Electricity at the macroscale and its microscopic origins

Paul Tangney

Department of Physics and Department of Materials,
Imperial College London

(Dated: April 10, 2024)

This work examines the relationships between electrical structures at the microscale and electrical structures at the macroscale. By *structures* I mean both physical structures, such as the spatial distributions of charge and potential, and the mathematical structures used to specify physical structures and to relate them to one another. I do not discuss magnetism and what little I say about energetics is incidental.

I define the fields that describe electrical macrostructure, and their rates of change, in terms of the microscopic charge density ρ, electric field $\mathcal{E}$, electric potential ϕ, and their rates of change. To deduce these definitions, I lay some foundations of a general theory of *structure homogenization*, meaning a theory of how any observable macroscopic field $\boldsymbol{\mathcal{V}}$ is related to spatial averages of its microscopic counterpart ν. An integral part of structure homogenization theory is the definition of macroscopic *excess fields* in terms of microscopic fields. The excess field of $\boldsymbol{\mathcal{V}} : \mathbb{R}^n \to \mathbb{R}$ on the boundary $\partial\Omega$ of a finite-measure subset Ω of $\mathbb{R}^n$ is the field $\boldsymbol{\sigma}_{\mathcal{V}} : \partial\Omega \to \mathbb{R}$ to which it is related by the generalized Stokes theorem, $\int_\Omega \boldsymbol{\mathcal{V}}\,\mathrm{d}\omega = \int_{\partial\Omega} \boldsymbol{\sigma}_{\mathcal{V}}\omega$, where ω and $\mathrm{d}\omega$ are volume forms on $\partial\Omega$ and Ω, respectively, and $\boldsymbol{\mathcal{V}}\,\mathrm{d}\omega \equiv \mathrm{d}\,(\boldsymbol{\sigma}_{\mathcal{V}}\omega)$. For example, the macroscopic volumetric charge density $\boldsymbol{\varrho}$ in a material Ω is related to the areal charge density $\boldsymbol{\sigma}$ on its surface by $\int_\Omega \boldsymbol{\varrho}\,\mathrm{d}^3\mathbf{r} = \int_{\partial\Omega} \boldsymbol{\sigma}\,\mathrm{d}^2\mathbf{s}$ and by $\boldsymbol{\varrho}\,\mathrm{d}^3\mathbf{r} \equiv \mathrm{d}\left(\boldsymbol{\sigma}\,\mathrm{d}^2\mathbf{s}\right)$. I derive an expression for $\boldsymbol{\sigma}_{\mathcal{V}}[\nu]$, which generalizes Finnis's expression for excess fields at the surfaces of crystals (e.g., surface charge density $\boldsymbol{\sigma}[\rho]$) to disordered microstructures.

I use homogenization theory to define the macroscopic potential $\boldsymbol{\Phi} \equiv \boldsymbol{\Phi}[\phi]$, electric field $\mathbf{E} \equiv \mathbf{E}[\mathcal{E}]$, and charge density $\boldsymbol{\varrho} \equiv \boldsymbol{\varrho}[\rho]$, and I define the macroscopic current density as $\mathbf{J} \equiv \dot{\boldsymbol{\sigma}}[\dot{\rho}]$. Using the microscopic theory, or *vacuum theory*, of electromagnetism as my starting point, I deduce that the relationships between these macroscopic fields are identical in form to the relationships between their microscopic counterparts. Without invoking quantum mechanics, I use the definitions $\mathbf{J} \equiv \dot{\boldsymbol{\sigma}}$ and $\boldsymbol{\sigma} \equiv \boldsymbol{\sigma}[\rho]$ to derive the expressions for so-called *polarization current* established by the *Modern Theory of Polarization.* I prove that the bulk-average electric potential, or *mean inner potential*, $\boldsymbol{\Phi}$, vanishes in a macroscopically-uniform charge-neutral material, and I show that when a crystal lattice lacks inversion symmetry, it does not imply the existence of macroscopic $\mathbf{E}$ or $\mathbf{P}$ fields in the crystal's bulk.

I point out that symmetry is scale-dependent. Therefore, if anisotropy of the microstructure does not manifest as anisotropy of the macrostructure, it cannot be the origin of a macroscopic vector field. Only anisotropy of the *macrostructure* can bestow directionality at the macroscale. The macroscopic charge density $\boldsymbol{\varrho}$ is isotropic in the bulks of most materials, because it vanishes at every point. This implies that, regardless of the microstructure ρ, a macroscopic electric field cannot emanate from the bulk. I find that all relationships between *observable* macroscopic fields can be expressed mathematically without introducing the polarization ($\mathbf{P}$) and electric displacement ($\mathbf{D}$) fields, neither of which is observable. Arguments for the existence of $\mathbf{P}$ and $\mathbf{D}$, and interpretations of them, have varied since they were introduced in the 19th century. I argue that none of these arguments and interpretations are valid, and that macroscale isotropy prohibits the existence of $\mathbf{P}$ and $\mathbf{D}$ fields.

CONTENTS

PREFACE

I have tried to make this manuscript as modular as possible, while preserving the logic of the narrative as a whole. My hope is that, if you have learned the gist of this work from the abstract and Introduction, you will find that many sections and subsections are reasonably self-contained.

Those interested only in the homogenization transformation that turns microstructure into macrostructure, should read Secs. VI and VII, and Appendix E. Those interested in everything except the homogenization transformation, and who are willing to trust the formulae derived in Sec. VII and presented in Appendix A, can safely skip those parts.

Those interested only in the *mean inner potential*, $\mathbf{\Phi}$, should read Sec. IV (particularly Sec. IV.F), Sec. XIII and Sec. XIV.

Those interested only in polarization current, $\mathbf{J}^{(p)}$, or the *Modern Theory of Polarization* should read Secs. IV, V, X, and XI.

Those interested only in the macroscopic electric field, $\mathbf{E}$, should read Sec. IV and Sec. XIII.

Those interested only in surface charge, $\boldsymbol{\sigma}$, should read Secs. VI.C, VII and X.

Those interested only in single particle states should read Secs. V.B, XI.B.1, and XII, and Appendices B, C, and D.

I would be grateful for critical feedback on any part of this work (`p.tangney@imperial.ac.uk`).

I. INTRODUCTION

A. Background

Most of the classical electromagnetic theory that is commonly described in textbooks was established in the 19th century before electrons were discovered or the existence of atoms had been confirmed (Buchwald and Fox, 2013; Heaviside, 1893; Lorentz, 1916; Maxwell, 1865, 1873, 1892). The constitutive relations, $\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P}$ and $\mathbf{H} = \mu_0^{-1}\mathbf{B} - \mathbf{M}$, between the macroscopic electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, the induced fields $\mathbf{P}$ and $\mathbf{M}$, and the auxiliary fields $\mathbf{D}$ and $\mathbf{H}$, are an important part of this theory. To deduce them, materials were approximated as continua at the macroscale and the polarization ($\mathbf{P}$) and magnetization ($\mathbf{M}$) densities were introduced to characterize how the state of the ether was altered by their presence. When the concept of an ether was abandoned, they were reinterpreted as linear electromagnetic responses of materials. However this appears to have been done *ad hoc* and without due concern for consistency with the nascent theory of material microstructure.

The *microscopic* or *vacuum* theory of electromagnetism rightly underpins our microscopic theory of material structure, composition, and energetics. The purpose of *macroscopic* electromagnetism, which reduces to microscopic electromagnetism when $\mathbf{P}$ and $\mathbf{M}$ vanish, is to provide a unified description of materials and electromagnetic fields at the macroscale. Therefore it should be underpinned by our mutually-consistent microscopic theories of materials physics and vacuum electromagnetism, and we should clearly understand the microscopic origins of $\mathbf{P}$ and $\mathbf{M}$. However development of the macroscopic theory was completed before many of the discoveries on which we base our microscopic understanding of materials were made, and it quickly became an established part of physics doctrine. Therefore it was not built on firm microscopic foundations and, unfortunately, it has never been reconciled fully and satisfactorily with microscopic physics.

Inconsistencies between the microscopic and macroscopic theories were not apparent to most scientists until after crystallography had come of age and it had become possible to compute materials' *microstructures*, by which I mean the statistical distributions, on the nanoscale, of their constituent charges and magnetic moments. It became obvious that we lacked precise and viable definitions of $\mathbf{P}$ and $\mathbf{M}$ when attempts were made to define them in terms of microstructures (Aizu, 1962; Landauer, 1960, 1981; Larmor, 1921; Littlewood, 1980; Littlewood and Heine, 1979; Martin, 1974; Resta, 1992; Tagantsev, 1991; Vogl, 1978; Woo, 1971). Neither $\mathbf{P}$ nor $\mathbf{M}$ is directly measureable, but definitions of observables attributed by classical electromagnetic theory to changes in their values also proved elusive. For example, it was not until the 1990s that researchers discovered how to calculate the so-called *polarization current*, $\mathbf{J}^{(p)}$, that flows through an insulating inversion-asymmetric crystal when it is uniformly perturbed by a stimulus, such as a strain or a change in temperature (King-Smith and Vanderbilt, 1993; Resta, 1993, 1994; Resta and Vanderbilt, 2007; Vanderbilt and King-Smith, 1993). When they did, they concluded that this current, which is what is directly measured in experiments that produce $\mathbf{P}-\mathbf{E}$ or $\mathbf{D}-\mathbf{E}$ hysteresis loops, has quantum mechanical origins and does not exist within classical physics. Therefore their *Modern Theory of Polarization (MTOP)* deviated substantially from both the revised (20th century) definition of $\mathbf{P}$ as a dipole moment density and the physical reasoning with which its use as a measure of dielectric response to a uniform field was justified.

Another elusive and hotly-debated definition was the areal density of charge at a surface or interface, $\boldsymbol{\sigma}$. This problem was solved, for crystals, by Finnis in 1998 (Finnis, 1998) and, although the discussion continued for a time, there appears to be agreement now that Finnis's definition is correct (Bristowe *et al.*, 2014, 2011; Goniakowski *et al.*, 2008; Goniakowski and Noguera, 2014; Noguera and Goniakowski, 2013; Resta and Vanderbilt, 2007; Stengel, 2011; Stengel and Vanderbilt, 2009; Vanderbilt, 2018). Unfortunately, the literature is far from clear on this point, because multiple equivalent definitions of $\boldsymbol{\sigma}$ have been proposed, and some works continue to make the unnecessary distinction between *free charge* and *bound charge*, and to express the latter's contribution to $\boldsymbol{\sigma}$ as $\boldsymbol{\sigma}^{\text{bound}} = \mathbf{P} \cdot \hat{n}$, where $\hat{n}$ is the unit surface normal and $\mathbf{P}$ is the polarization in the crystal's bulk. The definition is complicated further by the fact that the MTOP definition of $\mathbf{P}$ is multi-valued.

A third illustration of the tension that exists between the 20th century theory of material structure and 19th century electromagnetism, is the question of how to define and calculate the bulk macroscopic electric potential, $\mathbf{\Phi}$, from the microscopic charge density. This quantity, which is often called the *mean inner potential* (MIP), plays an important role in several areas of research, including theoretical electrochemistry and electron microscopy (Blumenthal *et al.*, 2017; Cendagorta and Ichiye, 2015; Gajdardziska-Josifovska and Carim, 1999; Gajdardziska-Josifovska *et al.*, 1993; Hörmann *et al.*, 2019; Ibers, 1958; Kathmann, 2021; Kathmann *et al.*, 2011; Leung, 2010; Madsen *et al.*, 2021; Miyake, 1940; Peng, 1999; Pratt, 1992; Rez *et al.*, 1994; Saldin and Spence, 1994; Sanchez and Ochando, 1985; Sokhan and Tildesley, 1997a,b; Spence, 1993, 1999; Stillinger and Ben-Naim, 1967; Wilson *et al.*, 1987, 1988, 1989; Yesibolati *et al.*, 2020). An exact general definition of it has not previously been found, but several approximations to it have been proposed and are in use (Kathmann *et al.*, 2011; Pratt, 1992; Saunders *et al.*, 1992; Sokhan and Tildesley, 1997b). Bethe derived one such expression by approximating the microstructure as a superposition of spherically-symmetric atomic charge densities (Bethe, 1928).

B. Motivations and objectives

Considered individually, the three examples cited above suggest, at the very least, that the connection between microscopic electromagnetism and macroscopic electromagnetism is subtle. However, considered collectively, the situation appears more serious, because the fields $\mathbf{J}^{(p)} \equiv \dot{\boldsymbol{\sigma}}$, $\boldsymbol{\sigma}$, and $\mathbf{\Phi}$, are all measureable elements of electricity at the macroscale. Therefore, these are all examples of attempts to bridge, or to partially fill, the same hole in existing physical theory, namely: We do not understand the relationship between electricity at the macroscale and electricity at the microscale well enough to express the fields that specify a material's *electrical macrostate* in terms of the fields that specify its *electrical microstate.*

An *electrical microstate*, $(\mathfrak{m}, \dot{\mathfrak{m}})$, is an *electrical microstructure* $\mathfrak{m}$ and its time derivative, $\dot{\mathfrak{m}} \equiv \partial\mathfrak{m}/\partial t$, at the same instant. An *electrical microstructure* is the most complete and detailed information pertaining to the instantaneous spatial distribution of charge and electric potential that could, in principle, exist. It could be specified by the wavefunction, density matrix, or position probability density function of the set of all particles, but for many purposes the information required is contained in the microscopic electric potential ϕ and the microscopic charge density $\rho \equiv -\nabla^2\phi$, in which case we say that the electrical microstate is $(\phi, \dot{\phi})$ or $(\rho, \dot{\rho})$. An *electrical macrostate*, $(\mathfrak{M}, \dot{\mathfrak{M}})$, is a specification of the spatial distributions of charge and electric potential at the macroscale, $\mathfrak{M}$, and their time derivatives, $\dot{\mathfrak{M}}$.

In each of the three examples discussed in Sec. I.A, a different line of reasoning was followed to derive an expression for $\mathbf{J}^{(p)} = \mathbf{J}^{(p)}[\dot{\mathfrak{m}}]$, $\boldsymbol{\sigma} = \boldsymbol{\sigma}[\mathfrak{m}]$, or $\mathbf{\Phi} = \mathbf{\Phi}[\mathfrak{m}]$. However, none of these lines of reasoning were pursued far enough to elucidate the relationship between macrostructure and microstructure fully, and with enough generality that what was learned could be applied, not only throughout electromagnetic theory, but far beyond it: in elasticity theory, meteorology, astrophysics, and countless other areas of research. For example, the MTOP did not provide an expression for $\mathbf{\Phi}[\mathfrak{m}]$, and Finnis did not derive an expression for $\mathbf{J}^{(p)}[\dot{\mathfrak{m}}]$ from his expression for $\boldsymbol{\sigma}[\mathfrak{m}]$.

Objective 1

My first objective is to reconcile the fundamental elements of our macroscale theory of electricity in materials with our mutually-compatible theories of electromagnetism and material structure at the microscale.

Objective 2

My second objective is to begin laying robust foundations on which a comprehensive theory of the relationship between physics at the microscale and physics at the macroscale can be built.

If you look around you, you will see surfaces, edges, and corners everywhere. Everything you see is a feature of the macrostructure, meaning that it is a blurred image of the microstructure at a surface, edge, or corner. You do not see the microstructure in its full horrendous complexity, and you do not notice that it changes from one femtosecond to the next. You see a relatively simple and relatively stable homogenized version of the microstructure.

There are many sources of imprecision, such as the diffraction limit, and it would be impossible for you to be aware of the full microstructure because, for example, a cubic molar sample of an element has $\sim 10^{16}$ atoms at each of its six faces, but the human brain only has $\sim 10^{11}$ neurons. Therefore homogenization of microstructure to form macrostructure is intrinsic to the act of observation.

Nevertheless, given a microstructure ν and access to an arbitrarily-powerful computer, it is not known how to calculate the macrostructure, or even what mathematical form it would take. The inconsistencies between Maxwell's macroscopic and microscopic theories of electromagnetism are only one of many important consequences of this gap in our understanding.

Therefore I begin to address the following question, which is of general importance to mathematical physics:

How can a macrostructure be expressed mathematically in terms of the microstructure underlying it?

This question leads quickly to a more fundamental question:

If the microstructure is a scalar field $\nu : \mathbb{R}^3 \to \mathbb{R}$*, what is the mathematical form of the macrostructure?*

As I will discuss in Sec. VI, qualitative differences exist between a macrostructure and a *base microstructure*, where I use the term *base microstructure* to mean a microstructure that is not itself the macrostructure arising from a structure on an even smaller length scale.

Objective 3

A third objective of this work is to emphasize how little of the physics of electricity in materials requires physical assumptions that are incompatible with classical physics.

Most textbooks on solid state physics or electronic structure theory do not clearly demarcate the features of mathematical representations of quantum mechanical microstates that, for fundamental physical reasons, are peculiar to quantum mechanics, from features that are perfectly consistent with a classical statistical microstate. For example, when we see a statistical state expressed as $\Psi \equiv \sqrt{p}e^{i\theta}$, where $p = p(\vec{r}_1, \vec{r}_2, \cdots)$ is a joint position probability density function, we often assume that quantum mechanics is being 'used'. However, it is perfectly valid to express the statistical state of a system of classical particles in this form, and it can be useful to do so. Having done so, the classical many-particle state Ψ can be expanded in a basis of single particle states, just as in quantum mechanics.

We largely base our physical intuition on what we observe at the human scale. If the blurred lines between classical and quantum physics were made more clear, we would have a better understanding of when we could apply our classical intuition to systems of quantum mechanical particles, and when our intuition was likely to fail us.

I begin to address this issue in the present work for two reasons. The first is that fulfilling my first objective, and relating my findings to the MTOP, requires me to survey many parts of electronic structure theory and solid state physics. Therefore I have the opportunity to point out that much of the mathematical infrastructure that we usually associate with quantum mechanics is perfectly consistent with classical physics.

The second reason is that there are claims in the literature on the MTOP that some of the observable quantities that I discuss in this work have quantum mechanical origins and do not have analogues within classical physics (Resta, 1993, 1994). It is important to examine these claims carefully.

Single particle states play a prominent role in the MTOP. Therefore I emphasize that there is nothing specific to quantum mechanics about Bloch functions (Bloch, 1929) and Wannier functions (Wannier, 1937). If the bulk of a crystal is represented in a torus, which is equivalent to using Born-von Kármán boundary conditions (Born and von Kármán, 1912), and if Ψ is a stationary statistical state resulting from a classical process that preserves the crystal's periodicity, it can be expanded in a basis of Bloch functions. Each set of Bloch functions can be transformed into an infinite number of sets of Wannier functions, which must include a maximally localized set (Ferreira and Parada, 1970; Marzari and Vanderbilt, 1997) and a minimally localized set.

The MTOP approach to calculating $\mathbf{J}^{(p)}$ gives exactly the right result when the charge density can be expressed in the form $\rho(\vec{r};\zeta) = \sum_i q_i |\phi_i(\vec{r};\zeta)|^2$, where ζ is the stimulus whose rate of change $\dot{\zeta}$ gives rise to $\mathbf{J}^{(p)}$, each ϕ_i is a single particle basis function with a smooth dependence on ζ, whose norm is independent of ζ, and each q_i is independent of ζ.

It is known that the MTOP can be applied to electrons in an insulator, because the electron density can be represented by a set of smoothly-evolving Bloch or Wannier states of fixed occupancies. However, it is not known which classical systems and processes it can be applied to, because the question of which charge densities admit such a representation has not be solved. I do not shed light on the answer to this representability problem, but I attempt to clarify the question.

C. Theoretical approach and outline of this work

I define the *homogenization transformation* that turns a microstructure into a macrostructure as a spatial averaging operation on a mesoscopic domain. This obvious approach, which appears physically reasonable, has been attempted many times before by many authors (de Groot and Vlieger, 1965; Frias and Smolyakov, 2012; de Groot and Vlieger, 1964; Kirkwood, 1936; de Lange and Raab, 2006; de Lange *et al.*, 2012; Mazur, 1957; Mazur and Nijboer, 1953; Raab and de Lange, 2005; Robinson, 1971; Roche, 2000; Russakoff, 1970; Schram, 1960). Some of these attempts are presented in well known textbooks (Ashcroft and Mermin, 1976; Jackson, 1998).

However none of these approaches have been adopted by the research community as foundations for the development of rigorous theory, because they do not lead to Maxwell's macroscopic theory of electricity. In Sec. II I explain why we should not be deterred by this: I outline the reasoning that Maxwell used to derive his macroscopic theory in order to demonstrate that his reasoning has been invalidated by what has since been learned about spacetime and material microstructure. Therefore we should not require unobservable elements of Maxwell's macroscopic theory, such as $\mathbf{P}$, to be elements of a macroscopic theory that is derived from, and consistent with, his vacuum theory of electromagnetism

In Sec. III, using the macroscopic polarization $\mathbf{P}$ as an example, I briefly explain some of the ways in which definitions of macroscopic fields have failed in the past.

In Sec. IV I argue that many of my conclusions, and many elementary aspects of electricity at the macroscale, are demands of symmetry or asymmetry. For example, any stimulus changes the microscopic charge density in the bulk of a crystal, to some degree. While ρ is changing, microscopic polarization current ($j^{(p)}$) flows, because $\partial\rho/\partial t = -\boldsymbol{\nabla}\cdot j^{(p)}$. Whether or not a net, or *macroscopic*, polarization current ($\mathbf{J}^{(p)}$) flows depends on the symmetry of the composite crystal+stimulus system: The component of $\mathbf{J}^{(p)}$ in direction $\hat{u}$ vanishes if there is a glide plane normal to $\hat{u}$, because then the sum,

$$j^{(p)}(\vec{r})\cdot\hat{u} + j^{(p)}(\hat{\mathfrak{g}}\,\vec{r})\cdot\hat{u},$$

of the contributions to $\mathbf{J}^{(p)}\cdot\hat{u}$ from an arbitrary point $\vec{r}$ and its image under the glide symmetry, $\hat{\mathfrak{g}}\,\vec{r}$, vanishes. On the other hand, if symmetry does not demand that $\mathbf{J}^{(p)}\cdot\hat{u}$ vanishes, it must be finite: There is a vanishing probability that, within each primitive unit cell, the positive contributions to $\mathbf{J}^{(p)}\cdot\hat{u}$ cancel the negative contributions *exactly* (i.e., to infinite precision) by chance. Therefore, either anisotropy demands that $\mathbf{J}^{(p)}$ is finite, or isotropy demands that it vanishes.

In Sec. V I discuss the *Modern Theory of Polarization*, and I derive the MTOP expression for $\mathbf{J}^{(p)}$ without invoking quantum mechanics.

In Sec. VI I explain in more detail what I mean by the prefixes *micro-* and *macro-*. I briefly explain some of the qualitative differences between a macrostructure and a base microstructure, and their mathematical and physical origins.

In Sec. VII I discuss the macroscopic *excess fields* that exist at surfaces, interfaces, edges, and line and point defects. Excess fields are the manifestations at the macroscale of *abrupt* changes of the microstructure, meaning changes that occur across microscopic distances. For example, the difference in microstructure between a material and vacuum manifests as a surface charge density, $\boldsymbol{\sigma}$. I derive expressions for macroscopic excess fields in terms of microscopic volumetric fields, which generalize Finnis's expression for surface excesses to nonperiodic microstructures.

In Sec. VIII I use spatial averaging of the microscopic charge density ρ to calculate its macroscopic counterpart $\boldsymbol{\varrho}$, and in Sec. X I define the surface charge density, $\boldsymbol{\sigma}$, as the integral of $\boldsymbol{\varrho}$ along a path that crosses the surface. This leads to Finnis's expression for the surface charge of a crystal, $\boldsymbol{\sigma}[\rho]$, and to my generalization of this expression to noncrystalline materials.

In Sec. XI I derive the MTOP expression for $\mathbf{J}^{(p)}$ again, but this time I derive it by defining it as $\mathbf{J}^{(p)} \equiv \dot{\boldsymbol{\sigma}} = \mathrm{d}\boldsymbol{\sigma}[\rho]/\mathrm{d}t$, where $\boldsymbol{\sigma}[\rho]$ is Finnis's formula.

In Sec. XII I point out that there does not exist a theoretical justification for interpreting the sets of single electron states that appear in the MTOP definitions of polarization current as chemically meaningful substructures of the electron density.

In Sec. XIII I prove that the macroscopic potential, or mean inner potential, $\boldsymbol{\Phi}$ vanishes in the bulk of any isolated material whose surface is locally charge neutral. Then I point out a flaw in the reasoning used by H. A. Lorentz to deduce that a macroscopic electric field exists in the bulk of a crystal whose microstructure is anisotropic. In Sec. XIV I discuss flaws in Bethe's derivation of an approximate expression for $\boldsymbol{\Phi}$, and I show that $\boldsymbol{\Phi}$ vanishes when these flaws are avoided.

This work comprises three interwoven strands, whose individual objectives are the three objectives outlined above. I conclude with a summary of each strand.

D. Notational conventions and some physical assumptions

I now explain a few non-standard notational conventions that I use throughout this work. More notation will be introduced as and when it is used. I also explain some of the assumptions that I make about materials' microstructures.

1. Macroscopic quantities

I use boldface type to distinguish macroscopic quantities from microscopic quantities throughout this work. For example, the macroscopic analogues of the microscopic charge density ρ, the microscopic electric potential ϕ, and the microscopic electric field $\mathcal{E}$, are denoted by $\boldsymbol{\varrho}$, $\boldsymbol{\Phi}$, and $\mathbf{E}$, respectively; and $\mathbf{r}$ and $\vec{r}$ denote points, positions, or displacements at the macroscale and the microscale, respectively.

2. Intervals

If u is any quantity with a continuous range of possible values (e.g., an x coordinate or an average of field ν), I use $\mathfrak{I}(u,\Delta)$, $\mathfrak{I}[u,\Delta)$ and $\mathfrak{I}(u,\Delta]$, and $\mathfrak{I}[u,\Delta]$ to denote interval subsets of this range, of width Δ and centered at u, which are open, half-open, and closed, respectively. I use the more conventional notation (u_1,u_2), $[u_1,u_2)$ and $(u_1,u_2]$, and $[u_1,u_2]$, to specify intervals by their end points. For example, $\mathfrak{I}(u,\Delta u] = (u-\Delta u/2, u+\Delta u/2]$ is an interval that is open at its lower boundary and closed at its upper boundary.

3. Length scales

The *microscale*, a, is the smallest length scale of relevance to the physics of materials, and the *macroscale*, L, is a much larger length scale, on which materials appear continuous, rather than particulate.

Whenever I use the prefixes *micro* and *macro*, it is implicit that there also exists a *mesoscale*, l, where $a \ll l \ll L$. The mesoscale is an intermediate length scale, which is orders of magnitude larger than a bond length, but small enough that all nonlinear contributions to the spatial variations of all macroscopic fields are negligible. The assumption that there exists a mesoscale is useful, and possibly necessary, for understanding the relationship between microstructures and macrostructures.

The statements $\Delta_a \sim a$, $\Delta_l \sim l$, and $\boldsymbol{\Delta}_L \sim L$, mean that Δ_a, Δ_l, and $\boldsymbol{\Delta}_L$ are distances or displacements which are microscopic, mesoscopic, and macroscopic, respectively. I will explain precisely what I mean by the terms *microscopic*, *mesoscopic*, and *macroscopic* in Sec. VI.

In Sec. VI, I will define the *macroscale infinitesimal* $|\mathbf{dx}|$, which is a lower bound on the magnitudes of displacements that are measurable at the macroscale. On the microscale I denote it by ε_x.

4. A *quasi*-one dimensional material

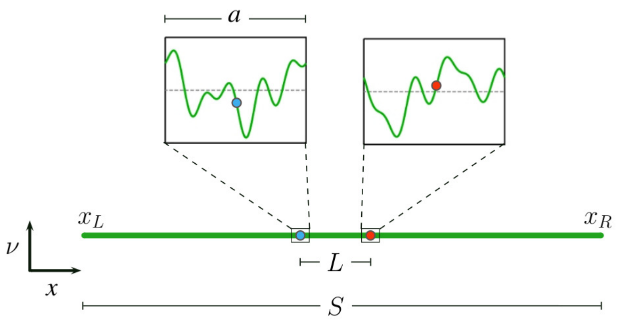

FIG. 1 Schematic of a *quasi* one dimensional material which is macroscopically-uniform but microscopically non-uniform.

It may be useful to consider the macroscopically-uniform material depicted schematically in Fig. 1. It can be viewed as a three dimensional material with a large aspect ratio and a microstructure, $\nu:\mathbb{R}\to\mathbb{R}$, which is only one dimensional because its value $\nu(x)$ at x is really an average of its three dimensional microstructure in the plane perpendicular to the page.

The planes perpendicular to the page at $x = x_L$ and $x = x_R = x_L + S$ bound the material in the $-\hat{x}$ and $\hat{x}$ directions, respectively, where S is the length of the material. Therefore $\nu(x)$ is negligible when $x \notin (x_L, x_R)$.

The bulk macrostructure is depicted as uniform in Fig. 1, which is not the general case. I do not make any assumptions about the macroscopic charge density $\boldsymbol{\varrho}$, except that it is differentiable *almost everywhere*. By this I mean that it is differentiable everywhere except where it changes nonlinearly on an interval of width $\varepsilon_x = |\mathbf{dx}|$.

