. A. A. Pogna, G. Cullero, E. Lidorikis, M. Polini, et al., ACS Photonics 6, 2841 (2019). URL https://doi.org/10.1021/acsphotonics.9b00928.

[22] J. Sabio, J. Nilsson, and A. H. Castro Neto, Phys. Rev. B 78, 075410 (2008). URL https://link.aps.org/doi/10.1103/PhysRevB.78.075410.

[23] D. E. Parker, T. Morimoto, J. Orenstein, and J. E. Moore, Phys. Rev. B 95, 045121 (2019), URL https://link.aps.org/doi/10.1103/PhysRevResearch.2.013069.

[24] Y. Shen, The Principles of Nonlinear Optics (Wiley, New York, 1984).

[25] R. Boyd, Nonlinear Optics (Academic Press, Cambridge, 2008).

[26] For numerical calculations we consider a two-dimensional Dirac material with realistic parameters for graphene. In particular we consider a honeycomb lattice with inter-atom distance $b = 1.42 \, \text{Å}$, with lattice parameter $a = \sqrt{3}b = 2.46 \, \text{Å}$ and nearest-neighbor hopping $t_0 \approx 3 \, \text{eV}$, corresponding to a Dirac velocity $v = \sqrt{3}t_0a/2h \approx 10^6 \, \text{m/s}$. The size of the unit-cell reads thus $S_{\text{cell}} = 3a^2/2$. We consider double spin and valley degeneracy $N_s = N_v = 2$. In order to preserve the size of the Brillouin zone $V_{\text{BZ}} \approx 8t_0^2/3\sqrt{3}a^2$, we introduce a momentum cut-off $k_F = 1.09 \, \text{Å}^{-1}$, so that $t_F = V_{\text{BZ}}$, defining thus an ultra-violet energy cut-off $W = hvk_F = 7.2 \, \text{eV}$ for the Dirac linear dispersion. Note that we have $t_0 = e^2/4h$ and $E_0 = \pi t_0/\sqrt{3}a = 22.0 \, \text{V/μm}$ is an ultra-strong characteristics electric field (see Supplementary Materials [45]).

[27] N. H. Shon and T. Ando, Journal of the Physical Society of Japan 67, 2421 (1998), https://doi.org/10.1143/JPSJ.67.2421, URL https://doi.org/10.1143/JPSJ.67.2421.

[28] A. Tsuneya, Y. Zheng, and S. Hidekatsu, Journal of...
of the Physical Society of Japan 71, 1318 (2002), https://doi.org/10.1143/JPSJ.71.1318, URL https://doi.org/10.1143/JPSJ.71.1318.

[40] H. Rostami and E. Cappelluti, Phys. Rev. B 96, 054205 (2017), URL https://link.aps.org/doi/10.1103/PhysRevB.96.054205.

[41] G. Leibbrandt, Rev. Mod. Phys. 47, 849 (1975), URL https://link.aps.org/doi/10.1103/RevModPhys.47.849.

[42] M. E. Peskin, An Introduction To Quantum Field Theory (CRC Press; 1 edition, 1995).

[43] G. Baym and L. P. Kadanoff, Phys. Rev. 124, 287 (1961), URL https://link.aps.org/doi/10.1103/PhysRev.124.287.

[44] L. P. Kadanoff and G. Baym, Quantum Statistical Mechanics (W.A. Benjamin; 1st edition, 1962).

[45] Supplementary Materials.

[46] We estimate the third-order dc conductivity as $\sigma_{dc}^{(3)} = \lim_{\omega \to 0} \sigma_{THG}^{(3)}(\omega)$. In Supplementary Materials [45] we plot the frequency dependence of the third-harmonic generation (THG) response function $\sigma_{THG}^{(3)}(\omega)$ for $\mu = 17$ THz and few representative values of $U$.

[47] The condition $g = 1$, in the regions where the third-order dc conductivity $\sigma_{dc}^{(3)}$ is negative should not be regarded as an onset of negative total dc conductivity, rather as a sign that the expansion at the third-order in $E$ starts to be a poor approximation and higher order terms in powers of $E$ must be included in the analysis.

[48] J. Horng, C.-F. Chen, B. Geng, C. Girit, Y. Zhang, Z. Hao, H. A. Bechtel, M. Martin, A. Zettl, M. F. Crommie, et al., Phys. Rev. B 83, 165113 (2011), URL https://link.aps.org/doi/10.1103/PhysRevB.83.165113.

[49] G. D. Mahan, Many-Particle Physics (Springer US; New York; 3 edition, 2000).
Supplemental Materials:
“Dominant Role of Two-Photon Vertex in Nonlinear Response of Dirac Materials”

Habib Rostami,1 Emmanuele Cappelluti2

1 Nordita, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden
2Istituto di Struttura della Materia, CNR, 34149 Trieste, Italy

Contents

References 5

S1. Ultra-violet cut-off and dimensional regularization 8

S2. Baym-Kadanoff derivations 10
A. Conserving linear response theory in Dirac systems 11
B. Conserving third order response theory in Dirac systems 12
1. Interaction induced two-photon vertex 13
2. Interaction induced three-photon vertex 13

S3. Analytical expressions for nonlinear conductivity 14
A. Different diagram contributions 14
B. Renormalization of the one-photon vertex 15
C. Renormalization of the two-photon vertex 16
D. Renormalization of the three-photon vertex 17

S4. Linear conductivity 18

S5. Nonlinear dc conductivity 20
A. Analytical derivation 20
B. Numerical evaluation of \( f_3(x, U) \) function in the full quantum theory 21

S6. Analytical continuation for the third order response function 22

S7. Numerical evaluation of the nonlinear optical and dc conductivities 24

S1. ULTRA-VIOLET CUT-OFF AND DIMENSIONAL REGULARIZATION

The introduction of a high-energy (ultra-violet) cut-off is an unavoidable requirement of Dirac models. There is however a relative large degree of freedom in the way how to introduce it, and particular care is needed in order to avoid spurious results and to preserve physical consistencies, like Ward’s identities, and gauge invariance. Dimensional regularization has proven to be a formidable tool to ensure that physical correctness is preserved [41, 42]. Here we show how such approach provides a consistent framework for evaluating self-energy and susceptibilities. In our work we employ dimensional regularization that endures gauge invariance. As a benchmark example, and for the sake of simplicity, we consider the evaluation of the disorder self-energy which displays a primary diverging integral.

We consider scattering on local impurity centers with density \( n_{\text{imp}} \) and potential \( V_{\text{imp}}(r) = \sum_i V_i \delta(r - R_i) \) where \( R_i \) are the coordinates of the lattice sites. We assume standard Born impurity correlations as \( \langle V_{\text{imp}}(r) \rangle = 0 \) and the effective scattering potential reads

\[
V(1, 2) = \langle V_{\text{imp}}(r_1)V_{\text{imp}}(r_2) \rangle_{\text{imp}} = \gamma_{\text{imp}}\delta(r_1 - r_2). \tag{S1}
\]

Note that the average \( \langle \ldots \rangle_{\text{imp}} \) is meant over all the impurity configurations and \( \gamma_{\text{imp}} = n_{\text{imp}}V_{\text{imp}}^2 \) in which \( n_{\text{imp}} = N_{\text{imp}}/N_{\text{cell}} \) (number of impurity centers per number of unit-cells) stands for the density of scattering centers and \( V_{\text{imp}} \) is the average strength of the scattering potential energy. The lowest-order self-consistent Born self-energy reads

\[
\Sigma(i\hbar \omega_n) = \gamma_{\text{imp}} \sum_k \hat{G}(k, i\hbar \omega_n) = \gamma_{\text{imp}} S \int \frac{d^2k}{(2\pi)^2} \hat{G}(k, i\hbar \omega_n). \tag{S2}
\]
Note that $S = N_{\text{cell}}S_{\text{cell}}$ is the system area with $S_{\text{cell}}$ being the unit-cell area. Due to the isotropic impurity scattering the self-energy spinor structure is trivial as $\Sigma(i\kappa_n) = \hat{\Sigma}(i\kappa_n)\Gamma$ and therefore the Green’s function can be explicitly written as follows

$$
\hat{G}(k, i\kappa_n) = \left[ z + \mu_0 - \Sigma(z) \right] \hat{I} + \hbar \nu \hat{\sigma} \cdot \mathbf{k} \left[ z + \mu_0 - \Sigma(z) \right]^{-1} (\hbar \nu k)^2
$$

(S3)

where $z = i\kappa_n$. Accordingly, we find

$$
\Sigma(z) = \frac{\gamma_{\text{imp}} N_{\text{cell}} S_{\text{cell}}}{(\hbar \nu)^2} \int \frac{d^D \ell}{(2\pi)^D} \frac{S(z)}{z^2 - \ell^2}
$$

