Optical bounds on many-electron localization

Ivo Souza[1,2], Richard M. Martin[3,4] and Massimiliano Stengel[5,6]

1 Centro de Física de Materiales, Universidad del País Vasco, 20018 San Sebastián, Spain
2 Ikerbasque Foundation, 48013 Bilbao, Spain
3 Department of Physics, University of Illinois at Urbana-Champaign,Urbana, Illinois 61801, USA
4 Department of Applied Physics, Stanford University, Stanford, California 94305, USA
5 Institut de Ciència de Materials de Barcelona (ICMAB-CSIC), Campus UAB, 08193 Bellaterra, Spain
6 ICREA-Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain

Abstract

We establish rigorous inequalities between different electronic properties linked to optical sum rules, and organize them into weak and strong bounds on three characteristic properties of insulators: electron localization length ℓ (the quantum fluctuations in polarization), electric susceptibility χ, and optical gap E_G. All-electron and valence-only versions of the bounds are given, and the latter are found to be more informative. The bounds on ℓ are particularly interesting, as they provide reasonably tight estimates for an ellusive ground-state property – the average localization length of valence electrons – from tabulated experimental data: electron density, high-frequency dielectric constant, and optical gap. The localization lengths estimated in this way for several materials follow simple chemical trends, especially for the alkali halides. We also illustrate our findings via analytically solvable harmonic oscillator models, which reveal an intriguing connection to the physics of long-ranged van der Waals forces.

Received Date
Accepted Date
Published Date

Contents

1 Introduction

The low-frequency electronic conductivity

$$\sigma_{aa}(\omega) = \mathrm{Re}\,\sigma_{aa}(\omega) + i\mathrm{Im}\,\sigma_{aa}(\omega) \tag{1}$$

displays sharply different behaviors in metals and in insulators. To characterize those behaviors one may define

$$D_{aa} = \pi \lim_{\omega\to 0} \omega\,\mathrm{Im}\,\sigma_{aa}(\omega)\,, \tag{2a}$$

$$\epsilon_0 \chi_{aa} = -\lim_{\omega\to 0} \omega^{-1}\,\mathrm{Im}\,\sigma_{aa}(\omega)\,, \tag{2b}$$

where ϵ_0 is the vacuum permittivity. The Drude weight D_{aa} is finite in metals and vanishes in insulators, whereas the clamped-ion electric susceptibility χ_{aa} is finite in insulators and diverges in metals. The $1/\omega$ divergence of $\mathrm{Im}\,\sigma_{aa}(\omega)$ in perfect conductors is due to the acceleration of free electrons under an applied electric field, and its linear decrease with ω in insulators reflects the polarization of bound electrons in reaction to the field.

In 1964, Kohn proposed electron localization as the essential property of the insulating state, and showed that it leads directly to its distinctive electrical behavior [1]. He argued that the ground-state wave function $\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ of an insulator in a periodic supercell breaks up into a sum of functions, $\Psi = \sum_M \Psi_M$, which are localized in disconnected regions of configuration space and have essentially vanishing overlap.

Kohn went on to show that the disconectedness of Ψ allows for the definition of an effective center-of-mass operator $\mathbf{X}/N$, even though the bare center-of-mass operator operator $(1/N)\sum_{i=1}^{N} \mathbf{r}_i$ is ill-defined under periodic boundary conditions. The operator $\mathbf{X}$ is based on sawtooth functions, whose discontinuities are placed in regions of configuration space where Ψ becomes exponentially small [2].

The importance of $\mathbf{X}$ can be seen from the fact that its ground-state expectation value yields the electronic contribution to the macroscopic electric polarization ($\mathbf{P}$),

$$\mathbf{P}_\mathrm{e} = -|e|\langle\mathbf{X}\rangle/V\,, \tag{3}$$

where V is the supercell volume. Thanks to the development of the modern theory of polarization, $\mathbf{P}$ is now understood as a fundamental bulk property of crystalline insulators, independent of surface termination modulo a discrete quantum of indeterminacy. In particular, within

a single-particle band picture, Eq. (3) reduces to a sum over the Wannier centers (Kohn's disconnected wave function pieces Ψ_M [1,2] can be viewed as "many-body Wannier functions"), or can be equivalently written as a Berry phase in momentum space [3, 4]. Crucially, this theory asserts that bulk polarization is a property of the wave function and not of the charge density – in line with Kohn's view on electron localization.

In addition, Kohn's center-of-mass operator allows for the definition of an electron localization tensor [5]

$$\langle r_a r_b \rangle_{\mathrm{c}} = \frac{1}{N} \left[\langle X_a X_b \rangle - \langle X_a \rangle \langle X_b \rangle \right], \tag{4}$$

where the subscript c (not a Cartesian index like a and b) stands for "cumulant moment." The diagonal entries of this tensor carry the interpretation of a localization length squared, averaged over the total number of electrons, along the corresponding direction:

$$\ell_a^2 = \langle r_a^2 \rangle_{\mathrm{c}} . \tag{5}$$

As in the case of $\mathbf{P}$, $\langle r_a r_b \rangle_{\mathrm{c}}$ also enjoys an elegant formulation in the framework of band theory, where it can be written as a quantum metric tensor [6] of the valence Bloch manifold [5,7,8], whose Cartesian trace is related to the Wannier spread [9]. First-principles studies of the localization length have been carried out for tetrahedrally coordinated semiconductors [10] and oxides [11].

Kohn did not directly relate the degree of wave function localization to any physical observable. An important step in that direction was taken shortly before the modern theory of polarization was developed. In Ref. [12], Kudinov proposed to quantify electron localization in insulators via the quantum fluctuations in the ground-state polarization [Eqs. (4) and (5)], connecting them to the optical absorption spectrum by means of a fluctuation-dissipation relation [13]. For a bulk crystal, such relation at zero temperature takes the form [5]

$$\langle r_a r_b \rangle_{\mathrm{c}} = \frac{\hbar}{\pi e^2 n_{\mathrm{e}}} \int_0^\infty d\omega \, \omega^{-1} \, \mathrm{Re}\, \sigma^{\mathrm{S}}_{ab}(\omega), \tag{6}$$

where $n_{\mathrm{e}} = N/V$ is the electron density, S indicates the symmetric part of the tensor, and the integral spans the positive-frequency optical absorption spectrum.[1] The trace $\langle r^2 \rangle_{\mathrm{c}} = \sum_a \ell_a^2$ of the localization tensor diverges in conductors by virtue of their nonzero DC Ohmic conductivity, while in insulators it remains finite.

The fluctuation-dissipation relation written above assumes a vanishing macroscopic electric field $\mathbf{E}$, as appropriate for transverse long-wave excitations. The needed generalization to accomodate more general electrical boundary conditions was given by Resta [14]. In particular, for longitudinal excitations where $\mathbf{E} = -\mathbf{P}/\epsilon_0$ ($\mathbf{D} = \mathbf{0}$), the fluctuation-dissipation relation becomes a sum rule for the energy-loss spectrum [14, 15],

$$\langle \widetilde{r_a r_b} \rangle_{\mathrm{c}} = \frac{\hbar \epsilon_0}{\pi e^2 n_{\mathrm{e}}} \int_0^\infty d\omega \, \mathrm{Im} \left[-\frac{1}{\epsilon(\omega)} \right]^{\mathrm{S}}_{ab} . \tag{7}$$

The quantum fluctuations encoded in the localization tensor clearly depend on the electrical boundary conditions, and it is only under $\mathbf{E} = \mathbf{0}$ as assumed in Eq. (6) that its trace discriminates between insulators and metals [14]. In the following, we will deal mostly with the transverse localization tensor; when referring to longitudinal quantities in Secs. 3 and 4.4, we will denote them with a tilde as done above.

[1] The conductivity tensor can be decomposed in three different ways: real and imaginary parts, $\mathrm{Re}\,\sigma$ and $\mathrm{Im}\,\sigma$; symmetric and antisymmetric parts, σ^{S} and σ^{A}; Hermitian and anti-Hermitian parts, σ^{H} and σ^{AH}. The Hermitian part of σ (and hence $\mathrm{Re}\,\sigma^{\mathrm{S}}$) is dissipative, while its anti-Hermitian part is reactive.

Length relations	References	Comments	Energy relations
$\ell \le \ell_{++}$	[5]	$\ell_{++}^2 \propto 1/E_G$	$E_L \ge E_G$
		Weak upper bound	
$\ell_- \le \ell$	[16,17]	$\ell_-^2 \propto \chi E_G/n_e$	$E_P^2/E_G \ge E_L$
		Lower bound	
		Sum-rule derivation in [17]	
$\ell \le \ell_+$	[15,18],	$\ell_+^2 \propto \sqrt{\chi/n_e}$	$E_L \ge E_P$
	this work	Strong upper bound	
$\ell_+ \le \ell_{++}$	This work	Equivalent to $\ell_- \le \ell_+, \ell_{++}$	$E_P \ge E_G$
$\ell_- \le \ell \le \ell_+ \le \ell_{++}$	This work	Chained inequalities	$E_P^2/E_G \ge E_L \ge E_P \ge E_G$

Table 1: Overview of the sum-rule inequalities on ℓ discussed in the present work. Those inequalities relate the electron localization length ℓ defined by Eq. (5) to the optical gap E_G, the clamped-ion electric susceptibility χ, and the electron density n_e. The works cited formulate the bounds for crystals; similar relations involving the electric polarizability of atoms had been previously established [19]. In Refs. [15,18] the relation $\ell \le \ell_+$ is formulated not for ℓ but for $2\pi n_e \ell^2$, which for uncorrelated band insulators gives the quantum metric traced over the filled bands: see Appendix A. The last column contains equivalent energy relations [17,19] expressed in terms of the localization gap E_L and the Penn gap E_P, which will be defined shortly [see Eq. (11)].