For example, as I explain in Sec. VI, a surface is assumed to have a width of less than ε_x at the microscale. This is not a physical assumption about the surface, but a consequence of the definition of ε_x. It means that the microscopic charge density ρ changes from being characteristic of the material's bulk to being characteristic of the vacuum above the surface (i.e., $\rho = 0$) on an interval that is smaller than the macroscale infinitesimal. Therefore, $\boldsymbol{\varrho}$ cannot be assumed to be differentiable or continuous at a surface.

5. Microstructure dimensionality

Whenever the argument of a microscopic field ν has an arrow over it (e.g., $\nu(\vec{r})$), its domain is implicitly three dimensional. When it does not (e.g., $\nu(x)$) its domain is implicitly one dimensional. When it has one argument with an arrow and one without an arrow (e.g., $\nu(u, \vec{s})$), its domain is three dimensional and the argument with the arrow ($\vec{s}$) denotes a vector in the plane perpendicular to the x axis at the x coordinate specified by the argument without an arrow (u).

6. Boundaries of a material and the bulk of a material

I denote the set of all values of $\vec{s}$ for which the point $\vec{r} = (x, \vec{s})$ is within the material by $\mathfrak{P}(x)$ and I denote the cross-sectional area of the material at x by $|\mathfrak{P}(x)|$.

In Sec. VI we will see that homogenization of microstructure by spatial averaging is tantamount to spatial compression, and that all points at the microscale within $\varepsilon_x/2$ of a material's boundary are mapped to the same locally-planar surface at the macroscale. The set of x values of points at the microscale that are mapped by homogenization to the material's left and right surfaces are $\mathbf{x}_L \equiv \mathfrak{I}(x_L, \varepsilon_x)$ and $\mathbf{x}_R \equiv \mathfrak{I}(x_R, \varepsilon_x)$, respectively.

At the microscale, $\mathbf{x}_L$ and $\mathbf{x}_R$ are *coincidence sets*, (see Sec. VI for a definition): they are the sets of all values of x that are indistinguishable from x_L and x_R, respectively, at the macroscale. However, at the macroscale they can be regarded and treated as coordinates, because any interval $\mathfrak{I}(x, \eta)$ for which $\eta \leq \varepsilon_x$ only contains points that cannot be distinguished from x by empirical means at the macroscale.

I will use $\mathfrak{B}$ to denote the set $[x_L + \varepsilon_x/2, x_R - \varepsilon_x/2]$ of x coordinates of points within the material's bulk. I denote the width of the bulk region by $S_{\mathfrak{B}} \equiv |\mathfrak{B}| = S - \varepsilon_x \approx S$. I will also use $\mathfrak{B}$ at the macroscale, where it denotes the open interval $(\mathbf{x}_L, \mathbf{x}_R)$. The *closure* $[\mathbf{x}_L, \mathbf{x}_R]$ of $(\mathbf{x}_L, \mathbf{x}_R)$ is the union of the bulk and the surfaces, which is the entire material.

7. Externally applied fields

I will primarily be concerned with materials that are either isolated or under the influence of constant or slowly-varying ($f \lesssim$ GHz $\iff \lambda \gtrsim 1\,\mathrm{m}$) externally-applied electromagnetic waves. I assume that the size (S) of the material object is much less than the wavelength (λ) and so, within the material, any external fields are effectively spatially-uniform. I also assume that the period of oscillation, $1/f \gtrsim 1$ ns, of the field is much longer than the charge density's relaxation time.

8. Microstructures of materials

To facilitate discussing materials in general terms I use a fairly general mathematical representation of a material. However I will assume that the net charge of the material is zero and, because I am not concerned with magnetism, I will assume that all particles have zero spin. For the purposes of this work, the only relevant characteristics of each particle are its charge and its mass. The relevance of particles' masses is that they determine how delocalized each particle is and that nuclei move much more slowly than electrons.

In quantum mechanics the state of an *isolated* thermally-disordered object at time t can be specified completely by the position probability density function of its N constituent particles,

$$p : \mathbb{R}^{3N+1} \to \mathbb{R}_+ ; \quad (\vec{r}_1, \cdots, \vec{r}_N, t) \mapsto p(\vec{r}_1, \cdots, \vec{r}_N, t).$$

The finite spatial and temporal precisions of all measurements, and the fundamentally-perturbative nature of the act of observation, mean that, even within classical physics, an observer's knowledge of the state of any physical system is a probability density function, rather than a set of precise values of positions and momenta.

When the only properties of interest to the observer are statistical properties of functions of particle positions, such as the expectation value and variance of the electric potential ϕ at a point, or the rates of change of those quantities, momenta can be integrated out. Therefore, both classically and quantum-mechanically the material's microstructure can be specified by a *continuous* probability density function $p(\vec{r}_1, \cdots, \vec{r}_N, t)$.

The probability density p could only be non-analytic if measurements had infinite precision in both space and time. However, even if it were non-analytic, every *delta distribution*, by which I mean a weighted sum of Dirac delta functions, is the limit of a smooth density. This means that there exists a smooth density arbitrarily close to any given delta distribution, from which properties of the delta distribution can be calculated to arbitrary precision if smoothness is required for the calculation. Therefore, I am not invoking quantum mechanics by using a position probability density function $p(\vec{r}_1 \cdots \vec{r}_N, t)$ to specify the structure of a material. I emphasize this point because it has sometimes been claimed that electric polarization $\mathbf{P}$ is a quantum phenomenon.

Until I discuss currents in Sec. XI I will assume that, because nuclei move slowly, the electrons can respond adiabatically to their motion. Therefore if, at a particular point in time, the subsystem of electrons is close to either a stationary state, such as its ground state, or a metastable state, it will remain close to this state as the state changes in response to the slowly-evolving confining potential from the nuclei. This means that, to a very good approximation, the time dependence of the number density of electrons, which I denote by $n(\vec{r},t)$, can be replaced by a parametric dependence on nuclear positions. I will not usually make this parametric dependence explicit, but I will omit t as an argument to $n(\vec{r})$ and to the microscopic charge density ρ whenever I am making this *adiabatic approximation.*

The quantity of primary interest in electrostatics at the microscale is a material's charge density function,

$$\rho(\vec{r}) = \sum_{i=1}^{N} q_i \int_{\mathbb{R}^3} \cdots \int_{\mathbb{R}^3} \delta(\vec{r} - \vec{r}_i) \times p(\vec{r}_1, \cdots, \vec{r}_N)\, \mathrm{d}^3 r_1 \cdots \mathrm{d}^3 r_N \qquad (1)$$

where q_i is the charge of particle i, $\vec{r}_i$ is its position, and δ is the Dirac delta distribution. From now on I will denote the position of the *nucleus* with index i by $\vec{R}_i$, to distinguish it from the positions of electrons. I will assume that, to a good approximation, ρ can be expressed in the form

$$\rho(\vec{r}) = \overbrace{-\,e\,n(\vec{r})}^{\rho^-(\vec{r})} + \overbrace{Ze \sum_{i\in\text{nuclei}} \tilde{\delta}(\vec{r} - \vec{R}_i)}^{\rho^+(\vec{r})} \qquad (2)$$

where $-e$ is the charge of an electron, and Z is the atomic number of the nuclei. For simplicity I will often assume that the material contains only one species of nucleus.

The function $\tilde{\delta}(\vec{r} - \vec{R}_i)$ is not quite the Dirac delta distribution, but a highly localized smooth probability density function for the position of nucleus i. In many situations, but not all, we can treat it mathematically as we would treat the Dirac delta distribution.

The energy of attraction between the nuclei and the electrons can be expressed as

$$(n, v^{\text{ext}}) \equiv \int_{\mathbb{R}^3} n(\vec{r}) v^{\text{ext}}(\vec{r})\, \mathrm{d}^3 r\,,$$

where v^{ext} is equal to $-e$ times the positive electric potential from the nuclei. In studies of the electronic subsystem at fixed nuclear positions, it is common to refer to v^{ext} as as the *external potential.*

For a one dimensional material aligned with the $x-$axis, such as the one depicted in Fig. 1, the analogue of Eq. 2 is

$$\begin{aligned}\rho(x) &= \rho^-(x) + \rho^+(x) \\ &= -e\,n(x) + Ze \sum_{i\in\text{nuclei}} \tilde{\delta}(x - X_i)\end{aligned} \qquad (3)$$

For most purposes, I will specfiy the (electrical) microstructure of the material as ρ or as (ρ^+, ρ^-)

II. MAXWELL'S THEORY OF THE ETHER

I now summarize some of the reasoning that led Maxwell to his macroscopic theory of electromagnetism. My purpose is to show that almost all of his reasoning is inconsistent with what has since been learned about materials and spacetime and that, in hindsight, the existence of **P** and **D** fields appears not to have a sound logical basis. The sources I have relied on most heavily are Maxwell, 1865, Maxwell, 1873, Maxwell, 1892, Heaviside, 1893, Lorentz, 1916, and Buchwald and Fox, 2013.

Maxwell believed that a *luminiferous ether* pervaded all matter and was the domain in which all electromagnetic processes occurred. He rejected the idea of electromagnetic action at a distance, believing instead that the ether was the medium by which, and through which, all forces between electrified bodies were exerted.

In vacuum he regarded the ether as isotropic, homogeneous, and with properties characterised by only two scalar constants, such as its permittivity ϵ_0 and its permeability μ_0, or either one and the speed of light c. He believed that the ether's properties were altered in the presence of matter, but that the effects of matter on electromagnetic phenomena were indirect and could, to a first approximation, be described by the changes $\epsilon_0 \mapsto \tilde{\epsilon}(\mathbf{r})$ and $\mu_0 \mapsto \tilde{\mu}(\mathbf{r})$ of the ether's characteristic constants from uniform scalars to tensor fields.

Maxwell used the term *electricity* in an abstract or vague sense and he likened electricity in the ether to elasticity in a solid. He regarded this analogy as so compelling that, on the basis of it, he was willing to impute to the ether the minimal set of physical properties necessary to make his theory internally consistent. He reasoned that, just as a slack rope or an unstrained rod cannot transmit forces between its two ends, the ether must be in a state of mechanical stress when electric forces are transmitted through it. Therefore, just as an elastically-deformed solid has a density of stored energy at each point, which is released when the deforming force is removed and the body resumes its original shape, he speculated that an electric force field was always accompanied by a displacement field *in the ether*, which stored potential energy.

As an alternative to specifying a deformed state of an elastic solid with a vector field whose value at each point is the point's displacement from equilibrium, one can describe it by a flux density vector field that is parallel to the direction of material flow at each point and has a magnitude equal to the total quantity of material, per unit area, that flowed through a small imaginary surface at the point during the deformation. Maxwell chose this latter approach to describe the state of the ether and the motion of electricity within it. One of his reasons was that certain fluxes, namely electric currents, were measureable, and measured fluxes were spatial averages, which could not be calculated using the former approach unless much more detailed information about the ether was available. For example, a rate of fluid flow cannot be calculated if one only knows the velocity of the fluid at each point; knowledge of its density is also required. Therefore Maxwell defined the confusingly-named electric displacement field $\mathbf{D}$ as a flux density. It specified how much electricity passed through each point, and in which direction, as an applied field that caused and maintained this displacement was 'switched on'. He regarded $\mathbf{D}$ as a specification of the electrically-deformed state of the ether, albeit one that was a spatial average of a more detailed microscopic flux density. He regarded the current density $\mathbf{J}$ as a rate of motion, or a velocity, of the ether that was driven by the electric force $\mathbf{E}$ and which changed the ether's displacement $\mathbf{D}$.

For reasons that remained mysterious to Maxwell, conductors lacked the restoring force that returned the $\mathbf{D}$ field in an electrically-deformed dielectric to its original state when the electric field supporting it was switched off. Therefore, instead of simply displacing, electricity flowed freely as a current. As it flowed, it dissipated some of the ether's energy into heat within the material; similarly, when a dielectric was placed in an electric field a transient current $\dot{\mathbf{D}}$ flowed and dissipated energy until the equilibrium displacement was reached. For energy to be conserved the energy stored in the ether by the displacement field had to be lower in the presence of a dielectric than it was in free space. It followed that, for the same electric force $\mathbf{E}$, $\mathbf{D}$ was different within a dielectric to its value of $\epsilon_0\mathbf{E}$ in vacuum. Its value in a dielectric was $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$, where $\mathbf{P}$ was known as the *electric polarization* of the dielectric.

Since $\mathbf{D}$ was different in dielectrics, it must change abruptly at a dielectric's boundaries. Maxwell viewed charge, not as a substance that can accumulate, but as a spatial discontinuity of $\mathbf{D}$. He did not understand a current to be a flow of charge but as a state of motion of the ether, which changed the $\mathbf{D}$ field, creating those discontinuities. So, although current did not transport charge, it caused it to exist.

Although there are some similarities between Maxwell's conception of electric polarization and more modern viewpoints, overall the physical picture described above bears little resemblance to modern conceptions of electromagnetism, spacetime, or the structures and compositions of materials. Maxwell's reasoning has become as obsolete as his conception of an ether is and both he and his contemporaries were alert to this eventuality. They regarded the properties he imputed to the ether as conjectures which would, when more was learned, either be confirmed and developed further, amended, or abandoned. Heaviside made his concern about the challenges the theory faced clear in 1893, more than a decade after Maxwell's death, when he wrote (Heaviside, 1893): "*Whether Maxwell's theory will last, as a sufficient and satisfactory primary theory upon which the numerous secondary developments required may be grafted, is a matter for the future to determine. Let it not be forgotten that Maxwell's theory is only the first step towards a full theory of the ether ; and, moreover, that no theory of the ether can be complete that does not fully account for the omnipresent force of gravitation.*"

Maxwell's theory should not be expected to make sense *conceptually* as a theory of material response, because he developed it as a theory of the ether. However, because he ensured that it reproduced all of the empirically-known mathematical relationships between macroscale observables, its *accuracy* as a macroscale tool is undiminished by the historical peculiarities of its mathematical form. It became confusing conceptually when the concept of an ether became obsolete, thereby stripping it of its logical foundations. More importantly, it fails when used beyond the macroscale domain for which it was constructed. At the microscale it conflicts with 20th century theories of material structure and composition; and this is why there has been so much confusion and debate about how to calculate macroscopic fields from microscopic ones.

III. HOW IS $\mathbf{P}$ DEFINED?

Many attempts have been made to reconcile Maxwell's auxiliary electric fields, $\mathbf{P}$ and $\mathbf{D}$, with modern conceptions of material structure and composition. Most have proposed definitions of $\mathbf{P}$ in terms of the microscopic charge density ρ. Once definitions of $\mathbf{E}$ and $\mathbf{P}$ are in hand, the definition of $\mathbf{D}$ follows from the constitutive relation quoted in the Introduction.

None of the proposed definitions of $\mathbf{P}$ are viable, to my knowledge, and I briefly explain their shortcomings in this section. I do not attempt to refute every paper directly, but I outline a few of the most common definitions of $\mathbf{P}$ and the reasons why they are unsatisfactory. The literature on the *Modern Theory of Polarization* (MTOP) (e.g., Resta and Vanderbilt, 2007), which is discussed in Sec. V, can be consulted for more discussion about shortcomings of pre-MTOP definitions of $\mathbf{P}$. I outline reasons to look beyond the MTOP definition of $\mathbf{P}$ in Sec. V.

1. Attempt 1

$\mathbf{P}$ has the dimensions of a dipole moment per unit volume and so it is natural to try to define it as such. Let us consider a material that occupies and fills a space $\Omega \subset \mathbb{R}^3$, whose volume is $|\Omega|$. Many authors have assumed, often tacitly, that Ω can be divided into microscopic partitions in some natural or 'right' way. For example, in a molecular material there might be a separate partition for each molecule. $\mathbf{P}$ is then defined as the macroscopic spatial average of the partitions' dipole moments divided by their volumes. This definition fails because, as Fig. 11 and Fig. 12 illustrate, there are an infinite number of ways to partition any material, which are equally justifiable theoretically, and each different set of partitions leads, in general, to a different magnitude and direction of $\mathbf{P}$.

2. Attempt 2

One could also define $\mathbf{P}$ as the dipole moment of the entire material divided by its volume, i.e.,

$$\mathbf{P} \equiv \frac{1}{|\Omega|} \int_{\Omega} \rho(\vec{r})\, \vec{r}\, \mathrm{d}^3 r \tag{4}$$

This definition is not satisfactory because it implies that $\mathbf{P}$ is not a property of the bulk, in general. To understand why, consider the dipole moment of a charge-neutral crystalline rod of length L and area of cross-section A, whose surfaces perpendicular to its long axis and carry charges of $+q$ and $-q$. For simplicity, let us suppose that the rod is carved from a perfect ionic crystal, immediately isolated so that its composition cannot change, and prevented from relaxing structurally. If L was large compared to the crystal's lattice constant the rod's dipole moment would be approximately equal to $qL\,\hat{n}$, where $\hat{n}$ is an outward unit normal to the surface of charge q. However, both the magnitude and the sign of q are determined by where along the rod's axis the bulk crystal was cleaved to form its surfaces. As illustrated in Fig. 12, two surfaces formed by cleaving a crystal along relatively-shifted parallel planes have different charges, in general. Therefore, by this definition $\mathbf{P} \approx qL\hat{n}/|\Omega| = (q/A)\,\hat{n}$ is mostly determined by the areal density of charge on the rod's surfaces $\boldsymbol{\sigma} \equiv q/A$ and it would vanish if the surfaces were neutralized. Therefore, Eq. 4 does not define a bulk property.

3. Attempt 3

One could consider basing a definition on either $-\boldsymbol{\nabla} \cdot \mathbf{P} = \boldsymbol{\varrho}$ or $-\boldsymbol{\nabla} \cdot \mathcal{P} = \rho$, where $\mathcal{P}$ is a microscale analogue of $\mathbf{P}$. This approach fails because, without boundary conditions, these equations only define $\mathbf{P}$ and $\mathcal{P}$ up to arbitrary constants; with boundary conditions, their values are determined by the charge at the material's surfaces, so they are not a property of the bulk.

4. Attempt 4

Finally, several well-known textbooks, including those by Jackson (Jackson, 1998) and Ashcroft and Mermin (Ashcroft and Mermin, 1976), use variants of a method refined by Russakoff (Russakoff, 1970) to define $\mathbf{P}$. They assume that the microscopic charge density can be expressed in the form $\rho(\vec{r}) = \sum_i \rho_i(\vec{r} - \vec{r}_i)$, where each ρ_i is a charge density that is localized around the origin and microscopic in extent, such that $\rho_i(\vec{r}) \approx 0$ when $|\vec{r}| \gg a$. For example, each ρ_i might be the distribution of a different molecule's charge.

The next step is to find the spatial average of ρ by convolving it with a smooth spherically-symmetric *averaging kernel* $\mu(\epsilon) : \mathbb{R}^3 \to \mathbb{R}_{\geq 0}$ whose width is proportional to ϵ and which has an integral of one, as follows:

$$\begin{aligned} \langle \rho; \mu \rangle_\epsilon (\vec{r}) &\equiv \int_{\mathbb{R}^3} \mu(\vec{u}; \epsilon)\, \rho(\vec{r} + \vec{u})\, \mathrm{d}^3 u \\ &= \sum_i \int_{\mathbb{R}^3} \mu(\vec{r} - \vec{r}_i - \vec{u}; \epsilon)\, \rho_i(\vec{u})\, \mathrm{d}^3 u\,, \end{aligned}$$

where I have changed the variable of integration and made use of μ's spherical symmetry. Taylor expanding $\mu(\epsilon)$ in each integrand to first order in $\vec{u}$ gives

$$\begin{aligned} \langle\, \rho\, \rangle (\vec{r}) \approx \sum_i \int_{\mathbb{R}^3} &\left[\mu(\vec{r} - \vec{r}_i; \epsilon) - \vec{u} \cdot \boldsymbol{\nabla}\mu(\vec{r} - \vec{r}_i; \epsilon)\right] \\ &\times \rho_i(\vec{u})\, \mathrm{d}^3 u \\ = \sum_i \mu(\vec{r} - \vec{r}_i; \epsilon)\, q_i &- \sum_i \boldsymbol{\nabla}\mu(\vec{r} - \vec{r}_i; \epsilon) \cdot \vec{d}_i \end{aligned}$$

where $q_i \equiv \int_{\mathbb{R}^3} \rho_i(\vec{r})\, \mathrm{d}^3 r$ and $\vec{d}_i \equiv \int_{\mathbb{R}^3} \rho_i(\vec{r} + \vec{u})\, \vec{u}\, \mathrm{d}^3 u$ are the net charge and the dipole moment of distribution ρ_i, respectively. If we identify $\langle \rho; \mu \rangle_\epsilon$ as the macroscopic charge density $\boldsymbol{\varrho}$ and express the second sum on the right hand side as $\boldsymbol{\nabla} \cdot \left(\sum_i \mu(\vec{r} - \vec{r}_i; \epsilon)\, \vec{d}_i\right)$, we find that

$$\boldsymbol{\varrho} = \boldsymbol{\varrho}^{\text{free}} - \boldsymbol{\nabla} \cdot \mathbf{P} = \boldsymbol{\varrho}^{\text{free}} + \boldsymbol{\varrho}^{\text{bound}}$$

where $\boldsymbol{\varrho}^{\text{bound}} \equiv -\boldsymbol{\nabla} \cdot \mathbf{P}$, and $\boldsymbol{\varrho}^{\text{free}}$ and $\mathbf{P}$ are volumetric densities of the molecules' net charges and dipole moments, respectively.

There are multiple fatal flaws in these definitions of $\boldsymbol{\varrho}^{\text{bound}}$, $\boldsymbol{\varrho}^{\text{free}}$, and $\mathbf{P}$: The value of each quantity depends sensitively on the value of ϵ, which is arbitrary, and both $\boldsymbol{\varrho}^{\text{bound}}$ and $\mathbf{P}$ vanish in the limit $\epsilon \to \infty$. Each field also depends on how ρ is partitioned into localized distributions ρ_i. However, even if the set $\{\rho_i\}$ was given, if $\boldsymbol{\varrho}^{\text{free}} \neq 0$ then the values of $\boldsymbol{\varrho}^{\text{bound}}$, $\boldsymbol{\varrho}^{\text{free}}$, and $\mathbf{P}$ would depend on the choice of origin. This is because the dipole moment of any charge distribution is origin dependent unless its net charge is zero.

IV. DEMANDS OF SYMMETRY AND ASYMMETRY

In this section I discuss some properties that we should expect a macrostructure to have, on symmetry grounds, if it is a spatial average of the microstructure.

I begin, in Sec. IV.A, by discussing consequences of the linearity of the spatial averaging operation $\nu \mapsto \langle \nu; \mu \rangle_\epsilon$. I discuss general macroscopic vector fields in Sec. IV.E, and I discuss the macroscopic potential $\boldsymbol{\Phi}$, the macroscopic polarization $\mathbf{P}$, and the polarization current $\mathbf{J}^{(p)}$ in Secs. IV.F, IV.G, and IV.H, respectively.

A. Linearity and the superposition principle

It is well known that, given two microscopic charge densities, ρ_1 and ρ_2, from which the microscopic electric fields $\mathcal{E}_1 \equiv \hat{\mathcal{E}}_\rho[\rho_1]$ and $\mathcal{E}_2 \equiv \hat{\mathcal{E}}_\rho[\rho_2]$ emanate, where $\hat{\mathcal{E}}_\rho$ is a functional; and given any two scalar constants $\omega_1, \omega_2 \in \mathbb{R}$; the following relation holds:

$$\hat{\mathcal{E}}_\rho[\omega_1\rho_1 + \omega_2\rho_2] = \omega_1\hat{\mathcal{E}}_\rho[\rho_1] + \omega_2\hat{\mathcal{E}}_\rho[\rho_2]. \tag{5}$$

Analogous relations hold for other functionals, such as $\hat{\phi}_\rho[\rho]$, $\hat{\mathcal{E}}_\phi[\phi]$, and $\hat{\rho}_\phi[\phi]$, which relate ϕ to ρ, $\mathcal{E}$ to ϕ and ρ to ϕ, respectively. The property of $\hat{\mathcal{E}}_\rho$ expressed by Eq. 5 is known as *linearity*, but in the context of electricity it is better known as the *principle of linear superposition* or simply the *superposition principle*. The superposition principle follows from the fact that derivatives and integrals are linear operations and the fact that $\mathcal{E} = -\boldsymbol{\nabla}\phi$ and $\rho/\epsilon_0 \equiv \boldsymbol{\nabla}\cdot\mathcal{E} = -\nabla^2\phi$ are both negative derivatives of ϕ.

A one dimensional spatial average has the general form

$$\langle \nu; \mu \rangle_\epsilon (x) \equiv \int_{\mathbb{R}} \nu(x')\mu(x'-x;\epsilon)\,\mathrm{d}x', \tag{6}$$

where ϵ is a parameter that is proportional to the width of the averaging kernel, $\mu(\epsilon)$. It is straightforward to use Eq. 6 to show that this is also a linear operation, i.e., $\langle \omega_1\nu_1 + \omega_2\nu_2; \mu \rangle_\epsilon (x) = \omega_1 \langle \nu_1; \mu \rangle_\epsilon (x) + \omega_2 \langle \nu_2; \mu \rangle_\epsilon (x)$, for any numbers $\omega_1, \omega_2 \in \mathbb{R}$ and any functions $\nu_1 = \nu_1(x)$ and $\nu_2 = \nu_2(x)$. The spatial averages in two and three dimensions are also linear operations.

B. Spatial averaging commutes with derivatives

It can be shown from Eq. 6 that spatial averages and derivatives commute. For example,

$$\partial_x^n \langle \nu; \mu \rangle_\epsilon \equiv \overbrace{\partial_x\partial_x\cdots\partial_x}^{n\text{ times}} \langle \nu; \mu \rangle_\epsilon = \langle \partial_x^n \nu; \mu \rangle_\epsilon \,,$$

where ∂_x is the partial derivative with respect to x. The analogous results for the gradient and laplacian in three dimensions, when $\nu = \nu(\vec{r})$, are $\boldsymbol{\nabla} \langle \nu; \mu \rangle_\epsilon = \langle \boldsymbol{\nabla}\nu; \mu \rangle_\epsilon$ and $\nabla^2 \langle \nu; \mu \rangle_\epsilon = \left\langle \nabla^2\nu; \mu \right\rangle_\epsilon$.

1. Relationships between $\boldsymbol{\Phi}$, $\mathbf{E}$, and $\boldsymbol{\varrho}$

Because spatial averaging commutes with derivatives, it would follow from defining macroscopic fields as spatial averages of their microscopic counterparts that the relationships between $\boldsymbol{\Phi}$, $\mathbf{E}$, and $\boldsymbol{\varrho}$ are the same as those between ϕ, $\mathcal{E}$, and ρ, i.e., $\mathbf{E} = -\boldsymbol{\nabla}\boldsymbol{\Phi}$ and $\boldsymbol{\varrho} = -\epsilon_0\nabla^2\boldsymbol{\Phi} = \epsilon_0\boldsymbol{\nabla}\cdot\mathbf{E}$. In Sec. VI we will identify $\boldsymbol{\Phi}$, $\mathbf{E}$, and $\boldsymbol{\varrho}$ with spatial averages of ϕ, $\mathcal{E}$, and ρ, respectively, but we will see that their definitions are a bit more complicated than, for example, $\boldsymbol{\Phi} \equiv \langle \phi; \mu \rangle_\epsilon$ for some averaging kernel μ and some width parameter ϵ. Nevertheless, they are spatial averages and the homogenization transformation is linear and the macroscopic counterpart of the derivative $\nu^{(1)}$ of ν is the derivative $\boldsymbol{\mathcal{V}}^{(1)}$ of the macroscopic counterpart of ν. Therefore homogenization does not create new fields, $\mathbf{P}$ and $\mathbf{D}$. Calculating $\boldsymbol{\Phi}$ and $\mathbf{E}$ from $\boldsymbol{\varrho}$ should be equivalent to first calculating ϕ and $\mathcal{E}$ from ρ and then spatially averaging them to find $\boldsymbol{\Phi}$ and $\mathbf{E}$, respectively.

C. Symmetry is scale-dependent

It appears to follow from the fact that derivatives and spatial averaging commute that symmetry is scale-dependent. For example, a crystal with microstructure ρ_{crystal} and a glass with microstructure ρ_{glass} can have exactly the same bulk macrostructure $\boldsymbol{\varrho}$; and they usually do because $\boldsymbol{\varrho} = 0$ in the bulk of any stable electromagnetically-isolated material whose surfaces are not charged.

The superposition principle implies that the macroscopic field emanating from the bulk of the crystal can be expressed as

$$\begin{aligned}\mathbf{E}^{\text{bulk}}_{\text{crystal}} &= \hat{E}_\rho[\rho_{\text{crystal}}] = \left\langle \hat{\mathcal{E}}_\rho[\rho_{\text{crystal}}] \right\rangle = \hat{\mathcal{E}}_\rho[\langle \rho_{\text{crystal}} \rangle] \\ &= \hat{\mathcal{E}}_\rho[\boldsymbol{\varrho}] = \hat{\mathcal{E}}_\rho[\langle \rho_{\text{glass}} \rangle] = \hat{E}_\rho[\rho_{\text{glass}}] = \mathbf{E}^{\text{bulk}}_{\text{glass}}\end{aligned}$$

where $\hat{E}_\rho$ is a linear functional of ρ, which satisfies $\hat{E}_\rho[\rho] = \hat{\mathcal{E}}_\rho[\boldsymbol{\varrho}]$, and for simplicity I am denoting the spatial average of each field ν simply as $\langle \nu \rangle$. It follows that neither a crystal's symmetry, nor any other characteristic of its microstructure that differs from the glass, alters the macroscopic electric field or the macroscopic electric potential emanating from its bulk.