(S4)

where $S(z) = z + \mu_0 - \Sigma(z)$. In arbitrary $D$ dimensions, we have

$$
\Sigma(z) = \frac{\gamma_{\text{imp}} N_{\text{cell}} S_{\text{cell}}}{(\hbar \nu)^D} \int \frac{d^D \ell}{(2\pi)^D} \frac{S(z)}{z^2 - \ell^2}.
$$

(S5)

Note that the above integral in $D$ dimensions can be solved in terms of Euler’s Gamma-function, $\Gamma_E(z)$, by utilizing the following identity [42]

$$
\int \frac{d^D \ell}{(2\pi)^D} \frac{1}{(r^2 + \Delta)^n} = \frac{1}{(4\pi)^{D/2}} \frac{\Gamma_E(n - D/2)}{\Gamma_E(n)} \left( \frac{1}{\Delta} \right)^{n - D/2}.
$$

(S6)

Therefore, we find

$$
\Sigma(z) = -U S(z) \frac{1}{(4\pi)^{D/2 - 1}(\hbar \nu)^{D - 2}} \Gamma_E\left( 1 - \frac{D}{2} \right) (-S(z)^2)^{D/2 - 1}
$$

(S7)

in which [37]

$$
U = \frac{\gamma_{\text{imp}} N_{\text{cell}} S_{\text{cell}}}{4\pi(\hbar \nu)^2} = \frac{N_{\text{imp}}}{4\pi\sqrt{3}} \left( \frac{V_{\text{imp}}}{t_0} \right)^2.
$$

(S8)

Now we set $D = d - \epsilon$ where $d = 2$ is the physical dimension and $\epsilon \to 0$. Note that $\Gamma_E(\epsilon/2) \approx 2/\epsilon$ for $\epsilon \to 0$ and

$$
\lim_{\epsilon \to 0} \frac{(X^2)^{-\epsilon/2}}{\epsilon^{\epsilon/2}} = \lim_{\epsilon \to 0} \frac{2}{\epsilon} \ln[X(2)] = \ln[W^2] - \ln[X^2] = \ln \left[ \frac{W^2}{X^2} \right].
$$

(S9)

FIG. S1: Numerical result for the imaginary and real parts of the self-energy. (a) The imaginary part of the self-energy at the Fermi surface $\Gamma = -\text{Im}[\Sigma(\omega = 0)]$ is shown versus bare chemical potential $\mu_0$. (b) The imaginary part of the self-energy at the Fermi surface $\Delta = \text{Re}[\Sigma(\omega = 0)]$ is shown versus bare chemical potential $\mu_0$. (c) $\Gamma$ is shown versus renormalized chemical potential $\mu = \mu_0 - \Delta(\mu)$. Different curves correspond to different values of $U$. 
Note that we use the prescription \( \lim_{\epsilon \to 0} \frac{1}{\epsilon} \equiv \ln[W] \) where \( W \) is the ultra-violet energy cut-off. Eventually, we obtain the following self-consistent formula for the self-energy

\[
\Sigma(z) = -US(z) \ln \left[ \frac{W^2}{S(z)^2} \right]. \tag{S10}
\]

After solving the above relation self-consistently, the real, \( \Delta = \text{Re}[\Sigma(\omega = 0)] \), and imaginary, \( \Gamma = -\text{Im}[\Sigma(\omega = 0)] \), parts of the self-energy at the Fermi surface are depicted in Fig. S1. For the undoped regime and at the Dirac point, we have \( \Sigma(0) = -i\Gamma(0) \) where

\[
\Gamma(0) = W \exp \left\{ -\frac{1}{2U} \right\}. \tag{S11}
\]

The above procedure of dimensional regularization is employed in similar way in the evaluation of other momentum integrals in this study.

### S2. BAYM-KADANOFF DERIVATIONS

Within the lowest-order self-consistent Born approximation, the self-energy correction induced by elastic impurity scattering reads

\[
\hat{\Sigma}(1, 2) = V(1, 2)\hat{G}(1, 2). \tag{S12}
\]

We use a shorthand notation \( 1 \equiv (r_1, t_1) \) for the space-time coordinate. Using Dyson recursive relation, the full field-dependent and interacting Green’s function is given in terms a field-dependent self-energy, \( \hat{\Sigma} \), and a bare Green’s function, \( \hat{G}_0 \), as follows

\[
\hat{G}(1, 1'; A) = \hat{G}_0(1, 1'; A) + \int_{2, 3} \hat{G}_0(1, 2; A)\hat{\Sigma}(2, 3; A)\hat{G}(3, 1'; A) \tag{S13}
\]

or equivalently we have

\[
\hat{G}^{-1}(1, 1'; A) = \hat{G}_0^{-1}(1, 1'; A) - \hat{\Sigma}(1, 1'; A). \tag{S14}
\]

The inverse of bare Green’s function reads

\[
\hat{G}_0^{-1}(1, 1'; A) = [i\partial_{t_1} - v\hat{\sigma} \cdot (-i\hbar \nabla_1 + eA(1))]\delta(1 - 1'). \tag{S15}
\]

We assume an external gauge field along y axis, \( A(1) = A(1)\hat{y} \). The one-photon current vertex is given in terms of variational derivative of bare Green’s function versus the gauge field:

\[
\hat{A}_1^{(0)}(1, 1'; 1'^n) = \frac{\delta \hat{G}_0^{-1}(1, 1')}{\delta A^{(1^s)}} \bigg|_{A \to 0} = -ev\hat{\sigma}_y\delta(1 - 1^n)\delta(1 - 1'). \tag{S16}
\]

The thermodynamic physical current, \( J(1; A) = J(1; A)\hat{y} \), in Dirac systems reads

\[
J(1; A) = -i \int_{1', 1'^*} \text{tr} \left[ \hat{A}_1^{(0)}(1, 1'; 1'^n)\hat{G}(1, 1'^*; A) \right]. \tag{S17}
\]

Note that \( 1'^s \equiv (r_1', t_1' = t_1 + 0^+) \) and “tr” stands for the “trace” operation over all spinor indexes i.e. \( \text{tr}[\hat{A}\hat{B}] = \sum_{\alpha\beta}[A_{\alpha\beta}B_{\beta\alpha}] \). Note the field operator \( \hat{\psi}_H(r, t) \) in the Heisenberg picture of \( \hat{\psi}(r) \) in the basis of full Hamiltonian \( \hat{H} \) which contains kinetics, light-matter and many-body interaction terms. The Baym-Kadanoff (or contour) Green’s function [43, 44] follows

\[
\hat{G}(1, 1'; A) = -i\{\hat{T}[\hat{\psi}_H(1)\hat{\psi}^+_H(1')] \}. \tag{S18}
\]

where \( \{\ldots\} \) stands for the thermodynamical average and \( \hat{T} \) is for the time-ordering operation.
A. Conserving linear response theory in Dirac systems

We define the linear susceptibility as follows

$$\chi^{(1)}(1, 2) = \delta J(1; A) \bigg|_{A=0},$$  \hspace{1cm} (S19)

Using Eq. (S17), the linear susceptibility reads

$$\chi^{(1)}(1, 2) = -i \int_{1', 1''} \text{tr} \left[ \hat{\Lambda}_1^{(0)}(1, 1', 1'') G(1, 1''; A) \right]_{A=0}$$  \hspace{1cm} (S20)

Using \( \hat{G}\hat{G}^{-1} = 1 \), we have

$$\frac{\delta \hat{G}(1, 1'; A)}{\delta A(2)} = - \int_{2, 3} \hat{G}(1, 2; A) \frac{\delta \hat{G}^{-1}(2, 3; A)}{\delta A(2)} \hat{G}(3, 1'; A)$$  \hspace{1cm} (S21)

Using Eq. (S14), we find

$$\frac{\delta \hat{G}^{-1}(1, 1'; A)}{\delta A(2)} = \frac{\delta \hat{G}^{-1}(1, 1'; A)}{\delta A(2)} - \frac{\delta \Sigma(1, 1'; A)}{\delta A(2)}$$  \hspace{1cm} (S22)

Since the self-energy depends on the external potential only through its dependence on the Green’s function, we can write down

$$\frac{\delta \Sigma(1, 1'; A)}{\delta A(2)} = \int_{3, 4} \hat{\Xi}(1, 1', 3, 4; A) \frac{\delta \hat{G}(3, 4; A)}{\delta A(2)}$$  \hspace{1cm} (S23)

We define the Bethe-Salpeter kernel as bellow

$$\hat{\Xi}(1, 2; 3, 4; A) = \frac{\delta \Sigma(1, 2; A)}{\delta \hat{G}(3, 4; A)}$$  \hspace{1cm} (S24)