Although Eq. (6) provides a way of extracting the transverse localization length ℓ from the optical absorption spectrum, we are not aware of any experimental work in that direction. As discussed in Ref. [17], an alternative is to estimate ℓ via rigorous upper and lower bounds involving readily-available experimental data: electron density n_e, electric susceptibility χ, and minimum optical gap E_G (see Table 1). This approach was used recently to estimate $2\pi n_e \ell_\alpha^2$ (the quantum metric of the filled bands) for a number of materials [15,20].

In this work, we employ a sum-rule approach [19] to estabish weak and strong bounds on ℓ, χ, and E_G. We give two formulations of the bounds – all electron and valence-only – and argue that the valence-only formulation, even if approximate, is more informative. This is confirmed by an explicit evaluation of the bounds on ℓ for a series of materials; the strong bound is found to be much tighter than the weak one, and the valence-only formulation reveals simple chemical trends. To illustrate the impact of long-ranged electrostatics on the polarization fluctuations [14], we apply our formalism to analytically solvable systems of harmonic oscillators. This exercise reveals an intriguing connection to the physics of van der Waals (dispersion) forces, and clarifies the central role of electron-electron correlation in the determination of the optical bounds.

The manuscript is organized as follows. In Sec. 2 the inverse moments of the optical absorption spectrum are introduced, the sum rules for the three leading moments are stated, and average optical gaps are defined. In Sec. 3, sum-rule inequalities are established for the inverse moments and for the average gaps; the latter are then organized into chained inequalities, from which various bounds on ℓ, χ, and E_G are deduced. In Sec. 4 those bounds are examined for several exactly-solvable models, including harmonic oscillator models coupled by dispersion interactions. In Sec. 5 the localization length ℓ is estimated for various materials using the all-electron and valence-only varieties of the bounds, and the observed trends are discussed. We conclude in Sec. 6 with a summary, and provide some accessory results in three appendices.

2 Sum rules and average gaps

For light with linear polarization along direction $\hat{\mathbf{n}}$, we define the inverse moments of the optical absorption spectrum at zero temperature as

$$I_p(\hat{\mathbf{n}}) = \frac{2}{\pi}\int_0^\infty d\omega\, \omega^{-p}\, \mathrm{Re}\,\sigma^{\mathrm{S}}_{ab}(\omega)\hat{n}_a\hat{n}_b\,, \tag{8}$$

where a summation over repeated Cartesian indices is implied; the integer p and the absorption strength $\mathrm{Re}\,\sigma^{\mathrm{S}}_{ab}(\omega)\hat{n}_a\hat{n}_b$ are both non-negative (the latter follows from the assumption that the unperturbed system is in the ground state), and the $2/\pi$ factor was included for convenience in writing the sum rules below. For simplicity we will assume cubic symmetry or higher so that $\sigma^{\mathrm{S}}_{ab} = \sigma^{\mathrm{S}}\delta_{ab}$, rendering I_p independent of $\hat{\mathbf{n}}$,

$$I_p = \frac{2}{\pi}\int_0^\infty d\omega\, \omega^{-p}\, \mathrm{Re}\,\sigma^{\mathrm{S}}(\omega)\,. \tag{9}$$

The inverse spectral moments with $p = 0, 1, 2$ satisfy

$$I_0 = \frac{e^2 n_{\mathrm{e}}}{m_{\mathrm{e}}} \equiv \epsilon_0\omega_{\mathrm{p}}^2\,, \tag{10a}$$

$$I_1 = \frac{2e^2}{\hbar} n_{\mathrm{e}}\ell^2\,, \tag{10b}$$

$$I_2 = \epsilon_0\chi \equiv \epsilon_0(\epsilon - 1)\,, \tag{10c}$$

where we have introduced the plasma frequency ω_{p} and the static electronic permittivity ϵ (often denoted as ϵ_∞). The above identities are respectively the oscillator-strength sum rule,[2] the fluctuation-dissipation relation of Eq. (6) [with ℓ^2 defined by Eq. (5)], and the electric-susceptibility sum rule; the corresponding relations for atomic systems are given in Ref. [19]. All three sum rules converge for insulators, while in metals I_1 and I_2 diverge as a result of the nonzero DC conductivity. Equations (10a) and (10c) follow from the Kramers-Krönig relations, which in the case of (10a) must be complemented by general arguments concerning the behavior of the permittivity at high frequencies [22].

In addition, we find it useful to define a "localization gap" E_{L} and a "Penn gap" E_{P} as [17]

$$E_{\mathrm{L}}^{-1} = \frac{\hbar^{-1}I_1}{I_0}\,, \qquad E_{\mathrm{P}}^{-2} = \frac{\hbar^{-2}I_2}{I_0}\,; \tag{11}$$

these expressions can be interpreted as average inverse excitation energies (or energies squared) weighted by the transition strength [19]. Using Eq. (10) and writing $\hbar^2/2m_{\mathrm{e}}$ as $a_0^2\,\mathrm{Ry}$ (a_0 is the Bohr radius and Ry is the Rydberg unit of energy), we obtain

$$E_{\mathrm{L}} = \frac{\hbar^2}{2m_{\mathrm{e}}\ell^2} \Leftrightarrow \left(\frac{\ell}{a_0}\right)^2 = \frac{\mathrm{Ry}}{E_{\mathrm{L}}} \tag{12}$$

and

$$\chi = \left(\frac{\hbar\omega_{\mathrm{p}}}{E_{\mathrm{P}}}\right)^2\,, \tag{13}$$

the latter being the standard definition of the Penn gap in semiconductor physics [23, 24].

As shown below and already indicated in Table 1, the inequalities of interest can be expressed concisely as relations among three characteristic energy scales of the band structure: optical gap E_{G} (the energy threshold for optical absorption), Penn gap, and localization gap.

[2] The oscillator-strength sum rule gets modified for tight-binding models [21], and also when nonlocal pseudopotentials are used in the context of first-principles calculations.

3 Sum-rule inequalities

From the fact that $\mathrm{Re}\,\sigma^{\mathrm{S}}(\omega) \geq 0$ one can readily establish two types of inequalities involving different spectral moments [19]. The first type are of the form

$$I_{p+q} \leq \frac{\hbar}{E_{\mathrm{G}}} I_{p+q-1} \leq \ldots \leq \left(\frac{\hbar}{E_{\mathrm{G}}}\right)^{q} I_p \,, \tag{14}$$

where $q > 0$; the second,

$$I_p^2 \leq I_{p-1} I_{p+1} \,, \tag{15}$$

follow from the Cauchy-Bunyakovsky-Schwarz inequality

$$\left(\int_0^\infty d\omega\, f(\omega) g(\omega)\right)^2 \leq \left(\int_0^\infty d\omega\, f(\omega)^2\right)\left(\int_0^\infty d\omega\, g(\omega)^2\right) \tag{16}$$

by setting $f(\omega) = \omega^{-(p-1)/2}\sqrt{\mathrm{Re}\,\sigma^{\mathrm{S}}(\omega)}$ and $g(\omega) = \omega^{-(p+1)/2}\sqrt{\mathrm{Re}\,\sigma^{\mathrm{S}}(\omega)}$. Both types of inequalities become saturated in the limit of a narrow absorption spectrum concentrated at E_{G} [19].

The average gaps introduced in Eq. (11) satisfy

$$E_{\mathrm{L}} \geq E_{\mathrm{P}} \geq E_{\mathrm{G}} \,, \qquad E_{\mathrm{P}}^2 \geq E_{\mathrm{G}} E_{\mathrm{L}} \,, \tag{17}$$

with the relation $E_{\mathrm{L}} \geq E_{\mathrm{P}}$ coming from Eq. (15) and the others from Eq. (14).[3] As expected, the average gaps E_{L} and E_{P} cannot be smaller than the minimum gap E_{G}.

Equation (17) allows to bracket E_{L} as $E_{\mathrm{P}}^2/E_{\mathrm{G}} \geq E_{\mathrm{L}} \geq E_{\mathrm{P}} \geq E_{\mathrm{G}}$ and E_{P} as $E_{\mathrm{L}}^2 \geq E_{\mathrm{P}}^2 \geq E_{\mathrm{G}} E_{\mathrm{L}} \geq E_{\mathrm{G}}^2$; combined with Eqs. (12) and (13), these chained inequalities yield

$$\frac{\epsilon_0 \chi E_{\mathrm{G}}}{2e^2 n_{\mathrm{e}}} \leq \ell^2 \leq \frac{\hbar}{2|e|}\sqrt{\frac{\epsilon_0 \chi}{m_{\mathrm{e}} n_{\mathrm{e}}}} \leq \frac{\hbar^2}{2 m_{\mathrm{e}} E_{\mathrm{G}}} \qquad \left(\ell_-^2 \leq \ell^2 \leq \ell_+^2 \leq \ell_{++}^2\right), \tag{18a}$$

$$\frac{4e^2 m_{\mathrm{e}} n_{\mathrm{e}} \ell^4}{\epsilon_0 \hbar^2} \leq \chi \leq \frac{2e^2 n_{\mathrm{e}} \ell^2}{\epsilon_0 E_{\mathrm{G}}} \leq \frac{\hbar^2 e^2 n_{\mathrm{e}}}{\epsilon_0 m_{\mathrm{e}} E_{\mathrm{G}}^2} \,. \tag{18b}$$

As already indicated in Table 1, we will refer to ℓ_- as the lower bound on ℓ, and to ℓ_+ and ℓ_{++} as the strong and weak upper bounds, respectively; the same terminology will be used for the bounds on χ. The weak upper bounds on ℓ [5] and on χ [25] reflect the intuitive notion that wide-gap materials tend to have more localized and less polarizable electrons.

The bounds on ℓ are particularly interesting, as they only involve parameters that are tabulated for many materials: electron density, electric susceptiblity, and optical gap. Since ℓ itself is not commonly measured, those bounds provide a simple and practical way of estimating its value. Note that the weak upper bound ℓ_{++} only depends on the inverse minimum gap; this is a delicate quantity, especially for narrow-gap semiconductors, and it is not representative of the entire spectrum (the nature of the electron system can be very different for materials with the same minimum gap). The localization length is instead a global property of the electron system, and the value of E_{G} is not its most relevant descriptor; for example, ℓ_{++} diverges in the same way for all materials as E_{G} is tuned to zero. We therefore expect ℓ_{++} to give a relatively poor estimate for ℓ in real systems. The strong upper bound ℓ_+ depends instead on χ and n_{e} via the average Penn gap, which is much more representative of the entire spectrum. As for the lower bound ℓ_-, it depends on both E_{G} and E_{P}; there is still some dependence on the minimum gap, but it is a smaller effect than for ℓ_{++}.