The only symmetries that manifest at the macroscale are symmetries of the macrostructure.

D. Nonlinear relationships and response functions

It is important to note that the superposition principle applies only to *linear* physical systems. A more accurate way to state this is that a linear physical system is defined to be a system for which the superposition principle applies.

The superposition principle could not apply to all of the quantities $\alpha(x)$, $\beta(x)$ and $\gamma(x)$ if they were related by $\alpha(x) = \beta(x)\gamma(x)$ because β and γ are not linearly related. For example, if $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 = \beta_1\gamma_1$ and $\alpha_2 = \beta_2\gamma_2$, then

$$\alpha = \alpha_1 + \alpha_2 = \beta_1\gamma_1 + \beta_2\gamma_2 \neq (\beta_1 + \beta_2)(\gamma_1 + \gamma_2).$$

This has important implications for material-specific response parameters, such as conductivities, which are not simply spatial averages of their microscopic counterparts.

For example, if a material's macroscopic response to an arbitrarily-weak applied field $\mathbf{E}_{\text{ext}}$ is $\boldsymbol{\Delta\varrho}$, and if the change $\boldsymbol{\Delta\sigma}$ in its surface charge is the integral of $\boldsymbol{\Delta\varrho}$ across the surface, then $\boldsymbol{\Delta\sigma} \propto \mathbf{E}_{\text{ext}}$. However, the constant of proportionality does not have an analogue at the microscale to which it can be related by spatial averaging.

In fact, because the material is macroscopically uniform at equilibrium before and after $\mathbf{E}_{\text{ext}}$ is switched on, $\boldsymbol{\varrho}$ vanishes in the bulk both with and without $\mathbf{E}_{\text{ext}}$. Therefore the macroscopic response to $\mathbf{E}_{\text{ext}}$ is not a response of the bulk, but a change of the excesses of charge at surfaces and interfaces. At the microscale, by contrast, the response of the material's equilibrium microstructure to $\mathbf{E}_{\text{ext}}$ is a change of the charge density at surfaces, interfaces, and at all points in the bulk.

The macrostructure of the bulk is indistinguishable from vacuum and the macroscale response to a uniform field manifests only at surfaces and other macroscopic heterogeneities. Therefore it may be better to view a macroscopic response function, such as a conductivity or permittivity, as a property of a pair of surfaces, rather than as a property of the material occupying the space between them.

For example, a macroscopic response function $\boldsymbol{\chi}_{12}(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \dot{\boldsymbol{\sigma}}_1, \dot{\boldsymbol{\sigma}}_2, \cdots)$ for a pair of surfaces, whose charge excesses are $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$, might describe the (change in the) rate of change $\dot{\boldsymbol{\sigma}}_1 = -\dot{\boldsymbol{\sigma}}_2$ when a field is applied, the rate of energy dissipation during this change, etc.. The bulk composition and microstructure would help to determine $\boldsymbol{\chi}_{12}$, but so would the microstructures and compositions of both surfaces.

E. Macroscopic vector fields

As discussed in Sec. IV.C, and as Fig. 1 illustrates, symmetry is scale-dependent: A material whose microstructure ρ is highly inhomogeneous and anisotropic can have a bulk macrostructure with local continuous translational symmetry,

$$\boldsymbol{\varrho}(\boldsymbol{x} + \boldsymbol{u}) = \boldsymbol{\varrho}(\mathbf{x}), \;\; \forall\, \mathbf{x} \in \mathfrak{B} \;\text{ and }\; \forall\, \boldsymbol{u} : \boldsymbol{x} + \boldsymbol{u} \in \mathfrak{B}, \quad (7)$$

which implies local isotropy,

$$\boldsymbol{\varrho}(\boldsymbol{x} + \boldsymbol{u}) = \boldsymbol{\varrho}(\boldsymbol{x} - \boldsymbol{u}), \;\; \forall\, \mathbf{x}, \boldsymbol{u} : \mathfrak{I}(\mathbf{x}, 2|\boldsymbol{u}|) \subset \mathfrak{B}. \quad (8)$$

By *local* I mean that for any $\mathbf{x} \in \mathfrak{B}$ there are limits to the magnitudes of $\boldsymbol{u}$ for which Eqs. 7 and 8 hold. Eq. 7 holds for $\boldsymbol{u} \in (-|\mathbf{x} - \mathbf{x}_L|, |\mathbf{x} - \mathbf{x}_R|)$ and Eq. 8 holds for $|\boldsymbol{u}| < \boldsymbol{u}_{\max}(\mathbf{x}) \equiv \min\{|\mathbf{x} - \mathbf{x}_L|, |\mathbf{x} - \mathbf{x}_R|\}$.

As noted in Sec. IV.C, the material cannot be stable unless its bulk is charge-neutral ($\boldsymbol{\varrho} = 0$) on average. Therefore it is macroscopically uniform and locally isotropic at each point $\mathbf{x}$. It follows that any observable directionality at $\mathbf{x}$ must be a consequence of the inequivalence of $\boldsymbol{\varrho}$ at distances larger than $\boldsymbol{u}_{\max}(\mathbf{x})$ in the two directions. Therefore, a macroscopic vector field whose value at each point is a linear functional of $\boldsymbol{\varrho}$ (or ρ), cannot emanate from the bulk of a macroscopically uniform material because the bulk macrostructure cannot bestow directionality. This implies that $\mathbf{E}$ vanishes and $\boldsymbol{\Phi}$ is constant in the bulk.

Similarly, the existence of $\mathbf{P}$ is attributed to the bulk microstructure lacking inversion symmetry. However, regardless of the microstructure, the bulk macrostructure has inversion symmetry and a macroscopic $\mathbf{P}$ field appears incompatible with that.

1. Current density

There is one macroscopic vector field in a uniform material's bulk that is, in a certain sense at least, retained by the homogenized theory. This is the current density $\mathbf{J}$, which is not observable at the macroscale *in the uniform bulk*, but whose consequences, namely, the rates of change of accumulations of charge at macroscopic heterogeneities (surfaces, defects, etc.), are both observable and bestow directionality to it.

The rate of change of an accumulation of charge at a macroscopic defect or interface can be calculated from $-\boldsymbol{\nabla}\cdot\mathbf{J}$ or from the difference $\Delta\mathbf{J}$ in the values of $\mathbf{J}$ on either side of an interface. Therefore, instead of defining $\mathbf{J}$ at each point in the bulk, at the macroscale we need only concern ourselves with $\Delta\mathbf{J} = \dot{\boldsymbol{\sigma}}$ or $-\boldsymbol{\nabla}\cdot\mathbf{J}|\mathbf{dx}| = \dot{\boldsymbol{\mathcal{Q}}}$, which are the rates of change of the excesses of charge at an interface and at a point, respectively.

It is not possible, in general, to deduce the direction of $\mathbf{J}$ from the rate of change of areal charge density $\boldsymbol{\sigma}$ at an interface. However if, in a closed system, one knows the rates of change of accumulations of charge on *every* source of macroscopic heterogeneity, one could calculate the magnitude and direction of the current density everywhere in the uniform regions surrounding and separating these heterogeneities. This is the essence of circuit theory.

My argument in this section is that symmetry is scale dependent and that, regardless of a material's microstructure, no directionality should be observable at the macroscale if the macrostructure is isotropic. $\mathbf{J}$ should not be regarded as either invalidating this argument or as an exception to this principle because $\mathbf{J}$ is not observable at the macroscale. Only its consequences, such as $\dot{\boldsymbol{\sigma}}$ and the rate of change of temperature $\dot{\mathbf{T}}$, are.

F. Mean inner potential, $\boldsymbol{\Phi}$

The arguments of the previous section, and the linear relationship between the mean inner potential $\boldsymbol{\Phi}$ and ρ and $\boldsymbol{\varrho}$, imply that $\boldsymbol{\Phi}$ vanishes if a material's surfaces are not charged. This contradicts a great deal of existing literature, including textbooks and many recent research articles (Bethe, 1928; Blumenthal *et al.*, 2017; Cendagorta and Ichiye, 2015; Gajdardziska-Josifovska *et al.*, 1993; Hörmann *et al.*, 2019; Kathmann, 2021; Kathmann *et al.*, 2011; Leung, 2010; Madsen *et al.*, 2021; Miyake, 1940; Pratt, 1992; Sanchez and Ochando, 1985; Sokhan and Tildesley, 1997a; Spence, 1993, 1999; Wilson *et al.*, 1987, 1988, 1989; Yesibolati *et al.*, 2020). However, those works often assume that Φ and the average potential *experienced by a particle* moving through the bulk are the same quantity, or approximate the latter as the former.

Even if a particle spends a very long time in a material, it does not sample space uniformly. It samples regions of positive electric potential, where the electron density is high, more than regions of negative potential. Furthermore, an electron is a *perturbing* probe of electric potential. It does not sample a material's equilibrium microscopic potential $\phi(\vec{r})$ because its presence at position $\vec{r}$ reduces the probability density of other electrons being in a neighbourhood of $\vec{r}$, which reduces the negative potential it experiences from other electrons.

Therefore if the spatial average of ϕ is zero, meaning that a non-perturbing probe would measure an average microscopic potential of zero if it sampled the entire space within a material uniformly, the average potential experienced by an electron would be positive. For the same reason, the average potential experienced by a diffusing cation is negative because it attracts electrons to it and repels nuclei and other cations.

The superposition principle helps to understand why $\boldsymbol{\Phi}$ vanishes in a material's bulk ($\Rightarrow -\boldsymbol{\nabla}\boldsymbol{\Phi} = 0$). It means that the potential emanating from any nucleus is the sum of the potentials emanating from its constituent protons. Therefore the spatial average of the potential inside a charge-neutral material can be expressed as the sum of the spatial averages of the potentials emanating from point particles of charge $+e$ (protons) and point particles of charge $-e$ (electrons). The spatial average of the potential from an electron is the negative of the spatial average of the potential from a proton. Therefore, since there are equal numbers of protons and electrons in a charge-neutral material, $\boldsymbol{\Phi}$ can be expressed as a sum of vanishing contributions from proton-electron pairs.

G. Bulk polarization, $\mathbf{P}$

In this section I examine some consequences of assuming that $\mathbf{P}$ can be expressed as some functional $\hat{P}_\rho$ of ρ. I do so under the assumption that $\mathbf{E} \equiv \hat{E}_\rho[\rho]$ and $\boldsymbol{\varrho} \equiv \hat{\varrho}_\rho[\rho]$ are both *linear* functionals of ρ. If they were not linear, either the superposition principle would not apply at the macroscale, or the $\rho \mapsto \boldsymbol{\varrho}$ homogenization transformation would not conserve net charge. They are linear if $\mathbf{E}$ and $\boldsymbol{\varrho}$ are spatial averages of $\mathcal{E}$ and ρ, respectively.

A polarization field is believed to exist in any crystal whose microstructure lacks inversion symmetry. Therefore $\hat{P}_\rho[\rho]$ must be a nonlinear functional because the superposition $\rho = \rho_1 + \rho_2$ of two inversion-symmetric crystal structures does not have inversion symmetry, in general. If $\hat{P}_\rho[\rho_1] = 0$, $\hat{P}_\rho[\rho_2] = 0$, and $\hat{P}_\rho[\rho] \neq 0$, then $\hat{P}_\rho[\rho_1 + \rho_2] \neq \hat{P}_\rho[\rho_1] + \hat{P}_\rho[\rho_2]$.

Let us assume that $\mathbf{D} \equiv \hat{D}_\rho[\rho] = \hat{D}^l_\rho[\rho] + \hat{D}^{nl}_\rho[\rho]$ and $\mathbf{P} \equiv \hat{P}_\rho[\rho] = \hat{P}^l_\rho[\rho] + \hat{P}^{nl}_\rho[\rho]$, where the superscripts 'l' and 'nl' identify linear and nonlinear parts of the functionals, respectively. It follows from the linearity of $\hat{E}_\rho$ and the relation $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} \Rightarrow \hat{D}_\rho = \epsilon_0\hat{E}_\rho + \hat{P}_\rho$ that $\hat{D}^{nl}_\rho = \hat{P}^{nl}_\rho$ and $\epsilon_0\hat{E}_\rho = \hat{D}^l_\rho - \hat{P}^l_\rho$.

Any microscopic charge density ρ can be written as a (possibly infinite) sum of inversion-symmetric charge densities; the Fourier series of a periodic ρ being one example. It follows that $\mathbf{P}^l \equiv \hat{P}^l_\rho[\rho] = 0$ for any ρ, that the linear part of $\mathbf{D}$ is $\mathbf{D}^l \equiv \hat{D}^l_\rho[\rho] = \epsilon_0\mathbf{E}$, and that its nonlinear part $\mathbf{D}^{nl}$ is simply $\mathbf{P}$.

With these constraints the set of relationships that constitute the macroscale theory of electricity can be expressed as

$$\varrho = \epsilon_0 \boldsymbol{\nabla}\cdot\mathbf{E}, \qquad \boldsymbol{\nabla}\cdot\mathbf{J} + \dot{\varrho} = 0$$
$$\boldsymbol{\nabla}\times\mathbf{H} = \mathbf{J} + \epsilon_0\dot{\mathbf{E}}, \qquad \boldsymbol{\nabla}\times\mathbf{E} = -\dot{\mathbf{B}}$$

where $\mathbf{J} \equiv \mathbf{J}^{(c)} + \dot{\mathbf{P}}$, $\mathbf{J}^{(c)}$ is the *conduction current*, and a dot denotes a partial time derivative at fixed position. The linear and nonlinear parts of $\dot{\mathbf{D}}$ appear separately in these equations as $\epsilon_0\dot{\mathbf{E}}$ and $\dot{\mathbf{P}}$, respectively, and the only purpose served by $\mathbf{P}$ is to define its time derivative, $\mathbf{J}^{(p)} = \dot{\mathbf{P}}$.

Let us consider two ways to proceed from here. The first is to follow convention by finding a way to define $\mathbf{P}$. This would lead to definitions of $\mathbf{D}$, ϱ^{bound}, ϱ^{free}, $\mathbf{J}^{\text{bound}}$, and $\mathbf{J}^{\text{free}}$, where $\mathbf{J}^{\text{bound}} \equiv \mathbf{J}^{(p)}$ and $\mathbf{J}^{\text{free}} \equiv \mathbf{J}^{(c)}$ are currents of bound charge density (ϱ^{bound}) and free charge density (ϱ^{free}), respectively. I denote them by $\mathbf{J}^{(p)}$ and $\mathbf{J}^{(c)}$ to avoid distinguishing between free and bound charges. Introducing these six fields to the theory does not make it any more predictive or useful. Furthermore, $\mathbf{P}$ and $\mathbf{D}$ are not observable; $\mathbf{J}^{\text{free}}$ is only observable when $\mathbf{J}^{\text{bound}} = 0$ and vice versa; and free charge is only observable where the net bound charge vanishes and vice versa.

A much simpler and less conventional way to proceed is to *not* introduce any unobservable quantities into the theory, but to find a way to calculate $\mathbf{J}^{(p)}$. We do not need to distinguish between different contributions to $\mathbf{J}$ in either Maxwell's equations or the contintuity equation if we are not distinguishing between ϱ^{free} and ϱ^{bound}. Therefore $\mathbf{J}^{(c)} + \dot{\mathbf{P}}$ can be replaced by $\mathbf{J}$ with the understanding that $\mathbf{J}$ is the net flow of charge from *all* mechanisms.

We are left with only the four equations above, which are identical in form to their counterparts at the microscale, and in which all three electrical quantities that appear in them (ϱ, $\mathbf{J}$, $\mathbf{E}$) are observables with clear and intuitive meanings. There is no downside to scrapping $\mathbf{P}$ and $\mathbf{D}$ and it circumvents many problems, such as the fact that we do not have a definition of $\mathbf{P}$, and the fact that, because $\mathbf{P}$ is nonlinear and $\mathbf{E}$ is linear, $\mathbf{P} \neq \epsilon_0\boldsymbol{\chi}\mathbf{E}$.

It is important to note that a key premise or conclusion of the *Modern Theory of Polarization* is that my central premise in this section, namely that $\mathbf{P}$ can be calculated from ρ, is false (Resta and Vanderbilt, 2007). It is claimed, instead, that $\mathbf{P}$ is a property of the phase θ of the material's wavefunction $\Psi = \sqrt{p}\, e^{i\theta}$. I will discuss this claim in Sec. V and Sec. XI.

H. Polarization current as a demand of anisotropy

Fig. 2 shows several stages in the evolution of the equilibrium charge density ρ in the bulks of three crystals as some stimulus ζ is applied uniformly to them. The stimulus might be a change in temperature, a strain, a displacement of one of the crystal's sublattices relative to the others, or anything else that changes a crystal's charge density uniformly throughout the bulk.

The charge density in Fig. 2 (c) has inversion symmetry, with two centers of inversion in each primitive unit cell, and it maintains this inversion symmetry as ζ changes and $\rho(x;\zeta)$ evolves. Clearly, the motion of charge relative to one of its centers of symmetry must be the same in the $-\hat{x}$ and $+\hat{x}$ directions. Therefore the existence of a net current is prohibited by symmetry. On the other hand, the only symmetries possessed by the charge densities in Figs. 2(a) and 2(b) are their periodicities. It is impossible that the motion of charge in the $-\hat{x}$ and $+\hat{x}$ directions is equitable, because those directions are inequivalent and two numbers cannot be equal to infinite precision except by reason of symmetry.

To tighten this argument let us consider the center of electron charge, C, of the six unit cell segment shown in Fig. 2(b). Let $\dot{c}(x_b)$ be the rate of change of the center of charge of the n_e electrons in the primitive cell of width a whose boundary is at x_b. There is no symmetry reason to expect $\dot{c}(x_b)$ to vanish; therefore it does not. The rate of change of C is the average of the rates of change of the centers of charge of the six cells, and therefore is also equal to $\dot{c}(x_b)$. The current per unit length associated with this motion of electrons is $-M_e\, e\, \dot{c}(x_b)/a$. To find the total current density the contribution from nuclei should be added to it.

Let us now turn to the material in Fig. 2 (c). However, instead of expressing $\dot{C}_e$ in terms of the rates of change of the centers of electron charge of the six complete cells delimited by vertical black lines, we can express it as the average over the five complete green-bordered cells and the remaining cell, which is shaded in grey and divided into two pieces on either side of the five. The centers of charge of the green cells are time-invariant, by symmetry, which implies that $\dot{C} = \dot{c}/6$, where c is the center of electron charge of the grey cell. Now suppose that the number of primitive cells in the segment was not six, but of order l/a, such that the segment was mesoscopic. The rate of change of its center of electron charge, and therefore the current, would be of order a/l and therefore negligible.

Just as a polarization current flows when a material without inversion symmetry is uniformly stimulated, it also flows when the crystal has inversion symmetry, but the stimulus that changes ρ breaks this symmetry. This is the case when, for example, the stimulus is an applied electric field or a non-uniform strain (flexoelectricity).

Crystals can be categorized based on what sorts of stimuli are capable of causing these transient polarization currents. An electric field induces a polarization current in any crystal, whereas only crystals that lack inversion symmetry tend to be pyroelectric, because temperature does not reduce a crystal's symmetry, in general, and so they remain inversion symmetric as they are heated. A larger set of crystal symmetries are compatible with piezoelectricity because uniform uniaxial strain can break inversion symmetry.

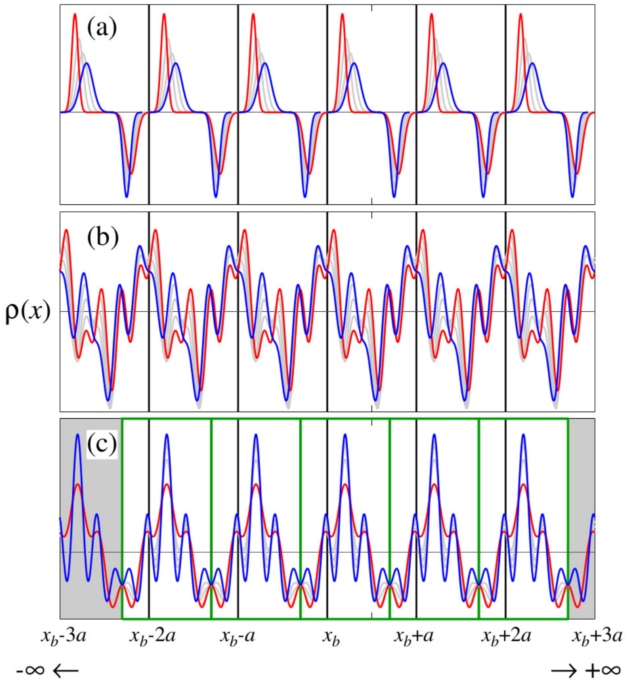

FIG. 2 Charge density ρ as function of position x in three one dimensional crystals. The crystals in (a) and (b) lack inversion symmetry, but the crystal in (c) has inversion symmetry, with two inversion centers per primitive cell Ω. In (c) a choice of primitive unit cell whose dipole moment is zero is outlined in green. If $\rho(x)$ changes continuously and uniformly between the densities plotted in red and blue, a macroscopic current flows in crystals (a) and (b) because the probability that the net movement of charge in direction $\hat{x}$ equals the net movement of charge in the inequivalent direction $-\hat{x}$ is zero. In (c) the symmetry of the crystal forbids a macroscopic flow of charge because the net movement of charge relative to an inversion center cannot differ between the two equivalent directions $\hat{x}$ and $-\hat{x}$. One way to see this is to note that in (c) most of space can be tiled with unit cells Ω whose dipole moment remains zero *throughout* the changing of the density. There remains only the two shaded regions of combined width $a = |\Omega|$ at the left and right boundaries of the chunk of bulk crystal comprised of $N_\Omega = 6$ primitive cells. In the limit of large N_Ω the change in the distance between x_b and the center of charge of the N_Ω cells, divided by their combined width $N_\Omega a$, vanishes. In (a) and (b) the current cannot vanish because the $\hat{x}$ and $-\hat{x}$ directions are inequivalent. An important question, which the MTOP solved for quantum systems, is how the current can be calculated from an evolving bulk microstructure, i.e., without knowing or calculating how much charge accumulates at surfaces. If the integrals q_1 and q_2 of the two peaks per unit cell in (a) remain constant, the current per unit length is simply $(q_1\dot{x}_1 + q_2\dot{x}_2)/a$, where $\dot{x}_1$ and $\dot{x}_2$ are the velocities with which the peaks move. However if the charge density is not organized into packets of fixed charge, as in (b), the definition of current is much less obvious.

V. MODERN THEORY OF POLARIZATION

A. Introduction

The *Modern Theory of Polarization* (MTOP), which was developed in the 1990s by Resta and Vanderbilt and their collaborators (King-Smith and Vanderbilt, 1993; Resta, 1993, 1994; Resta and Vanderbilt, 2007; Vanderbilt and King-Smith, 1993), solved the problem of how to define the polarization current $\mathbf{J}^{(p)}$ of an insulating material in terms of the evolving electronic structure *of its bulk*, and provided a widely-applicable method of calculating it. I emphasize that the MTOP defines $\mathbf{J}^{(p)}$ in terms of the changing *bulk* microstructure because the alternative, namely defining it as $\mathbf{J}^{(p)} = \dot{\boldsymbol{\sigma}}$, would be of much less practical use: calculations of it would require knowledge of the time-dependent surface microstructure, making accurate calculations of it intractable in many or most cases due to the number of electrons in the surface region. By contrast, the MTOP can be used to calculate the electronic contribution to $\dot{\boldsymbol{\sigma}}$ from a simulation that takes explicit account of only the electrons in a single primitive unit cell in the bulk of a crystal. Therefore the method can be applied to a wide range of materials and has a low computational overhead. In Sec. V I present an explanation of the MTOP definition of polarization current that is based on the relation $\mathbf{J} = \dot{\boldsymbol{\sigma}}$ and on Finnis's expression for the (macroscale) surface excess of charge, $\boldsymbol{\sigma}$, in terms of the microscopic densities of positive and negative charge. This perspective makes it clear that the MTOP definition of polarization current, in its original form, is basically exact.

To clarify what I mean by the MTOP providing an *exact* definition of $\mathbf{J}^{(p)}$, let us suppose that ζ is one or more parameters on which a material's electron density n has a continuous dependence. Let us suppose that this density is known exactly as $\zeta(t)$ changes slowly, and that it has a ζ-dependent decomposition into packets $n(x;\zeta) = \sum_i n_i(x;\zeta)$ such that each packet $n_i(\zeta)$ evolves smoothly with ζ and each packet's integral $-q_i/e$ is in-

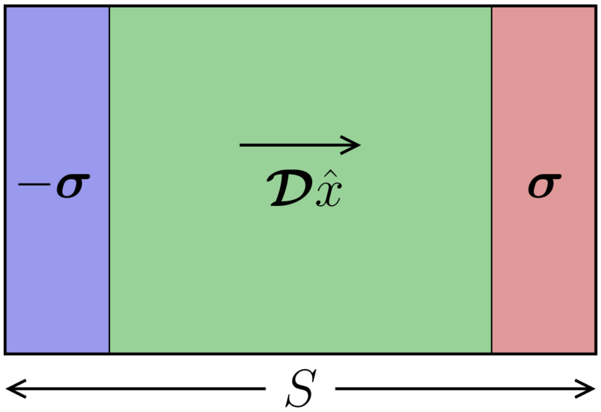

FIG. 3 See Sec. V.

dependent of ζ. Then the polarization current density in the bulk is given *exactly* by

$$\mathbf{J}^{(p)} = \frac{1}{|V|}\sum_{x_i\in V} q_i\dot{x}_i = \frac{\dot\zeta}{|V|}\frac{\mathrm{d}}{\mathrm{d}\zeta}\left(\sum_{x_i\in V} q_i x_i\right), \tag{9}$$

where the sum is over all x_i within a region V that is representative of the bulk. If the material is a crystal, V can be a single primitive unit cell, Ω. In an amorphous system V must be large enough to sample all relevant features of the microstructure and, by expanding it to improve the sampling of the bulk, $\mathbf{J}^{(p)}$ can be calculated to any desired precision.

The physical idea underpinning Eq. 9 is simple, and easiest to understand for the isolated nonconducting crystal depicted in Fig. 3. $\boldsymbol{\mathcal{D}}\hat{x}$ is the crystal's dipole moment and there are two planar surfaces perpendicular to $\hat{x}$ that carry excesses of charge. The width ε_x of the surface regions is exaggerated. In reality it would be many orders of magnitude smaller than the width of the crystal, S. Therefore $S_{\mathfrak{B}} \equiv |\mathfrak{B}| \approx S$. Between the two charged surfaces, the crystal's shape and cross-sectional area $|\mathfrak{P}|$ is uniform.

Now let us suppose that the electron density in the bulk can be expressed as a superposition of N_Ω identical electron densities $n_\Omega(\vec{r};\zeta)$, such that each primitive unit cell Ω contains the center of exactly one of them, i.e., $n(\vec{r};\zeta) = \sum_i n_\Omega(\vec{r}-\vec{A}_i;\zeta)$, $\forall \vec{r}\in\mathfrak{B}$, where the sum is over all lattice vectors $\vec{A}_i$ for which the center of $n_\Omega(\vec{r}-\vec{A}_i;\zeta)$ is in the bulk, $\mathfrak{B}$. The charge neutrality of the crystal's bulk implies that ρ_Ω, defined as the sum of the delta charge distributions of the nuclei in a particular primitive cell and the electron charge density $-e\,n_\Omega$ whose center is in the same cell, is a charge neutral distribution with a well defined (origin independent) dipole moment d_Ω. When the sum of all N_Ω of the distributions ρ_Ω centered in different cells is subtracted from the total charge density ρ, the result is zero in the bulk and its integrals along $\hat{x}$ at the two surfaces are σ_s and $-\sigma_s$, which are not, in general, equal to $\boldsymbol\sigma$ and $-\boldsymbol\sigma$, respectively (see Sec. X). If, for simplicity, we assume that the widths of all surfaces parallel to $\hat{x}$ can be neglected, the volume of the crystal is $S|\mathfrak{P}| = (S_{\mathfrak{B}}+S_s)\,|\mathfrak{P}| = N_\Omega|\Omega| + S_s|\mathfrak{P}|$, where S_s is the sum of the widths of the surface regions normal to $\hat{x}$, and $S_s|\mathfrak{P}|/(N_\Omega|\Omega|) \sim a/S < a/L \ll a/l \ll 1$. The crystal's dipole moment is $\boldsymbol{\mathcal{D}} = N_\Omega d_\Omega + \sigma_s|\mathfrak{P}|(N_\Omega|\Omega|/|\mathfrak{P}| + S_s)$ (its direction $\hat{x}$ is assumed not to change and is omitted) and the spatial average, over the entire crystal, of the current density flowing within it is

$$\begin{aligned}\mathbf{J}^{(p)} = \dot{\boldsymbol\sigma} = \dot{\boldsymbol{\mathcal{P}}} &= \frac{\dot{\boldsymbol{\mathcal{D}}}}{N_\Omega\Omega + S_s|\mathfrak{P}|}\\ &= \frac{N_\Omega\dot{d}_\Omega + \dot\sigma_s(N_\Omega|\Omega| + S_s|\mathfrak{P}|)}{N_\Omega|\Omega| + S_s|\mathfrak{P}|}\\ &= \frac{\dot{d}_\Omega}{|\Omega|} + \dot\sigma_s + \mathcal{O}(a/S) = \frac{\dot{d}_\Omega}{|\Omega|} + \mathcal{O}(a/S),\end{aligned}$$

where $\boldsymbol{\mathcal{P}} \equiv \boldsymbol{\mathcal{D}}/(N_\Omega|\Omega| + S_s|\mathfrak{P}|)$ is the crystal's dipole moment per unit of its volume, and $\dot\sigma_s$ vanishes because there is no conduction current and the crystal is isolated. The term $\mathcal{O}(a/S)$ on the right hand side is negligible and neglecting it leads to the simple expression $\mathbf{J}^{(p)} = \dot{d}_\Omega/|\Omega|$ for the bulk polarization current, which, in one dimension, is equivalent to Eq. 9 when the set $\{(x_i,q_i)\}$ is the union of two subsets: one containing the positions and charges of the nuclei, and the other containing the centers and $-e$ times the integrals of a set of packets of electron density whose sum is n_Ω and whose integrals do not change as ζ changes. Note that this derivation also neglects contributions to $\dot{\boldsymbol{\mathcal{D}}}$ from redistributions of charge within the surface regions because changes of surface polarization should not be counted as contributing to the bulk polarization current.