Note that \( \hat{G}(1, 1'; A)|_{A=0} = \hat{G}(1, 1') \) and \( \hat{\Xi}(1, 2; 3, 4; A)|_{A=0} = \hat{\Xi}(1, 2; 3, 4) \). For our self-energy model within the self-consistent Born approximation given in Eq. (S12), we have

$$\hat{\Xi}(1, 2; 3, 4) = V(1, 2) \delta(1 - 3) \delta(2 - 4)$$  \hspace{1cm} (S25)

Therefore, using Eq. (S21) the self-consistent Bethe-Salpeter relation for the one-photon vertex function follows

$$\hat{\Lambda}_1(1, 1'; 2) = \hat{\Lambda}_1^{(0)}(1, 1'; 2) + V(1, 1') \int_{5, 6} \hat{G}(1, 5) \hat{\Lambda}_1(5, 6; 2) \hat{G}(6, 1')$$  \hspace{1cm} (S26)

Note that the dressed one-photon vertex function are defined as

$$\hat{\Lambda}_1(1, 1'; 2) = \frac{\delta \hat{G}^{-1}(1, 1'; A)}{\delta A(2)}|_{A=0}$$  \hspace{1cm} (S27)

For the sake of simplicity, we extend the definition of space-time parameter to include also spinor indexes as \( \hat{I} = (r_1, t_1, s_1) \) and, for instance, we can drop `\( \frac{\delta \sigma}{\delta A} \)` symbol in \( \Lambda_1 \hat{G} \) instead use \( \Lambda_1 G \) since the spinor multiplication is taken into account when we replace \( \int dI \sum_{s_1} \rightarrow \int dI \). Moreover we use shorthand notation for \( C = \int dI A(\ldots, 1) B(1, \ldots) \) as \( C = AB \). We use this compact notation from now on

$$\chi^{(1)}_{\alpha \beta}(1, 2) = i \text{Tr} \left[ \hat{\Lambda}_1^{(0)}(1) GA_1(2) G \right]$$  \hspace{1cm} (S28)

$$\Lambda_1(2) = \Lambda_1^{(0)}(2) + V G \Lambda_1(2)$$  \hspace{1cm} (S29)

Note that \( \text{Tr} \ldots \) in the above formula stands for the sum over un-contracted spinor index. In the compact notation 1 and 2 are space-time symbols.
Third-order response function is given by
\[
\chi^{(3)}(1, 2, 3, 4) = \frac{1}{3!} \left. \frac{\delta^3 J(1; \mathbf{A})}{\delta A(2) \delta A(3) \delta A(4)} \right|_{\mathbf{A} = 0} = -\frac{i}{3!} \text{Tr} \left[ \Lambda^{(0)}_1(A) \frac{\delta^3 G(1, 1^+; \mathbf{A})}{\delta A(2) \delta A(3) \delta A(4)} \right]_{\mathbf{A} = 0} \quad (S30)
\]
Using \( GG^{-1} = 1 \) and Eq. (S21), we evaluate the third derivative of the Green’s function versus vector potential component,
\[
\frac{\delta^3 G}{\delta A(2) \delta A(3) \delta A(4)} = -\frac{\delta^2 G}{\delta A(2) \delta A(3)} G G^{-1} G - \frac{\delta^2 G}{\delta A(3) \delta A(4)} G G^{-1} G + \frac{\delta^2 G}{\delta A(2) \delta A(4)} G G^{-1} G - \frac{\delta^2 G}{\delta A(2) \delta A(3) \delta A(4)} G \quad (S31)
\]
Using Eq. (S27), we have
\[
\frac{\delta^3 G}{\delta A(2) \delta A(3) \delta A(4)} = -\frac{\delta^2 G}{\delta A(2) \delta A(3)} G^{-1} G - \frac{\delta^2 G}{\delta A(3) \delta A(4)} G^{-1} G + \frac{\delta^2 G}{\delta A(2) \delta A(4)} G^{-1} G - \frac{\delta^2 G}{\delta A(2) \delta A(3) \delta A(4)} G \quad (S32)
\]
Once more we use \( GG^{-1} = 1 \) and perform second derivative of the Green’s function which leads
\[
\frac{\delta^2 G}{\delta A(2) \delta A(3)} = -\frac{\delta G}{\delta A(2)} G^{-1} G + \frac{\delta G}{\delta A(2)} G G^{-1} G - \frac{\delta G}{\delta A(2)} G^{-1} G + \frac{\delta G}{\delta A(2)} G \quad (S33)
\]
The two-photon vertex function is defined as follows
\[
\Lambda_2(2, 3) = \frac{\delta^2 G^{-1}}{\delta A(2) \delta A(3)} \quad (S34)
\]
Using Eq. (S21), Eq. (S27) and Eq. (S34), we find
\[
\frac{\delta^2 G}{\delta A(2) \delta A(3)} = -GA_2(2, 3) G + \sum_{P(2, 3)} GA_1(2) G A_1(3) G \quad (S35)
\]
where \( P(2; 3) \) stands for the permutation between 2 \( \leftrightarrow \) 3. The three-photon vertex function reads
\[
\Lambda_3(2, 3, 4) = \frac{1}{2} \frac{\delta^2 G^{-1}}{\delta A(2) \delta A(3) \delta A(4)} \quad (S36)
\]
By plugging Eq. (S34), Eq. (S35), and Eq. (S36) in Eq. (S32), we obtain
\[
\frac{\delta^3 G}{\delta A(2) \delta A(3) \delta A(4)} = -2GA_3(2, 3, 4) G - \sum_{P(2, 3, 4)} GA_1(2) G A_1(3) G A_1(4) G + \frac{1}{2} \sum_{P(2, 3, 4)} GA_2(2, 3) G A_1(4) G \\
+ \frac{1}{2} \sum_{P(2, 3, 4)} GA_1(4) G A_2(2, 3) G \quad (S37)
\]
where \( P(2; 3, 4) \) stands for all six permutations among 2, 3, and 4 space-time coordinates. Therefore, the third-order response function reads
\[
\chi^{(3)}(1, 2, 3, 4) = \frac{2}{3!} \text{Tr} \left[ \Lambda^{(0)}_1(A) G A_3(2, 3, 4) G \right] + \frac{i}{3!} \sum_{P(2; 3, 4)} \text{Tr} \left[ \Lambda^{(0)}_1(A) G A_2(2, 3) G A_1(4) G \right] \\
- \frac{1}{2 \cdot 3!} \sum_{P(2; 3, 4)} \text{Tr} \left[ \Lambda^{(0)}_1(A) G A_2(2, 3) G A_1(4) G \right] - \frac{1}{2 \cdot 3!} \sum_{P(2; 3, 4)} \text{Tr} \left[ \Lambda^{(0)}_1(A) G A_2(2, 3) G A_1(4) G \right] \quad (S38)
\]
1. Interaction induced two-photon vertex

The two-photon vertex function is defined as follows

\[ \Lambda_2(2, 3) = \frac{\delta^2 G^{-1}}{\delta A(2) \delta A(3)} = \frac{\delta^2 G_0^{-1}}{\delta A(2) \delta A(3)} - \frac{\delta^2 \Sigma}{\delta A(2) \delta A(3)}. \]  

(S39)

The second derivative of the bare Green’s function vanishes in Dirac systems which implies the absence of bare two-photon vertex. However, an interaction induced two-photon vertex function is obtained owing to the field-dependent self-energy. Since the self-energy depends on the external field only through the dependence on the Green’s function, we have

\[ \Lambda_2(2, 3) = -\frac{\delta^2 \Sigma}{\delta A(2) \delta A(3)} = -V \frac{\delta^2 G}{\delta A(2) \delta A(3)}. \]  

(S40)

Using Eq. (S35), we obtain

\[ \Lambda_2(2, 3) = \Lambda_2^{(0)}(2, 3) + V \Lambda_2(2, 3)G \]  

(S41)

where

\[ \Lambda_2^{(0)}(2, 3) = -\sum_{\mathcal{P}(2:3)} V G \Lambda_1(2) G \Lambda_1(3) G. \]  

(S42)

2. Interaction induced three-photon vertex

The three-photon vertex function is defined as follows

\[ \Lambda_3(2, 3, 4) = \frac{1}{2} \frac{\delta^3 G^{-1}}{\delta A(2) \delta A(3) \delta A(4)} = \frac{1}{2} \frac{\delta^3 G_0^{-1}}{\delta A(2) \delta A(3) \delta A(4)} - \frac{1}{2} \frac{\delta^3 \Sigma}{\delta A(2) \delta A(3) \delta A(4)}. \]  

(S43)

The third derivative of the bare Green’s function vanishes in Dirac systems which implies the absence of bare three-photon vertex. However, an interaction induced three-photon vertex function is obtained owing to the field-dependent self-energy. Since the self-energy depends on the external field only through the dependence on the Green’s function, we have

\[ \Lambda_3(2, 3, 4) = -\frac{1}{2} \frac{\delta^3 \Sigma}{\delta A(2) \delta A(3) \delta A(4)} = -\frac{1}{2} V \frac{\delta^3 G}{\delta A(4) \delta A(3) \delta A(2)}. \]  