[3] For an alternative derivation of $E_{\mathrm{L}} \geq E_{\mathrm{P}}$ (or $I_1^2 \leq I_0 I_2$), see Refs. [18, 19]. In the notation of Ref. [19], that relation reads $m_0^2 \leq m_{-1} m_1$.

The relations in Eq. (17) can also be arranged as $E_G \le E_P^2/E_L \le E_P \le E_L$ to place bounds on the optical gap,

$$E_G \le \frac{2e^2 n_e \ell^2}{\epsilon_0 \chi} \le \hbar|e|\sqrt{\frac{n_e}{m_e \epsilon_0 \chi}} \le \frac{\hbar^2}{2m_e \ell^2}\,. \tag{19}$$

It is significant that there are several upper bounds, but no lower bound. This is consistent with the existence of electronic systems without an energy gap that are strict insulators [1,2].

Although we have been focusing on transverse long-wave modes, similar results hold for longitudinal ones [15,25]. As shown in Appendix B, the only changes to Eqs. (18) and (19) are

$$E_G \to \tilde{E}_G\,, \qquad \ell \to \tilde{\ell}\,, \qquad \chi \to 1-\epsilon^{-1}\,, \tag{20}$$

where $\tilde{E}_G$ is the minimum energy for long-wave longitudinal excitations (plasmon gap), and $\tilde{\ell}$ is the localization length associated with the longitudinal quantum fluctuations described by Eq. (7). The lower and strong upper bounds on $\tilde{\ell}^2$ are written in Ref. [15] (in terms of $2\pi n_e \tilde{\ell}^2$), and the weak upper bound on $1-\epsilon^{-1}$ is given in Ref. [25].

In closing, we comment on the applicability of the above relations to Chern insultors (CIs). The general character of the sum rules in Eq. (10) suggests that the inequalities deduced from them remain valid for CIs. The subtlety is that CIs occupy a middle ground between metals and ordinary insulators [4], and the I_1 and I_2 sum rules diverge for metals. On the other hand, all three sum rules involve the symmetric (time even) part of the optical conductivity, whereas the distinction between ordinary and Chern insulators rests with the antisymmetric (time odd) part; from this we can conclude that the inequalities obtained above do apply to CIs, even if such materials fall outside the scope of Kohn's theory of the insulating state. Indeed, while the total Wannier spread diverges in a CI, its gauge invariant part proportional to ℓ^2 remains finite [26], consistent with the weak upper bound on ℓ^2. Likewise, the weak upper bound on χ implies that the susceptibility remains finite in CIs, even though the concept of spontaneous polarization requires special care [27]. We note that there exist additional sum rules [8] and inequalities [28–32] involving the time-odd part of the optical conductivity and the Chern invariant; such relations fall outside the scope of the present work.

4 Analytically solvable models

To build intuition on the bounds obtained above, we will now apply them to several models that can be treated analytically. For the first few examples dealing with finite systems, we introduce a polarizability per electron via the relation $\mathbf{d} = N\alpha\mathbf{E}_0$; here $\mathbf{d}$ is the dipole moment induced on the N-electron system by the applied electric field $\mathbf{E}_0$. To use the bulk relations (18) and (19), we place the system in a periodic supercell. In the limit where the supercell dimensions far exceed those of the system, the applied field $\mathbf{E}_0$ generates a macroscopic field $\mathbf{E} = \mathbf{E}_0$ in the effective medium; from $\mathbf{P} = \epsilon_0\chi\mathbf{E} = \mathbf{d}/V$ we get $\chi = n_e\alpha/\epsilon_0 + \mathcal{O}(V^{-1})$, where the additional terms (originating from the Clausius-Mossotti relation, see Sec. 4.4) vanish in the assumed limit of large V. Plugging this expression for χ into Eqs. (18) and (19) gives

$$\frac{\alpha E_G}{2e^2} \le \ell^2 \le \frac{\hbar}{2|e|}\sqrt{\frac{\alpha}{m_e}} \le \frac{\hbar^2}{2m_e E_G}\,, \tag{21a}$$

$$\frac{4e^2 m_e \ell^4}{\hbar^2} \le \alpha \le \frac{2e^2\ell^2}{E_G} \le \frac{\hbar^2 e^2}{m_e E_G^2}\,, \tag{21b}$$

$$E_G \le \frac{2e^2\ell^2}{\alpha} \le \frac{\hbar|e|}{\sqrt{m_e\alpha}} \le \frac{\hbar^2}{2m_e\ell^2}\,. \tag{21c}$$

At this point we make contact with known results for atoms and molecules. The strong upper bound on α, with E_G replaced by a mean excitation energy ΔE and $3e^2\ell^2$ expressed as the dipole fluctuation $\langle d^2\rangle_c$, becomes

$$\alpha \approx \frac{2\langle d^2\rangle_c}{3\Delta E}\,. \tag{22}$$

This estimate for the polarizability is discussed in Ref. [33], where its relation to the fluctuation-dissipation theorem is also mentioned. That textbook also gives an estimate for α in terms of the weak upper bound in Eq. (21b), invoking the oscillator-strength sum rule.

4.1 Hydrogen atom

Introducing the polarizability volume $\alpha' = \alpha/4\pi\epsilon_0$ [33], Eq. (21) becomes

$$\frac{1}{4}\frac{\alpha'}{a_0^3}\frac{E_G}{\mathrm{Ry}} \le \frac{\ell^2}{a_0^2} \le \frac{1}{2}\sqrt{\frac{\alpha'}{a_0^3}} \le \frac{\mathrm{Ry}}{E_G}\,, \tag{23a}$$

$$4\frac{\ell^4}{a_0^4} \le \frac{\alpha'}{a_0^3} \le 4\frac{\mathrm{Ry}}{E_G}\frac{\ell^2}{a_0^2} \le 4\frac{\mathrm{Ry}^2}{E_G^2}\,, \tag{23b}$$

$$\frac{E_G}{\mathrm{Ry}} \le 4\frac{a_0\ell^2}{\alpha'} \le 2\sqrt{\frac{a_0^3}{\alpha'}} \le \frac{a_0^2}{\ell^2}\,, \tag{23c}$$

where every fraction is dimensionless. For the nonrelativistic hydrogen atom we have [34,35]

$$E_G = 0.75\mathrm{Ry}\,, \qquad \alpha' = 4.5a_0^3\,, \qquad \ell^2 = a_0^2\,, \tag{24}$$

which plugged into Eq. (23) gives

$$\frac{27}{32} \le \frac{\ell^2}{a_0^2} = 1 \le \sqrt{\frac{9}{8}} \le \frac{4}{3}\,, \tag{25a}$$

$$4 \le \frac{\alpha'}{a_0^3} = 4.5 \le \frac{16}{3} \le \left(\frac{8}{3}\right)^2\,, \tag{25b}$$

$$\frac{E_G}{\mathrm{Ry}} = 0.75 \le \frac{8}{9} \le \frac{2}{\sqrt{4.5}} \le 1\,. \tag{25c}$$

The lower and strong upper bounds on α' are given in Ref. [19], and the latter is also discussed in Ref. [34] and in other textbooks.

Taking the average of the lower and strong upper bounds on ℓ^2 and on α' produces the reasonably accurate estimates $\ell^2 \approx 0.952a_0^2$ and $\alpha' \approx 4.(6)a_0^3$. The estimates $\ell^2 \approx 1.089a_0^2$ and $\alpha' \approx 5.(5)a_0^3$ obtained by taking the average of the lower and weak upper bounds are much less accurate, especially for α'.

We also note that for the hydrogen atom the strong upper bound ℓ_+ is closer to ℓ than the lower bound ℓ_-. This supports the notion that ℓ_+, being based solely on the average Penn gap, is more representative of the entire absorption spectrum than ℓ_-, which also depends on the minimum gap. Further evidence that ℓ_+ tends to track ℓ more closely than ℓ_- will be presented in Sec. 5 for crystalline materials.

4.2 Isotropic harmonic oscillator

For an electron trapped in an isotropic harmonic potential of frequency ω_0 the parameters are [33,34]

$$E_G = \hbar\omega_0\,, \quad \alpha = \frac{e^2}{m_e\omega_0^2} \equiv a_0\,, \quad \ell^2 = \frac{\hbar}{2m_e\omega_0} \equiv \ell_0^2\,, \tag{26}$$

saturating all the inequalities in Eq. (21). This can be understood from the selection rules for the harmonic oscillator: as the only allowed dipole transition from the ground state is to the first excited state, the entire spectral weight is at E_G, producing the saturation [19].

4.3 Van der Waals dimer model

So far we have only discussed one-electron systems. To analyze the effect of electron correlations, we now consider a system of two identical harmonic oscillators 1 and 2 separated by $\mathbf{R}$. We think of these oscillators as vibrating electrical dipoles in which the $+e$ charges (ions) are held in the position of equilibrium while the $-e$ charges (electrons) vibrate about these equilibrium positions, their displacements being $\mathbf{r}_1$ and $\mathbf{r}_2$. In the limit where $r_1, r_2 \ll R$, this provides a simple model for the van der Waals interaction [33,36].