I describe the MTOP definition of $\mathbf{J}^{(p)}$ as 'exact' to emphasize that the partitioning of the electron density into packets (n_Ω) does not constitute an approximation. Furthermore, although the integral of each packet tends to be either one or two in applications of the MTOP, it is clear from the derivation of Eq. 9 in sections X.F and XI, which is based on the homogenization theory of Secs. VI and VII, that any method of partitioning the electron density into moving packets *of fixed charge* would give the same result, regardless of whether or not the packets' charges are integer multiples of $-e$. The question is: how can $n(\zeta)$ be partitioned into packets that evolve smoothly as ζ change, but whose integrals are independent of ζ? It is not clear that this is always possible for a distribution of any species of particle evolving smoothly via any physical process. However, we do know, from Kohn-Sham theory, that it is possible for the ground state density of electrons in an insulator when the external potential from the nuclei and externally-applied fields, $v^{\mathrm{ext}}(\zeta)$, varies smoothly with ζ (Kohn and Sham, 1965; van Leeuwen, 2003).

The MTOP approach is to use the Kohn-Sham construction (Kohn and Sham, 1965), whereby $n(\zeta)$ is expressed as a sum of the densities of the integer-occupied orthogonal eigenstates of an independent-electron Hamiltonian $\hat{h}(\zeta)$. In an insulator, the N_e-dimensional Hilbert space spanned by the occupied eigenstates, $\mathcal{H}_\zeta$, changes smoothly with ζ, i.e., it rotates within the infinite-dimensional Hilbert-Lebesgue space, $\mathcal{L}^2(\mathbb{R}^{\mathfrak{d}N_e})$, of which it is a subspace, where $\mathfrak{d}$ is the dimensionality of the system. There is a charge of $-e$ associated with each of the real space eigenfunctions $\varphi_i(\vec{r};\zeta) \equiv \langle \vec{r}|\varphi_i(\zeta)\rangle$ of $\hat{h}(\zeta)$ and their centers, $\vec{r}_i$, move during this rotation. Therefore a current flows as ζ changes, which is the polarization current. The value of the electronic contribution to $\mathbf{J}^{(p)}$ is given, in one dimension, by Eq. 9 with $q_i = -e$. Its generalization to three dimensions is trivial, but the remainder of this section will focus on the 1-d case and so $\varphi_i = \varphi_i(x;\zeta)$ and the center of φ_i will be denoted by x_i instead of $\vec{r}_i$.

As discussed in Sec. XI and the caption of Fig. 13, it is the space $\mathcal{H}_\zeta$ that determines $n(\zeta)$, not its basis of eigenstates. Therefore we are free to use a different basis to calculate $\mathbf{J}^{(p)}$ and it is common to transform the set $\{\varphi_i(\zeta)\}$ to one for which the density packets $\{|\varphi_i(\zeta)|^2\}$ are highly localized. However, this step is unnecessary, in principle: the centers of the eigenstates can be used directly in Eq. 9.

An alternative is to Fourier transform the eigenfunctions and to compute x_i as the expectation value of the Fourier transformed position operator $\hat{x}^{(k)} \equiv i\,\partial/\partial k$. This is equivalent to Fourier transforming $\hat{h}(\zeta)$ and finding the centers of the eigenfunctions of its Fourier transform, $\mathfrak{F}\left[\hat{h}(\zeta)\right]$. For example, the center of $n_i(x) = |\varphi_i(x)|^2$, when $\langle\varphi_i|\varphi_i\rangle = 1$, is

$$\begin{aligned} x_i &\equiv \int \mathrm{d}x\, n_i(x)x = \int \mathrm{d}x\, \varphi_i^*(x)\left(x\varphi_i(x)\right) \\ &= \frac{1}{2\pi}\int \mathrm{d}x \int \mathrm{d}k\, e^{ikx}\mathfrak{F}\left[x\varphi_i\right](k)\int \mathrm{d}q\, \tilde{\varphi}_i^*(k+q)e^{-i(k+q)x} \\ &= \frac{1}{2\pi}\int \mathrm{d}k\, i\frac{\partial\tilde{\varphi}_i(k)}{\partial k}\int \mathrm{d}q\left(\int \mathrm{d}x\, e^{-iqx}\right)\tilde{\varphi}_i^*(k+q) \\ &= \int \mathrm{d}k\, \tilde{\varphi}_i^*(k)\left(i\frac{\partial}{\partial k}\right)\tilde{\varphi}_i(k) = \int \mathrm{d}k\, \tilde{\varphi}_i^*(k)\,\hat{x}^{(k)}\,\tilde{\varphi}_i(k). \end{aligned}$$

Despite my use of suggestive notation, there is no quantum mechanical content in what I have just done. It is a purely mathematical procedure. Any classical probability density $n_i(x)$ can be expressed as the square modulus of a Hilbert space vector $\sqrt{n_i(x)}e^{i\theta(x)}$; and this is not only a valid way to represent a classically-evolving probability density function, it is necessary to represent it this way, in general, to ensure that it is a smooth function of time and position (see Appendix B).

The phase $\theta(x)$ is often irrelevant and arbitrary when studying the stationary states of an isolated system, but becomes important when dividing a system into subsystems (e.g., nuclei and electrons and/or surface and bulk) or when studying the time dependence of a nonequilibrium probability density.

B. The bulk subsystem

In theoretical solid state physics it is common to study a material's bulk subsystem, meaning the material without its surfaces, by placing it in a torus. For a 1-d material the bulk would be represented in a 1-torus $\mathbb{T}$, whose circumference is the width, $S_{\mathfrak{B}}$, of the material's bulk. Representing bulk wavefunctions in $\mathbb{T}$ is equivalent to enforcing periodic *Born-von Kármán* boundary conditions on wavefunctions in $\mathbb{R}$ (Born and von Kármán, 1912). In textbooks on solid state physics these conditions are usually discussed in the context of crystalline materials, but they are also commonly used for amorphous materials. The bulks of amorphous materials can be represented in $\mathbb{T}$ by ensuring that $S_{\mathfrak{B}}$ is large enough to contain a representative sample of the material, where 'representative' means that the sample's microstructure shares, or approximates, whatever statistical characteristics of the true material are relevant to the properties of the material that are being studied.

Suppose that we have found an orthonormal set of N_e states $\psi_i(x;\zeta)$, which vary smoothly with ζ, and such that the material's electron density as a function of ζ is given *everywhere* (bulk and surfaces) by $\sum_{i=1}^{N_e}|\psi_i(x;\zeta)|^2$. Then we can use Eq. 9 to calculate the average of $\mathbf{J}^{(p)}$ over the entire material and we can define any number of eigenoperators $\hat{\mathfrak{h}}(\zeta) \equiv \sum_{i=1}^{N_e}\epsilon_i\,|\psi_i(\zeta)\rangle\langle\psi_i(\zeta)|$ for this basis and density, with each eigenoperator corresponding to a particular choice of the N_e eigenvalues $\{\epsilon_i\}$. The operator in question is usually a single-electron Hamiltonian, and I will refer to it as such, but this is not necessary. When it is a single-electron Hamiltonian, the eigenvalues are energies (dimensionally) and the N_e eigenstates $|\psi_i(\zeta)\rangle$ contributing to the density are those with the lowest eigenvalues. I make no assumption about the eigenvalues, but I order the eigenstates such that it is those with indices (i) between 1 and N_e that contribute to the density.

Let us assume that the material's surfaces do not carry any net charge, and let us replace $\hat{\mathfrak{h}}$, which is the single-electron Hamiltonian of the material in the position (x) basis, with a Hamiltonian $\hat{h}$ that is bulk-like everywhere in space, and which becomes equal to $\hat{\mathfrak{h}}$ in the center of the material in the limit $S_{\mathfrak{B}} \to \infty$. Therefore the new Hamiltonian approximates the true Hamiltonian in the bulk, but differs greatly from it at surfaces and outside the material, where it remains bulk-like. Next we assume that each eigenfunction $\psi_i(x)$ of $\hat{\mathfrak{h}}$ can be expressed as $\psi_i(x) = \varphi_i(x)e^{i\theta_i}w(x) + \Delta\psi_i(x)$, for some x-independent θ_i, where $\varphi_i(x)$ is an eigenfunction of $\hat{h}$, $\Delta\psi_i(x)$ is negligible in the bulk, and $w(x) > 0$ is almost one in the bulk and drops rapidly to zero beyond the bulk. Therefore, if the material was a crystal, the eigenfunction's density $|\psi_i(x)|^2$ would be almost periodic in the bulk, it would deviate from periodicity near surfaces, and vanish outside the crystal.

We are now in a position to move from $\mathbb{R}$ to $\mathbb{T}$. We do so by enforcing the Born-von Kármán condition that φ_i has $S_{\mathfrak{B}}$-periodicity.

1. Calculating $\mathbf{J}^{(p)}$ in $\mathbb{T}$

Let us assume that we have found a set of N_e $S_{\mathfrak{B}}$-periodic orthonormal functions $\varphi_i(\zeta) \in \mathcal{L}^2(\mathbb{R})$ such that the density in $\mathbb{T}$ is $\sum_{i=1}^{N_e} |\varphi_i(x;\zeta)|^2$. The $S_{\mathfrak{B}}$-periodicity of φ_i means that we can express it as the Fourier series $\varphi_i(x;\zeta) \equiv \sum_{g\in 2\pi\mathbb{Z}/S_{\mathfrak{B}}} \tilde{\varphi}_i(g;\zeta)e^{igx}$, where $\tilde{\varphi}(g;\zeta) \equiv S_{\mathfrak{B}}^{-1}\int_0^{S_{\mathfrak{B}}} \varphi_i(x;\zeta)e^{-igx}\,\mathrm{d}x$. Let us normalize each φ_i to one in $\mathbb{T}$, i.e., $\int_{\mathbb{T}} |\varphi_i(x;\zeta)|^2\,\mathrm{d}x = 1$, where $\int_{\mathbb{T}} \equiv \int_0^{S_{\mathfrak{B}}}$. This implies that the normalization of $\tilde{\varphi}_i$ is $\sum_g |\tilde{\varphi}_i(g;\zeta)|^2 = 1/S_{\mathfrak{B}}$. Now we can express the center of $\varphi_i(\zeta)$ in $\mathbb{T}$ as

$$\begin{aligned} x_i(\zeta) &\equiv \int_{\mathbb{T}} x\,|\varphi_i(x;\zeta)|^2\,\mathrm{d}x = \sum_g \tilde{\varphi}_i^*(g;\zeta)\sum_{g'}\tilde{\varphi}_i(g+g';\zeta)\int_{\mathbb{T}} xe^{ig'x}\,\mathrm{d}x \\ &= -i\sum_g \tilde{\varphi}_i^*(g;\zeta)\sum_{g'}\tilde{\varphi}_i(g+g';\zeta)\left[\lim_{h\to 0}\frac{1}{h}\int_{\mathbb{T}}\left(e^{i(g'+h)x} - e^{ig'x}\right)\mathrm{d}x\right] = -i\sum_g \tilde{\varphi}_i^*(g;\zeta)\lim_{h\to 0}\left(\frac{\tilde{\varphi}_i(g-h;\zeta) - \tilde{\varphi}_i(g;\zeta)}{h}\right), \end{aligned}$$

where I have used the fact that the integrals of $\exp\{ig'x\}$ and $\exp\{i(g'+h)x\}$ vanish unless $g' = 0$ and $g' = -h$, respectively. Defining $h \equiv 2\pi/S_{\mathfrak{B}}$ allows us to take the $h \to 0$ limit by taking the $S_{\mathfrak{B}} \to \infty$ limit, and also ensures that the sum includes the value $g' = -h$. In this limit the sum over g becomes an integral over the set $\mathbb{R}^*$ of all real wavenumbers, i.e., $\sum_g \to (S_{\mathfrak{B}}/2\pi)\int_{\mathbb{R}^*}\mathrm{d}g$, and we get

$$\begin{aligned} x_i(\zeta) &= i\frac{S_{\mathfrak{B}}}{2\pi}\int_{\mathbb{R}^*}\tilde{\varphi}_i^*(g;\zeta)\partial_g\tilde{\varphi}_i(g;\zeta)\,\mathrm{d}g \\ \Rightarrow \frac{\mathrm{d}x_i}{\mathrm{d}\zeta} &= i\frac{S_{\mathfrak{B}}}{2\pi}\int_{\mathbb{R}^*}\Big[\partial_\zeta\tilde{\varphi}_i^*(g;\zeta)\partial_g\tilde{\varphi}_i(g;\zeta) \\ &\qquad\qquad + \tilde{\varphi}_i^*(g;\zeta)\partial_\zeta\partial_g\tilde{\varphi}_i(g;\zeta)\Big]\,\mathrm{d}g \\ &= -\frac{S_{\mathfrak{B}}}{2\pi}\times 2\,\mathrm{Im}\left\{\int_{\mathbb{R}^*}\partial_\zeta\tilde{\varphi}_i^*(g;\zeta)\partial_g\tilde{\varphi}_i(g;\zeta)\,\mathrm{d}g\right\} \end{aligned}$$

where I have integrated by parts and used the fact that $\lim_{g\to\infty}\varphi_i(\pm g;\zeta) = 0$ to reach the final expression. If we substitute this into Eq. 9, we find that the polarization current density is

$$\mathbf{J}^{(p)} = \frac{e\dot{\zeta}}{\pi}\sum_i \mathrm{Im}\left\{\int_{\mathbb{R}^*}\partial_\zeta\tilde{\varphi}_i^*(g;\zeta)\partial_g\tilde{\varphi}_i(g;\zeta)\,\mathrm{d}g\right\}. \quad (10)$$

Now let us assume that the material represented in $\mathbb{T}$ is a crystal, which means that $\hat{h}(\zeta)$ is $|\Omega|$-periodic as well as being $S_{\mathfrak{B}}$-periodic. Therefore, let us express the Fourier series expansion of φ_i as a nested sum over reciprocal lattice vectors $G \in 2\pi\mathbb{Z}/|\Omega|$ and vectors in the set $\Omega^* \equiv \{k \in 2\pi\mathbb{Z}/S_{\mathfrak{B}} : 0 \leq k < 2\pi/|\Omega|\}$, i.e.,

$$\begin{aligned} \varphi_i(x;\zeta) &= \sum_{k\in\Omega^*} e^{ikx}\sum_{G\in 2\pi\mathbb{Z}/|\Omega|}\tilde{\varphi}_i(G+k;\zeta)e^{iGx} && (11) \\ &= \sum_{k\in\Omega^*} e^{ikx}u_i(k,x;\zeta) = \sum_{k\in\Omega^*} b_i(k,x;\zeta) && (12) \end{aligned}$$

where $u_i(k,x;\zeta) \equiv \sum_G \tilde{\varphi}_i(G+k;\zeta)e^{iGx}$ is a Fourier series and therefore $|\Omega|$-periodic, and $b_i(k,x;\zeta) \equiv e^{ikx}u_i(k,x;\zeta)$. Now, because $\hat{h}(\zeta)$ is $|\Omega|$-periodic, if $\varphi_i(x;\zeta)$ is an eigenfunction of $\hat{h}(\zeta)$, then $\varphi_i(x - m|\Omega|;\zeta)$, where $m \in \{1,\cdots,N_\Omega - 1\}$, is also an eigenfunction of $\hat{h}(\zeta)$ with the same eigenvalue, $\epsilon_i(\zeta)$. Therefore, using the $|\Omega|$-periodicity of u_i, we can write

$$\begin{aligned} \hat{h}(\zeta)\varphi_i(x - m|\Omega|;\zeta) &= \sum_{k\in\Omega^*} e^{-ikm|\Omega|}\hat{h}(\zeta)\left(e^{ikx}u_i(k,x;\zeta)\right) \\ &= \epsilon_i(\zeta)\sum_{k\in\Omega^*} e^{-ikm|\Omega|}e^{ikx}u_i(k,x;\zeta) \\ \Rightarrow \sum_{k\in\Omega^*} e^{-ikm|\Omega|}&\Big[\hat{h}(\zeta)b_i(k,x;\zeta) - \epsilon_i(\zeta)b_i(k,x;\zeta)\Big] = 0 \end{aligned}$$

If we multiply this by $e^{ik'm|\Omega|}$, where $k' \in \Omega^*$, and sum over all N_Ω values of m between 0 and $N_\Omega - 1$, we find that only the $k = k'$ term survives and that if $b_i(k,x;\zeta) \equiv e^{ikx}u_i(x,k;\zeta)$ does not vanish everywhere in $\mathbb{T}$, it is an eigenfunction of $\hat{h}(\zeta)$ with eigenvalue $\epsilon_i(\zeta)$. Furthermore, the eigenfunctions b_i that do not vanish everywhere must all have the same eigenvalue ϵ_i. Otherwise φ_i would not have eigenvalue ϵ_i. Eigenvalues are only degenerate when symmetry demands it, and there is nothing in our construction that demands that $b_i(k+\Delta k, x; \zeta)$ and $b_i(k, x; \zeta)$ are related by symmetry for $k+\Delta k, \Delta k \in \Omega^* \setminus \{0\}$. Therefore, in general, they have different eigenvalues and, in general, only one term contributes to the sum in Eq. 12. Therefore, by assuming only that each eigenfunction φ_i of $\hat{h}$ is still an eigenfunction of $\hat{h}$ after it has been translated by a lattice vector, we have found that nondegenerate eigenfunctions of a $|\Omega|$-periodic Hamiltonian in $\mathbb{T}$ have the form $b_i(k, x; \zeta) \equiv e^{ikx} u_i(x, k; \zeta)$ for a particular value of $k \in \Omega^*$.

Note that there is some flexibility in how we define u_i. For example, given any reciprocal lattice vector G_0, we can use the definition of u_i as a Fourier series to write

$$\begin{aligned} b_i(k+G_0, x; \zeta) &= e^{i(k+G_0)x} u_i(k+G_0, x; \zeta) \\ &= e^{i(k+G_0)x} \sum_{G \in 2\pi\mathbb{Z}/|\Omega|} \tilde{\varphi}(G+k+G_0; \zeta) e^{iGx} \\ &= e^{ikx} \sum_{G \in 2\pi\mathbb{Z}/|\Omega|} \tilde{\varphi}(G+G_0+k; \zeta) e^{i(G+G_0)x} \\ &= e^{ikx} \sum_{\tilde{G} \in 2\pi\mathbb{Z}/|\Omega|} \tilde{\varphi}(\tilde{G}+k; \zeta) e^{i\tilde{G}x}, \end{aligned}$$

where $\tilde{G} = G_0 + G$. Therefore, we are to free to choose to define u_i such that $e^{iGx} u_i(k+G, x; \zeta) = u_i(k, x; \zeta)$ for any reciprocal lattice vector G. If we make this choice, the functions b_i are periodic in reciprocal space, i.e.,

$$b_i(k+G, x; \zeta) = b_i(k, x; \zeta) = e^{ikx} u_i(k, x; \zeta), \; \forall G \in \frac{2\pi\mathbb{Z}}{|\Omega|}$$

Because each nondegenerate eigenfunction b_i is associated with one particular value of $k \in \Omega^*$, it makes sense to replace the label i and the functional dependence of b_i on k with a composite label αk, where k is one of the N_Ω elements of Ω^* and α is an index that runs over all eigenstates associated with the same value of k. In other words, the eigenfunctions of $\hat{h}$ can be expressed as

$$b_{\alpha k}(x; \zeta) = e^{ikx} u_{\alpha k}(x; \zeta) \tag{13}$$

where $u_{\alpha k}$ is $|\Omega|$-periodic and $b_{\alpha k}(x; \zeta)$ may be chosen to be $|\Omega^*|$-periodic. Functions of this form are known as *Bloch functions* and are discussed in most textbooks on condensed matter physics.

Let us now return to Eq. 10 and replace $\sum_i$ with $\sum_\alpha \sum_{k \in \Omega*}$ and $\tilde{\varphi}_i$ with the Fourier transform $\tilde{b}_{\alpha k}$ of $b_{\alpha k}$. Eq. 13 implies that $\tilde{b}_{\alpha k}(g) = \tilde{u}_{\alpha k}(g-k)$ and the $|\Omega|$-periodicity of $u_{\alpha k}$ means that this vanishes unless $g - k = G$ for some reciprocal lattice vector $G \in 2\pi\mathbb{Z}/|\Omega|$. Therefore, if we express $\int_{\mathbb{R}^*} \mathrm{d}g$ as $\sum_G \int_{\Omega^*} \mathrm{d}k'$, the only nonvanishing term in the sum over k is $k = k'$. Putting all of these ingredients together, we find that

$$\mathbf{J}^{(p)} = \frac{e\dot{\zeta}}{\pi} \sum_\alpha^{\text{occ}} \text{Im} \left\{ \int_{\Omega^*} \mathrm{d}k \sum_G \partial_\zeta \tilde{u}^*_{\alpha k}(G; \zeta) \partial_k \tilde{u}_{\alpha k}(G; \zeta) \right\},$$

where the sum over α is restricted to 'occupied' states (those that contribute to $n(x; \zeta)$). This can also be expressed in the form

$$\mathbf{J}^{(p)} = \frac{e\dot{\zeta}}{\pi} \sum_\alpha^{\text{occ}} \text{Im} \left\{ \int_{\Omega^*} \mathrm{d}k \, \langle \partial_\zeta u_{\alpha k} | \partial_k u_{\alpha k} \rangle \right\}. \tag{14}$$

Eq. 14 is one of the most commonly-quoted forms of the MTOP definition of $\mathbf{J}^{(p)}$.

I emphasize, again, that I have not made any use of quantum mechanics in this derivation or anywhere in Sec. V. I assumed that the density could be represented as $n(x; \zeta) = \sum_i |\varphi_i(x; \zeta)|^2$, where the functions φ_i are mutually orthogonal and normalized to one. However they could also have different normalizations as long as $n(x; \zeta)$ was normalized to N_e and the normalization of each φ_i did not vary with ζ. Such a representation of n is possible for at least some classically-generated probability densities. For example, when the nuclei are treated as classical particles, their delta distribution can be represented in this way. All of what followed from the assumption that a representation of this form existed was simply mathematics in a torus with a 'Hamiltonian' defined by $\hat{h}(\zeta) \equiv \langle x | \left(\sum_i |\varphi_i(\zeta)\rangle\langle\varphi_i(\zeta)| \right) | x \rangle$.

The MTOP can be applied to a system of classical identical particles in an evolving steady state $n(x; \zeta)$ that can be expressed as a sum of densities from orthogonal functions φ_i whose normalizations do not change as ζ changes. Mathematically, this means that the MTOP is applicable whenever n can be expressed as a sum of contributions from eigenstates of a self-adjoint one-particle operator, such as the solutions of any Sturm-Liouville equation, and when the subset of eigenstates contributing to n does not change suddenly as ζ changes; rather, each contributing eigenstate must evolve smoothly with ζ, while preserving its normalization.

C. Interpretation of the MTOP

I have derived the MTOP expression for $\mathbf{J}^{(p)}$ via a very different route to the one originally used to derive it, and I have shown that it is more widely applicable than previously believed. Previous derivations have used quantum mechanical perturbation theory, which is nothing more than a Taylor expansion of a microstate about a stationary state (i.e., an eigenstate) of a Hermitian operator. The applicability of such an expansion is not restricted to quantum mechanical systems.

An alternative derivation of Eq. 9, based on the homogenization theory proposed in Secs. VI and VII, is presented and discussed in Sec. X.F and Sec. XI.

Although I have concluded that the MTOP definition of polarization current is correct, and exact, I am not satisfied that the MTOP solves the problem discussed in Sec. III and I do not agree with one interpretation of the MTOP; namely, that polarization is a fundamentally quantum mechanical phenomenon (Resta, 1993, 1994) and a property of the phase of a material's wavefunction.

The wavefunction of an isolated material at equilibrium does not have a relevant phase. It is real, apart from an arbitrary and irrelevant constant phase factor. Therefore if $\mathbf{P}$ were *exclusively* a property of a wavefunction's phase, $\mathbf{P}$ would only exist in an isolated material when it was being observed. This particular strain of quantum weirdness may be palatable to some, but it should be noted that delving more deeply into it leads to difficult philosophical questions.

Note that, by applying the kinetic energy operator to a wavefunction $\Psi = \sqrt{p}e^{i\theta}$ whose phase $\theta(\{x_i\})$ is nontrivial, the phase can be shown to be equivalent to a potential $\propto \sum_i |\partial\theta/\partial x_i|^2$. Although this potential is positive, potentials are only defined up to a constant applied to the entire physical system of interest. This means that any positive or negative potential is either equivalent to, or arbitrarily close to, the combination of a phase and this irrelevant constant. Therefore, when studying the bulk subsystem, the interaction between the bulk and surface subsystems could be expressed as a phase. Then $\mathbf{J}^{(p)} = \dot{\boldsymbol{\sigma}}$ would manifest as a time-dependence of this phase, which is equivalent to a changing interaction with the surface. However this is not the rationale used to relate $\mathbf{J}^{(p)}$ to a phase in the MTOP.

Furthermore, it has sometimes been implied or stated that $\mathbf{P}$ 'is', rather than 'can be', a property of the phase and that this means that it does not have an analogue within classical physics. This interpretation of the mathematics cannot be correct because the only assumption that I made in Sec. V.B to derive Eq. 14, which is identical to the MTOP expression for $\mathbf{J}^{(p)}$, is that the electron density n can be represented by a sum of contributions from smoothly-evolving mutually-orthogonal functions in an $\mathcal{L}^2$ Hilbert-Lebesgue space.

The 'quantum' MTOP definition of $\mathbf{P}$ was developed to solve the problem discussed in Sec. III and the problem of how to calculate the currents demanded by asymmetry (Sec. IV.H). However, I did not invoke quantum mechanics to explain those problems because they are problems that exist within classical physics. This implies that, despite the development of the MTOP, there was more to understand because classical mechanics and classical statistical mechanics have never been derived from quantum mechanics. The *correspondence principle* is conjecture: Quantum mechanics is not known to be more general than classical statistical mechanics, it is assumed to be.

Classical physics is arguably no less internally-consistent than quantum physics. The problem with classical physics is not internal inconsistency but that, in its current state of development, it does not agree with experimental observations of very small or sensitive systems. Therefore, with the exception of a disagreement with experiment, any physical problem that can be stated within the classical realm, and which does not expose an internal inconsistency in classical physics, must have a solution within the classical realm.

The well-defined problem of how to calculate the polarization current from the changing charge densities in Fig. 2 is an excellent example. There exist classical processes that result in the equilibrium distribution of charges (e.g., ions in solution) changing. In most cases the charges are either too numerous or too sensitive to the act of observation to allow the net flow of charge to be calculated within classical deterministic mechanics, and one must turn to statistical mechanics. How can the current resulting from an adiabatic evolution of a classical equilibrium statistical state be calculated? The MTOP, as it was originally derived and interpreted, does not answer this question directly, but if a classically evolving charge density can be decomposed into packets whose integrals are constant, it answers it indirectly: $\mathbf{J}^{(p)}$ can be calculated from Eq. 9.

My approach to deriving Eq. 9 in Sections X.C and XI is conceptually very simple. In Secs. VI and VII I use a systematic and unbiased approach to structure homogenization to derive expressions for interfacial excesses and changes in macroscopic fields across interfaces. From these I derive an expression for $\boldsymbol{\sigma}$ in terms of the microstructure (a generalization of Finnis's expression), which leads to an expression for $\mathbf{J} \equiv \dot{\boldsymbol{\sigma}}$ that is slightly more general than the MTOP expression, but otherwise equivalent to it.

I do not find any reason to retain the concept of a $\mathbf{P}$ field, and it appears to violate macroscale symmetry. However, as I will discuss in Sec. X.E.1, if one chooses to retain it, and if one also demands that $\mathbf{P}$ determines the excess charge on a pristine surface of a perfect crystal, which is devoid of extrinsic charges, then, even within classical physics, $\mathbf{P}$ must be quantized in the same way as the MTOP prescribes.