(S44)

Using Eq. (S37), we obtain

\[ \Lambda_3(2, 3, 4) = \Lambda_3^{(0)}(2, 3, 4) + V \Lambda_3(2, 3, 4)G \]  

(S45)

where

\[ \Lambda_3^{(0)}(2, 3, 4) = \frac{1}{2} \sum_{\mathcal{P}(2:3:4)} V \left[ G \Lambda_1(2) G \Lambda_1(3) G \Lambda_1(4) G - \frac{1}{2} G \Lambda_2(2, 3) G \Lambda_1(4) G - \frac{1}{2} G \Lambda_1(4) G \Lambda_2(2, 3) G \right]. \]  

(S46)

To summarise, the formal derivation given in the current Section guides us in constructing a conserving diagrammatic theory for the third-order response function in Dirac systems. Lengthy mathematical relations for different contributions to the nonlinear response function, i.e. Eq. (S38), and the multi-photon vertex functions, i.e. Eqs. (S29),(S41),(S42),(S45), and (S46), are graphically illustrated in Feynman diagrams depicted in Fig. 1 of the main text. Quantitative evaluation of these diagrams are explicitly discussed with great details in the next Section.
S3. ANALYTICAL EXPRESSIONS FOR NONLINEAR CONDUCTIVITY

We present here the analytical expression of the third-order optical response function which in the Matsubara space can be formally written as:

\[ \chi_{\text{THG}}^{(3)}(m_1, m_2, m_3) = \frac{1}{\beta} \sum_n P(n, n + m_1, n + m_2, n + m_1 + m_2 + m_3), \]  

(S47)

where \( n \) stands for a short notation \( n = i\omega_n \) (\( \omega_n \) being a fermionic frequency), and \( m = i\omega_m \) (\( \omega_m \) being a bosonic frequency). For the third-harmonic generation we have \( m_1 = m_2 = m_3 = m \) and

\[ \chi_{\text{THG}}^{(3)}(m) = \frac{1}{\beta} \sum_n P(n, n + m, n + 2m, n + 3m). \]  

(S48)

After a straightforward algebra we perform the Matsubara summation and analytic continuation as \( i\omega_m \rightarrow \hbar \omega + i0^+ \), we obtain (see Section S6)

\[ \chi_{\text{THG}}^{(3)}(\omega) = \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left[ n_F(\epsilon) P_{\text{ARRR}}^{RRRR} - n_F(\epsilon + 3\hbar\omega) P_{\text{AAAA}}^{AAAA} + (n_F(\epsilon + \hbar\omega) - n_F(\epsilon)) P_{\text{ARRR}}^{ARRR} + (n_F(\epsilon + 2\hbar\omega) - n_F(\epsilon + \hbar\omega) - n_F(\epsilon + 2\hbar\omega)) P_{\text{AAAA}}^{AAAR} \right]. \]  

(S49)

Note that \( P_{\text{ARRR}}^{RRRR}(\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3) \) with \( \epsilon_j = \epsilon + j\omega + i\eta_j \) means that all frequency arguments are in the retarded channel (i.e. \( \eta_j \rightarrow 0^+ \)) while \( P_{\text{AAAA}}^{AAAA} \) implies that the first argument is in the advanced channel, \( \eta_1 \rightarrow -0^+ \), but the other are retarded, \( \eta_{j \neq 0} \rightarrow 0^+ \). Therefore, we have \( P_{\text{AAAA}} = (P_{\text{RRRR}})^* \). By knowing the response function, the third-harmonic optical conductivity reads

\[ \sigma_{\text{THG}}^{(3)}(\omega) = i \frac{\chi_{\text{THG}}^{(3)}(\omega)}{\omega^3}. \]  

(S50)

A. Different diagram contributions

Using Baym-Kadanoff analysis summarised in the previous section, we construct diagrams for the third-order response function in Dirac materials as illustrated in Fig. 1 of the main text. As depicted in Fig. 1a of the main text, the \( P \)-function contains four main different contributions, \( P = P_1 + P_2 + P_3 + P_4 \), associated respectively with square (\( P_1 \)), triangles (\( P_2, P_3 \)) and bubble (\( P_4 \)) diagrams. More explicitly we can write:

\[ P_1(z_0, z_1, z_2, z_3) = \left( \frac{e^4 v^2 N_f}{2\pi \hbar^2} \right) Q_1(z_0, z_1) Q_1(z_1, z_2) Q_1(z_2, z_3) \Omega_1(z_0, z_1, z_2, z_3) \]  

(S51)

The sum over spin and valley indices just leads to an overall degeneracy factor \( N_f = N_s N_v \), where \( N_s = 2 \) and \( N_v = 2 \). Note that \( \Omega_1(z_0, z_1, z_2, z_3) \) is the bare square diagram in the absence of vertex renormalization,

\[ \Omega_1(z_0, z_1, z_2, z_3) = \frac{2im_p}{2U} \sum_k \text{Tr} \left[ \hat{\sigma}_y \hat{G}(k, z_0) \hat{\sigma}_y \hat{G}(k, z_1) \hat{\sigma}_y \hat{G}(k, z_2) \hat{\sigma}_y \hat{G}(k, z_3) \right]. \]  

(S52)

and the one-photon vertex \( \Lambda_1(z_1, z_2) = (-ev\hat{\sigma}_y) \Lambda_1^{(0)}(z_1, z_2) \) where \( \Lambda_1^{(0)} = 1 \) is the bare one-photon vertex function and \( Q_1(z_1, z_2) \) is the one-photon Bethe-Salpeter renormalization factor which is given in the following subsection. In similar way we can write the contributions of the two triangles diagrams, namely

\[ P_2(z_0, z_1, z_2, z_3) = \left( \frac{e^4 v^2 N_f}{2\pi \hbar^2} \right) Q_1(z_2, z_3) Q_2(z_0, z_2) \Lambda_2^{(0)}(z_0, z_1, z_2) \Omega_2(z_0, z_2, z_3), \]  

(S53)

where

\[ \Omega_2(z_0, z_2, z_3) = -\frac{\gamma_{\text{imp}}}{2U} \sum_k \text{Tr} \left[ \hat{\sigma}_y \hat{G}(k, z_0) \hat{G}(k, z_2) \hat{\sigma}_y \hat{G}(k, z_3) \right], \]  

(S54)
Note that $\hat{\Lambda}_3(z_0, z_1, z_2) = (-e\nu\gamma)\Lambda_2^{(0)}(z_0, z_1, z_2)Q_2(z_0, z_2)$ is the two-photon vertex function where $\Lambda_2^{(0)}(z_0, z_1, z_2)$ is the unrenormalized two-photon vertex function and $Q_2(z_0, z_2)$ is the two-photon Bethe-Salpeter renormalization factor (see subsections below). We have also the further triangle diagram:

$$P_3(z_0, z_1, z_2, z_3) = \left(\frac{e^2\nu^2}{2\pi\hbar^2}\right) Q_1(z_0, z_1)Q_2(z_1, z_3)\Lambda_2^{(0)}(z_1, z_2, z_3)\Omega_3(z_0, z_1, z_3) .$$

where

$$\Omega_3(z_0, z_1, z_3) = -\frac{\gamma_{\text{imp}}}{2\nu} \sum_k \text{Tr}[\sigma_y \hat{G}(k, z_0)\sigma_y \hat{G}(k, z_1)\hat{G}(k, z_3)] ,$$

Finally we make the bubble term explicit:

$$P_4(z_0, z_1, z_2, z_3) = \left(\frac{e^2\nu^2}{2\pi\hbar^2}\right) \Lambda_3^{(0)}(z_0, z_1, z_2, z_3)Q_3(z_0, z_2)X_1(z_0, z_1),$$

where $\hat{\Lambda}_3(z_0, z_1, z_2, z_3) = (-e\nu\gamma_j)\Lambda_3^{(0)}(z_0, z_1, z_2, z_3)Q_3(z_0, z_3)$ is the three-photon vertex function with $\Lambda_3^{(0)}(z_0, z_1, z_2, z_3)$ being the unrenormalized three-photon vertex function and $Q_3(z_0, z_3)$ is the three-photon Bethe-Salpeter renormalization factor (see subsections below). The explicit expressions of $\Lambda_3^{(0)}(z_0, z_1, z_2)$, $\Lambda_2^{(0)}(z_0, z_1, z_2, z_3)$, and $Q_{n=1,2,3}(z_1, z_2)$, $\Omega_1(z_0, z_1, z_2, z_3)$, $\Omega_2(z_0, z_2, z_3)$, and $\Omega_3(z_0, z_1, z_3)$ are provided in the next subsections.