The interaction term is

$$H_{12} = \frac{e^2}{4\pi\epsilon_0}\left[\frac{1}{R} + \frac{1}{|\mathbf{R}+\mathbf{r}_1-\mathbf{r}_2|} - \frac{1}{|\mathbf{R}+\mathbf{r}_1|} - \frac{1}{|\mathbf{R}-\mathbf{r}_2|}\right]. \tag{27}$$

In the approximation $r_1, r_2 \ll R$ we expand Eq. (27) to obtain in lowest order

$$H_{12} \simeq \frac{e^2}{4\pi\epsilon_0} r_{1a} r_{2b}\left(\frac{\delta_{ab}}{R^3} - 3\frac{R_a R_b}{R^5}\right), \tag{28}$$

which is in the form of a dipole-dipole interaction. Orienting the Cartesian frame such that $\mathbf{R} = R\hat{\mathbf{x}}$ leads to

$$\begin{aligned} H_{12} &\simeq \frac{e^2}{4\pi\epsilon_0}\left[-\frac{2}{R^3}x_1x_2 + \frac{1}{R^3}y_1y_2 + \frac{1}{R^3}z_1z_2\right] \\ &\equiv H_{12}^{\parallel} + H_{12}^{\perp}, \end{aligned} \tag{29}$$

where $H_{12}^{\parallel}$ denotes the first term and $H_{12}^{\perp}$ the other two.

For oscillations along $\mathbf{R}$ the only surviving term in Eq. (29) is $H_{12}^{\parallel}$, and we recover the 1D model of Ref. [36]. Denoting by H_0 the Hamiltonian of the two uncoupled oscillators, $H_0 + H_{12}^{\parallel}$ is diagonalized by the transformation

$$x_{\pm} = \frac{1}{\sqrt{2}}(x_1 \pm x_2) \tag{30}$$

together with a similar transformation for the momenta, resulting in two decoupled oscillators with frequencies

$$\omega_{\pm}^{\parallel} = \omega_0\sqrt{1 \mp \frac{2\alpha_0'}{R^3}}, \tag{31}$$

where α_0' is the polarizability volume of a single oscillator.

For unrestricted 3D oscillations, the interaction term is given by the full Eq. (29). Now instead of two modes we have six modes. By following through the same derivation, we can split also the y and z modes into symmetric and antisymmetric combinations with frequencies

$$\omega_{\pm}^{\perp} = \omega_0\sqrt{1 \pm \frac{\alpha_0'}{R^3}}; \tag{32}$$

thus, for transverse oscillations the symmetric modes have higher frequency than the antisymmetric ones.

In the 3D model the parameters E_G, α, and ℓ^2 are anisotropic, carrying labels $\parallel$ or $\perp$. To evaluate the $\parallel$ components, note that the interaction with a field $\mathbf{E}_{\parallel} = E\hat{\mathbf{x}}$ is described by

$eE(x_1+x_2)=\sqrt{2}eEx_+$, and that $\ell_\parallel^2$ is defined via (4) in terms of $X\equiv x_1+x_2=\sqrt{2}x_+$. This means that only the symmetric mode participates, and with a simple calculation one finds that the three parameters are obtained by replacing ω_0 with $\omega_+^\parallel$ in Eq. (26),

$$E_{\rm G}^\parallel=\hbar\omega_+^\parallel\simeq\hbar\omega_0\left(1-\alpha_0'/R^3\right), \tag{33a}$$

$$\alpha_\parallel=\frac{e^2}{m_{\rm e}\left(\omega_+^\parallel\right)^2}\simeq\alpha_0\left(1+2\alpha_0'/R^3\right), \tag{33b}$$

$$\ell_\parallel^2=\frac{\hbar}{2m_{\rm e}\omega_+^\parallel}\simeq\ell_0^2\left(1+\alpha_0'/R^3\right). \tag{33c}$$

The $\perp$ components are obtained by sending $\omega_+^\parallel\to\omega_+^\perp$ and $\alpha_0'\to-\alpha_0'/2$ in these expressions.

In conclusion, the van der Waals interaction reduces the optical gap and increases both the polarizability and the localization length in the axial direction of the dimer, and the opposite happens in the perpendicular directions. As the antisymmetric modes are dipole inactive, the entire spectral weight for light polarized along $\mathbf{R}$ or perpendicularly to it is concentrated at a single frequency $\hbar\omega_+^\parallel$ or $\hbar\omega_+^\perp$, respectively. In both cases the bounds in Eq. (21) remain saturated, just like for a single oscillator.

We emphasize that the explicit treatment of electron correlations is essential to obtain a qualitatively correct physical picture. For example, it is easy to show that the fluctuation-dissipation sum rule fails if the electron-electron interaction is treated at the mean-field level, e.g., within Hartree-Fock (HF) theory. Within HF, the dielectric susceptibility of the system of interacting oscillators is described exactly; nonetheless, the localization length is unaffected by the interaction and corresponds to that of the isolated monomer. This implies that the correct description of the macroscopic polarization fluctuations goes hand in hand with the ability of the theory to capture dispersion interactions between isolated bodies.

4.4 Van der Waals crystal model

As an extension of the dimer model, we now consider a periodic array of oscillators coupled by dipole-dipole interactions. The potential energy reads

$$U=\frac{1}{2}m_{\rm e}\omega_0^2\sum_{\mathbf{R}}\left|\mathbf{r}^{\mathbf{R}}\right|^2+\frac{e^2}{4\pi\epsilon_0}\sum_{\mathbf{R}}\sum_{\mathbf{R}'\neq\mathbf{0}}\frac{r_a^{\mathbf{R}}r_b^{\mathbf{R}+\mathbf{R}'}}{2}\left(\frac{\delta_{ab}}{R'^3}-3\frac{R'_aR'_b}{R'^5}\right), \tag{34}$$

where $\mathbf{r}^{\mathbf{R}}$ denotes the displacement of an electron away from its equilibrium position $\mathbf{R}$, taken to be a point on a Bravais lattice. A similar model was discussed in Ref. [25]; the only difference is that the positive charges, instead of being point charges placed at the lattice points, are smeared into a uniform background.

4.4.1 Dynamical matrix

As in the dimer model, the electrons are assumed to be strongly localized in the sense that the quantum fluctuations are small compared to the separation between the ions. The resulting potential bears many similarities to the form that appears in the context of lattice vibrations; we will therefore borrow the same terminology in discussing the relevant contributions to the electronic Hamiltonian.

To determine the normal modes of the system we first evaluate the force-constant matrix

$$D_{a\mathbf{R},b\mathbf{R}'}\equiv\frac{\partial^2U}{\partial r_a^{\mathbf{R}}\partial r_b^{\mathbf{R}'}}=D_{a\mathbf{0},b\mathbf{R}'-\mathbf{R}} \tag{35}$$

to find

$$D_{a\mathbf{0},b\mathbf{R}} = m_e\omega_0^2\delta_{ab}\delta_{\mathbf{R0}} + \frac{e^2}{4\pi\epsilon_0}(1-\delta_{\mathbf{R0}})\left(\frac{\delta_{ab}}{R^3} - 3\frac{R_aR_b}{R^5}\right), \tag{36}$$

and then convert it into a dynamical matrix using

$$D_{ab}(\mathbf{q}) = \frac{1}{m_e}\sum_{\mathbf{R}} D_{a\mathbf{0},b\mathbf{R}}\, e^{-i\mathbf{q}\cdot\mathbf{R}}. \tag{37}$$

The result is

$$D_{ab}(\mathbf{q}) = \omega_0^2\delta_{ab} + C_{ab}(\mathbf{q}), \tag{38}$$

where

$$C_{ab}(\mathbf{q}) = \frac{e^2/m_e}{4\pi\epsilon_0}\sum_{\mathbf{R}\neq\mathbf{0}} e^{-i\mathbf{q}\cdot\mathbf{R}}\left(\frac{\delta_{ab}}{R^3} - 3\frac{R_aR_b}{R^5}\right). \tag{39}$$

To carry out the above lattice sum it is convenient to work in reciprocal space, where the interaction can be recast as a rapidly converging Ewald summation,

$$C_{ab}(\mathbf{q}) = \frac{e^2/m_e}{4\pi\epsilon_0}\left[\frac{4\pi}{\Omega}\sum_{\mathbf{G}}{}' \frac{K_aK_b}{K^2}e^{-\frac{K^2\sigma^2}{4}} - \delta_{ab}\frac{4}{3\sqrt{\pi}\sigma^3}\right], \qquad \mathbf{K} = \mathbf{G}+\mathbf{q}, \tag{40}$$

with Ω the volume of a primitive cell. The primed sum excludes the divergent $\mathbf{G}+\mathbf{q}=\mathbf{0}$ term, and the second term removes the self-interaction of the dipole in the origin cell; the result is independent of the Ewald parameter σ provided that $\sigma \ll R$ for all $\mathbf{R}\neq\mathbf{0}$.

By diagonalizing the 3×3 matrix $C_{ab}(\mathbf{q})$ at every point in the Brillouin zone, we have rewritten the problem as a set of independent oscillators. In particular, we have three modes at each $\mathbf{q}$ that are characterized by a frequency

$$\omega_i^2(\mathbf{q}) = \omega_0^2 + \lambda_i(\mathbf{q}), \tag{41}$$

where $\lambda_i(\mathbf{q})$ are the eigenvalues of $C_{ab}(\mathbf{q})$.

In Appendix C, we calculate the zero-point energy of this model by collecting the contributions from all normal modes across the Brillouin zone.

4.4.2 Long-wave limit

The $\mathbf{q}\to\mathbf{0}$ limit is particularly relevant to our discussion, since it corresponds to the collective displacement of the electronic center of mass. For a cubic lattice we find two TO modes and one LO mode where

$$\lambda_{\mathrm{TO}} = -\frac{1}{3}\omega_p^2, \qquad \lambda_{\mathrm{LO}} = \frac{2}{3}\omega_p^2, \tag{42}$$

with $\omega_p^2 = e^2/\epsilon_0 m_e\Omega$. To obtain this result note that the matrix $C(\mathbf{q})$ is traceless so that $2\lambda_{\mathrm{TO}}+\lambda_{\mathrm{LO}}=0$, and that $\lambda_{\mathrm{LO}} = \lambda_{\mathrm{TO}} + \omega_p^2$, where ω_p^2 is the contribution from the $\mathbf{G}=\mathbf{0}$ term in Eq. (40) when $\mathbf{q}\to\mathbf{0}$.