VI. MACROSTRUCTURE AS HOMOGENIZED MICROSTRUCTURE

A. Introduction

In this section I present elements of a classical theory of *structure homogenization*, meaning a theory of how microstructures determine macrostructures, and of the resulting natures of macrostructures. In later sections I will use this theory to elucidate the relationships between the microscopic fields ϕ, $\mathcal{E}$, and ρ and their macroscopic counterparts $\mathbf{\Phi}$, $\mathbf{E}$, and $\boldsymbol{\varrho}$.

The most obvious and intuitive approach to deriving electromagnetic theory at the macroscale from Maxwell's vacuum theory is to define macroscopic fields as spatial averages of their microscopic counterparts. However, despite this having been tried many times and in many ways (Ashcroft and Mermin, 1976; Bethe, 1928; de Groot and Vlieger, 1965; de Groot and Vlieger, 1964; Jackson, 1998; Kamenetskii, 1998; Kirkwood, 1936, 1940; Mazur, 1957; Mazur and Nijboer, 1953; Miyake, 1940; Robinson, 1971; Russakoff, 1970; Schram, 1960; Vinogradov and Aivazyan, 1999; van Vleck, 1937; Wilson *et al.*, 1987), we lack a fundamental understanding of the relationship between microstructure and macrostructure.

There are two main reasons why previous attempts have not succeeded, or have not succeeded fully. The first is that more fields appear in Maxwell's macroscopic theory than appear in his vacuum theory. Therefore, for example, the fields $\mathbf{P}$ and $\mathbf{D}$ cannot be defined as spatial averages of their counterparts at the microscale because they do not have counterparts at the microscale.

The second reason is that, to define one field as the spatial average of another, it is necessary to introduce one or more parameters specifying the size and shape of the region of space that is averaged over, and the distribution of weights with which points in this region contribute to the average. The dependences of macroscopic fields on these parameters has been interpreted as a fatal flaw in their definitions as spatial averages.

However, this non-uniqueness should not be interpreted as a fatal flaw, but as intrinsic to the nature of macrostructure, which is not determined by microstructure alone. It is determined, in part, by the relationship between the observer and the microstructure: A macroscopic field is a microscopic field observed on a large length scale. What is observed depends on how large that scale is. Therefore any definition of a macroscopic field in terms of its microscopic counterpart must depend on a parameter that specifies it.

In addition to the scale on which the underlying microscopic field is observed, a macroscopic field depends on the perspective from which the microstructure is observed, the apparatus with which it is observed, and the fields that mediate the observation. Therefore, when defining reproducibly-measureable macrostructure it is necessary to choose which of these influences are incorporated into the definition, and which of them are left to observers as apparatus-specific corrections. For simplicity and generality, I will assume that the only parameter on which the definition of macrostructure depends is the smallest distance, $|\mathbf{dx}|$ across which changes in macroscopic fields are observable at the macroscale. All other observer-specific influences are left as corrections to be applied when a specific observation is compared to the theory built on this one-parameter definition.

Perhaps unsurprisingly, there is a trade-off between the precisions to which gradients of macroscopic fields are defined and the spatial precision at the macroscale, $|\mathbf{dx}|$. When the uncertainty in position is small, the uncertainty in the gradient of a macroscopic field is large, and vice-versa, with the product of these uncertainties being proportional to the uncertainty in the value of the macroscopic field itself. This uncertainty principle is derived in Sec. VI.I, but I begin discussing its consequences for the nature of macrostructure in Sec. VI.C, so that readers understand why defining a macroscopic field in terms of its microscopic counterpart is only one of two primary objectives of this structure homogenization theory. An equally important objective, which is arguably more important from a practical perspective, is to define macroscopic *excess fields*, such as surface charge densities, $\boldsymbol{\sigma}$.

I begin by outlining the simplifying assumptions that I make about the nature of the microstructure.

B. Assumed properties of the microstructure, ν

Here, and throughout this work, I will assume that the microstructure only varies significantly on the microscale and on the macroscale, and that the macroscale is many orders of magnitude larger than the microscale. For example, the microscale might be the nanometer scale, and the smallest distance across which variations of macroscopic fields are observed might be $|\mathbf{dx}| \sim 1\,\text{mm}$.

I will assume that the physical system of interest does not contain any material with a microstructure that varies significantly on intermediate length scales. For example, the microstructure of wood varies on every length scale between $\sim 1\,\text{nm}$ and the scale ($\sim 10\,\text{m}$) of the tree from which it was harvested (Toumpanaki *et al.*, 2021). When it is observed on any length scale in this ten order-of-magnitude range, its surface is observed to have texture. Therefore the theory presented herein does not apply to wood. It applies when texture is only observable on the microscale and on the macroscale. The texture observed on the macroscale is caused by *excess fields* at interfaces between regions with different microstructures, or by changes of the microstructure that occur gradually across macroscopic distances.

For simplicity, I will assume that the microstructure is specified by a single scalar field, $\nu : \mathbb{R}^n \to \mathbb{R}$, and I will denote its counterpart on the macroscale by $\boldsymbol{\mathcal{V}}$. Although the value of n is three, the results derived for the $n = 1$ case can be applied in three dimensions by defining $\nu(x)$ to be the average of a function of (x, y, z) on the $y - z$ plane at x. Therefore, to make discussions of the main physical ideas simpler, in this section I will mainly focus on the one dimensional microstructure $\nu : \mathbb{R} \to \mathbb{R}$.

I make the following assumption about how observations and measurements occur at the macroscale.

***Physical assumption* 1**: At the macroscale, all positions, distances, and displacements are deduced from measurements of $\boldsymbol{\mathcal{V}}$.

Generalizing the theory to the case of multiple pairs $(\nu_i, \boldsymbol{\mathcal{V}}_i)$ of microscopic fields and their macroscopic counterparts is straightforward. Therefore, in the context of electricity, this is the assumption that all measurements and observations at the macroscale are measurements and observations of the electric potential, $\boldsymbol{\Phi}$, its derivatives $\boldsymbol{\Phi^{(1)}} = -\mathbf{E}$ and $\boldsymbol{\Phi^{(2)}} = -\boldsymbol{\varrho}$, and/or their excesses (see Sec. VI.C.1 and Sec. VII).

A microstructure ν may fluctuate, to some degree, on every length scale. However, as discussed above, the theory proposed in this section applies under the assumption that ν may vary significantly on the macroscale L and on the microscale a, which are widely separated, but its variations on any intermediate scale, or *mesoscale*, l are negligible. Generalizing the theory to microstructures that vary on three or more widely-separated length scales appears straightforward, but generalizing it to microstructures that vary on all length scales does not.

To define the microscale and the macroscale more precisely, I introduce the distances $\mathfrak{a}$ and $\mathfrak{L}$. Roughly-speaking, $\mathfrak{L}$ is the largest distance such that nonlinearities in the variation of $\boldsymbol{\mathcal{V}}$ across distances less than $\mathfrak{L}$ are negligible; and $\mathfrak{a}$ is orders of magnitude smaller than $\mathfrak{L}$, but large enough that every interval in the set $\{\mathfrak{I}(x, \mathfrak{a}) : x \in \operatorname{dom} \nu\}$, whose center x is not in vacuum, contains many local extrema of ν.

These conditions do not define $\mathfrak{a}$ and $\mathfrak{L}$ uniquely, but it will not be necessary to define them uniquely. Their primary purpose will be to define a mesoscopic distance $|\mathbf{dx}|$, which is the macroscale infinitesimal, and is much larger than any reasonable choice of $\mathfrak{a}$ and smaller than any reasonable choice of $\mathfrak{L}$. At the microscale I will denote $|\mathbf{dx}|$ by ε_x.

I define the microscale, the mesoscale, and the macroscale in terms of $\mathfrak{a}$ and $\mathfrak{L}$ as follows:

$$\begin{aligned} \eta_a \sim a &\iff \eta_a < \mathfrak{a} \\ \eta_l \sim l &\iff \mathfrak{a} < \eta_l < \mathfrak{L} \\ \eta_L \sim L &\iff \eta_L > \mathfrak{L}. \end{aligned}$$

My physical assumption that all fluctuations of the microstructure whose wavelengths are between $\mathfrak{a}$ and $\mathfrak{L}$ have negligible amplitudes can be expressed in terms of the Fourier transform $\tilde{\nu}(k)$ of $\nu(x)$ as follows.

***Physical assumption* 2**:

$$\int_0^{k_L} |\tilde{\nu}(k)|^2 \, \mathrm{d}k \gg \int_{k_L}^{k_a} |\tilde{\nu}(k)|^2 \, \mathrm{d}k \ll \int_{k_a}^{\infty} |\tilde{\nu}(k)|^2 \, \mathrm{d}k \,,$$

where $k_L \equiv 2\pi/\mathfrak{L} \lll k_a \equiv 2\pi/\mathfrak{a}$.

Obviously, if this assumption holds true for particular values of $\mathfrak{a}$ and $\mathfrak{L}$, it also holds true for larger values of $\mathfrak{a}$ and smaller values of $\mathfrak{L}$, as long as $\mathfrak{a} < \mathfrak{L}$.

C. The nature of macrostructure

I now outline some important features of macrostructure, which are consequences of finite spatial precision at the macroscale.

At the macroscale we observe homogenized microstructure. Under the two simplifying assumptions presented in Sec. VI.B, the macrostructures of the bulks of materials are uniform, except, possibly, for the existence of macroscopic point, line, or locally-planar defects. If there are no such defects, all materials appear uniform (textureless) at the macroscale and we only perceive a difference between the bulks of two materials on either side of an interface via observable properties of their *microstructures*, such as their colours.

A material's colour is determined by the microscopic wavelengths of electromagnetic radiation with which its microstructure interacts. Since the wavevectors of this radiation are normal to the plane on which the spatial average that produces the observed macrostructure from the microstructure is performed, we observe the radiation, rather than the averages of its electric and magnetic fields along its axis of propagation, which would vanish.

We have assumed, via Physical Assumption 2, that differences in texture are not observable at the macroscale. Physical Assumption 1 implies that differences in colour are irrelevant to macroscale electricity. For example, colour differences cannot be detected as changes in the distributions of electric potential or charge at the macroscale. Therefore the only observable structure at the macroscale is the network of interfaces separating otherwise-indistinguishable regions of uniformity, and any macroscopic defects within these otherwise-uniform regions. Interfaces and other macroscopic heterogeneities are observable because, and only if, they carry observable excess fields.

1. Excess fields

At the macroscale, an excess field can be understood from Stokes' theorem. For example, the net charge within a material Ω is

$$\int_{\Omega} \boldsymbol{\varrho}(\mathbf{r})\,\mathrm{d}^3\mathbf{r} = \int_{\partial\Omega} \boldsymbol{\sigma}(\mathbf{s})\,\mathrm{d}^2\mathbf{s}\,, \tag{15}$$

where the integral on the right hand side is an integral of the areal charge density, $\boldsymbol{\sigma}$, on the material's surface, $\partial\Omega$. In this case, $\boldsymbol{\sigma}$ is an excess field: it is the *surface excess* of field $\boldsymbol{\varrho}$.

Stokes' theorem also holds at the microscale if point charges are treated as singularities and the sets of zero measure in which they reside are omitted from the integrals. However, in that case both integrals in the microscale analogue of Eq. 15 vanish, and we must take a different route to understand surface excesses.

A further complication is that surfaces are ill-defined at the microscale: A surface's microstructure differs, to some degrees, from the microstructures of both the bulk and the atmosphere or vacuum above the surface. Its difference with respect to the bulk microstructure lessens *gradually* with increasing depth below the surface, so there does not exist a depth below which the material is bulk-like and above which it is not. Therefore surfaces are ill-defined regions of indeterminate widths at the microscale.

The surface $\partial\Omega$ of material Ω that appears in Eq. 15 is well defined because, although the thickness of the surface region is indeterminate at the microscale, it is less than $\varepsilon_x \equiv |\mathbf{dx}|$. Therefore the depths of any two points within it differ by less than ε_x, which implies that, at the macroscale, the distance between the atmosphere outside a material and the material's bulk is $|\mathbf{dx}|$. Surfaces are well-defined at the macroscale because every curve that crosses the surface exactly once contains exactly one surface point. In other words, locally, surfaces are *literally* planar at the macroscale.

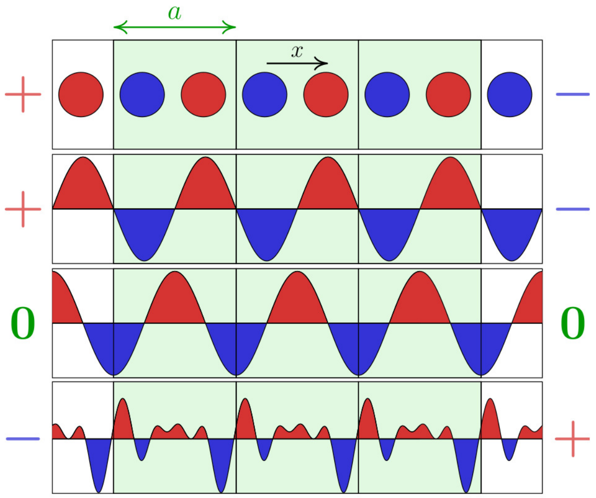

FIG. 4 *Excess fields*: Each of the four vertically-stacked panels is a schematic plot of the microscopic charge density $\rho(x)$ of a different one dimensional material, with positive and negative charge(s) coloured red and blue, respectively. If, at each value of x, we calculate the average, $\langle\rho\rangle_a(x)$, of all charge within a distance $a/2$ of x, we find that it vanishes everywhere in the green-shaded 'bulk' of each material, but is finite in the white surface regions. At each surface, the integral of $\langle\rho\rangle_a(x)$ over all points that are not in the bulk, but are within a distance a of it, is the surface's excess of charge, $\boldsymbol{\sigma}$. The symbols $+$, $-$, and 0 next to each surface indicate whether $\boldsymbol{\sigma}$ is positive, negative, or zero, respectively. The macroscopic analogue, $\boldsymbol{\varrho}$, of ρ is defined, to a finite precision ε_{ϱ}, as its mesoscale average $\bar{\rho}$. $\boldsymbol{\varrho}$ vanishes everywhere in the bulk, but not at interfaces, in general. Therefore, because spatial averaging conserves charge, the excess charge at an interface, $\boldsymbol{\sigma}$, is simply the integral of $\boldsymbol{\varrho}$ across it.

Figure 4 illustrates, from a microscopic perspective, why surfaces carry excess fields: The average of a microscopic charge density, ρ, vanishes in the green-shaded bulk of the one-dimensional materials depicted, but not at points near a surface, in general. For example, consider a point very close to the left-hand edge of one of the unshaded surface regions on the left-hand side. The macroscopic charge density, $\boldsymbol{\varrho}$, vanishes in the green-shaded bulk because the contribution to the spatial average at a point x from all points $\{x+u : 0<u<a/2\}$ to its right is cancelled exactly by the contribution from all points $\{x-u : 0<u<a/2\}$ to its left. However, this cancellation cannot happen at a point near the left-most edge of one of the surfaces on the left-hand side, because most of the points to its left are in vacuum, where ρ vanishes.

Therefore, $\boldsymbol{\varrho}$ vanishes in the green-shaded bulk, but does not necessarily vanish in the unshaded surface regions. Since the widths of these regions are much less than $\varepsilon_x = |\mathbf{dx}|$, each one corresponds to a single point at the macroscale. This one-dimensional example illustrates that $\boldsymbol{\varrho}$ vanishes everywhere in the bulk and it vanishes in the vacuum outside the material, but it is finite, in general, at a surface.

Therefore, in general, the value of a macroscopic field $\boldsymbol{\mathcal{V}}$ on any curve that intersects a surface or interface has a jump discontinuity or a removable discontinuity at the point of intersection, $\mathbf{x}_s$. The easiest way to deal with $\boldsymbol{\mathcal{V}}$ being discontinuous at $\mathbf{x}_s$ is by treating all surface points separately: they can be omitted from the domain of $\boldsymbol{\mathcal{V}}$ and an excess field, $\boldsymbol{\sigma}_{\mathcal{V}}(\mathbf{x}_s) \equiv \boldsymbol{\mathcal{V}}(\mathbf{x}_s)|\mathbf{dx}|$, can be defined

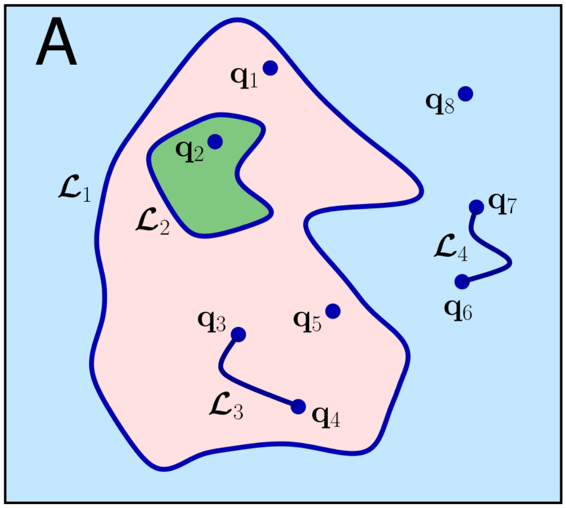

FIG. 5 *A 2D macrostructure*: An electrical microstructure in $A \subset \mathbb{R}^2$ is an areal charge density $\sigma : A \to \mathbb{R}; (y, z) \mapsto \sigma(y, z)$, which may be the excess of a volumetric charge density $\rho(x, y, z)$ on a surface or interface normal to the x-axis. An excess field (see Fig 4 and Sec. VII) is created at an interface whenever a higher dimensional microstructure is homogenized along an axis normal to the interface. The *macrostructure* of A consists of continua that are punctuated with, and separated by, subspaces s_i of dimensions zero or one, on which excess fields are defined. If s_i has dimension one its excess field is a linear charge density $\boldsymbol{\mathcal{L}}_i$; and if it has dimension zero its excess field is a point charge $\mathbf{q}^{(i)}$.

on each surface, interface, or locally-planar defect.

More generally, excess fields describe macroscopically-observable accumulations whose manifestations at the macroscale are best described by distributions whose domains are manifolds of dimension $n < 3$. For example, the excess of charge on the i^{th} surface or interface, which is a two dimensional manifold (or simply *2-manifold*), $\mathcal{M}_i^{\langle 2\rangle}$, is an areal density of charge, $\boldsymbol{\sigma}_i : \mathcal{M}_i^{\langle 2\rangle} \to \mathbb{R}$; the j^{th} linear charge density, $\boldsymbol{\mathcal{L}}_j : \mathcal{M}_j^{\langle 1\rangle} \to \mathbb{R}$, is a charge excess defined on a curve (the 1-manifold, $\mathcal{M}_j^{\langle 1\rangle}$); and the k^{th} macroscopic point charge, $\mathbf{q}_k : \mathcal{M}_k^{\langle 0\rangle} \to \mathbb{R}$, is a charge excess defined at a point (the 0-manifold, $\mathcal{M}_k^{\langle 0\rangle}$). The macrostructure arising from a two-dimensional microscopic charge density is illustrated in Fig. 5.

What this means is that, despite the microstructure being defined by a single field ν, there is more to the macrostructure than the single field $\boldsymbol{\mathcal{V}}$. The macrostructure comprises $\boldsymbol{\mathcal{V}}$, where $\operatorname{dom}\boldsymbol{\mathcal{V}} \equiv \mathbb{R}^3 \setminus \left(\{\mathcal{M}_i^{\langle 2\rangle}\} \cup \{\mathcal{M}_j^{\langle 1\rangle}\} \cup \{\mathcal{M}_k^{\langle 0\rangle}\}\right)$, and the set $\{\boldsymbol{\sigma}_{\mathcal{V},i}\} \cup \{\boldsymbol{\mathcal{L}}_{\mathcal{V},j}\} \cup \{\mathbf{q}_{\mathcal{V},k}\}$ of excess fields.

We will see that $\boldsymbol{\varrho}$ vanishes in the bulk of any stable material, and that $\boldsymbol{\Phi}$ and $\mathbf{E}$ vanish in a material's bulk if the material is isolated and its surfaces are charge-neutral. Therefore, defining a material's electrical macrostructure entails finding a way to define excess fields in terms of the microstructure, ν. Finnis solved this problem for periodic microstructures (Finnis, 1998), such as those plotted in Fig. 4, and I generalize his work to non-periodic microstructures in Sec. VII.

D. Assumed properties of $\boldsymbol{\mathcal{V}}$

Defining $\boldsymbol{\mathcal{V}}$ in terms of ν is trickier than it first appears, so it is useful to list and discuss the properties that $\boldsymbol{\mathcal{V}}$ is assumed to have. There are three of them, which I list below and discuss in Secs. VI.D.1, VI.D.3, and VI.D.2.

***Physical assumption* 3**: $\boldsymbol{\mathcal{V}}$ is reproducibly measurable.

***Physical assumption* 4**: ν fluctuates microscopically about $\boldsymbol{\mathcal{V}}(\mathbf{x})$ at each $\mathbf{x} \in \operatorname{dom}\boldsymbol{\mathcal{V}}$.

***Physical assumption* 5**: $\boldsymbol{\mathcal{V}}$ is differentiable, except, possibly, on a set of zero measure.

1. Reproducible measurability of $\boldsymbol{\mathcal{V}}$

$\boldsymbol{\mathcal{V}}$ is measurable by a blunt probe or as an average of the values of ν measured by many sharp probes whose locations cannot be controlled or known to microscopic precisions.

For example, when you look at a surface, light enters your eye from many closely-spaced points of the surface's microstructure, but each ray enters at a different angle and with a different intensity and a different frequency, in general. The contributions from all of the rays merge to produce an image of homogenized microstructure in your mind. This is the macrostructure. The merger that homogenizes the microstructure occurs in many stages, involves many different mechanisms, and occurs at many different locations along the path from the surface to your eye to your brain.

When I say that $\boldsymbol{\mathcal{V}}$ is *reproducibly* measurable, I mean that when a particular spatial resolution $|\mathbf{dx}|$ is chosen, and shared by all repeated measurements, the value of $\boldsymbol{\mathcal{V}}$ at each point can be defined independently of any measuring technique or apparatus. This means that, although a measured value of $\boldsymbol{\mathcal{V}}$ at a point always contains artefacts of the method used to measure it, if the magnitudes of these artefacts could be made sufficiently small, or if corrections could be applied to remove them, any two measurements of $\boldsymbol{\mathcal{V}}$ at the same point would both yield values that were *both* consistent with the microstructure ν, *and* with the definition of $\boldsymbol{\mathcal{V}}$ in terms of ν.

Despite the complexity of the processes that turn microstructure into macrostructure, I base the homogenization theory presented herein on the following assumption.

***Physical assumption* 6**: Any accurate measurement of $\boldsymbol{\mathcal{V}}$ at precisely the point $\mathbf{x}$ is a measurement of a weighted spatial average of ν on a mesoscopic domain centered at a point that is macroscopically-indistinguishable from $\mathbf{x}$, i.e., a point in the interval $\mathfrak{I}(\mathbf{x}, |\mathbf{dx}|)$.

2. Definition of "*fluctuates microscopically*"

When I say that ν *fluctuates microscopically*, or that ν is a *microscopic quantity* or a *microscopic function*, I mean that ν fluctuates on length scale a. This means that every extremum of ν is within a distance $\mathfrak{a}$ of another extremum of ν. Defining the statements "*ν fluctuates microscopically about υ at* $\mathbf{x}$" and "*ν fluctuates microscopically about $\boldsymbol{\mathcal{V}}$ at* $\mathbf{x}$" is more difficult, so I defer discussing them until Sec. VI.H.

3. Macroscale differentiability

The purpose of a macroscopic field theory is to describe changes over macroscopic distances. Therefore if $\boldsymbol{\mathcal{V}}$ fluctuated microscopically it would, effectively, be nondifferentiable. For example, if the value of $\boldsymbol{\Delta}\boldsymbol{\mathcal{V}}_1 \equiv \boldsymbol{\mathcal{V}}(\mathbf{x}+\boldsymbol{h}_l+\boldsymbol{h}_a) - \boldsymbol{\mathcal{V}}(\mathbf{x})$ differed significantly from the value of $\boldsymbol{\Delta}\boldsymbol{\mathcal{V}}_2 \equiv \boldsymbol{\mathcal{V}}(\mathbf{x}+\boldsymbol{h}_l) - \boldsymbol{\mathcal{V}}(\mathbf{x})$, when $|\mathbf{dx}| < |\boldsymbol{h}_l| < \mathfrak{L}$ and $|\boldsymbol{h}_a| < \mathfrak{a}$, then it would not be possible to approximate $\boldsymbol{\mathcal{V}}(\mathbf{x}+\boldsymbol{h}_l)$ with a truncated Taylor expansion of $\boldsymbol{\mathcal{V}}$ about $\mathbf{x}$ containing few terms.

$\boldsymbol{\mathcal{V}}$ being differentiable at the macroscale means that, given any point $\mathbf{x} \in \operatorname{dom}\boldsymbol{\mathcal{V}}$, the values of $\left(\boldsymbol{\mathcal{V}}(\mathbf{x}+\boldsymbol{h}) - \boldsymbol{\mathcal{V}}(\mathbf{x})\right)/\boldsymbol{h}$ and $\left(\boldsymbol{\mathcal{V}}(\mathbf{x}) - \boldsymbol{\mathcal{V}}(\mathbf{x}-\boldsymbol{h})\right)/\boldsymbol{h}$ are equal in the limit $|\boldsymbol{h}| \to |\mathbf{dx}|$, *to within the precisions to which they are defined.*

E. Spatial averages

I will define $\boldsymbol{\mathcal{V}}$ *in terms of* (not *as*) spatial averages of ν of the form

$$\langle \nu; \mu \rangle_\epsilon (x) \equiv \left(\nu * \mu(\epsilon)\right)(x) \equiv \int_{\mathbb{R}} \nu(x')\mu(x'-x;\epsilon)\,\mathrm{d}x' ,$$

where the parameter ϵ of the *averaging kernel*, $\mu(\epsilon)$, is twice its standard deviation, i.e.,

$$\int_{\mathbb{R}} u^2\,\mu(u;\epsilon)\,\mathrm{d}u = \left(\frac{\epsilon}{2}\right)^2 .$$

I will assume that $\mu(\epsilon)$ has three other properties for every value of ϵ. The first property is

$$\int_{\mathbb{R}} \mu(u;\epsilon)\,\mathrm{d}u = 1,$$

which implies that the homogenization of ν is conservative. The second property is

$$\int_{\mathbb{R}} u\,\mu(u;\epsilon)\,\mathrm{d}u = 0,$$

which implies that the value of $\langle \nu; \mu \rangle_\epsilon$ at x is a weighted average of ν from points whose weighted-average position is x. The third property is

$$\mu(u; s\epsilon) = s^{-1}\mu(u/s;\epsilon), \tag{16}$$

which simply means that the effect of changing ϵ is to scale μ without changing its shape or its integral.

These properties do not place strong or unphysical constraints on the form of $\mu(\epsilon)$, because any function with a well-defined mean and standard deviation can be normalized and translated to give a function whose integral is one and whose mean is zero. That function can be identified as $\mu(1)$ and Eq. 16 can be used to define the narrower or wider function $\mu(u;\epsilon) \equiv \mu(u/\epsilon;1)/\epsilon$.

The reason for giving μ a parametric dependence on its width is that it makes it easier to discuss separately the effects on $\langle \nu; \mu \rangle_\epsilon$ of varying the width and the shape of $\mu(\epsilon)$. For example, it follows from Eq. 16 that the n^{th} derivative of μ satisfies

$$\mu^{(n)}(su; s\epsilon) = \frac{\mu^{(n)}(u;\epsilon)}{s^{n+1}}. \tag{17}$$

Therefore, whereas the average magnitude of $\mu(\epsilon)$ scales as $1/\epsilon$, the average magnitude of its first derivative scales as $1/\epsilon^2$, and higher-order derivatives decay even faster as ϵ increases. An important implication of this is that the shape of μ has less of an influence on the value of $\langle \nu; \mu \rangle_\epsilon$ as ϵ increases.

Note that if $\mu(u;\epsilon)$ has all of the properties discussed above, then so does the function $\mu(-u;\epsilon)$. Therefore the general form of the spatial averages considered in this work may also be expressed as

$$\langle \nu; \mu \rangle_\epsilon (x) \equiv \int_{\mathbb{R}} \nu(x+u)\mu(u;\epsilon)\,\mathrm{d}u\,. \tag{18}$$

1. Schwartz and non-Schwartz averaging kernels

For some averaging kernels, $\mu(u;\epsilon)$, there exist values of m such that their rates of decay in the limits $u \to \pm\infty$ are slower than $1/|u|^m$. Other kernels, such as Gaussians, decay faster than any power law. A smooth function that decays faster than any power law is known as a *Schwartz function*, so I will refer to kernels of the first and second types as *non-Schwartz kernels* and *Schwartz kernels*, respectively.

Non-Schwartz kernels tend to describe relatively-direct and physical weightings of microstructures, such as the shape of a blunt probe or the decay of light intensity with distance. Schwartz kernels tend to arise from disorder and uncertainty: They describe limiting cases of homogenizing physical processes, such as the $N \to \infty$ limit of the combined effects of N homogenizing influences, each of which can be described by non-Schwartz kernels.