### B. Renormalization of the one-photon vertex

The one-photon vertex renormalization is diagrammatically depicted in Fig. 1b of the main text and it reads

$$\hat{\Lambda}_1(p, p + q; n, n + m) - \hat{\Lambda}_1^{(0)}(p, p + q; n, n + m) = \gamma_{\text{imp}} \sum_k \hat{G}(k, n)\hat{\Lambda}_1(k, n + q; n, n + m)\hat{G}(k + q, n + m) ,$$

where $m \equiv i\omega_n$ and $n \equiv i\kappa_n$ stand for the bosonic and fermionic Matsubara frequencies, respectively. Note that in the integrand we have shifted the dummy momentum $k$ as $k + p \rightarrow k$ and therefore we can see that vertex correction does not depend on the fermion momentum $p$. For the optical (or dipole) approximation we have $q = 0$. Therefore, the Bethe-Salpeter relation for the one-photon vertex function reads

$$\hat{\Lambda}_1(n, n + m) = \hat{\Lambda}_1^{(0)} + \gamma_{\text{imp}} \sum_k \hat{G}(k, n)\hat{\Lambda}_1(n, n + m)\hat{G}(k + m) .$$

Note that we have $\hat{\Lambda}_1^{(0)} = \delta \hat{G}_{\text{0}}^{-1}/\delta A|_{A=0} = -e\nu\gamma_j$ where $\hat{G}_0$ stands for the non-interacting Green’s function. We assume the following ansatz for the vertex function

$$\hat{\Lambda}_1(n, n + m) = a \hat{I} + b\sigma_x + (c + \nu)\sigma_y + d\sigma_z .$$

Using the fact that the integral of odd-function of $k$ is zero, we obtain $a = b = d = 0$ and eventually the following result for the vertex function $\hat{\Lambda}_1 = (-e\nu\gamma_j)\Lambda_1$ where

$$\Lambda_1(n, n + m) = \frac{\hat{\Lambda}_1^{(0)}}{1 - UX_1(n, n + m)} .$$

Note that $\Lambda_1^{(0)} = 1$ and

$$X_1(n, n + m) = \frac{\gamma_{\text{imp}}}{2\nu} \sum_k \text{Tr}[\sigma_y \hat{G}(k, n)\sigma_y \hat{G}(k, n + m)] .$$

Using dimensional regularization, we find the following formula for the $X_1$ function:

$$X_1(n, n + m) = \frac{S(n)S(n + m)}{S(n)^2 - S(n + m)^2} \ln \left[ \frac{S(n + m)^2}{S(n)^2} \right] .$$

The above equation can be straightforwardly generalized in the generic complex space by replacing $n \rightarrow i\omega_n \rightarrow z_1, n + m \rightarrow i\omega_n + i\omega_m \rightarrow z_2$. Eq. (S61) defines the one-photon vertex renormalization factor:

$$Q_1(z_1, z_2) = \frac{1}{1 - UX_1(z_1, z_2)} .$$

Obviously $X_1(z_1, z_2) = X_1(z_2, z_1)$.
C. Renormalization of the two-photon vertex

Similar to the one-photon vertex case, it can be shown that the two-photon vertex function is independent of the fermionic momentum, $p$. Moreover, for the optical limit we can neglect the photon momentum $q$.

The self-consistent Bethe-Salpeter relation for the two-photon vertex function is depicted in Fig. 1c of the main text and it reads

$$\hat{\Lambda}_2(n, n + m, n + 2m) = \hat{\Lambda}_2^{(0)}(n, n + m, n + 2m) + \gamma_{\text{imp}} \sum_k \hat{G}(k, n) \hat{\Lambda}_2(n, n + m, n + 2m) \hat{G}(k, n + 2m).$$  \hspace{1cm} (S65)

In the non-interacting Dirac system the “bare” two-photon vertex function is zero, $\hat{\Lambda}_2^{(0)} \propto \delta^2 \hat{G}_0^{-1}/\delta \hat{A}^2|_{\Lambda \to 0} = 0$, due to the linear momentum dependence of the Hamiltonian. However, due to interaction the unrenormalized two-photon vertex $\hat{\Lambda}_2^{(0)}$ is finite given by the following relation (see Fig. 1e of the main text)

$$\hat{\Lambda}_2^{(0)}(n, n + m, n + 2m) = -\gamma_{\text{imp}} \sum_k \hat{G}(k, n) \hat{\Lambda}_2(n, n + m) \hat{G}(k, n + m) \hat{\Lambda}_2(n, n + m, n + 2m) \hat{G}(k, n + 2m).$$  \hspace{1cm} (S66)

From now on we adopt the short-hand notation $z_j = n + jm$ with $j = 0, 1, 2, 3$. We find $\hat{\Lambda}_2^{(0)} = (-ev\hat{\sigma}_y)^2 \hat{\Lambda}_2^{(0)}$ with $\hat{\sigma}_y^2 = 1$ and

$$\hat{\Lambda}_2^{(0)}(z_0, z_1, z_2) = Q_1(z_0, z_1) Q_1(z_1, z_2) U Z(z_0, z_1, z_2),$$  \hspace{1cm} (S67)

in which $Q_1(z_j, z_j)$ is the one-photon renormalization factor defined in the previous subsection, and where

$$Z(z_0, z_1, z_2) = \frac{-\gamma_{\text{imp}}}{2U} \sum_k \text{Tr} \left[ \hat{G}(k, z_0) \hat{\sigma}_y \hat{G}(k, z_1) \hat{\sigma}_y \hat{G}(k, z_2) \right].$$  \hspace{1cm} (S68)

By performing the momentum integration using the dimensional regularization, we obtain

$$Z(z_0, z_1, z_2) = \frac{S(z_1)}{S(z_0) - S(z_2)} \left\{ \frac{S(z_0)}{S(z_0)^2 - S(z_1)^2} \ln \left( \frac{S(z_1)^2}{S(z_0)^2} \right) + \frac{S(z_2)}{S(z_2)^2 - S(z_1)^2} \ln \left( \frac{S(z_1)^2}{S(z_2)^2} \right) \right\}. \hspace{1cm} (S69)$$

In a compact form we can write

$$Z(z_0, z_1, z_2) = \frac{X_2(z_0, z_1) - X_2(z_1, z_2)}{S(z_0) - S(z_2)}. \hspace{1cm} (S70)$$

By solving the the self-consistent Bethe-Salpeter relation for the two-photon vertex given in Eq. (S65), we obtain $\hat{\Lambda}_2 = (-ev\hat{\sigma}_y)^2 \hat{\Lambda}_2$ with

$$\Lambda_2(z_0, z_1, z_2) = Q_2(z_0, z_2) \Lambda_2^{(0)}(z_0, z_1, z_2),$$  \hspace{1cm} (S71)

in which $Q_2(z_0, z_2)$ is the two-photon Bethe-Salpeter renormalization factor

$$Q_2(z_0, z_2) = \frac{1}{1 - UX_2(z_0, z_2)}.$$  \hspace{1cm} (S72)

and where

$$X_2(z, z') = \frac{-\gamma_{\text{imp}}}{2U} \sum_k \text{Tr} \left[ \hat{G}(k, z) \hat{G}(k, z') \right].$$  \hspace{1cm} (S73)

We explicitly obtain

$$X_2(z, z') = \frac{1}{S(z) - S(z')} \left\{ S(z) \ln \left( -\frac{W^2}{S(z)^2} \right) - S(z') \ln \left( -\frac{W^2}{S(z')^2} \right) \right\}. \hspace{1cm} (S74)$$

Note that $X_2(z, z') = X_2(z', z)$. Using the self-energy relation Eq. (S10) we have

$$S(z) \ln \left( -\frac{W^2}{S(z)^2} \right) = -\frac{\Sigma(z)}{U}. \hspace{1cm} (S75)$$
Therefore, the two-photon renormalization factor reads (Eq. (6) of the main text)

\[ Q_2(z, z') = \frac{1}{1 - UX_2(z, z')} = \frac{S(z) - S(z')}{{z - z'}}. \]  

(S76)

It is useful to evaluate of \( Q_{2RR}^2(\epsilon, \epsilon) \) which is given in Eq. (7) of the main text. Using the above relation in the retarded-retarded (RR) channel and employing the self-energy relation Eq. (S10), we obtain

\[ Q_{2RR}^2(\epsilon, \epsilon) = \frac{dS(\epsilon)}{d\epsilon} = 1 - \frac{d\Sigma(\epsilon)}{d\epsilon} \]

\[ = 1 + US'(\epsilon) \ln \left( -\frac{W^2}{S(\epsilon)^2} \right) - US(\epsilon) \frac{d\ln[S(\epsilon)^2]}{d\epsilon} \]

\[ = 1 + US'(\epsilon) \ln \left( -\frac{W^2}{S(\epsilon)^2} \right) - 2US'\epsilon). \]  

(S77)

Therefore, we have

\[ 1 + 2U - U \ln \left( -\frac{W^2}{S(\epsilon)^2} \right) \]  