The dielectric susceptibility and permittivity are readily given in terms of the TO mode frequency,

$$\chi = \frac{\omega_p^2}{\omega_{\mathrm{TO}}^2}, \qquad \epsilon = 1+\chi\,. \tag{43}$$

Then, based on the above, we can quickly verify that the following results hold,

$$\epsilon = \frac{\omega_{\mathrm{LO}}^2}{\omega_{\mathrm{TO}}^2}, \qquad \frac{\epsilon-1}{\epsilon+2} = \frac{\alpha_0}{3\epsilon_0\Omega}. \tag{44}$$

The first result is the Lyddane-Sachs-Teller relation [36], valid for a single-mode dielectric. The second is the Clausius-Mossotti relation [36], linking the macroscopic permittivity to the molecular polarizability α_0.

For the TO modes we have

$$E_G = \hbar\omega_{TO}\,, \qquad \chi = \frac{\omega_p^2}{\omega_{TO}^2}\,, \qquad \ell^2 = \frac{\hbar}{2m_e\omega_{TO}}\,. \tag{45}$$

When plugged into Eqs. (12) and (13) these parameters give $E_L = E_P = E_G$, saturating all the bounds in Eqs. (18) and (19). The parameters for the LO modes are

$$\tilde{E}_G = \hbar\omega_{LO}\,, \qquad 1-\epsilon^{-1} = \frac{\omega_p^2}{\omega_{LO}^2}\,, \qquad \tilde{\ell}^2 = \frac{\hbar}{2m_e\omega_{LO}}\,, \tag{46}$$

and again the corresponding bounds, obtained by modifying Eqs. (18) and (19) according to Eq. (20), are saturated.

5 Real materials

Starting from experimental data, we have evaluated the bounds on ℓ in Eq. (18a) for a number of materials. To visualize the results, it is helpful to bring that equation to the form[4]

$$\frac{\mathrm{Ry}}{E_P}\sqrt{\frac{E_G}{\mathrm{Ry}}} \le \frac{\ell}{a_0} \le \sqrt{\frac{\mathrm{Ry}}{E_P}} \le \sqrt{\frac{\mathrm{Ry}}{E_G}}\,, \tag{47}$$

which suggests plotting the data as shown schematically in Fig. 1. Given a data point (large blue dot), the range $[\ell_-, \ell_+]$ in units of a_0 is obtained by drawing horizontal and vertical line segments from it to the diagonal dashed line; its projection on that line (small black dot) yields

$$\ell \approx (\ell_+ + \ell_-)/2\,, \tag{48}$$

which we will refer to as the "strong bound" estimate, as opposed to the "weak bound" estimate obtained by replacing ℓ_+ with ℓ_{++} in the expression above.

In the following, we use Eqs. (47) and (48) to estimate the electron localization length in different classes of materials. The needed experimental data are the optical gap (the lowest energy for optical absorption), the electron density, and the clamped-ion electric susceptibility; the last two enter via Eq. (13) for the Penn gap.

5.1 Rocksalt alkali halides

Figure 2 shows the results obtained for alkali halides with the rocksalt structure. Consider first the top panels, where E_P was calculated from the total electron density n_e including inner core electrons. Such "all-electron" bounds inevitably provide average localization lengths that include those tight inert states; as a consequence, the bounds are rather loose not only on the left panel (weak bound) but also on the right panel (strong bound). On the left panel the upper bound is independent of n_e, and hence it is insensitive to the different localization lengths of valence and core electrons. This is not the case for the right panel, where the tighter upper bound containing n_e narrows down the estimates for ℓ; nevertheless, the data points are still quite far from the diagonal.

[4] To obtain Eq. (47), combine the energy relations in the last line of Table 1 with the second form of Eq. (12).

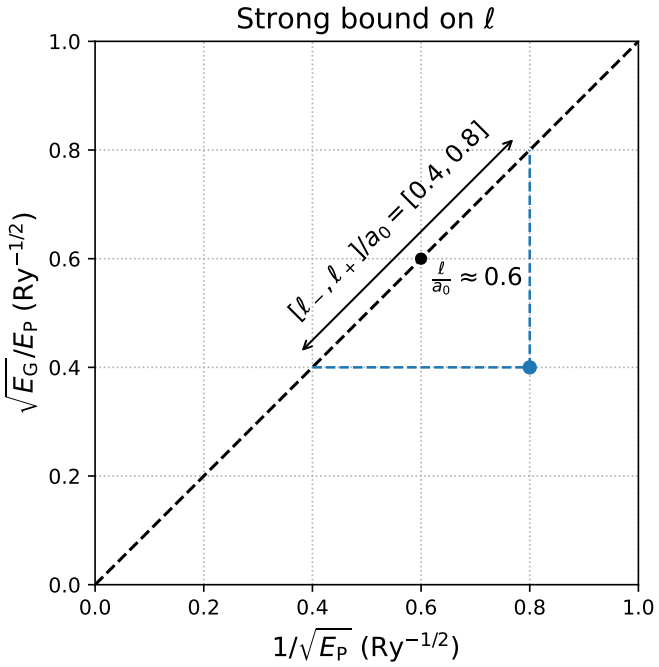

Figure 1: Schematic representation of the strong bound on ℓ in Eq. (47). For the weak bound, replace $E_P \to E_G$ on the horizontal axis; since $E_G \leq E_P$, the data point (blue dot) will move to the right, resulting in a wider range $[\ell_-, \ell_{++}]$.

To rationalize the results for the strong bound, note that

$$\frac{\ell_+ - \ell_-}{a_0} = \sqrt{\frac{\mathrm{Ry}}{E_P}}\left(1 - \sqrt{\frac{E_G}{E_P}}\right), \tag{49}$$

and thus the range $[\ell_-, \ell_+]$ gets tighter and tighter as E_P gets closer to E_G. Since $E_P \propto \sqrt{n_e}$, the inclusion of core electrons goes in the opposite direction, and the range $[\ell_-, \ell_+]$ tends to increase as we move down the periodic table. This can be seen in the top-right panel of Fig. 2, where the distance from the diagonal line increases from the fluorides to the chlorides, from these to the bromides, and from these to the iodides.

It would be much more relevant for physical properties if one could estimate the average localization length of the valence electrons only. Here, we take the simple approach of replacing E_P in Eq. (47) with a valence Penn gap calculated from the valence electron density.[5] The bounds on ℓ obtained in this manner are presented in the bottom panels of Fig. 2. As a result of discarding the core electrons the data points move closer to diagonal line (the bounds get tighter), and their projections on that line move further up (the average localization lengths increase). Most interestingly, simple trends emerge in this valence-only formulation, with ℓ increasing from the lighter to the heavier halogens; this agrees with the intuition on chemical bonding in strongly ionic crystals [38]. The trend is most visible in the bottom right panel displaying the strong bound. The valence-only values for ℓ_-, ℓ_+, and ℓ_{++} are compiled in Table 2.

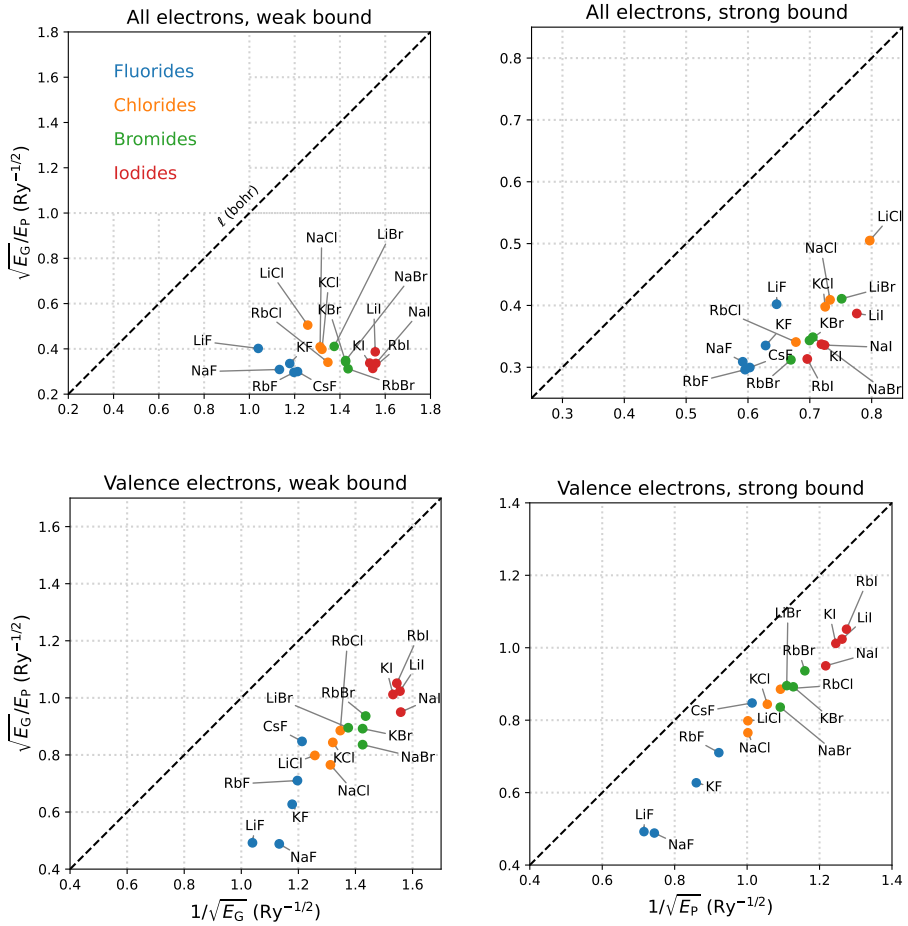

Figure 2: Bounds on ℓ for the rocksalt alkali halides, plotted using the scheme outlined in Fig. 1. The weak and strong bounds are represented on the left and right panels, respectively, while the top and bottom panels show all-electron and valence-electron results, respectively, with E_P defined accordingly in each case. Note that each panel has its own scale.