For example, the *central limit theorem* demonstrates how a Gaussian distribution arises from a very large number of contributions from independent random variables whose distributions do not necessarily decay faster than a power law (see, for example, Riley *et al.*, 2006, Chapter 30).

Non-Schwartz kernels make it easier to illustrate the complications that arise from defining macroscopic fields as spatial averages of their microscopic counterparts, and *top-hat kernels* are among the simplest of non-Schwartz kernels. Therefore I will introduce top-hat kernels in Sec. VI.E.2 and use them to illustrate an important consequence of Eq. 17; namely, the fact that $\langle \nu ; \mu \rangle_\epsilon$ depends less and less on the shape of μ as ϵ increases. Then, in Sec. VI.F, I will use top-hat kernels to illustrate why precision is finite at the macroscale.

2. Top-hat kernels

A very simple non-Schwartz kernel is the top-hat function,

$$\mu\left(u; \epsilon/\sqrt{3}\right) = T(u\,; \epsilon) \equiv \begin{cases} 0 & \text{if } |\,u\,| > \epsilon/2 \\ 1/(2\epsilon) & \text{if } |\,u\,| = \epsilon/2 \\ 1/\epsilon & \text{if } |\,u\,| < \epsilon/2. \end{cases} \qquad (19)$$

This function is discontinuous at $u = \pm\epsilon/2$, but I will sometimes use it when I require μ to be differentiable. In those cases I use it with the understanding that I am using a differentiable function that approximates the top-hat function arbitrarily closely.

I will refer to the average with a top-hat kernel as a *simple spatial average*, I will denote it by $\langle \nu \rangle_\epsilon (x)$, and although I will express it as

$$\langle \nu \rangle_\epsilon (x) \equiv \frac{1}{\epsilon} \int_{-\epsilon/2}^{\epsilon/2} \nu(x+u)\,\mathrm{d}u\,, \qquad (20)$$

I do so with the understanding that whenever $\langle \nu \rangle_\epsilon (x)$ is required to be a differentiable function of either x or ϵ, it is implied that the spatial average is defined by Eq. 18, with a top-hat kernel $\mu(\epsilon/\sqrt{3})$ whose corners are arbitrarily sharp, but differentiable.

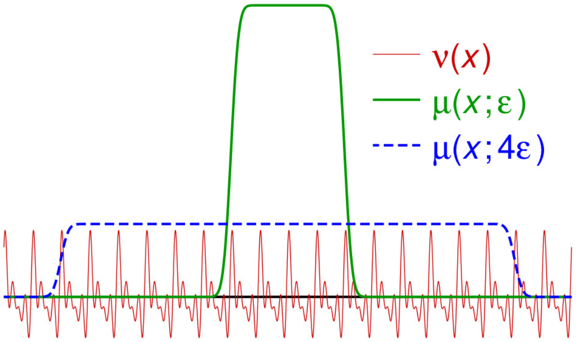

FIG. 6 A periodic function $\nu(x)$ and two smooth 'top-hat' averaging kernels, $\mu(\epsilon)$ and $\mu(4\epsilon)$, which differ only by the value of the parameter determining their widths. Increasing the kernel's width increases the number of periods of ν that contribute to the average. The derivative of μ is only nonzero at the edges. Therefore, the rate at which it decays to zero becomes less and less significant to the average of ν as μ is widened. This illustrates a result that applies to a much wider class of kernels than top-hat kernels: the average is independent of the kernel's shape in the $\epsilon \to \infty$ limit.

Figure 6 depicts a periodic microstructure and two differentiable top-hat averaging kernels that might be used to find its spatial average. One of the kernels is four times the width of the other, but both are almost constant almost everywhere: their derivatives are only finite near where they decay to zero.

The wider kernel averages four times more periods of the microstructure than the narrower kernel, but the weighting it applies to each one is smaller by a factor of four. Therefore, any difference between the averages calculated with the two kernels is a result of them applying different non-uniform weights to points near where they decay to zero.

As ϵ increases, the contribution to the average of points where the derivative $\mu^{(1)}(\epsilon)$ is non-negligible becomes smaller relative to the contribution from points where $\mu^{(1)}(\epsilon)$ almost vanishes. This illustrates the fact that the average magnitudes of derivatives of $\mu(\epsilon)$ decay faster as ϵ increases than the average of $\mu(\epsilon)$ does (Eq. 17). Therefore it illustrates the fact that the shape of μ becomes increasingly irrelevant to the value of $\langle \nu ; \mu \rangle_\epsilon (x)$ as ϵ increases.

In much of what follows I will assume that ϵ is large enough that the shape of μ is irrelevant and, for simplicity, I will only consider simple spatial averages.

F. Why $\boldsymbol{\mathcal{V}} \equiv \langle \nu \rangle_{\epsilon}$ fails as a definition

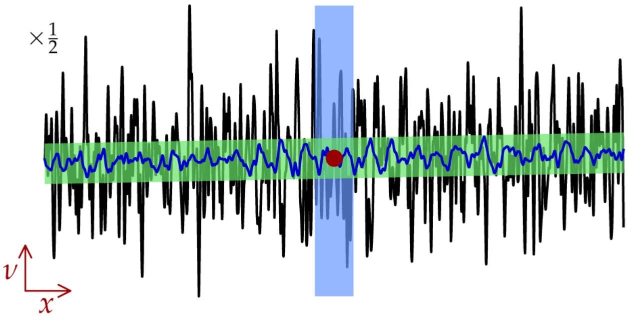

FIG. 7 The black curve is a microstructure, $\nu(x)$, after its fluctuations have been scaled by $\frac{1}{2}$ for visibility. The blue line within the green band is its simple spatial average $\langle \nu \rangle_{\epsilon}$ when ϵ is equal to the horizontal width of the vertical blue band. The red dot is the average of $\nu(x)$ over all points x within the blue band. The green band is the range of values that $\langle \nu \rangle_{\epsilon}$ takes on the segment of its domain shown. Increasing the width, ϵ, of the blue band reduces the width of the green band, but its width only vanishes in the limit $\epsilon \to \infty$.

Figure 7 is a plot of a microstructure ν and its spatial average, $\langle \nu \rangle_{\epsilon}(x)$. The range of values taken by $\langle \nu \rangle_{\epsilon}$ in a microscopic neighbourhood of the red dot is indicated by the almost-horizontal green band. As ϵ increases, the microscopic fluctuations of $\langle \nu \rangle_{\epsilon}(x)$, and therefore the width of the green band, reduce in magnitude as $1/\epsilon$ (Eq. 16). However they do not vanish. They vanish only in the limit $\epsilon \to \infty$, which is the limit in which the average is performed over the entirety of ν's domain. Therefore it is the limit in which $\langle \nu \rangle_{\epsilon}$ has the same value at every point, meaning that all structure has been lost.

Now let us assume that $\boldsymbol{\mathcal{V}}(x) \equiv \langle \nu \rangle_{\epsilon}(x)$, for some finite value of ϵ, so that the reasons why this definition fails become clear.

One reason why it fails is that two points x_1 and x_2, which are separated by a microscopic distance $|x_1 - x_2| < \mathfrak{a}$, would be indistinguishable at the macroscale. Measurements of $\boldsymbol{\mathcal{V}}(x_1)$ and $\boldsymbol{\mathcal{V}}(x_2)$ would differ, despite appearing to have been performed at the same macroscale point. Therefore $\boldsymbol{\mathcal{V}}$ is not reproducibly-measureable *at the macroscale*.

Another reason why it fails is that the finite difference derivative $\left(\boldsymbol{\mathcal{V}}(x+h) - \boldsymbol{\mathcal{V}}(x)\right)/h$ depends sensitively on x and h and fluctuates microscopically as a function of each one, as illustrated in Fig. 8. Therefore $\boldsymbol{\mathcal{V}}$ is not differentiable *at the macroscale*, because its derivative does not converge with respect to h while h is still macroscopic or mesoscopic. It does not converge until h is much smaller than the microscopic distances between successive extrema of $\boldsymbol{\mathcal{V}}$.

Both of these problems can be resolved by defining $\boldsymbol{\mathcal{V}}(\mathbf{x})$ to be the set of *all* values that would be measured at the same macroscale point, $\mathbf{x}$. This is a set of all spatial averages of ν centered at points in an interval whose width is the lower bound, $\varepsilon_x = |\mathbf{dx}|$ on distances that are observable at the macroscale.

Since $|\mathbf{dx}|$ is the limit of spatial precision at the macroscale, the most precise measurements of $\boldsymbol{\mathcal{V}}$ are either performed with microscopically-blunt probes (radii $\gtrsim \varepsilon_x$), or with sharper probes whose positions can only be controlled or known to within an interval of width ε_x.

Therefore if ε_x is large enough that $\langle \nu; \mu \rangle_{\varepsilon_x}$ is independent of the shape of μ, the set of all measured values of $\boldsymbol{\mathcal{V}}(\mathbf{x})$ is the set of values of $\langle \nu \rangle_{\varepsilon_x}(x)$ at microscale points x that are within an interval of width $|\mathbf{dx}|$ centered at $\mathbf{x}$. This set is an interval, $\mathfrak{I}(\overline{\boldsymbol{\mathcal{V}}}(\mathbf{x}), \varepsilon_{\mathcal{V}}(\mathbf{x}))$. Therefore $\boldsymbol{\mathcal{V}}(\mathbf{x})$ is only defined to a precision, $\varepsilon_{\mathcal{V}}(\mathbf{x})$, that is finite.

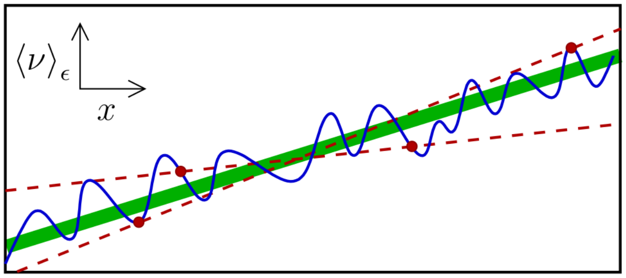

FIG. 8 The blue curve, which is the simple average, $\langle \nu \rangle_{\epsilon}$, of ν, fluctuates microscopically about the set of values plotted with a green band. The magnitudes of these fluctuations can be reduced by increasing ϵ, but no matter how small the fluctuations are, if they are finite, the finite-difference derivative of the blue curve differs from the slope of the green band, to some degree, for most choices of the two red points used to calculate it.

G. The macroscale infinitesimal, $|\mathbf{dx}|$

In this section I illustrate the fact that if $\boldsymbol{\mathcal{V}}$ is only defined to a finite precision, $\varepsilon_{\mathcal{V}}$, the value of $\varepsilon_{\mathcal{V}}$ imposes a lower bound on the macroscale spatial precision, $|\mathbf{dx}|$. I will then make the following assumption.

***Physical assumption* 7**: The only limit on spatial precision at the macroscale, $|\mathbf{dx}|$, is the limit imposed by the finite precision, $\varepsilon_{\mathcal{V}}$, to which $\boldsymbol{\mathcal{V}}$ can be defined.

In other words, I will neglect all other sources of spatial imprecision in order to isolate and investigate imprecisions and uncertainties that are intrinsic to acts of observation in which the observer inhabits a length scale that is orders of magnitude larger than $\mathfrak{a}$.

Figure 9 illustrates why measurements of $\boldsymbol{\mathcal{V}}$ cannot conclusively distinguish between $\mathbf{x}_1$ and $\mathbf{x}_2$ if

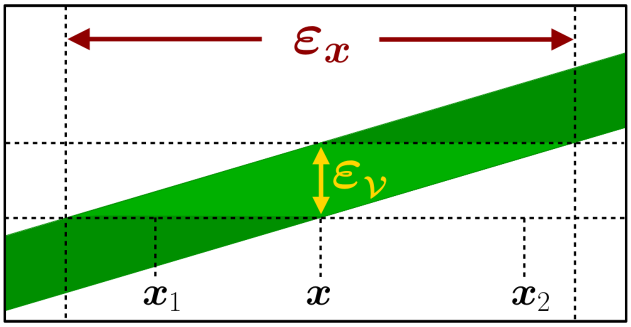

FIG. 9 Plot of $\boldsymbol{\mathcal{V}}$ versus $\mathbf{x}$ when the precision $\varepsilon_{\mathcal{V}}$ to which $\boldsymbol{\mathcal{V}}$ is defined is finite. $\varepsilon_{\mathcal{V}}$ is finite if the results of repeated accurate measurements of $\boldsymbol{\mathcal{V}}$ at the same macroscale point are not all equal, but only within $\varepsilon_{\mathcal{V}}$ of one another. It means that if the slope $\boldsymbol{\mathcal{V}}^{(1)}$ was known, measurements of $\boldsymbol{\mathcal{V}}$ could not be used to distinguish between two points, $\mathbf{x}_1$ and $\mathbf{x}_2$, reliably and conclusively, or to measure the distance between them. Therefore, when measurements and observations at the macroscale are mediated by macroscopic fields, there is an unavoidable imprecision $\varepsilon_x \equiv |\mathbf{dx}|$ in positions, distances, and displacements.

$|\mathbf{x}_1 - \mathbf{x}_2| < \varepsilon_{\mathcal{V}}/\left|\boldsymbol{\mathcal{V}}^{(1)}\right|$. In Sec. VI.I I will present a discussion, for an arbitrary choice of the averaging kernel μ, of the relationship between ε_x, $\varepsilon_{\mathcal{V}}$, and the precision ε_p to which the derivative $\boldsymbol{\mathcal{V}}^{(1)}$ of $\boldsymbol{\mathcal{V}}$ is defined.

The macroscale infinitesimal $|\mathbf{dx}|$ is the smallest distance between empirically-distinguishable points, i.e.,

$$|\mathbf{dx}| \equiv \inf\Big\{|\boldsymbol{\Delta}\mathbf{x}| : |\mathbf{x}_1 - \mathbf{x}_2| > |\boldsymbol{\Delta}\mathbf{x}|/2 \Rightarrow \mathbf{x}_1 \neq \mathbf{x}_2, \\ \forall \mathbf{x}_1, \mathbf{x}_2 \in \operatorname{dom} \boldsymbol{\mathcal{V}}\Big\}.$$

This definition implies that distances smaller than $|\mathbf{dx}|$ do not have meaning at the macroscale. However, they do have meaning at the microscale, where $|\mathbf{dx}|$ is denoted by ε_x. For simplicity, this definition also assumes that the value of $|\mathbf{dx}|$ is the same everywhere in dom $\boldsymbol{\mathcal{V}}$.

At the microscale, $|x_1 - x_2| < \varepsilon_x/2 = |\mathbf{dx}|/2$ does not imply that $x_1 = x_2$. Therefore each point $\mathbf{x}$ at the macroscale corresponds to an interval of width ε_x at the microscale. I denote the midpoint of this interval by $\bar{x}(\mathbf{x})$ and I refer to the interval as the *coincidence set* of $\bar{x}(\mathbf{x})$. Mathematically, it is defined as

$$[\bar{x}(\mathbf{x})]_{\overset{L}{\sim}} \equiv \left\{x : x \overset{L}{\sim} \bar{x}(\mathbf{x})\right\} = \mathfrak{I}(\bar{x}(\mathbf{x}), \varepsilon_x),$$

where *macroscale coincidence*, $\overset{L}{\sim}$, which is nontransitive and therefore not an equivalence, is defined by

$$x_1 \overset{L}{\sim} x_2 \iff |x_1 - x_2| < \varepsilon_x/2.$$

The one-to-many relationship between points at the macroscale and points at the microscale has important implications for the nature of macrostructure, which have already been discussed in Sec. VI.C. It implies that the transition from the microscale to the macroscale can be viewed as a compression of space, which shrinks all microscopic distances to zero, resulting in surfaces and interfaces becoming literally planar, locally.

H. Mutually-consistent values of $\varepsilon_{\mathcal{V}}$ and ε_x

It seems natural to say that ν fluctuates microscopically about its spatial average. However the dependence of $\langle\nu\rangle_\epsilon(x)$ on ϵ means that the spatial average of ν at x is not unique, and its dependence on x means that the sets of all spatial averages on intervals of width less than ϵ centered at different macroscopically-coincident points are different. In other words, if $|x_1 - x_2| > 0$, then, in general, and notwithstanding the fact that $x_1 \overset{L}{\sim} x_2$,

$$\left\{\langle\nu\rangle_\eta(x_1) : 0 < \eta < \epsilon\right\} \neq \left\{\langle\nu\rangle_\eta(x_2) : 0 < \eta < \epsilon\right\}.$$

As discussed in Sec. VI.F, the set of all possible accurately- and precisely-measured values of $\boldsymbol{\mathcal{V}}(\mathbf{x})$ is

$$\boldsymbol{\mathcal{V}}(\mathbf{x}) \equiv \left\{\langle\nu\rangle_{\varepsilon_x}(x) : x \in [\bar{x}(\mathbf{x})]_{\overset{L}{\sim}}\right\}. \tag{21}$$

However, on its own, this does not constitute a definition of $\boldsymbol{\mathcal{V}}(\mathbf{x})$ because $\varepsilon_{\mathcal{V}}$ determines ε_x, so we cannot define $\varepsilon_{\mathcal{V}}$ in terms of ε_x.

Furthermore, we must take care to satisfy the requirement that ν fluctuates microscopically about $\boldsymbol{\mathcal{V}}(\mathbf{x})$ at $\mathbf{x}$. If we choose an arbitrary mesoscopic value of ε_x, and then use Eq. 21 as the definition of the set of values about which ν fluctuates microscopically at $\mathbf{x}$, this requirement may not be satisfied. For example, as Fig. 9 illustrates, if $\left|\boldsymbol{\mathcal{V}}^{(1)}\right|$ is large enough, the sets

$$\left\{\langle\nu\rangle_{\varepsilon_x}(x - \varepsilon_x/2 + u) : 0 < u < \mathfrak{a}\right\} \tag{22}$$

and

$$\left\{\langle\nu\rangle_{\varepsilon_x}(x + \varepsilon_x/2 - u) : 0 < u < \mathfrak{a}\right\} \tag{23}$$

do not intersect. Therefore, although $x_1 \in [\bar{x}(\mathbf{x})]_{\overset{L}{\sim}}$ is required to imply that $\langle\nu\rangle_{\varepsilon_x}(x_1)$ is among the set of values of $\boldsymbol{\mathcal{V}}$ that might be measured at $\mathbf{x}$, if ε_x is not chosen carefully, and if the phrase 'ν *fluctuates microscopically about* υ *at* x' is defined to mean

$$\upsilon \in \left\{\langle\nu\rangle_{\varepsilon_x}(x + u) : -\mathfrak{a}/2 < u < \mathfrak{a}/2\right\},$$

$\langle\nu\rangle_{\varepsilon_x}(x_1)$ may not be a value about which ν fluctuates microscopically at another point $x_2 \in [\bar{x}(\mathbf{x})]_{\overset{L}{\sim}}$.

To remedy this problem we should define this phrase without referring to ε_x, and then choose ε_x such that

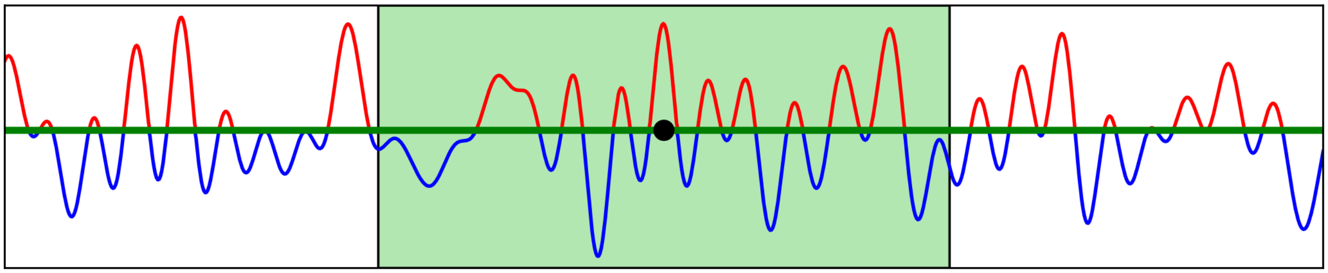

FIG. 10 Schematic. The red and blue curve represents a microstructure and the horizontal green line is its average on the green-shaded interval whose center is marked by a black spot. Although it seems natural to say that a microstructure *fluctuates microscopically* about its average, its average is a microscopic function of both the width and the position of the averaging interval. This means that the green line moves up and down as either one of them changes, and that the distances between successive extrema of these fluctuations of the average are microscopic. There does not exist a unique value about which a microstructure fluctuates microscopically, in general. because there is no reason to choose one interval width over another, or to choose to center the interval at a particular point instead of one a microscopic distance away,

$\boldsymbol{\mathcal{V}}(\mathbf{x})$, as defined by Eq. 21, is a *subset* of the set of values about which ν fluctuates microscopically at *every* $x \in [\bar{x}(\mathbf{x})]_{\underset{\sim}{L}}$.

I now propose possible definitions of the phrases '*fluctuates microscopically about υ at x*' and '*fluctuates microscopically about υ at* $\mathbf{x}$,' which I have not justified rigorously. I present them to illustrate the difficulties with circular definitions, and because they might be useful as a starting point for the development of rigorously-justified definitions that seamlessly link microstructure to what is measured and observed at the macroscale.

Definition. *ν fluctuates microscopically about υ at x* if and only if there exists a microscopic interval centered at x on which the average of ν is υ, i.e., if and only if

$$\upsilon \in A_{\mathfrak{a}}[\nu](x) \equiv \text{int}\left\{\langle\nu\rangle_{\eta}(x) : \eta < \mathfrak{a}\right\}, \tag{24}$$

In this expression 'int' denotes the interior of the set, meaning the set without its boundary points. Although it is not necessary to define $A_{\mathfrak{a}}$ to be an open set here, in more rigorous investigations of its properties I have found it useful or necessary to define it as open.

Definition. *ν fluctuates microscopically about υ at* $\mathbf{x}$ if and only if

$$\upsilon \in B_{\mathfrak{a}}^{\varepsilon_x}[\nu](\bar{x}(\mathbf{x})) \equiv \bigcap_{x' \in \mathfrak{I}(\bar{x}(\mathbf{x}),\varepsilon_x)} A_{\mathfrak{a}}[\nu](x').$$

Definition. *ν fluctuates microscopically about* $\boldsymbol{\mathcal{V}}$ if and only if

$$\boldsymbol{\mathcal{V}}(\mathbf{x}) \subseteq B_{\mathfrak{a}}^{\varepsilon_x}[\nu](\bar{x}(\mathbf{x})), \quad \forall \mathbf{x} \in \text{dom}\, \boldsymbol{\mathcal{V}}.$$

Note that Eq. 21 implies that

$$\boldsymbol{\mathcal{V}}(\mathbf{x}) \subseteq \bigcup_{x' \in \mathfrak{I}(\bar{x}(\mathbf{x}),\varepsilon_x)} \mathfrak{cl}\, A_{\varepsilon_x}[\nu](x'),$$

where $A_{\varepsilon_x}[\nu](x') \equiv \text{int}\{\langle\nu\rangle_{\eta}(x') : \eta < \varepsilon_x\}$, and the *closure operator*, $\mathfrak{cl}$, closes a set by adding its boundary points to it.

The following assumption specifies the domain of validity of the three definitions proposed above.

***Physical assumption* 8**: There exist values of $\mathfrak{a}$ and ε_x such that $\mathfrak{a} \ll \varepsilon_x < \mathfrak{L}$, and such that

$$\boldsymbol{\mathcal{V}}(\mathbf{x}) \equiv \left\{\langle\nu\rangle_{\varepsilon_x}(x) : x \in [\bar{x}(\mathbf{x})]_{\underset{\sim}{L}}\right\} \subseteq B_{\mathfrak{a}}^{\varepsilon_x}[\nu](\bar{x}(\mathbf{x})),$$

at every point $\mathbf{x} \in \text{dom}\, \boldsymbol{\mathcal{V}}$.

Note that microstructures with perfect periodicities are pathological in various ways, but they are also unphysical because there always exists some degree of disorder. Even the periodicity of the time average of a crystal's microstructure on an interval of length τ only has perfect periodicity in the limit $\tau \to \infty$. Therefore I propose, as a conjecture, that Physical Assumption 8 holds true for a useful subset of physical (disordered) microstructures, which satisfy the first seven physical assumptions stated earlier in Sec. VI.

I. Uncertainty principle

I have discussed the case of a simple average, with a top-hat kernel, in some detail. The purpose of this section is to discuss, in more general terms, how the shape and width of the averaging kernel determine unavoidable uncertainties in measured values of $\boldsymbol{\mathcal{V}}$ and its derivative, $\boldsymbol{\mathcal{V}}^{(1)}$. By analogy with Eq. 21, the set of values of $\boldsymbol{\mathcal{V}}^{(1)}$ that could be measured at a point $\mathbf{x}$ is

$$\boldsymbol{\mathcal{V}}^{(1)}(\mathbf{x}) \equiv \left\{\langle\partial_x\nu\rangle_{\varepsilon_x}(x) : x \in [\bar{x}(\mathbf{x})]_{\underset{\sim}{L}}\right\}, \tag{25}$$

and I denote the width of this interval by ε_p, where the subscript 'p' is intended to be reminiscent of momentum in quantum mechanics.

I will now derive a relationship between $\varepsilon_{\mathcal{V}}$, ε_p, and ε_x. For simplicity I will denote the spatial average, $\langle \nu; \mu \rangle_{\varepsilon_x}$, simply as $\langle \nu \rangle$, I will denote the average of the average as $\langle\langle \nu \rangle\rangle \equiv \left\langle \langle \nu; \mu \rangle_{\varepsilon_x}; \mu \right\rangle_{\varepsilon_x}$, etc..

To a first approximation, the uncertainty in the value of $\boldsymbol{\mathcal{V}}^{(1)}$ can be quantified by the variance of $\langle \partial_x \nu \rangle$. Let us use the fact that spatial derivatives commute with spatial averaging to write

$$\begin{aligned}\langle \partial_x \nu \rangle - \langle\langle \partial_x \nu \rangle\rangle &= \partial_x \left[\langle \nu \rangle - \langle\langle \nu \rangle\rangle \right] = \partial_x \langle \nu - \langle \nu \rangle \rangle \\ &= \partial_x \langle \Delta\nu \rangle = \partial_x \left(\mu(\varepsilon_x) * \Delta\nu \right) = \mu(\varepsilon_x) * \left(\partial_x \Delta\nu \right),\end{aligned}$$

where $\Delta\nu(x) \equiv \nu(x) - \langle \nu \rangle (x)$. In more explicit notation, this can be expressed as

$$\langle \partial_x \nu \rangle (x) - \langle\langle \partial_x \nu \rangle\rangle (x) = \int_{\mathbb{R}} \mu(u; \varepsilon_x) \partial_x \Delta\nu(x+u) \, \mathrm{d}u \,.$$

Let us replace $\int_{\mathbb{R}}$ with $\int_{-\ell}^{\ell}$, where the value of $\ell = \ell(\varepsilon_x)$ has been chosen such that

$$\left| \int_{-\ell}^{\ell} \mu(u; \varepsilon_x) \partial_x \Delta\nu(x+u) \, \mathrm{d}u \right. \\ \left. - \lim_{\ell \to \infty} \int_{-\ell}^{\ell} \mu(u; \varepsilon_x) \partial_x \Delta\nu(x+u) \, \mathrm{d}u \right|$$

is negligible. If we also replace $\mu(u; \varepsilon_x)$ with its Taylor expansion about $u = 0$, we find

$$\begin{aligned}\langle \partial_x \nu \rangle - \langle\langle \partial_x \nu \rangle\rangle = \mu(0; \varepsilon_x) &\int_{-\ell}^{\ell} \partial_x \Delta\nu(x+u) \, \mathrm{d}u \\ &+ \sum_{m=1}^{\infty} \frac{\mu^{(m)}(0; \varepsilon_x)}{m!} \int_{-\ell}^{\ell} u^m \partial_x \Delta\nu(x+u) \, \mathrm{d}u \,.\end{aligned}$$

The integrals appearing in the sum on the right hand side can be expressed as

$$\begin{aligned}&\int_{-\ell}^{\ell} u^m \partial_x \Delta\nu(x+u) \, \mathrm{d}u \\ &\qquad = \int_0^{\ell} u^m \left[\partial_x \Delta\nu(x+u) \pm \partial_x \Delta\nu(x-u) \right] \mathrm{d}u \,,\end{aligned}$$

where ± is + when m is even and − when m is odd. In both cases there is partial cancellation, which reduces the magnitudes of the integrals by a factor of about $1/\sqrt{2}$. We know from Eq. 17 that when ε_x is large the m^{th} derivative of the kernel, $\mu^{(m)}(0; \varepsilon_x)$, scales as $1/{\varepsilon_x}^m$. Furthermore, if μ is symmetric, then $\mu^{(1)}(0; \varepsilon_x) = 0$ and the $m = 1$ term vanishes.