(S78)

Consequently, we arrive at

\[ S'(\epsilon) = \frac{1}{1 + 2U - U \ln \left( -\frac{W^2}{S(\epsilon)^2} \right)}. \]  

(S79)

Using Eq. (S10), we obtain

\[ 1 - U \ln \left( -\frac{W^2}{S(\epsilon)^2} \right) = 1 + \frac{\Sigma(\epsilon)}{S(\epsilon)} = \frac{\mu_0 + \epsilon}{S(\epsilon)}. \]  

(S80)

Therefore, we find the following relation which is given in Eq. (7) of the main text.

\[ Q_{2RR}^R(\epsilon, \epsilon) = S'(\epsilon) = \frac{S(\epsilon)}{2US(\epsilon) + \mu_0 + \epsilon}. \]  

(S81)

**D. Renormalization of the three-photon vertex**

Similar to the case of two-photon case, the impurity scattering induces a finite three-photon vertex as defined in Fig. 1f of the main text. Accordingly we find \( \hat{A}_3^{(0)} = (-ev\delta_g)^3 \hat{A}_3^{(0)} \) with

\[ \hat{A}_3^{(0)}(z_0, z_1, z_2, z_3) = M_1(z_0, z_1, z_2, z_3) + M_2(z_0, z_1, z_2, z_3) + M_3(z_0, z_1, z_2, z_3) \]  

(S82)

where

\[ M_1(z_0, z_1, z_2, z_3) = U\Omega_1(z_0, z_1, z_2, z_3)Q_1(z_0, z_1)Q_1(z_1, z_2)Q_1(z_2, z_3), \]  

(S83)

\[ M_2(z_0, z_1, z_2, z_3) = U\Omega_2(z_0, z_1, z_2, z_3)\hat{A}_2^{(0)}(z_0, z_1, z_2, z_3)Q_1(z_2, z_3)Q_2(z_0, z_2), \]  

(S84)

\[ M_3(z_0, z_1, z_2, z_3) = U\Omega_3(z_0, z_1, z_2, z_3)\hat{A}_2^{(0)}(z_1, z_2, z_3)Q_1(z_0, z_3)Q_2(z_1, z_3). \]  

(S85)

Here \( Q_1(z, z), Q_2(z, z) \) are the Bethe-Salpeter one- and two-photon renormalization functions, respectively. The explicit expression for \( \Omega_1 \) function is given by

\[ \Omega_1(z_0, z_1, z_2, z_3) = u_1(z_0, z_1, z_2, z_3) \ln \left( \frac{S(z_0)^2}{S(z_1)^2} \right) + u_2(z_0, z_1, z_2, z_3) \ln \left( \frac{S(z_0)^2}{S(z_2)^2} \right) + u_3(z_0, z_1, z_2, z_3) \ln \left( \frac{S(z_0)^2}{S(z_3)^2} \right) \]  

(S86)
where
\[ u_1(z_0, z_1, z_2, z_3) = \frac{S(z_0)S(z_1)S(z_2)S(z_3) + [S(z_0)S(z_2) + S(z_1)S(z_3)]S(z_1)^2}{[S(z_1)^2 - S(z_0)^2][S(z_1)^2 - S(z_2)^2][S(z_1)^2 - S(z_3)^2]} , \quad (S87) \]
\[ u_2(z_0, z_1, z_2, z_3) = \frac{S(z_0)S(z_1)S(z_2)S(z_3) + [S(z_0)S(z_2) + S(z_1)S(z_3)]S(z_2)^2}{[S(z_2)^2 - S(z_0)^2][S(z_2)^2 - S(z_1)^2][S(z_2)^2 - S(z_3)^2]} , \quad (S88) \]
\[ u_3(z_0, z_1, z_2, z_3) = \frac{S(z_0)S(z_1)S(z_2)S(z_3) + [S(z_0)S(z_2) + S(z_1)S(z_3)]S(z_3)^2}{[S(z_3)^2 - S(z_0)^2][S(z_3)^2 - S(z_1)^2][S(z_3)^2 - S(z_2)^2]} . \quad (S89) \]

Similarly, one can obtain
\[ \Omega_2(z_0, z_2, z_3) = Z(z_2, z_3, z_0) , \quad (S90) \]
\[ \Omega_3(z_0, z_1, z_3) = Z(z_3, z_0, z_1) . \quad (S91) \]

Finally, the Bethe-Salpeter renormalization of the three-photon vertex function gives:
\[ \Lambda_3(z_0, z_1, z_2, z_3) = Q_3(z_0, z_3)A_3^{(0)}(z_0, z_1, z_2, z_3) . \quad (S92) \]

where
\[ Q_3(z_1, z_2) = \frac{1}{1 - UX_3(z_1, z_2)} \quad (S93) \]

Note that \( X_3(z_1, z_2) = X_1(z_1, z_2) \).

### S4. Linear Conductivity

Linear response function is obtained after performing a Matsubara summation as follows
\[ \chi^{(1)}(m) = \frac{1}{\beta} \sum_n P(n, n + m) . \quad (S94) \]

Note that \( \beta = 1/k_BT \) where \( k_B \) is the Boltzmann constant and \( T \) stands for the electronic temperature. The \( P \)-function is analytical in the complex plain except two branch cuts at \( \epsilon \) and \( \epsilon - m \) where \( \epsilon \) span over whole real axes. After performing the summation and an analytical continuation as \( m \rightarrow \hbar\omega + i0^+ \), we find [49]
\[ \chi^{(1)}(\omega) = \int_{-\infty}^{\infty} \frac{d\epsilon}{2\pi} \left\{ [n_F(\epsilon) - n_F(\epsilon + \hbar\omega)]P^{RR}(\epsilon, \epsilon + \hbar\omega) - n_F(\epsilon)P^{RR}(\epsilon, \epsilon + \hbar\omega) + n_F(\epsilon + \hbar\omega)P^{AA}(\epsilon, \epsilon + \hbar\omega) \right\} \quad (S95) \]

where \( n_F(x) = 1/(e^{\beta x} + 1) \) is the Fermi-Dirac distribution function. Note that “R” and “A” superscripts stand for the retarded and advanced, respectively. Accordingly, we have \( P^{RR}(\epsilon, \epsilon + \hbar\omega) = [P^{AA}(\epsilon, \epsilon + \hbar\omega)]^* = P(\epsilon + i0^+, \epsilon + \hbar\omega + i0^+) \) and \( P^{AR}(\epsilon, \epsilon + \hbar\omega) = P(\epsilon - i0^+, \epsilon + \hbar\omega + i0^+) \). The linear optical conductivity reads
\[ \sigma^{(1)}(\omega) = i \frac{\chi^{(3)}(\omega)}{\omega} . \quad (S96) \]

For the 2D Dirac model, the \( P \)-function reads
\[ P(n, n + m) = \frac{N_f e^2}{2\pi \hbar^2} \frac{X_1(n, n + m)}{1 - UX_1(n, n + m)} . \quad (S97) \]

Note that the sum over spin and valley index just leads to an overall degeneracy factor \( N_f = N_s N_v \) where \( N_s = 2 \) and \( N_v = 2 \). The linear \( dc \)-conductivity follows
\[ \sigma^{(1)}_{dc} = \frac{4e^2}{\pi^2} \left[ \frac{X_1^{AR}(0, 0)}{1 - UX_1^{AR}(0, 0)} - \frac{X_1^{RR}(0, 0)}{1 - UX_1^{RR}(0, 0)} \right] . \quad (S98) \]
where $\sigma_0 = e^2 / 4\hbar$. The retarded self-energy can be decomposed into its real and imaginary parts $\Sigma(\epsilon) = \Delta(\epsilon) - i\Gamma(\epsilon)$ with $\Delta(\epsilon)$ and $\Gamma(\epsilon) > 0$ being odd and even real functions, respectively. Using this notation we find $X_1^{RR}(0,0) = -1$, $X_1^{AR}(0,0) = w(x)$ with $x = \mu/\Gamma(\mu)$ in which the renormalized chemical potential is given by $\mu = \mu_0 - \Delta(\mu)$, and $w(x)$ reads

$$w(x) = \frac{1 + x^2}{2x} \arctan \left[ \frac{1 - x^2}{1 + x^2} \cdot \frac{2x}{1 + x^2} \right]. \quad (S99)$$

Therefore, we find

$$\sigma^{(1)}_{dc} = \sigma_0 f_1 \left( \frac{\mu}{\Gamma(\mu)}; U \right) \quad (S100)$$

where

$$f_1(x; U) = \frac{4}{\pi^2} \left\{ \frac{1}{1 + U} + \frac{w(x)}{1 - Uw(x)} \right\}. \quad (S101)$$