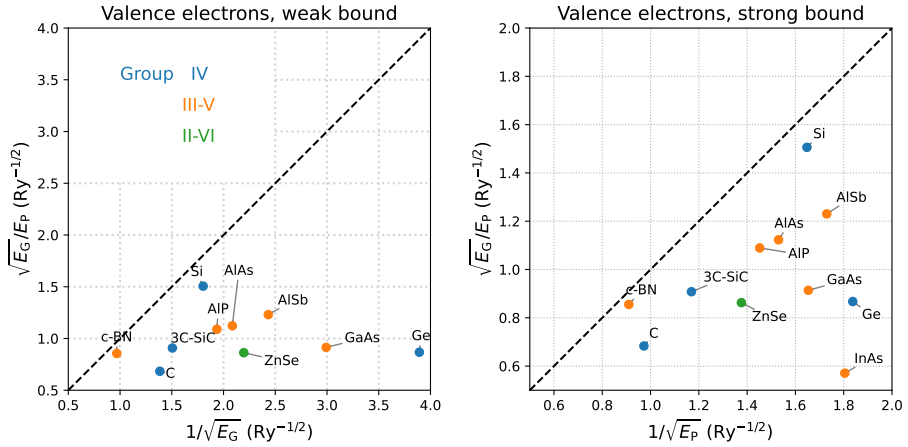

Figure 3: Same as the bottom two panels of Fig. 2, but for materials with the diamond or the zincblende structure. On the left panel, the data point for InAs is out of bounds.

Crystal	a (Å)	ϵ	E_G (eV)	$\ell_-(a_0)$	$\ell_+(a_0)$	$\ell_{++}(a_0)$
LiF	4.02	1.96	12.6	0.40	0.65	1.04
NaF	4.62	1.74	10.6	0.31	0.59	1.13
KF	5.35	1.85	9.8	0.34	0.63	1.18
RbF	5.64	1.96	9.5	0.30	0.60	1.20
CsF	6.01	2.16	9.25	0.30	0.60	1.21
LiCl	5.13	2.78	8.6	0.50	0.80	1.26
NaCl	5.64	2.34	7.9	0.41	0.73	1.31
KCl	6.29	2.19	7.8	0.40	0.72	1.32
RbCl	6.58	2.19	7.5	0.34	0.68	1.35
LiBr	5.50	3.17	7.2	0.41	0.75	1.37
NaBr	5.97	2.59	6.7	0.35	0.70	1.43
KBr	6.60	2.34	6.7	0.34	0.70	1.43
RbBr	6.58	2.19	7.5	0.31	0.67	1.44
LiI	6.00	3.80	5.62	0.39	0.78	1.56
NaI	6.47	2.93	5.6	0.34	0.72	1.56
KI	7.07	2.62	5.8	0.34	0.72	1.53
RbI	7.34	2.59	5.7	0.31	0.70	1.54

Table 2: Bounds on ℓ for the valence electrons in rocksalt alkali halides, estimated from experimental data: lattice constant a (the density of valence electrons is $n_e = 32/a^3$, corresponding to eight valence electrons per formula unit), electronic permittivity ϵ, and optical gap E_G. The values for a and ϵ are from Ref. [38], and those for E_G correspond to the lowest absorption peaks in Ref. [39]; the exceptions are LiF and LiI, for which E_G are the excitonic gaps reported in Refs. [40] and [41], respectively. Note that lattice constants are quoted in Angstroms while localization bounds are in atomic units.

Crystal	a (Å)	ϵ	E_G (eV)	$\ell_-(a_0)$	$\ell_+(a_0)$	$\ell_{++}(a_0)$
C	3.57	5.7	7.1	0.68	0.97	1.38
Si	5.43	11.97	4.19	1.51	1.65	1.80
Ge	5.66	16.00	0.90	0.87	1.84	3.89
3C-SiC	4.36	6.38	6.0	0.91	1.17	1.51
c-BN	3.62	4.46	14.5	0.86	0.91	0.97
AlP	5.46	7.5	3.63	1.09	1.45	1.94
AlAs	5.66	8.2	3.13	1.12	1.53	2.08
AlSb	6.14	10.24	2.3	1.23	1.73	2.43
GaAs	5.65	10.86	1.52	0.91	1.65	2.99
InAs	6.06	12.37	0.42	0.57	1.80	5.71
ZnSe	5.68	5.7	2.82	0.86	1.38	2.20

Table 3: Same as Table 2, but for materials with the diamond or the zincblende structure. We assume four valence electrons per atom on average, so that $n_e = 32/a^3$. The experimental data is from Ref. [42], where E_G is the direct gap.

5.2 Tetrahedrally-coordinated materials

Figure 3 and Table 3 show the valence-only results obtained for materials with the diamond or the zincblende structure from groups IV, III-V, and II-VI in the periodic table. The trends are not as uniform as in the case of the halides because there is a larger range of gaps and susceptibilities. Nevertheless, one observes that the values of ℓ estimated from Eq. (48) tend to decrease with increasing ionicity, e.g., along the isoelectronic series Si → AlP and Ge → GaAs → ZnSe, as also found in Ref. [10]; this is consistent with the intuition that ionic bonding yields more localized electrons than covalent bonding [38]. Accordingly, the estimated localization lengths in Table 3 tend to be larger than those in Table 2 for the strongly ionic alkali halides.

How well do the ℓ values estimated from experimental data via Eq. (48) compare with those obtained from first-principles calculations? To address this question, in Fig. 4 we compare them with the *ab initio* values reported in Ref. [10]. The correlation is quite satisfactory, although the theoretical values tend to be somewhat larger. To explain this trend, one could invoke the band gap underestimation in density functional theory, which may well lead to a systematic overestimation of the calculated localization lengths. Since expressing ℓ as the average of ℓ_- and ℓ_+ is an approximation, however, it is difficult to draw definitive conclusions, especially for cases like Ge where ℓ_- and ℓ_+ are rather different. Yet, it is interesting to note that the upper bounds in Fig. 4 essentially fall on the diagonal in all cases, which means that they closely match the available theoretical data. (The lower bounds, involving the minimum gap, display a much larger scatter.) This gives further credit to our earlier statements that ℓ_+ is indeed a more robust indicator of the polarization fluctuation amplitude compared to ℓ_-.

6 Conclusions

The use of sum-rule inequalities to estimate the electronic polarizability has a long tradition in atomic and molecular physics [19,33]. The extension of those ideas to crystals and to other physical properties is not as well established, and the results are scattered in the literature. In this work, we provided a unified perspective on several sum-rule inequalities for bulk sys-

[5] Such replacement assumes that valence-only versions of the sum rules in Eq. (10) can be formulated, which requires the excitation energies of core electrons to be well separated in energy from those of valence electrons [22, 37].

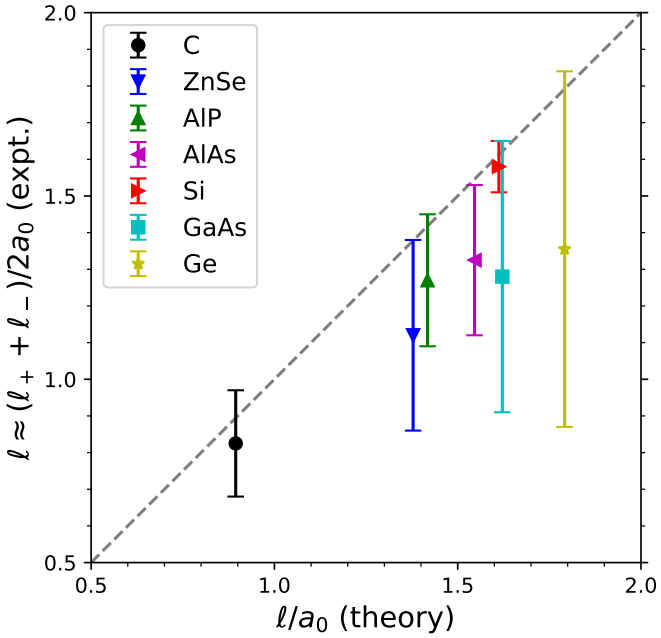

Figure 4: Comparison, for tetrehedrally-cordinated materials, between the ℓ values for valence electrons estimated from experimental data (Table 3), and those calculated from first principles in Ref. [10] using a pseudopotential method. The error bars indicate the range $[\ell_-, \ell_+]$.

tems, and organized them into chained inequalities bracketing three electronic properties of insulators: localization length ℓ, static susceptibility χ, and optical gap E_G. As they are based on exact sum rules, those inequalities remain valid for correlated, disordered, and topological insulators, and in the presence of relativistic effects including spin-orbit coupling. As an application, we used them to estimate ℓ (a ground-state property) from readily available experimental data on the response properties χ and E_G, together with the electron density. By focusing on the valence electrons, we obtained meaningful estimates for their average localization length that follow simple chemical trends.

The study of several exactly solvable models, from the hydrogen atom to isolated and coupled oscillators, provided useful insights. In particular, the coupled oscillator models illustrated how the fluctuation-dissipation relation breaks down at the mean-field level and critically requires an explicit treatment of dynamical correlations.

Acknowledgements

We are indebted to Morrel H. Cohen for an unpublished collaboration with one of us (I. S.), on which Problem 22.5 of Ref. [17] was based. We also wish to thank Liang Fu, Yugo Onishi, and Raffaele Resta for stimulating discussions. M. S. thanks the CCQ at the Flatiron Institute for hospitality while this research was carried out. The Flatiron Institute is a division of the Simons Foundation.

Funding information Work by I. S. was supported by Grant No. PID2021-129035NB-I00 funded by MCIN/AEI/10.13039/501100011033. Work by M. S. was supported by the State Investigation Agency through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2023-001263-S), and from Generalitat de Catalunya (Grant No. 2021 SGR 01519).

A Polarization and localization in band insulators

In Eqs. (3) and (4), the electronic polarization and the electron localization tensor were written down for a generic bulk insulator (possibly correlated and/or disordered) using Kohn's center-of-mass operator. Alternatively, those expressions can be recast in terms of the Berry phase and quantum metric defined by the change in the many-body ground state under twisted boundary conditons [5, 43]. Here we specialize to the single-particle picture, and review the corresponding formulas for uncorrelated crystalline insulators.