Therefore, to a first approximation, or in the limit of large ε_x, the variance of the slope of $\boldsymbol{\mathcal{V}}$ is

$$\begin{aligned}\left(\frac{\varepsilon_p}{2} \right)^2 \equiv \left\langle \left(\boldsymbol{\Delta \mathcal{V}^{(1)}} \right)^2 \right\rangle &= \left\langle \left(\langle \partial_x \nu \rangle - \langle\langle \partial_x \nu \rangle\rangle \right)^2 \right\rangle \\ &\approx \mu(0; \varepsilon_x)^2 \left\langle \left(\Delta\nu(x+\ell) - \Delta\nu(x-\ell) \right)^2 \right\rangle \\ &= \mu(0; \varepsilon_x)^2 \Big[\left\langle \Delta\nu(x+\ell)^2 \right\rangle + \left\langle \Delta\nu(x-\ell)^2 \right\rangle \\ &\qquad - 2 \left\langle \Delta\nu(x+\ell) \Delta\nu(x-\ell) \right\rangle \Big]\end{aligned}$$

If we now assume that, for the purpose of calculating $\langle \Delta\nu(x+\ell) \Delta\nu(x-\ell) \rangle$, the values of $\Delta\nu(x+\ell)$ and $\Delta\nu(x-\ell)$ can be treated as independent random variables with means of zero and variances of $(\varepsilon_{\mathcal{V}}/2)^2$, then $\langle \Delta\nu(x+\ell) \Delta\nu(x-\ell) \rangle$ vanishes and we get

$$\left(\frac{\varepsilon_p}{2} \right)^2 \approx 2\mu(0; \varepsilon_x)^2 \left(\frac{\varepsilon_{\mathcal{V}}}{2} \right)^2 \Rightarrow \varepsilon_x \varepsilon_p \approx r_\mu \varepsilon_{\mathcal{V}}, \qquad (26)$$

where $r_\mu \equiv \sqrt{2}\mu(0; \varepsilon_x)\varepsilon_x \sim 1$, is dimensionless and with a value that depends on the shape of μ. If μ is Gaussian, then $\mu(0; \varepsilon_x) = (1/\varepsilon_x)\sqrt{2/\pi}$ and $r_\mu = 2/\sqrt{\pi}$. If μ is a top-hat, then $\mu(0; \varepsilon_x) = 1/\left(\varepsilon_x \sqrt{3} \right)$ and $r_\mu = \sqrt{2/3}$. If $\sigma_x \equiv \varepsilon_x/2$, $\sigma_{\mathcal{V}} \equiv \varepsilon_{\mathcal{V}}/2$, and $\sigma_p \equiv \varepsilon_p/2$, Eq. 26 can be expressed as

$$\begin{aligned}\sigma_x \sigma_p &= r_\mu \sigma_{\mathcal{V}}/2 && \text{(General kernel)} \\ \sigma_x \sigma_p &= \sigma_{\mathcal{V}}/\sqrt{\pi} && \text{(Gaussian kernel)} \\ \sigma_x \sigma_p &= \sigma_{\mathcal{V}}/\sqrt{6} && \text{(Top-hat kernel)}\end{aligned}$$

These relations imply that there is a trade-off between macroscale spatial precision and the uncertainty in $\boldsymbol{\mathcal{V}}^{(1)}$. When microscopic fluctuations of ν are large, $\varepsilon_{\mathcal{V}} = 2\sigma_{\mathcal{V}}$ is large, and $\varepsilon_x \varepsilon_p = 4\sigma_x \sigma_p$ is large.

J. Summary of the fundaments of homogenization theory

In this section I have discussed some of the fundamental features of the homogenization transformation that turns microstructure into macrostructure. I have pointed out that macrostructure cannot be defined uniquely, because it depends on the scale, $\varepsilon_x = |\mathbf{dx}|$, at which the microstructure is observed, and which defines the smallest distance, $|\mathbf{dx}|/2$, between mutually-distinguishable points at the macroscale.

However the value of $|\mathbf{dx}|$ cannot be chosen to be arbitrarily small if distances and displacements are measured with macroscopic fields. This is because ε_x both determines, and is bounded from below by, the finite precisions, $\varepsilon_{\mathcal{V}}$ and ε_p, to which macroscopic fields and their derivatives, respectively, are defined. Therefore ε_x, $\varepsilon_{\mathcal{V}}$, and ε_p are all interrelated, and can be interpreted either as unavoidably-finite precisions or as measures of unavoidable uncertainty. In Sec. VI.I I derived uncertainty relations which imply that reducing ε_x increases the uncertainty ε_p in derivatives of the macroscopic field used to measure distances and displacements.

I have not presented a rigorously-justified relationship between ε_x and $\varepsilon_{\mathcal{V}}$, for an arbitrary microstructure, ν, which satisfies my physical assumptions. In part, this is because any such definition would have to be accompanied by further physical assumptions, which specified more precisely the set of microstructures to which it would apply. However, I have highlighted some of the difficulties that must be overcome to devise rigorously-justified definitions, and I have proposed a definition that I have found to be viable for a useful subset of microstructures that satisfy my physical assumptions. I will present these numerical and theoretical findings elsewhere.

The domain of validity of my proposed definitions is not relevant to the two most fundamental conclusions of Sec. VI. The first of these is that homogenization introduces unavoidable uncertainty at the macroscale, making spatial precision, and the precisions to which macroscopic fields are defined, finite. The second is that finite spatial precision has important implications for the nature of macrostructure, which I discussed in Sec. VI.C, and will discuss further in Sec. VII.

Briefly, it means that surfaces and interfaces, which do not exist in a well-defined sense at the microscale, are created by the homogenization transformation. When they are created they carry excess fields, in general, and these fields are an integral component of macrostructure. In fact, because the macroscopic charge density $\boldsymbol{\varrho}$ vanishes in the bulks of stable materials, in the context of electricity it can be the case that the excess fields, and the 2-, 1-, and 0-manifolds they inhabit, *are* the macrostructure.

Therefore the task of laying foundations for a homogenization theory that defines macrostructure in terms of microstructure, ν, is far from complete. Completing the foundations entails defining excess fields in terms of ν. This is the subject of Sec. VII.

It is straightforward to generalize the theory presented in this section to systems in which the microstructure varies significantly on three or more widely-separated length scales. In that case homogenization proceeds in stages from the *base microstructure*, on the smallest length scale, to the *apex macrostructure*, on the largest length scale. The base microstructure is the only microstructure that is not also a macrostructure determined by a microstructure on a smaller length scale, and the apex macrostructure is the only macrostructure that is not also a microstructure which determines a macrostructure on a larger length scale.

It may not be straightforward to adapt the theory presented in this section to materials whose structures vary significantly on every length scale, such as wood (Toumpanaki *et al.*, 2021). However, such an adaptation may not be useful because, unlike most artificial materials, wood does not appear to be locally homogeneous when observed with either the naked eye or a microscope at any level of magnification.

VII. EXCESS FIELDS

A. Introduction

As discussed in Sec. VI, macrostructure essentially consists of smoothly-varying continua interspersed with heterogeneitites such as defects, interfaces, and inclusions.

An inclusion is an embedded region whose microstructure differs from that of its host. If all of its dimensions are smaller than $\varepsilon_x = \mathbf{dx}$, an inclusion is a singular point in the volumetric macroscopic field $\boldsymbol{\mathcal{V}}$ and therefore a macroscale point defect. Similarly, a macroscale line [planar] defect is an inclusion that is larger than ε_x along only one [two] of its dimensions. Macroscale defects may require special treatments when applying the macroscale theory, but their macrostructures can be calculated from the microstructures by reasonably-straightforward application of the three-dimensional mesoscale averaging operation.

Macroscale defects should not be confused with their microscale counterparts. For example, consider vacancies and impurities in crystals, which are microscale point defects. Although they may be charged, and therefore may contribute directly to the microscopic charge density ρ, and they may perturb the arrangement of atoms, thereby indirectly changing ρ, their concentrations are usually high enough and/or their effects on averages of ρ small enough, that they can be regarded as just another feature of the microstructure. Their presence in a crystal does not alter the relationship between $\boldsymbol{\varrho}$ and ρ in most cases.

Exceptionally, microscopic defects might increase the upper bound, $\mathfrak{a}$, on distances regarded as microscopic so much that it becomes comparable to $\mathfrak{L}$, thereby rendering Physical Assumpion 8, and much of the theory presented in Sec. VI invalid. However, I restrict attention to systems in which $\mathfrak{a} \ll \mathfrak{L}$.

If all of an inclusion's dimensions are much larger than ε_x, the curvature of its macroscale boundary with the host material will be negligible on the mesoscale. Therefore, the inclusion is simply another macroscale material whose boundary with its host is locally flat. On either side of that boundary $\boldsymbol{\mathcal{V}}$ is differentiable and the boundary itself can be treated like any other mesoscopically-planar interface. As discussed in Sec. VI, not only does

$\boldsymbol{\mathcal{V}}$ tend to be discontinuous at interfaces, but interfaces carry excesses $\boldsymbol{\sigma}_{\mathcal{V}}$ of $\boldsymbol{\mathcal{V}}$, in general, which play important roles in physics at the macroscale.

I am trying to emphasise that, for a large class of systems, and from a practical perspective, the only ingredient of a mutually-consistent description of macrostructure and microstructure that we lack is the relationship between microstructure ν and the macroscale interfacial excess $\boldsymbol{\sigma}_{\mathcal{V}}$. The purpose of this section is to derive this relationship. Specifically, I derive expressions for excesses at surfaces (interfaces with a vacuum), which are trivial to generalize to interfaces by treating them as adjoined surfaces.

1. Notation

In this section, and henceforth, I will assume that $\boldsymbol{\mathcal{V}}$ is a single-valued field at the macroscale, which is a mesoscale spatial average of ν that I will often denote by $\bar{\nu}$. At the microscale I will denote the midpoint of interval $\boldsymbol{\mathcal{V}}(\mathbf{x}) \equiv \mathfrak{I}(\overline{\boldsymbol{\mathcal{V}}}(\mathbf{x}), \varepsilon_{\mathcal{V}})$, by $\overline{\mathcal{V}}(\bar{x}(\mathbf{x}))$, and I will assume that $\overline{\boldsymbol{\mathcal{V}}}(\mathbf{x})$ is the average of $\overline{\mathcal{V}}(x)$ over all $x \in [\bar{x}(\mathbf{x})]_{\underset{\sim}{L}}$.

As in Sec. VI.I, $\ell = \ell(\varepsilon_x) > \varepsilon_x$ will denote the mesoscopic width of the domain of a mesoscale spatial average. Increasing its value makes the approximation

$$\begin{aligned}\bar{\nu}(x) \equiv \langle \nu; \mu \rangle_{\varepsilon_x}(x) &\equiv \int_{\mathbb{R}} \nu(x+u)\mu(u;\varepsilon_x)\,\mathrm{d}u \\ &\approx \int_{-\ell/2}^{\ell/2} \nu(x+u)\mu(u;\varepsilon_x)\,\mathrm{d}u\end{aligned}$$

an arbitrarily close one. I introduce this finite width to help with derivations and I do not attach physical meaning to it.

In Fig. 9, while considering the example of a top-hat kernel, I defined $\varepsilon_x = |\mathbf{dx}|$ such that if the distance between two macroscale points was greater than $|\mathbf{dx}|/2$, they could be distinguished from one another, *with certainty*, by macroscale measurements. However, real measurements do not provide certainty - only probabilities and degrees of certainty.

In this section I will consider spatial averages with an arbitrary kernel, but to avoid cluttering and complicating the theory and discussion, I will not discuss probabilities. I will continue to refer to ε_x as the macroscopic spatial precision, and as the width of an interface at the microscale, but with the understanding that $\varepsilon_x/2$ is now the standard deviation of the position probability density function, $\mu(\varepsilon_x)$. In other words, I will continue to use precise non-probabilistic language and mathematics, while cognisant of the fact that this preciseness is unjustified. For example, I will continue to regard the coincidence set $[\bar{x}(\mathbf{x})]_{\underset{\sim}{L}}$ of $\mathbf{x}$ as a well-defined set of points, and I will continue to discuss an interface as having the precisely-defined width, ε_x.

If this sloppiness introduces doubt about the validities of the derivations that follow, this doubt can be removed by strengthening our physical assumptions about the nature of microstructure: We can assume that $\mathfrak{a}/\varepsilon_x$ and $\varepsilon_x/\mathfrak{L}$ are both so small that the shape of μ has a negligible influence on the macrostructure (see Eq. 17). Then μ is effectively a top-hat kernel. To achieve further comfort, by reverting to perfect consistency with our discussion of top-hat kernels in Sec. VI, we could mentally replace every instance of $\mu(\varepsilon_x)$ in what follows with $\mu(\varepsilon_x/\sqrt{3})$.

B. Surface excesses in three dimensions

Far from an interface, the relationship between the macrostructure and the microstructure is as described, but not rigorously and precisely defined, in Sec. VI. Homogenizing the interface region presents new problems as a consequence of the fundamental difference between interfaces at the macroscale and interfaces at the microscale, which I briefly discussed in Sec. VI.C.1. These are that interfaces are ill-defined at the microscale, because their widths are indeterminate, whereas at the macroscale they are well-defined two dimensional manifolds which carry excess fields.

In Sec. VI.C.1 we considered excesses of one dimensional microstructures. Let us begin our discussion of excesses of three dimensional microstructures by considering an excess field on a surface that is perpendicular to the $\mathbf{x}-$axis and whose macroscale $\mathbf{x}-$coordinate is $\mathbf{x}_L$ (see Fig. 11). Let $x_L \equiv \bar{x}(\mathbf{x}_L)$. Above the surface, by which I mean $\mathbf{x} < \mathbf{x}_L$ at the macroscale and $x < x_L - \varepsilon_x/2$ at the microscale, there is vacuum, meaning that both ν and $\boldsymbol{\mathcal{V}}$ are zero. Let us assume that $\boldsymbol{\mathcal{V}}$ is also zero far below the surface, but that the average $\bar{\nu}_{yz}(x)$ of ν on the plane parallel to the surface at x does not vanish for every x. This means that the mesoscale average of ν only vanishes if

$$\begin{aligned}\int_0^{\ell/2} &\bar{\nu}_{yz}(x+u)\mu(u;\varepsilon_x)\,\mathrm{d}u \\ &= -\int_{-\ell/2}^{0} \bar{\nu}_{yz}(x+u)\mu(u;\varepsilon_x)\,\mathrm{d}u\,,\end{aligned}$$

and, in general, neither of these integrals is zero. Therefore, unless the average of ν vanishes on all planes parallel to the surface, $\boldsymbol{\mathcal{V}}$ can only vanish if the contributions to it from different depths below the surface cancel one another.

It follows that if we create a surface perpendicular to the $\mathbf{x}-$axis by removing all material from one side of an imaginary plane passing through the bulk, the removal of this material disrupts the cancellation that causes $\boldsymbol{\mathcal{V}}$ to vanish. Therefore, within $\pm\varepsilon_x/2$ of the surface, $\boldsymbol{\mathcal{V}}$ does not vanish. The integral of $\boldsymbol{\mathcal{V}}(\mathbf{x},\mathbf{y},\mathbf{z})$ between any point above the surface, $\mathbf{x}_1 < \mathbf{x}_L - |\mathbf{dx}|/2$, and any point below it, $\mathbf{x}_2 > \mathbf{x}_L + |\mathbf{dx}|/2$, is

$$\int_{\mathbf{x}_1}^{\mathbf{x}_L-|\mathbf{dx}|/2}\boldsymbol{\mathcal{V}}(\mathbf{x},\mathbf{y},\mathbf{z})\mathbf{dx}+\int_{\mathbf{x}_L-|\mathbf{dx}|/2}^{\mathbf{x}_L+|\mathbf{dx}|/2}\boldsymbol{\mathcal{V}}(\mathbf{x},\mathbf{y},\mathbf{z})\mathbf{dx} + \int_{\mathbf{x}_L+|\mathbf{dx}|/2}^{\mathbf{x}_2}\boldsymbol{\mathcal{V}}(\mathbf{x},\mathbf{y},\mathbf{z})\mathbf{dx} = \boldsymbol{\mathcal{V}}(\mathbf{x}_L,\mathbf{y},\mathbf{z})|\mathbf{dx}| \neq 0$$

The first and the third integrals are zero because $\boldsymbol{\mathcal{V}}(\mathbf{x},\mathbf{y},\mathbf{z}) = 0$ if $|\mathbf{x}-\mathbf{x}_L| > |\mathbf{dx}|/2$. Therefore, at position $(\mathbf{y},\mathbf{z})$ on the surface plane, the excess of $\boldsymbol{\mathcal{V}}$ is $\boldsymbol{\sigma}_{\mathcal{V}}(\mathbf{y},\mathbf{z}) \equiv \boldsymbol{\mathcal{V}}(\mathbf{x}_L,\mathbf{y},\mathbf{z})|\mathbf{dx}|$. The surface average, $\bar{\boldsymbol{\sigma}}_{\mathcal{V}}$, of $\boldsymbol{\sigma}_{\mathcal{V}}$ is $\overline{\boldsymbol{\mathcal{V}}}_{yz}(\mathbf{x_L})|\mathbf{dx}|$, where $\overline{\boldsymbol{\mathcal{V}}}_{yz}$ is the macroscale counterpart of $\bar{\nu}_{yz}$, meaning its average on a mesoscopic two dimensional domain.

C. Calculating interfacial excesses from the microstructure

To address the question of how excess fields can be calculated from the microstructure, ν. let us continue to assume that the x-axis is normal to the surface and, to simplify the notation by keeping the problem one-dimensional, let us assume that $\nu(x)$ is the average of some other microscopic quantity on the plane parallel to the surface at x. As before, let us assume that $\boldsymbol{\mathcal{V}} = 0$ in the bulk. An obvious starting point is to define the microscale surface excess, σ_ν, as

$$\sigma_\nu(x_b) = \int_{x_L-\varepsilon_x/2}^{x_b} \nu(x)\ \mathrm{d}x \tag{27}$$

where $\nu(x) = 0$ if $x < x_L - \varepsilon_x/2$, and x_b is a point deep below the surface ('b'='bulk').

To see that $\boldsymbol{\sigma}_{\mathcal{V}} \equiv \sigma_\nu(x_b)$ is not a good definition of the macroscale surface excess, consider the example depicted in Fig. 11. In this example, the material could be a three dimensional crystal and ν the average of the charge density over planes parallel to the surface. The value of ν is zero everywhere except at a discrete set of $x-$values, corresponding to lattice planes, on which it is either $+1$ or -1. Therefore, calculating $\sigma_\nu(x_b)$ is as simple as counting these charges from $x < x_L$ to $x = x_b$. By inspection, we find that $\sigma_\nu(x_1) = 0$ and $\sigma_\nu(x_2) = +1$, where x_1 and x_2 are the positions indicated in Fig. 11. If one continues counting beyond $x = x_2$, the value of σ_ν continues to jump between 0 and +1 and it never converges.

The problem with defining $\boldsymbol{\sigma}_{\mathcal{V}} \equiv \sigma_\nu(x_b)$ is twofold. First, identifying $\sigma_\nu(x_b)$ as the surface excess appears to imply that $x < x_b$ is the surface region and $x > x_b$ is the bulk. However, as discussed in Sec. VI.C.1, there is no clear boundary between surface and bulk at the microscale and so the "surface region" is ill-defined. Second, although $\boldsymbol{\mathcal{V}}$ vanishes in the bulk, the same is not true of ν, and any integral of a microscopic quantity is a microscopic function of its upper and lower bounds of integration. Therefore, $\sigma_\nu(x_b)$ fluctuates microscopically as x_b is varied.

This simple example, which is typical rather than pathological, illustrates how interfaces being ill-defined at the microscale can be troublesome when one attempts to calculate macroscale properties of interfaces from the microstructure. It also underscores the importance of a careful understanding of the relationship between microscale physics and macroscale physics.

To deduce the relationship between $\boldsymbol{\sigma}_{\mathcal{V}}$ and ν, consider the following two slightly-different lines of reasoning. The first is to define $\boldsymbol{\sigma}_{\mathcal{V}}$ as the mesoscale average of $\sigma_\nu(x_b)$. This means that, instead of terminating the integral at a single plane (at x_b) we take an average over an ensemble of terminating planes. This was the approach taken by Finnis (Finnis, 1998), who appears to have been the first to solve the problem of calculating what he called *thermodynamic excesses* of charge and other quantities at interfaces. In Finnis, 1998 he reasoned that, by averaging over terminating planes, "*we can reconcile the atomistic picture, in which excesses appear to oscillate on the atomic length scale as a function of the region size, with the thermodynamic picture*". He used this approach to derive an expression for the surface charge in crystals. One purpose of Sec. VII.H is to generalize his result to noncrystalline materials.

The second line of reasoning, which may seem more natural in the present context, begins with the fact that surfaces, and therefore surface excesses, are only well-defined at the macroscale. Therefore, $\boldsymbol{\sigma}_{\mathcal{V}}$ must be the integral of $\boldsymbol{\mathcal{V}}$ across the surface, i.e., $\boldsymbol{\sigma}_{\mathcal{V}} \equiv \int_{\mathbf{x}_1}^{\mathbf{x}_2} \boldsymbol{\mathcal{V}}(\mathbf{x})\,\mathbf{dx}$, where $\mathbf{x}_1 < \mathbf{x}_L$ and $\mathbf{x}_2 > \mathbf{x}_L$. This integral must converge to the surface excess because $\boldsymbol{\mathcal{V}} = 0$ in the bulk and above the surface, and because the spatial averaging operation is conservative, by virtue of μ being normalized to one. It is straightforward to show that this viewpoint and Finnis's thermodynamic viewpoint are equivalent, because the spatial average of an integral of ν is equal to the integral of the spatial average of ν.

D. Changes of macroscale quantities across interfaces

To calculate the change in $\boldsymbol{\mathcal{V}}$ between a point $\mathbf{x}_2$ on one side of an interface and a point $\mathbf{x}_1$ on the other, one could simply calculate $\boldsymbol{\mathcal{V}}(\mathbf{x}_1)$ and $\boldsymbol{\mathcal{V}}(\mathbf{x}_2)$ from the microstructure, ν. However, this is not always the easiest approach. For example, if $\nu^{(1)}$ is known, but ν is not, it might be easier to recognise that the change of $\boldsymbol{\mathcal{V}}$ across the interface is the interfacial excess of $\boldsymbol{\mathcal{V}}^{(1)}$. Therefore,

$$\boldsymbol{\mathcal{V}}(\mathbf{x}_2) - \boldsymbol{\mathcal{V}}(\mathbf{x}_1) = \boldsymbol{\sigma}_{\mathcal{V}^{(1)}} \equiv \int_{\mathbf{x}_1}^{\mathbf{x}_2} \boldsymbol{\mathcal{V}}^{(1)}(\mathbf{x})\,\mathbf{dx}. \tag{28}$$

There are many important physical systems in which ν is related to a source function, ψ, by the Poisson equation, $\nu^{(2)} = \psi$. Since differentiation and spatial averaging commute, their macroscale counterparts have the same relationship, $\boldsymbol{\mathcal{V}}^{(2)} = \boldsymbol{\Psi}$. When $\boldsymbol{\mathcal{V}}$ has different values on either side of the interface, but is constant on both sides, the step change in its value across the interface can be calculated from ψ by integrating Eq. 28 by parts and substituting the Poisson equation to give

$$\boldsymbol{\mathcal{V}}(\mathbf{x}_2) - \boldsymbol{\mathcal{V}}(\mathbf{x}_1) = -\int_{\mathbf{x}_1}^{\mathbf{x}_2} \mathbf{x}\,\boldsymbol{\Psi}(\mathbf{x})\,\mathbf{dx} \tag{29}$$

In Sec. XIII we will find that Eq. 29 is useful way for calculating the change in the macroscopic potential $\boldsymbol{\Phi}$ across an interface, and therefore for calculating the *mean inner potential* (Bethe, 1928; Cendagorta and Ichiye, 2015; Gajdardziska-Josifovska *et al.*, 1993; Kathmann, 2021; Kathmann *et al.*, 2011; Leung, 2010; Miyake, 1940; Pratt, 1992; Sanchez and Ochando, 1985; Sokhan and Tildesley, 1997a; Spence, 1993; Wilson *et al.*, 1987, 1988, 1989; Yesibolati *et al.*, 2020).

For the purposes of calculating the interfacial excesses, and step-changes in macroscopic quantities across interfaces, that are required in this work about electricity, we only need to deduce relationships between the right-hand-sides of Eqs. 28 and 29 and the microstructures ν and ψ, respectively. Mathematically, the problem at hand is to find simple and general expressions for

$$\left\langle \int_{x_1}^{x_2} \nu(x)\,\mathrm{d}x\,;\mu \right\rangle_{\varepsilon_x} \quad \text{and} \quad \left\langle \int_{x_1}^{x_2} x\,\nu(x)\,\mathrm{d}x\,;\mu \right\rangle_{\varepsilon_x} \tag{30}$$

in terms of ν, where the spatial average is performed over the upper bound, x_2, of the integrals at a fixed value of x_1. Once these expressions are in hand, it will be straightforward to find expressions for the averages of these integrals over x_1 or over both x_1 and x_2.

E. Mesoscale averages of integrals

The goal of this section is to deduce general relationships between the microstructure and the integral averages in Expression 30, which are equivalent to the right-hand-sides of Eqs. 28 and 29, apart from the appearance of $\boldsymbol{\Psi}$ instead of $\boldsymbol{\mathcal{V}}$ in the latter.

Calculating the integral of $\overline{\mathcal{V}}$, which is constant or linear on both sides of an interface, is straightforward. Therefore, let us define $\Delta\nu(x) \equiv \nu(x) - \overline{\mathcal{V}}(x)$, and instead calculate the mesoscale average over x_b of

$$\mathcal{S}_r^{[\Delta\nu]}(x_b) \equiv \int_{x_L}^{x_b} x^r\,\Delta\nu(x)\,\mathrm{d}x \tag{31}$$

for $r = 0$ and 1. I will denote these averages by $\boldsymbol{\mathcal{S}}_0^{[\boldsymbol{\Delta}\boldsymbol{\mathcal{V}}]}$ and $\boldsymbol{\mathcal{S}}_1^{[\boldsymbol{\Delta}\boldsymbol{\mathcal{V}}]}$. The reasons for replacing ν by $\Delta\nu$ are that for $r = 0$ the derivation is made easier by the fact that $\Delta\nu$ fluctuates microscopically about zero, and that the average of $\mathcal{S}_1^{[\nu]}(x_b)$ does not converge with respect to x_b unless ν fluctuates about zero.

Apart from those stated and discussed in Sec. VI, we will not make any assumptions about the microstructure. Therefore, our goal of deriving *generally-applicable* expressions for $\boldsymbol{\mathcal{S}}_r^{[\boldsymbol{\Delta}\boldsymbol{\mathcal{V}}]}$ in terms of ν is only possible if $\boldsymbol{\mathcal{S}}_r^{[\boldsymbol{\Delta}\boldsymbol{\mathcal{V}}]}$ can be related to some calculable characteristics of the microstructure. Guided by Finnis's expression for the surface excess (Finnis, 1998), we will characterize the microstructure using moments and moment densities. This is explained in Sec. VII.F and Sec. VII.G.

F. Partitioning space into microscopic intervals

To characterize the microstructure in the mesoscopic neighbourhood of x_b, let us assume that the macrostructure is either uniform or linearly-varying in this neighbourhood. Let us partition an interval of width ℓ centered near x_b into a set of $2M$ contiguous microscopic subintervals, or *microintervals*, demarcated by the set of points

$$\Pi(x_b, \ell) \equiv \{x_m : m \in \mathbb{Z},\, |m| \leq M,\, x_0 \equiv x_b,\, x_{m+1} > x_m,\\ x_{m+1} - x_m < \alpha,\, x_{-M} = x_b - \ell/2,\, |x_M - x_b - \ell/2| < \mathfrak{a}\}$$

Notice that, although $|x_M - x_{-M}| = \ell(\varepsilon_x)$, the midpoint of (x_{-M}, x_M) is displaced from x_b by a microscopic distance. The reason for this will soon become clear. The microinterval designated *'interval m'* and denoted by $\mathfrak{I}_m \equiv \mathfrak{I}(\bar{x}_m, \Delta_m)$ has midpoint

$$\bar{x}_m \equiv \begin{cases} \frac{1}{2}\,(x_{m+1} + x_m) & \text{if } m < 0 \\ \frac{1}{2}\,(x_{m-1} + x_m) & \text{if } m > 0, \end{cases} \tag{32}$$

and width

$$\Delta_m \equiv \begin{cases} x_{m+1} - x_m & \text{if } m < 0 \\ x_m - x_{m-1} & \text{if } m > 0. \end{cases} \tag{33}$$

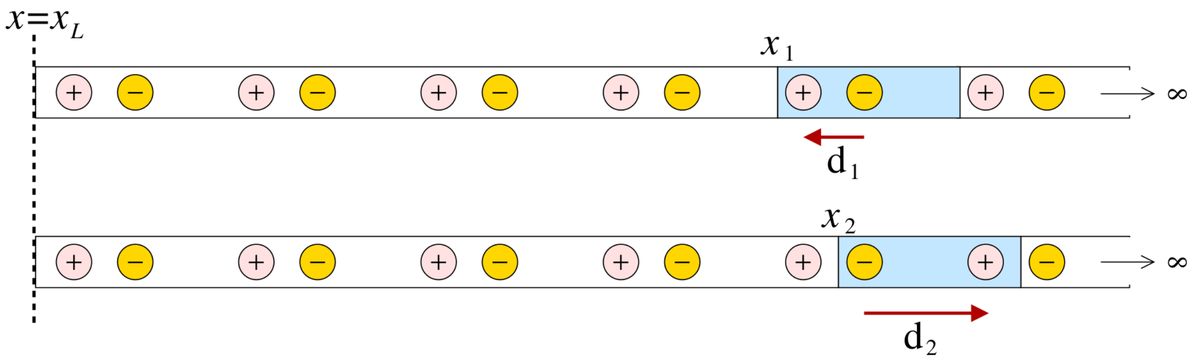

FIG. 11 Cartoons depicting the surface (at $x = x_L$) of a one dimensional material. The sums $\sigma_s(x_1)$ and $\sigma_s(x_2)$ of the charges between the surface at $x = x_L$ and x_1 and x_2, respectively, are very different. Therefore neither can be identified as the surface charge. The blue shaded regions are two different equally-valid choices of the unit cell of the bulk crystal, which have very different dipole moments $d(x_1 + \frac{a}{2})$ and $d(x_2 + \frac{a}{2})$, where $d(x)$ is the dipole moment of the unit cell centered at x. It was shown by Finnis(Finnis, 1998) that the excess surface charge is $\sigma = \sigma_s(x_1) + \frac{d_1}{a} = \sigma_s(x_2) + \frac{d_2}{a}$. Although the choice of unit cell to describe the periodicity of the crystal changes the "dipole moment density" $\mathcal{P}(x) = d(x)/a$ at every point, the excess surface charge is independent of this choice and is well defined.