The functional dependence of $f_1(x; U)$ on $x$ and $U$ is illustrated in Fig. S2. In the constant-$\Gamma$ model, $\Sigma = -i\Gamma$, we have

$$f_1(x) = \frac{4[1 + w(x)]}{\pi^2}. \quad (S102)$$

The asymptotic form of $f_1(x)$ for small and large $x$ follows

$$f_1(x) \approx \frac{8}{\pi^2} \left[ 1 + \frac{x^2}{3} \right], \quad x \ll 1, \quad (S103)$$

$$f_1(x) \approx \frac{2}{\pi} \left( x + \frac{1}{x} \right), \quad x \gg 1. \quad (S104)$$
S5. NONLINEAR DC CONDUCTIVITY

A. Analytical derivation $f_3(x)$ function in the constant-$\Gamma$ model

Nonlinear dc conductivity is given by

$$
\sigma_{dc}^{(3)} = -\lim_{\omega \to 0} \frac{\text{Im}[\chi^{(3)}_{THG}(\omega)]}{\omega^3}
$$

(S105)

where the imaginary part of nonlinear (third-harmonic) response function follows

$$
\text{Im}[\chi^{(3)}_{THG}(\omega)] = \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi} \left\{ (n_F(\epsilon + 3h\omega) - n_F(\epsilon)) \text{Re}[P^{ARRR}_{RRRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \\
- (n_F(\epsilon + h\omega) - n_F(\epsilon)) \text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \\
- (n_F(\epsilon + 2h\omega) - n_F(\epsilon + h\omega)) \text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \\
- (n_F(\epsilon + 3h\omega) - n_F(\epsilon + 2h\omega)) \text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \right\}.
$$

(S106)

For small frequency we can expand the Fermi function as follows

$$
n_F(\epsilon + h\omega) - n_F(\epsilon) = h\omega \partial_n n_F(\epsilon) + \frac{(h\omega)^2}{2} \partial_n^2 n_F(\epsilon) + \frac{(h\omega)^3}{6} \partial_n^3 n_F(\epsilon) + \ldots
$$

\[= -(h\omega) \delta(\epsilon) - \frac{(h\omega)^2}{2} \partial_\epsilon \delta(\epsilon) - \frac{(h\omega)^3}{6} \partial_\epsilon^2 \delta(\epsilon) + \ldots. \]

(S107)

Notice that at zero temperature we have $n_F(\epsilon) = \Theta(-\epsilon)$ that is the Heaviside step function and its first derivative $\partial_n n_F(\epsilon) = -\delta(\epsilon)$ which stands for the Dirac delta function. After integration over $\epsilon$, we find

$$
\text{Im}[\chi^{(3)}_{THG}(\omega)] \approx -(h\omega) h_1(\omega) + \frac{(h\omega)^2}{2} h_2(\omega) - \frac{(h\omega)^3}{6} h_3(\omega)
$$

(S108)

where we define

$$
h_1(\omega) = 3\text{Re}[P^{ARRR}_{RRRR}(0, h\omega, 2h\omega, 3h\omega)] - \text{Re}[P^{ARRR}_{ARRR}(0, h\omega, 2h\omega, 3h\omega)] \\
- \text{Re}[P^{ARRR}_{ARRR}(0, h\omega, 2h\omega, 3h\omega)] - \text{Re}[P^{ARRR}_{ARRR}(0, h\omega, 2h\omega, 3h\omega)]
$$

(S109)

$$
h_2(\omega) = \lim_{\epsilon \to 0} \frac{\partial}{\partial \epsilon} \left\{ 9\text{Re}[P^{ARRR}_{RRRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] - \text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \\
- 3\text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] - 5\text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \right\}
$$

(S110)

$$
h_3(\omega) = \lim_{\epsilon \to 0} \frac{\partial^2}{\partial \epsilon^2} \left\{ 27\text{Re}[P^{ARRR}_{RRRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] - \text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \\
- 7\text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] - 19\text{Re}[P^{ARRR}_{ARRR}(\epsilon, \epsilon + h\omega, \epsilon + 2h\omega, \epsilon + 3h\omega)] \right\}
$$

(S111)

In the absence of vertex correction, we have

$$
P(z_0, z_1, z_2, z_3) = \left( \frac{e^4 V^2 N_f}{2\pi h^2} \right) \Omega_4(z_0, z_1, z_2, z_3).
$$

(S112)

We assume $\Sigma(\epsilon) = \mp i\Gamma$ where $\Gamma > 0$ is a phenomenological constant and $-/+$ stands for the retarded (R) and advanced (A) channel, respectively. For the case of $h\omega \ll \mu, \Gamma$, it is legitimate to expand the integrand for small $\omega$. In the constant-$\Gamma$ model, the contribution from $h_3(\omega)$ exactly cancels that of $h_3(\omega)$. Eventually, the nonlinear dc conductivity in the constant-$\Gamma$ model reads

$$
\sigma_{dc}^{(3)} = \left( \frac{e^4 V^2 N_f h^3}{4\pi^2 k^2 h_0^4} \right) \frac{t_0^4}{\Gamma^4} f_3 \left( \frac{\mu}{\Gamma} \right)
$$

(S113)
Considering $h \nu = \sqrt{3} t_0 a / 2$ and $N_f = 4$, we have
\[
\left\{ \frac{e^4 v^2 N_f h^3}{(2\pi)^2 h^2 t_0^4} \right\} = \frac{\sigma_0}{E_0^2}
\] (S114)
with $\sigma_0 = e^2/4h$ and
\[
E_0 = \frac{\pi t_0}{\sqrt{3} ea}
\] (S115)
where with $t_0 \approx 3eV$ and $a \approx 0.246nm$ we find $E_0 \approx 22V/nm$. The universal $f_3(x)$ function reads
\[
f_3(x) = \frac{3}{16} \left[ \frac{1}{x^4} - \frac{1}{x^2} \right] + \frac{2}{3} \left[ \frac{1}{1 + x^2} \right]^3 + \frac{1}{16} \left[ \frac{5}{x^2} - \frac{8}{1 + x^2} - \frac{3}{x^4} \right] u(x).
\] (S116)
There is a sign change in $f_3(x)$ at $x_0 \approx 0.655$. It is good to check the following asymptotic cases :
\[
f_3(x) \approx \frac{2}{3} \left( 1 - \frac{33}{x} \right), \quad x \ll 1,
\] (S117)
\[
f_3(x) \approx -\frac{3\pi}{32} \left( \frac{1}{x} - \frac{2}{3x^2} \right), \quad x \gg 1.
\] (S118)

**B. Numerical evaluation of $f_3(x, U)$ function in the full quantum theory**

The third-harmonic generation (THG) conductivity is given by
\[
\sigma_{\text{THG}}^{(3)}(\omega) = \frac{\lambda_{\text{THG}}^{(3)}(\omega)}{\omega^3} = \left( \frac{e^4 v^2 N_f}{2\pi h^2} \right) \frac{1}{\omega^3} \int_{-\infty}^{\infty} \frac{d\epsilon}{2\pi} \text{Re}[K(\epsilon, \omega)]
\] (S119)
where
\[
K(\epsilon, \omega) = \left( \frac{e^4 v^2 N_f}{2\pi h^2} \right)^{-1} \left\{ n_F(\epsilon) P_{\text{RRRR}}^{\text{PRRR}} - n_F(\epsilon + 3h\omega) P_{\text{AAAA}}^{\text{AAAA}} + (n_F(\epsilon + h\omega) - n_F(\epsilon)) P_{\text{ARRR}}^{\text{ARRR}} + (n_F(\epsilon + 2h\omega) - n_F(\epsilon + h\omega)) P_{\text{AAPR}}^{\text{AAPR}} + (n_F(\epsilon + 3h\omega) - n_F(\epsilon + 2h\omega)) P_{\text{AARR}}^{\text{AARR}} \right\}.
\] (S120)
Therefore, we obtain
\[
\sigma_{\text{THG}}^{(3)}(\omega) = \frac{\sigma_0}{E_0^2} \frac{t_0^4}{(h\omega)^3} \int_{-\infty}^{\infty} d\epsilon \text{Re}[K(\epsilon, \omega)].
\] (S121)
We evaluate the third-order dc conductivity as dc limit of the third-harmonic conductivity,
\[
\sigma_{\text{dc}}^{(3)} = \lim_{\omega \to 0} \sigma_{\text{THG}}^{(3)}(\omega) = \frac{\sigma_0}{E_0^2} \frac{t_0^4}{\Gamma(\mu)} \int_{-\infty}^{\infty} d\epsilon \text{Re}[K(\epsilon, \omega)].
\] (S122)
where the universal $f_3$ function can be evaluated numerically by using the following relation
\[
f_3 \left( \frac{\mu}{\Gamma(\mu)} ; U \right) = \lim_{\omega \to 0} \frac{\Gamma(\mu)^4}{(h\omega)^3} \int_{-\infty}^{\infty} d\epsilon \text{Re}[K(\epsilon, \omega)].
\] (S123)
S6. ANALYTICAL CONTINUATION FOR THE THIRD ORDER RESPONSE FUNCTION