The electronic polarization of a band insulator takes the form of a Berry phase of the cell-periodic Bloch states in momentum space [3, 4],

$$\mathbf{P}_\mathrm{e} = \frac{-|e|}{(2\pi)^3} \int d^3k \sum_{n=1}^{J} \mathbf{A}_{nn}(\mathbf{k}) ; \tag{A.1}$$

here $\mathbf{A}_{mn}(\mathbf{k}) = i\langle u_{m\mathbf{k}}|\boldsymbol{\nabla}_\mathbf{k} u_{n\mathbf{k}}\rangle$ is the Berry connection matrix, the integral is over the first Brillouin zone (BZ), and the summation is over the valence bands. Alternatively, $\mathbf{P}_\mathrm{e}$ can be written as [3, 4]

$$\mathbf{P}_\mathrm{e} = \frac{-|e|}{\Omega} \sum_{n=1}^{J} \langle \mathbf{r} \rangle_n , \tag{A.2}$$

where $\langle \mathbf{r} \rangle_n$ is the center of charge of a Wannier function constructed for band n. This expression remains valid for a disordered insulator, in which case $\Omega, J \to \infty$.

The localization tensor can be obtained from the fluctuation-dissipation relation in Eq. (6). Using the Kubo-Greenwood formula for the optical conductivity, one finds the sum rule [5, 8]

$$\int_0^\infty d\omega\, \omega^{-1}\, \mathrm{Re}\, \sigma^\mathrm{S}_{ab}(\omega) = \frac{\pi e^2}{(2\pi)^3 \hbar} \int d^3k \sum_{n=1}^{J} g_{nn,ab}(\mathbf{k}) ; \tag{A.3}$$

on the right-hand side, g is the quantum metric tensor [6] of the valence manifold [4, 9],

$$g_{mn,ab}(\mathbf{k}) = \frac{1}{2}\langle \partial_a u_{m\mathbf{k}}|Q(\mathbf{k})|\partial_b u_{n\mathbf{k}}\rangle + \frac{1}{2}\langle \partial_b u_{m\mathbf{k}}|Q(\mathbf{k})|\partial_a u_{n\mathbf{k}}\rangle . \tag{A.4}$$

Here $\partial_a = \partial/\partial k_a$, and $Q(\mathbf{k}) = 1 - \sum_{n=1}^{J} |u_{n\mathbf{k}}\rangle\langle u_{n\mathbf{k}}|$ is the projection operator for the conduction states. Inserting Eq. (A.3) in Eq. (6) gives

$$\langle r_a r_b \rangle_\mathrm{c} = \frac{1}{(2\pi)^3 n_\mathrm{e}} \int d^3k \sum_{n=1}^{J} g_{nn,ab}(\mathbf{k}) , \tag{A.5}$$

which expresses the bulk localization tensor as a ground-state quantity, without any reference to the excitation spectrum.

For a one-dimensional (1D) insulator the tensor $\langle r_a r_b \rangle_\mathrm{c}$ reduces to a scalar ℓ^2, and Eq. (A.5) can be written in terms of maximally-localized Wannier functions as

$$\ell^2 = \frac{1}{J} \sum_{n=1}^{J} \left[\langle x^2 \rangle_n - \langle x \rangle_n^2 \right] , \tag{A.6}$$

which follows from the relation between the BZ integral of the metric and the quadratic Wannier spread [9]. Thus, in 1D the localization tensor is equal to the average spread of the maximally-localized Wannier functions. More generally, in d dimensions its trace $\sum_{a=1}^{d} \ell_a^2$ equals the gauge-invariant part of the average Wannier spread, which for $d > 1$ is smaller than the actual spread in any gauge [9].

In summary, electronic polarization is related to the Wannier centers of the valence bands, and the electron localization length squared (polarization fluctuations) gives a lower bound to the average Wannier spread.

B Longitudinal optical bounds

Here, we outline the extension to long-wave longitudinal modes of the analysis carried out in Secs. 2 and 3 for transverse modes; a related discussion is given in Refs. [15, 25]. We first define the moments of the energy-loss spectrum as

$$M_p = \frac{2}{\pi}\int_0^\infty d\omega\, \omega^p \,\mathrm{Im}\left[-\frac{1}{\epsilon(\omega)}\right]. \tag{B.1}$$

The moments with $p = 1, 0, -1$ satisfy the relations

$$M_1 = \omega_\mathrm{p}^2, \tag{B.2a}$$

$$M_0 = \frac{2e^2}{\hbar\epsilon_0} n_\mathrm{e} \tilde{\ell}^2, \tag{B.2b}$$

$$M_{-1} = 1 - \epsilon^{-1}, \tag{B.2c}$$

where ϵ^{-1} stands for $\epsilon^{-1}(0)$. These are respectively the longitudinal counterpart of the oscillator-strength sum rule (10a) [22], the longitudinal fluctuation-dissipation relation (7), and the longitudinal counterpart of the Kramers-Krönig relation (10c) [44].

Next, we introduce average gaps for longitudinal excitations by analogy with Eqs. (11-13),

$$\tilde{E}_\mathrm{L} = \frac{\hbar M_1}{M_0}, \qquad \tilde{E}_\mathrm{P}^2 = \frac{\hbar M_1}{\hbar^{-1} M_{-1}}, \tag{B.3}$$

$$\tilde{E}_\mathrm{L} = \frac{\hbar^2}{2 m_\mathrm{e} \tilde{\ell}^2} \Leftrightarrow \left(\frac{\tilde{\ell}}{a_0}\right)^2 = \frac{\mathrm{Ry}}{\tilde{E}_\mathrm{L}} \tag{B.4}$$

$$1 - \epsilon^{-1} = \left(\frac{\hbar\omega_\mathrm{p}}{\tilde{E}_\mathrm{P}}\right)^2. \tag{B.5}$$

Since the loss function $\mathrm{Im}\left[-\epsilon^{-1}(\omega)\right]$ is non-negative, one can immediately write down inequalities anlogous to those in Eqs. (14), (15), and (17),

$$M_{p-q} \le \frac{\hbar}{\tilde{E}_\mathrm{G}} M_{p-q+1} \le \ldots \le \left(\frac{\hbar}{\tilde{E}_\mathrm{G}}\right)^q M_p \tag{B.6}$$

($\tilde{E}_\mathrm{G}$ is the plasmon gap),

$$M_p^2 \le M_{p-1} M_{p+1}, \tag{B.7}$$

and

$$\tilde{E}_\mathrm{L} \ge \tilde{E}_\mathrm{p} \ge \tilde{E}_\mathrm{G}, \qquad \tilde{E}_\mathrm{P}^2 \ge \tilde{E}_\mathrm{G}\tilde{E}_\mathrm{L}. \tag{B.8}$$

Finally, by forming the chained inequalities

$$\tilde{E}_\mathrm{P}^2/\tilde{E}_\mathrm{G} \ge \tilde{E}_\mathrm{L} \ge \tilde{E}_\mathrm{P} \ge \tilde{E}_\mathrm{G}, \tag{B.9a}$$

$$\tilde{E}_\mathrm{L}^2 \ge \tilde{E}_\mathrm{P}^2 \ge \tilde{E}_\mathrm{G}\tilde{E}_\mathrm{L} \ge \tilde{E}_\mathrm{G}^2, \tag{B.9b}$$

$$\tilde{E}_\mathrm{G} \le \tilde{E}_\mathrm{P}^2/\tilde{E}_\mathrm{L} \le \tilde{E}_\mathrm{P} \le \tilde{E}_\mathrm{L} \tag{B.9c}$$

and combining them with Eqs. (B.4) and (B.5), we obtain weak and strong bounds on $\tilde{\ell}^2$, $1-\epsilon^{-1}$ and $\tilde{E}_\mathrm{G}$, respectively. Those bounds are given by Eqs. (18) and (19), with the replacements indicated in Eq. (20).

C Zero-point energy of the van der Waals crystal model

In this appendix we return to the van der Waals crystal model of Sec. 4.4, and calculate its zero-point energy in two different ways. First we use a Brillouin-zone integral,

$$E = \frac{\hbar}{2}\frac{\Omega}{(2\pi)^3}\int d^3q \sum_i \omega_i(\mathbf{q}) . \tag{C.1}$$

To verify that the normalization factors are correct, note that in the absence of interactions we recover the correct result for the isolated 3D oscillator,

$$E_0 = \frac{3}{2}\hbar\omega_0 . \tag{C.2}$$

The interaction is regarded as a small perturbation, so we can Taylor-expand the square root of Eq. (41) for $\omega_i^2(\mathbf{q})$,

$$\omega_i = \sqrt{\omega_0^2 + \lambda_i} \simeq \omega_0 + \frac{1}{2}\frac{\lambda_i}{\omega_0} - \frac{1}{8}\frac{\lambda_i^2}{\omega_0^3} . \tag{C.3}$$

As the $C(\mathbf{q})$ matrix defined by Eq. (39) is traceless for all $\mathbf{q}$, the second term above drops out from Eq. (C.1). The leading correction is then given by the third term,

$$\Delta E = -\frac{\hbar}{16\omega_0^3}\frac{\Omega}{(2\pi)^3}\int d^3q \sum_i \lambda_i^2(\mathbf{q}) . \tag{C.4}$$

Overall, the interaction energy is negative and in view of Eq. (42) it appears to scale as Ω^{-2}, which at first sight seems consistent with van der Waals. This is confirmed by a numerical evaluation of Eq. (C.4) for a simple-cubic lattice (Fig. 5), which shows a Ω^{-2} behavior for ΔE in the limit of a dense $\mathbf{q}$ mesh.