The summation of the fermionic Matsubara frequency \( n \) is performed by the contour integration technique.

\[
B(m) = \frac{1}{\beta} \sum_n P(n, n + m, n + 2m, n + 3m) .
\]  

(S124)

The \( P \)-function contains four branch cut in complex plane. The Matsubara summation is performed on a contour with four cuts at \( \epsilon, \epsilon - m, \epsilon - 2m, \) and \( \epsilon - 3m \), see Fig. S3. Note that \( \epsilon \) runs over the entire real axes. Therefore, we write

\[
B(m) = \oint \frac{dz}{2\pi i} n_F(z) P(z, z + m, z + 2m, z + 3m)
\]

\[
= \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon + i\eta) P(\epsilon + i\eta, \epsilon + i\eta + m, \epsilon + i\eta + 2m, \epsilon + i\eta + 3m) - (\eta \to -\eta) \right)
\]

\[
+ \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon - m + i\eta) P(\epsilon - m + i\eta, \epsilon + i\eta + m, \epsilon + i\eta + 2m) - (\eta \to -\eta) \right)
\]

\[
+ \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon - 2m + i\eta) P(\epsilon - 2m + i\eta, \epsilon - m + i\eta, \epsilon + i\eta + m) - (\eta \to -\eta) \right)
\]

\[
+ \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon - 3m + i\eta) P(\epsilon - 3m + i\eta, \epsilon - 2m + i\eta, \epsilon - m + i\eta, \epsilon + i\eta) - (\eta \to -\eta) \right)
\]

(S125)

where \( n_F(z) = 1/(1+e^{\beta z}) \) is the Fermi-Dirac distribution function. Note that \( m \) stands for external bosonic Matsubara (imaginary) frequency. Since \( \eta \ll |m| \), we have \( m + i\eta \to m \) which implies

\[
B(m) = \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon + i\eta) P(\epsilon + i\eta, \epsilon + m, \epsilon + 2m, \epsilon + 3m) - (\eta \to -\eta) \right)
\]

\[
+ \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon - m) P(\epsilon - m, \epsilon + m, \epsilon + 2m) - (\eta \to -\eta) \right)
\]

\[
+ \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon - 2m) P(\epsilon - 2m, \epsilon - m, \epsilon + m) - (\eta \to -\eta) \right)
\]

\[
+ \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left( n_F(\epsilon - 3m) P(\epsilon - 3m, \epsilon - 2m, \epsilon - m, \epsilon + i\eta) - (\eta \to -\eta) \right)
\]

(S126)
Note that for bosonic frequency $m$ we have $n_F(\epsilon - m) = n_F(\epsilon)$ and therefore we find

$$B(m) = \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P(\epsilon + i\eta, \epsilon + m, \epsilon + 2m, \epsilon + 3m) - (\eta \to -\eta) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P(\epsilon - m, \epsilon + i\eta, \epsilon + m, \epsilon + 2m) - (\eta \to -\eta) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P(\epsilon - 2m, \epsilon - m, \epsilon + i\eta, \epsilon + m) - (\eta \to -\eta) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P(\epsilon - 3m, \epsilon - 2m, \epsilon - m, \epsilon + i\eta) - (\eta \to -\eta) \} \ . \quad (S17)$$

We do analytical continuation as $m \to \hbar \omega + i\eta$:

$$B(\omega) = \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{RRRR}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega)$$

$$- P_{RRRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{ARRR}^{\mathbb{A}}(\epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) - P_{ARRR}^{\mathbb{A}}(\epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega, \epsilon + 2\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{ARRR}^{\mathbb{A}}(\epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega) - P_{ARRR}^{\mathbb{A}}(\epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{AAAR}^{\mathbb{A}}(\epsilon - 3\hbar \omega, \epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon)$$

$$- P_{AAAR}^{\mathbb{A}}(\epsilon - 3\hbar \omega, \epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon) \} \ . \quad (S128)$$

Since $\eta \to 0^+$, we have

$$B(\omega) = \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{RRRR}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega)$$

$$- P_{RRRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{ARRR}^{\mathbb{A}}(\epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) - P_{ARRR}^{\mathbb{A}}(\epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega, \epsilon + 2\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{ARRR}^{\mathbb{A}}(\epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega) - P_{ARRR}^{\mathbb{A}}(\epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon, \epsilon + \hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{AAAR}^{\mathbb{A}}(\epsilon - 3\hbar \omega, \epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon)$$

$$- P_{AAAR}^{\mathbb{A}}(\epsilon - 3\hbar \omega, \epsilon - 2\hbar \omega, \epsilon - \hbar \omega, \epsilon) \} \ . \quad (S129)$$

Note that “R” and “A” superscript stand for the retarded and advanced, respectively. We shift $\epsilon$ in such a way that all $P$-function arguments is $(\epsilon, \epsilon + \hbar \omega, \epsilon + 2\hbar \omega, \epsilon + 3\hbar \omega)$:

$$B(\omega) = \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{RRRR}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega)$$

$$- P_{RRRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{ARRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega)$$

$$- P_{ARRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{ARRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega)$$

$$- P_{ARRR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) \}$$

$$+ \int_{-\infty}^{+\infty} \frac{de}{2\pi i} n_F(\epsilon) \{ P_{AAAR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega)$$

$$- P_{AAAR}^{\mathbb{A}}(\epsilon, \epsilon + \hbar \omega, \epsilon + \hbar \omega, \epsilon + 3\hbar \omega) \} \ . \quad (S130)$$
Eventually, we obtain

\[
B(\omega) = \int_{-\infty}^{+\infty} \frac{d\epsilon}{2\pi i} \left\{ \left[ n_F(\epsilon) P^{RRRR} - n_F(\epsilon + 3h\omega) P^{AAAA} \right] + (n_F(\epsilon + h\omega) - n_F(\epsilon)) P^{ARRR} \right. \\
+ (n_F(\epsilon + 2h\omega) - n_F(\epsilon + h\omega)) P^{AARR} + (n_F(\epsilon + 3h\omega) - n_F(\epsilon + 2h\omega)) P^{AAAR} \right\}. 
\]

(S131)

S7. NUMERICAL EVALUATION OF THE NONLINEAR OPTICAL AND DC CONDUCTIVITIES

In Fig. S4, we plot the frequency dependence of the third-harmonic generation (THG) response function \(\sigma^{(3)}_{\text{THG}}(\omega)\) for \(\mu = 17\) THz and few representative values of \(U\). For comparison we show also the THG optical response \(\sigma^{(3)}_{\text{THG}}(\omega)\) for non-interacting electrons which is null for \(\hbar\omega \leq 2\mu/3\). Note the change of sign of \(\sigma^{(3)}_{\text{dc}}\) varying the scattering strength.

![Graph showing the frequency dependence of the third-harmonic optical conductivity versus frequency.](image)

**FIG. S4:** Real part of third-harmonic optical conductivity versus frequency in comparison with the non-interacting result. Note that the chemical potential is set \(\mu = 17\) THz, and \(\sigma^{(3)}_0 = \sigma_0/E_0^2\).

In Fig. S5, we illustrate the universal scaling of \(f_3\) versus \(U\) in the quantum regime for different values of \(x = \mu/\Gamma(\mu) < 1\). As seen the slop of the curves in the log-log scale plot does not strongly depends on the value of \(x\) which support the validity of Eq. (5) given in the main text.

In Fig. S6, we show the phase diagram for the constant-\(\Gamma\) model. As it is seen this phase diagram is completely different from that of the full quantum theory which is given in Fig. 3c of the main text. We can see only one sign-change in the constant-\(\Gamma\) model in contrast to that of full quantum theory which gives two sing-changes. Unlike the full quantum theory, the constant-\(\Gamma\) model predicts a positive nonlinear correction in the quantum regime.
FIG. S5: Log-log scale plot for the absolute value of the universal $f_3(\mu, U)$ function versus $U$ at $x = 0.1, 0.4$ and $0.7$ which are respectively depicted in panel a, b and c. Different lines correspond to the individual contribution of diagrams in Fig. 1a of the main text.

FIG. S6: Colormap plot for $g = |\sigma^{(3)}_{dc} E^2/\sigma^{(1)}_{dc}|$ factor with $E = 1\text{mV}/\text{nm}$ versus chemical potential $\mu$ and scattering rate $\Gamma$ in the constant-$\Gamma$ model, $\Sigma = -i\Gamma$. The sign of $\sigma^{(3)}_{dc}$ is written on the plot where the sign-switch border is highlighted by a dashed red line. Green and blue dotted lines stand for the contour lines with $g = 1$ and $g = 0.1$, respectively.