As further validation, we have computed the same energy as a real-space sum of pair interactions. We start from the interaction energy of the 3D dimer model of Sec. 4.3, which is obtained by expanding Eqs. (31) and (32) according to Eq. (C.3). The result [33]

$$\Delta E_{12} = -\frac{3}{4}\left(\frac{\alpha_0'}{R^3}\right)^2 \hbar\omega_0 , \tag{C.5}$$

which is enhanced by a factor of 3/2 relative to that of the 1D dimer model [36], leads to a crystal energy of

$$\Delta E = -\hbar\left(\frac{e^2/m_e}{4\pi\epsilon_0}\right)^2 \frac{3}{8\omega_0^3}\sum_{\mathbf{R}\neq\mathbf{0}}\frac{1}{R^6} . \tag{C.6}$$

(Note the additional factor of 1/2 to avoid double counting of the pair interactions.) As shown in Fig. 5, the converged value of this real-space summation agrees with that of the reciprocal-space summation (C.4). The plotted quantity is $\Delta E/(\omega_0^{-3}\Omega^{-2})$ in Hartee atomic units, and its converged value is precisely $-(3/8)A_6$, where

$$A_6 \equiv \sum_{i,j,k}{}' (i^2+j^2+k^2)^{-3} \simeq 8.40192 \tag{C.7}$$

(with $i = j = k = 0$ excluded) is a lattice sum tabulated by Lennard-Jones and Ingham [45].

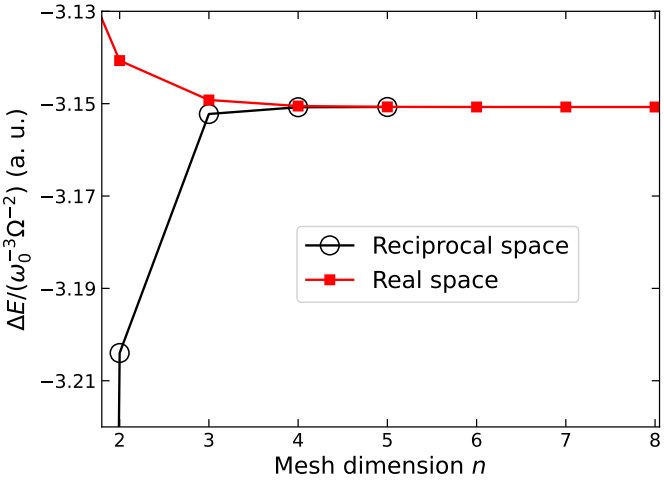

Figure 5: Convergence of the reciprocal- and real-space sums for the dispersion interaction energy using meshes of dimension $2n$. The plotted values correspond to Eqs. (C.4) and (C.6) for a simple-cubic lattice, in units of $\omega_0^{-3}\Omega^{-2}$ using Hartree atomic units (a. u.).

References

[1] W. Kohn, *Theory of the insulating state*, Phys. Rev. **133**, A171 (1964), doi:10.1103/PhysRev.133.A171.

[2] W. Kohn, *Metals and Insulators*, In C. DeWitt and R. Balian, eds., *Many-Body Physics*. Gordon and Breach (1968).

[3] R. D. King-Smith and D. Vanderbilt, *Theory of polarization of crystalline solids*, Phys. Rev. B **47**, 1651 (1993), doi:10.1103/PhysRevB.47.1651.

[4] D. Vanderbilt, *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators*, Cambridge University Press, doi:10.1017/9781316662205 (2018).

[5] I. Souza, T. Wilkens and R. M. Martin, *Polarization and localization in insulators: Generating function approach*, Phys. Rev. B **62**, 1666 (2000), doi:10.1103/PhysRevB.62.1666.

[6] J. P. Provost and G. Vallee, *Riemannian structure on manifolds of quantum states*, Commun. Math. Phys. **76**, 289 (1980), doi:10.1007/BF02193559.

[7] R. Resta and S. Sorella, *Electron localization in the insulating state*, Phys. Rev. Lett. **82**, 370 (1999), doi:10.1103/PhysRevLett.82.370.

[8] I. Souza and D. Vanderbilt, *Dichroic f-sum rule and the orbital magnetization of crystals*, Phys. Rev. B **77**, 054438 (2008), doi:10.1103/PhysRevB.77.054438.

[9] N. Marzari and D. Vanderbilt, *Maximally localized generalized Wannier functions for composite energy bands*, Phys. Rev. B **56**, 12847 (1997), doi:10.1103/PhysRevB.56.12847.

[10] C. Sgiarovello, M. Peressi and R. Resta, *Electron localization in the insulating state: Application to crystalline semiconductors*, Phys. Rev. B **64**, 115202 (2001), doi:10.1103/PhysRevB.64.115202.

[11] M. Veithen, X. Gonze and P. Ghosez, *Electron localization: Band-by-band decomposition and application to oxides*, Phys. Rev. B **66**, 235113 (2002), doi:10.1103/PhysRevB.66.235113.

[12] E. K. Kudinov, *Difference between insulating and conducting states*, Sov. Phys. Solid State **33**, 1299 (1991), doi:10.48550/arXiv.cond-mat/9902361.

[13] L. D. Landau and E. M. Lifshitz, *Statistical Physics, Part 1*, Pergamon, New York (1980).

[14] R. Resta, *Polarization fluctuations in insulators and metals: New and old theories merge*, Phys. Rev. Lett. **96**, 137601 (2006), doi:10.1103/PhysRevLett.96.137601.

[15] Y. Onishi and L. Fu, *Quantum weight* (2024), 2406.06783.

[16] C. Aebischer, D. Baeriswyl and R. M. Noack, *Dielectric Catastrophe at the Mott Transition*, Phys. Rev. Lett. **86**, 468 (2001), doi:10.1103/PhysRevLett.86.468.

[17] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge, 1st edition edn. (2004).

[18] N. Verma and R. Queiroz, *Instantaneous response and quantum geometry of insulators* (2024), 2403.07052.

[19] M. Traini, *Electric polarizability of the hydrogen atom: a sum rule approach*, European J. Phys. **17**, 30 (1996), doi:10.1088/0143-0807/17/1/006.

[20] I. Komissarov, T. Holder and R. Queiroz, *The quantum geometric origin of capacitance in insulators*, Nature Commun. **15**, 4621 (2024), doi:10.1038/s41467-024-48808-x.

[21] M. Graf and P. Vogl, *Electromagnetic fields and dielectric response in empirical tight-binding theory*, Phys. Rev. B **51**, 4940 (1995), doi:10.1103/PhysRevB.51.4940.

[22] F. Wooten, *Optical Properties of Solids*, Academic Press (1972).

[23] D. R. Penn, *Wave-Number-Dependent Dielectric Function of Semiconductors*, Phys. Rev. **128**, 2093 (1962), doi:10.1103/PhysRev.128.2093.

[24] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, Springer, 4th edition edn. (2010).

[25] Y. Onishi and L. Fu, *Universal relation between energy gap and dielectric constant* (2024), 2401.04180.

[26] T. Thonhauser and D. Vanderbilt, *Insulator/Chern-insulator transition in the Haldane model*, Phys. Rev. B **74**, 235111 (2006), doi:10.1103/PhysRevB.74.235111.

[27] S. Coh and D. Vanderbilt, *Electric Polarization in a Chern Insulator*, Phys. Rev. Lett. **102**, 107603 (2009), doi:10.1103/PhysRevLett.102.107603.

[28] R. Roy, *Band geometry of fractional topological insulators*, Phys. Rev. B **90**, 165139 (2014), doi:10.1103/PhysRevB.90.165139.

[29] S. Peotta and P. Törmä, *Superfluidity in topologically nontrivial flat bands*, Nat. Communs. **6**, 8944 (2015), doi:10.1038/ncomms9944.

[30] T. Ozawa and B. Mera, *Relations between topology and the quantum metric for Chern insulators*, Phys. Rev. B **104**, 045103 (2021), doi:10.1103/PhysRevB.104.045103.

[31] Y. Onishi and L. Fu, *Fundamental bound on topological gap*, Phys. Rev. X **14**, 011052 (2024), doi:10.1103/PhysRevX.14.011052.

[32] Y. Onishi and L. Fu, *Topological bound on structure factor* (2024), 2406.18654.

[33] P. Atkins and R. Friedman, *Molecular Quantum Mechanics*, Oxford University Press, 4th edn. (2005).

[34] J. J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley (1994).

[35] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Pergamon (1977).

[36] C. Kittel, *Introduction to Solid State Physics*, Wiley, 8th edition edn. (2004).

[37] J. A. Van Vechten, *Quantum Dielectric Theory of Electronegativity in Covalent Systems. I. Electronic Dielectric Constant*, Phys. Rev. **182**, 891 (1969), doi:10.1103/PhysRev.182.891.

[38] N. Ashcroft and N. D. Mermin, *Solid State Physics*, Saunders College Publishing (1976).

[39] K. Teegarden and G. Baldini, *Optical Absorption Spectra of the Alkali Halides at 10°K*, Phys. Rev. **155**, 896 (1967), doi:10.1103/PhysRev.155.896.

[40] D. Roessler and W. Walker, *Electronic spectrum of crystalline lithium fluoride*, J. Phys. Chem. Solids **28**(8), 1507 (1967), doi:https://doi.org/10.1016/0022-3697(67)90280-6.

[41] R. Z. Bachrach, *The optical absorption of lithium iodide*, Phys. Lett. A **30**(5), 318 (1969), doi:https://doi.org/10.1016/0375-9601(69)91018-4.

[42] O. Madelung, *Semiconductors: Data Handbook*, Springer (2004).

[43] G. Ortiz and R. M. Martin, *Macroscopic polarization as a geometric quantum phase: Many-body formulation*, Phys. Rev. B **49**, 14202 (1994), doi:10.1103/PhysRevB.49.14202.

[44] M. Dressel and G. Grüner, *Electrodynamics of Solids*, Cambridge University Press (2003).

[45] J. E. Lennard-Jones and A. E. Ingham, *On the calculation of certain crystal potential constants, and on the cubic crystal of least potential energy*, Proc. R. Soc. Lond. A **107**, 636 (1925), doi:10.1098/rspa.1925.0047.