Now we can write the mesoscale average of ν as the following sum of integrals over microintervals:

$$\begin{aligned}\bar{\nu}(x_b) &= \int_{-\ell/2}^{\ell/2} \mu(x_b - x; \varepsilon_x)\,\nu(x)\,\mathrm{d}x \\ &= \sum_m \int_{-\Delta_m/2}^{\Delta_m/2} \mu(x_b - \bar{x}_m - u; \varepsilon_x)\,\nu(\bar{x}_m + u)\,\mathrm{d}u \\ &= \sum_m \Delta_m \mu(x_b - \bar{x}_m; \varepsilon_x)\,\langle \nu \rangle_{\Delta_m}(\bar{x}_m), \end{aligned} \tag{34}$$

where $\sum_m$ denotes summation over all $2M$ microintervals and we have used the fact that the change of μ across each microinterval is negligible when $\mathfrak{a}/\varepsilon_x$ is sufficiently small.

Now let us place one further constraint on $\Pi(x_b, \ell)$, which explains why x_b is not the midpoint of (x_{-M}, x_M): The microinterval boundary points are chosen such that $\langle \Delta\nu \rangle_{\Delta_m}(\bar{x}_m) = 0$ for all m, which implies that $\bar{\nu}(x) = \overline{\mathcal{V}}(x)$. This choice is possible because $\Delta\nu$ fluctuates microscopically about zero everywhere in $\mathfrak{I}(x_b, \ell)$. Therefore, starting from x_b, the nearest point $x_1 > x_b$ such that the average of $\Delta\nu$ on $[x_b, x_1]$ is zero must be a microscopic distance Δ_1 away. The nearest point $x_2 > x_1$ such that $\langle \Delta\nu \rangle_{\Delta_2}(\bar{x}_2) = 0$ is a microscopic distance Δ_2 away, and so on. Having chosen a set $\Pi(x_b, \ell)$ for which the average of $\Delta\nu$ on each microinterval vanishes, Equation 34 becomes

$$\begin{aligned}\bar{\nu}(x_b) &= \sum_m \Delta_m \mu(x_b - \bar{x}_m; \varepsilon_x)\overline{\mathcal{V}}(\bar{x}_m) \\ &= \int_{-\infty}^{\infty} \mu(x_b - x; \varepsilon_x)\overline{\mathcal{V}}(x)\,\mathrm{d}x = \overline{\mathcal{V}}(x_b),\end{aligned}$$

which is independent of μ, as expected from Eq. 17 in the limit that $\mathfrak{a}/\varepsilon_x$ vanishes.

G. Characterising microstructure with moment distributions of microscopic intervals

Let us characterise the microstructure in interval m by the set of moment densities

$$\mathcal{M}^{\langle n\rangle}_{\Delta\nu}(\bar{x}_m, \Delta_m) \equiv \frac{1}{\Delta_m}\int_{-\Delta_m/2}^{\Delta_m/2} \Delta\nu(\bar{x}_m + u)\,u^n\,\mathrm{d}u\,,$$

where $n = 0, 1, 2$, etc.. The zeroth moment density is simply the average of $\Delta\nu$ on interval $[x_m, x_{m+1}]$, i.e.,

$$\mathcal{M}^{\langle 0\rangle}_{\Delta\nu}(\bar{x}_m, \Delta_m) = \langle \Delta\nu \rangle_{\Delta_m}(\bar{x}_m).$$

Each moment density $\mathcal{M}^{\langle n\rangle}_{\Delta\nu}(\bar{x}_m, \Delta_m)$ can be considered a microscopic quantity because its value fluctuates microscopically as a function of $\bar{x}_m$, at fixed Δ_m, and as a function of Δ_m, at fixed $\bar{x}_m$. Therefore, the set of all moment densities depends strongly on the choice of set $\Pi(x_b, \ell)$, which is, to a large extent, arbitrary.

Let us define the mesoscale average $\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b)$ of $\mathcal{M}^{\langle n\rangle}_{\Delta\nu}(\bar{x}_m, \Delta_m)$ as follows.

$$\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b) \equiv \left\langle \mathcal{M}^{(1)}_{\nu}; \mu \right\rangle^*_{\varepsilon_x}(x_b) \equiv \frac{1}{\ell}\sum_m \Delta_m \mathcal{M}^{(n)}_{\nu}(\bar{x}_m, \Delta_m).$$

I have introduced the notation $\langle\ .\ \rangle^*$ to denote a particular kind of spatial average - one which cannot be calculated by a continuous integral. It is a weighted average, over a discrete and finite set of values, each of which is calculated on a different microinterval from the set $\{\mathfrak{I}_m \equiv \mathfrak{I}(\bar{x}_m, \Delta_m)\}$ that partitions $\mathfrak{I}\,(x_{-M} + x_M)/2, \ell)$.

In general, the average moment densities $\bar{\mathcal{M}}^{(n)}_{\nu}$ can depend strongly on the choice of $\Pi(x_b, \ell)$ and so they are not physically very meaningful. Nevertheless, we will see that it is possible to derive useful expressions that relate them to macroscopic observables, and which are valid for any choice of $\Pi(x_b, \ell)$ that satisfies the conditions specified above.

In a crystal whose periodicity along the $x-$axis is a (i.e., $\nu(x+a) = \nu(x),\ \forall x$), the $\langle\ .\ \rangle^*$ average is unnecessary because, by choosing $\Delta_m = a,\ \forall m$, all microintervals are identical and so

$$\begin{aligned}\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b) &= \mathcal{M}^{\langle n\rangle}_{\Delta\nu}(x_b + a/2, a)\\ &= \int_{-a/2}^{a/2} \Delta\nu(x_b + a/2)\, u^n\ \mathrm{d}u\,.\end{aligned}$$

H. Surface excess

Equation 31 can be written as

$$\mathcal{S}^{[\Delta\nu]}_r(x_b) = \int_{x_L}^{\infty} x^r\, \Delta\nu(x) F(x - x_b)\,\mathrm{d}x$$

where $F(x) = 1 - H(x) = H(-x)$ is one for $x < 0$ and zero for $x > 0$, and $H(x) \equiv \frac{\mathrm{d}}{\mathrm{d}x}\max\{x, 0\}$ is the Heaviside step function. The mesoscale average of $\mathcal{S}^{[\Delta\nu]}_r(x_b)$ is

$$\begin{aligned}\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_r &\equiv \int_{-\infty}^{\infty} \mu(x' - x_b; \varepsilon_x)\left(\int_{x_L}^{x'} x^r\, \Delta\nu(x)\,\mathrm{d}x\right)\mathrm{d}x'\\ &= \int_{-\ell/2}^{\ell/2} \mu(u; \varepsilon_x)\left(\int_{x_L}^{x_b - \ell/2} x^r\, \Delta\nu(x)\,\mathrm{d}x\right)\mathrm{d}u\\ &+ \int_{-\ell/2}^{\ell/2} \mu(u; \varepsilon_x)\left(\int_{x_b-\ell/2}^{x_b + u} x^r\, \Delta\nu(x)\,\mathrm{d}x\right)\mathrm{d}u\\ &= \bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{r,s}(x_b) + \bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{r,b}(x_b)\end{aligned}$$

where we have assumed that $x_b > x_L + \ell/2$ and we have split the mesoscale average, $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_r$, into the sum of a 'surface' term, $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{r,s}(x_b)$, and a 'bulk' term, $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{r,b}(x_b)$, which can also be expressed as

$$\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{r,s}(x_b) \equiv \int_{x_L}^{x_b - \ell/2} x^r\, \Delta\nu(x)\,\mathrm{d}x \tag{35}$$

and

$$\bar{\mathcal{S}}^{[\Delta\nu]}_{r,b}(x_b) = \int_{-\ell/2}^{\ell/2} (x_b + u)^r\, \Delta\nu(x_b + u)\mathcal{F}_\mu(u; \varepsilon_x)\,\mathrm{d}u\,, \tag{36}$$

where $\mathcal{F}_\mu(u; \varepsilon_x) \equiv \int_{-\infty}^{\infty} F(u - u')\mu(u'; \varepsilon_x)\,\mathrm{d}u'$ decays smoothly from a value of almost one at $u = -\ell/2$ to almost zero at $u = \ell/2$. Both its average value and its value at $u = 0$ are one half and its derivative is $\mathcal{F}^{(1)}_\mu(u; \varepsilon_x) = -\mu(u; \varepsilon_x)$. The split of $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_r$ into bulk and surface terms is not unique: both terms are microscopic functions of x_b, which is an arbitrarily-chosen point in the bulk. However, we will find that their sum is independent of x_b.

Now let us split the integral in Eq. 36 into a sum of integrals over the microintervals, $\mathfrak{I}_m \equiv \mathfrak{I}(\bar{x}_m, \Delta_m)$. We can again exploit the slowness of the variation of μ and $\mathcal{F}_\mu$ on the microscale, when $\mathfrak{a}/\varepsilon_x$ is very small, to replace $\mathcal{F}_\mu(x - x_b; \varepsilon_x)$ in each microinterval by its Taylor expansion about the microinterval midpoint. If $\mathfrak{a}/\ell$ is sufficiently small, we can discard the second and higher-order terms, which involve first- and higher-order derivatives of μ. Therefore, we get

$$\begin{aligned}\bar{\mathcal{S}}^{[\Delta\nu]}_{r,b}(x_b) = \sum_m \Bigg[&\mathcal{F}_\mu(\bar{x}_m - x_b; \varepsilon_x)\int_{-\Delta_m/2}^{\Delta_m/2} (\bar{x}_m + u)^r \Delta\nu(\bar{x}_m + u)\,\mathrm{d}u\\ &- \mu(\bar{x}_m - x_b; \varepsilon_x)\int_{-\Delta_m/2}^{\Delta_m/2} u(\bar{x}_m + u)^r \Delta\nu(\bar{x}_m + u)\,\mathrm{d}u\Bigg]\end{aligned} \tag{37}$$

1. Case I: $\boldsymbol{\mathcal{S}}_0^{[\boldsymbol{\Delta\nu}]}$

Setting $r = 0$ in Eq. 37 gives

$$\begin{aligned}\bar{\mathcal{S}}^{[\Delta\nu]}_{0,b}(x_b; \varepsilon_x) &= \sum_m \Delta_m\, \mathcal{F}_\mu(\bar{x}_m - x_b; \varepsilon_x)\, \mathcal{M}^{\langle 0\rangle}_{\Delta\nu}(\bar{x}_m, \Delta_m)\\ &- \sum_m \Delta_m\, \mu(\bar{x}_m - x_b; \varepsilon_x)\, \mathcal{M}^{\langle 1\rangle}_{\Delta\nu}(\bar{x}_m, \Delta_m)\end{aligned}$$

Assuming that the microstructure is the same everywhere in a mesoscopic neighbourhood of x_b, the average of the microintervals' n^{th} moment density on every sufficiently-wide subinterval of $[x_b - \ell/2, x_b + \ell/2]$ should be the same and equal to $\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b)$ in the limit $\mathfrak{a}/\varepsilon_x \to 0$. Therefore the first term on the right hand side is simply equal to $(\ell/2)\bar{\mathcal{M}}^{\langle 0\rangle}_{\Delta\nu} = (\ell/2)\overline{\Delta\nu}(x_b)$, and

$$\begin{aligned}\bar{\mathcal{S}}^{[\Delta\nu]}_{0,b}(x_b; l) &= \frac{\ell}{2}\,\overline{\Delta\nu}(x_b) - \left\langle \mathcal{M}^{\langle 1\rangle}_{\Delta\nu};\mu\right\rangle^*_{\varepsilon_x}(x_b)\\ &= \int_{x_b-\ell/2}^{x_b}\nu(x)\,\mathrm{d}x - \bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b).\end{aligned} \tag{38}$$

Adding $\bar{\mathcal{S}}^{[\Delta\nu]}_{0,s}(x_b)$ and identifying the macroscopic quantity

$$\boldsymbol{\mathcal{S}}^{[\boldsymbol{\Delta\mathcal{V}}]}_0(\mathbf{x}_b) = \int_{\mathbf{x}_L}^{\mathbf{x}_b}\boldsymbol{\Delta\mathcal{V}}(\mathbf{x})\,\mathbf{dx}$$

as $\bar{\mathcal{S}}^{[\Delta\nu]}_0 = \bar{\mathcal{S}}^{[\Delta\nu]}_{0,s}(x_b)+\bar{\mathcal{S}}^{[\Delta\nu]}_{0,b}(x_b)$, we find that

$$\boldsymbol{\mathcal{S}}^{[\boldsymbol{\Delta\mathcal{V}}]}_0(\mathbf{x}_b) = \int_{x_L}^{x_b}\Delta\nu(x)\,\mathrm{d}x - \bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b) \tag{39}$$

Note that $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{0,s}(x_b) = \mathcal{S}^{[\Delta\nu]}_0(x_b)$, which suggests that $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{0,b}(x_b) = -\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b)$ can be viewed as a correction to $\mathcal{S}^{[\Delta\nu]}_0(x_b)$ that removes its sensitivity to x_b.

2. Case II: $\boldsymbol{\mathcal{S}}^{[\boldsymbol{\Delta\mathcal{V}}]}_1$ when $\boldsymbol{\mathcal{V}}(\boldsymbol{x_b}) = \mathbf{0}$

Returning to Eq. 37, setting $r = 1$, and using the fact that, for all m,

$$\mathcal{M}^{\langle0\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m) = \frac{1}{\Delta_m}\int_{x_m^-}^{x_m^+}\Delta\nu(x)\,\mathrm{d}x = \overline{\Delta\nu}(x_b) = 0,$$

we find that

$$\begin{aligned}\bar{\mathcal{S}}^{[\Delta\nu]}_{1,b}(x_b;\varepsilon_x) = \sum_m \Delta_m\Big\{&\mathcal{M}^{\langle1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\Big[\mathcal{F}_\mu(\bar{x}_m - x_b;\varepsilon_x)\\ &- \bar{x}_m\,\mu(\bar{x}_m - x_b;\varepsilon_x)\Big] - \mathcal{M}^{\langle2\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\,\mu(\bar{x}_m-x_b;\varepsilon_x)\Big\}\end{aligned}$$

As in Sec. VII.H.1, the $\mathcal{M}^{\langle n\rangle}_{\Delta\nu}\,\mathcal{F}_\mu$ term on the right hand side is equal to $(\ell/2)\,\bar{\mathcal{M}}^{(n)}_{\nu}(x_b)$, with $n = 1$ in this case.

$$\begin{aligned}\bar{\mathcal{S}}^{[\Delta\nu]}_{1,b}(x_b;\varepsilon_x) &= \frac{\ell}{2}\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b) - \bar{\mathcal{M}}^{\langle2\rangle}_{\Delta\nu}(x_b)\\ &-\sum_m \Delta_m\,\mu(\bar{x}_m - x_b;\varepsilon_x)\,\bar{x}_m\,\mathcal{M}^{\langle1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\end{aligned} \tag{40}$$

The third term on the right hand side is $-\left\langle x\,\mathcal{M}^{\langle1\rangle}_{\Delta\nu};\mu\right\rangle^*_{\varepsilon_x}(x_b)$. Subtracting $x_b\left\langle\mathcal{M}^{\langle1\rangle}_{\Delta\nu};\mu\right\rangle^*_{\varepsilon_x}(x_b)$ from the first term and adding it to the third term gives

$$\begin{aligned}\bar{\mathcal{S}}^{[\Delta\nu]}_{1,b}(x_b;\varepsilon_x) &= -\left(x_b - \frac{\ell}{2}\right)\left\langle\mathcal{M}^{\langle1\rangle}_{\Delta\nu};\mu\right\rangle^*_{\varepsilon_x}(x_b)\\ &- \left\langle\mathcal{M}^{\langle2\rangle}_{\Delta\nu};\mu\right\rangle^*_{\varepsilon_x}(x_b) - \left\langle(x-x_b)\mathcal{M}^{\langle1\rangle}_{\Delta\nu};\mu\right\rangle^*_{\varepsilon_x}(x_b)\end{aligned}$$

It can be shown that the uniformity of the microstructure on the mesoscale implies that the third term scales like $\mathfrak{a}/\varepsilon_x$ when $\varepsilon_x \gg \mathfrak{a}$. This is because the distribution of microinterval moment densities is the same on either side of x_b, but the sign of $(x - x_b)$ is different. Therefore, the contributions to this term from $(x_b - \ell/2, x_b)$ and $(x_b, x_b + \ell/2)$ cancel one another. Assuming that $\varepsilon_x \gg \mathfrak{a}$, we get

$$\bar{\mathcal{S}}^{[\Delta\nu]}_{1,b}(x_b;\varepsilon_x) = -\left(x_b - \frac{\ell}{2}\right)\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b) - \bar{\mathcal{M}}^{\langle2\rangle}_{\Delta\nu}(x_b) \tag{41}$$

Now, because $\mathcal{M}^{\langle0\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m) = 0$, we can write

$$\begin{aligned}\int_{x_b-\ell/2}^{x_b} x\,\Delta\nu(x)\,\mathrm{d}x &= \sum_{\bar{x}_m<x_b}\int_{-\Delta_m/2}^{\Delta_m/2}u\,\Delta\nu(\bar{x}_m+u)\,\mathrm{d}u\\ &= \sum_{\bar{x}_m<x_b}\Delta_m\mathcal{M}^{\langle1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m) = \frac{\ell}{2}\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b)\end{aligned}$$

Therefore, adding $\bar{\mathcal{S}}^{[\Delta\nu]}_{1,s}(x_b;\varepsilon_x)$ to Eq. 41 gives

$$\begin{aligned}\boldsymbol{\mathcal{S}}^{[\boldsymbol{\Delta\mathcal{V}}]}_1(\mathbf{x}_b) &= \int_{\mathbf{x}_L}^{\mathbf{x}_b}\mathbf{x}\,\boldsymbol{\Delta\mathcal{V}}(\mathbf{x})\,\mathbf{dx}\\ &= \int_{x_L}^{x_b}x\,\Delta\nu(x)\,\mathrm{d}x - x_b\,\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}(x_b) - \bar{\mathcal{M}}^{\langle2\rangle}_{\Delta\nu}(x_b)\end{aligned} \tag{42}$$

As with $\boldsymbol{\mathcal{S}}^{[\boldsymbol{\Delta\mathcal{V}}]}_0$, we can write $\boldsymbol{\mathcal{S}}^{[\boldsymbol{\Delta\mathcal{V}}]}_1$ as an x_b-independent sum of an x_b-dependent surface term, $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{1,s}(x_b)$, which is simply the original microscopically-varying integral $\mathcal{S}^{[\Delta\nu]}_1(x_b)$, and an x_b-dependent bulk term, $\bar{\mathcal{S}}^{\langle\Delta\nu\rangle}_{1,s}(x_b)$, which can be viewed as a correction that removes the strong dependence on the arbitrarily-chosen position x_b.

3. Idempotency of the mesoscale average

In Sec. VI.I we assumed, implicitly, that the mesoscale averaging operation is not idempotent. This allowed us to deduce that there is a trade-off between spatial precision/uncertainty and the precision/uncertainty of macroscopic fields and their derivatives. However, the uncertainty relations were derived under a 'first approximation', and are far from exact. Throughout Sec. VII we have assumed that we are much closer to the limit $\mathfrak{a}/\varepsilon_x \to 0$, and therefore closer to the limit in which the averaging operation is idempotent. Bearing this in mind, let us consider one important consequence of idempotency.

Idempotency of the averaging operation would allow the following deduction to be made about the mesoscale averages, $\boldsymbol{\mathcal{M}}^{\langle\mathbf{1}\rangle}_{\boldsymbol{\Delta\nu}}$ and $\boldsymbol{\mathcal{M}}^{\langle\mathbf{2}\rangle}_{\boldsymbol{\Delta\nu}}$, of $\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}$ and $\bar{\mathcal{M}}^{\langle2\rangle}_{\Delta\nu}$, respectively.

$$\begin{aligned}\left\langle\bar{\mathcal{S}}^{[\Delta\nu]}_0\right\rangle_{\varepsilon_x} = \bar{\mathcal{S}}^{[\Delta\nu]}_0 &\Rightarrow \boldsymbol{\mathcal{M}}^{\langle\mathbf{1}\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b) \equiv \left\langle\bar{\mathcal{M}}^{\langle1\rangle}_{\Delta\nu}\right\rangle_{\varepsilon_x}(x_b) = 0\\ \left\langle\bar{\mathcal{S}}^{[\Delta\nu]}_1\right\rangle_{\varepsilon_x} = \bar{\mathcal{S}}^{[\Delta\nu]}_1 &\Rightarrow \boldsymbol{\mathcal{M}}^{\langle\mathbf{2}\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b) \equiv \left\langle\bar{\mathcal{M}}^{\langle2\rangle}_{\Delta\nu}\right\rangle_{\varepsilon_x}(x_b) = 0\end{aligned}$$

The finding that both $\boldsymbol{\mathcal{M}}^{\langle 1\rangle}_{\boldsymbol{\Delta\nu}}$ and $\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}$ are zero would have some very important consequences. Therefore, guided by the knowledge that they vanish in the idempotent limit ($\mathfrak{a}/\varepsilon_x \to 0,\ \varepsilon_x/\mathfrak{L} \to 0$), I show that they vanish without assuming idempotency in Appendix E. The importance of them vanishing will become clear in Sec. XIII. The idempotency limit is the limit $\mathfrak{a}/\varepsilon_x \to 0,\ \varepsilon_x/\mathfrak{L} \to 0$, whereas the limit in which they vanish is $\mathfrak{a}/\varepsilon_x \to 0$.

4. Mesoscale average over the lower limit of an integral

Either by following similar procedures to those that led to Eq. 39 and Eq. 42, or by invoking symmetry, one can find the following expressions for mesoscale averages of integrals in which the average is performed over the lower bound, x_b, of the integrals from x_b to x_r, where $x_r > x_b$.

$$\int_{\mathbf{x}_b}^{\mathbf{x}_R} \boldsymbol{\Delta\mathcal{V}}(\mathbf{x})\,\mathbf{dx} = \left\langle \int_{x_b}^{x_r} \Delta\nu(x)\,\mathrm{d}x\,;\mu\right\rangle_{\varepsilon_x}(x_b) = \int_{x_b}^{x_r} \Delta\nu(x)\,\mathrm{d}x + \bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b) \tag{43}$$

$$\int_{\mathbf{x}_b}^{\mathbf{x}_R} \mathbf{x}\,\boldsymbol{\Delta\mathcal{V}}(\mathbf{x})\,\mathbf{dx} = \left\langle \int_{x_b}^{x_r} x\,\Delta\nu(x)\,\mathrm{d}x\,;\mu\right\rangle_{\varepsilon_x}(x_b) = \int_{x_b}^{x_r} x\,\Delta\nu(x)\,\mathrm{d}x + x_b\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b) + \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b) \tag{44}$$

VIII. CHARGE DENSITY ($\boldsymbol{\varrho}$) AND DIPOLE MOMENT DENSITY ($\boldsymbol{\mathcal{P}}$)

In this section I consider the mesoscale averages of charge and dipole moment densities. For simplicity I define the mesoscale average as the simple average introduced and discussed in detail in Sec. VI.

A. Charge density

If we ignore the microscale variability of the mesoscale average, $\bar{\rho}$, of ρ, and the consequent uncertainty, ε_ϱ, in the value of the macroscopic charge density, its definition is simply

$$\boldsymbol{\varrho}(\mathbf{x}) \equiv \bar{\rho}(x) = \frac{1}{\ell}\int_{-\ell/2}^{\ell/2} \rho(x+u)\ \mathrm{d}u \tag{45}$$

where $\ell \gg \mathfrak{a}$.

In the bulk of a crystal, ℓ can be chosen to be an integer multiple of the periodicity, a, where $\rho(x+a) = \rho(x),\ \forall\, x \in \mathfrak{B}$. It is then easy to show that

$$\boldsymbol{\varrho}(\mathbf{x}_b) = \bar{\rho}(x_b) = \frac{1}{a}\int_{-a/2}^{a/2} \rho(x_b+u)\ \mathrm{d}u\,.$$

This vanishes if the crystal is charge-neutral, as expected of $\boldsymbol{\varrho}$. In amorphous materials, if ρ fluctuates microscopically about zero, it is always possible to find microscopic displacements, $\eta_1 \sim a$ and $\eta_2 \sim a$, such that $\int_{-\ell/2+\eta_1}^{\ell/2+\eta_2} \rho(x_b+u)\ \mathrm{d}u = 0$. By expressing the integral in Eq. 45 as $\int_{-\ell/2}^{\ell/2} = \int_{-\ell/2}^{-\ell/2+\eta_1} + \int_{-\ell/2+\eta_1}^{\ell/2+\eta_2} - \int_{\ell/2}^{\ell/2-\eta_2}$ it is straightforward to show that $\bar{\rho}(x_b) = 0 + \mathcal{O}(\mathfrak{a}/\varepsilon_x) \approx 0$. Therefore, to within the finite precision ε_ϱ with which $\boldsymbol{\varrho}$ can be defined, $\boldsymbol{\varrho}$ vanishes in the bulk of any material that is stable when it is electromagnetically isolated, and whose surfaces are locally charge neutral.

B. Dipole moment density

$\mathbf{P}$ has the dimensions of a dipole moment per unit volume, area, and length in three, two, and one dimensions, respectively. Therefore, to define $\mathbf{P}$ within the bulk of a one dimensional material, it seems natural to start from the quantity

$$\mathcal{P}(x,\epsilon) \equiv \frac{1}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2} \rho(x+u)\,u\ \mathrm{d}u\,, \tag{46}$$

which is the dipole moment per unit length of $\mathfrak{I}(x,\epsilon) \subset \mathfrak{B}$ with respect to an origin at x. $\mathcal{P}(x,\epsilon)$ is strongly dependent on both x and ϵ and so it is difficult to attach physical meaning to it. However, it is clearly a microscopically-varying quantity and its mesoscale average is

$$\begin{aligned}\bar{\mathcal{P}}(x_b) &= \frac{1}{\ell}\int_{x_b-\ell/2}^{x_b+\ell/2}\left(\frac{1}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2}\rho(x+u)\,u\ \mathrm{d}u\right)\ \mathrm{d}x \\ &= \frac{1}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2} u\left(\frac{1}{\ell}\int_{x_b-\ell/2}^{x_b+\ell/2}\rho(x+u)\ \mathrm{d}x\right)\ \mathrm{d}u \\ &= \frac{1}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2} u\,\bar{\rho}(x_b+u)\,\mathrm{d}u \\ &= \frac{\bar{\rho}(x_b)}{\epsilon}\int_{-\epsilon/2}^{\epsilon/2} u\ \mathrm{d}u + \mathcal{O}(\mathfrak{a}/\varepsilon_x) \approx 0,\end{aligned} \tag{47}$$

where, by using $\bar{\rho}(x_b + u) = \bar{\rho}(x_b) + \mathcal{O}(\mathfrak{a}/\varepsilon_x)$, I am assuming that $\bar{\rho}$ fluctuates microscopically but does not change systematically on length scale ϵ. Therefore, the mesoscale average $\bar{\mathcal{P}}$ of $\mathcal{P}$ is negligible when $\mathfrak{a}/\varepsilon_x$ is sufficiently small, regardless of the value of $\bar{\rho}$.

This result generalises to three dimensions, where it can be shown that each Cartesian component of the mesoscale average of the dipole moment per unit volume of a region of arbitrary shape scales like $\mathfrak{a}/\varepsilon_x$. This is a generalisation to non-crystalline materials of the well known result that, in a crystal, the average over all choices of unit cell of the dipole moment per unit cell is zero (Resta and Vanderbilt, 2007).

These results suggest that $\boldsymbol{\mathcal{P}} \equiv \bar{\mathcal{P}}$ is not a useful macroscopic quantity with which to characterise the bulk of a material because it does not distinguish between different mesoscopically-uniform materials, or even between a material and empty space. We can only identify the macroscopic polarization $\mathbf{P}$ as $\boldsymbol{\mathcal{P}}$ if we are willing to accept that $\mathbf{P} = 0$ in *every* mesoscopically-uniform material, regardless of its microstructure.

IX. INTERLUDE

In the sections that follow I discuss several quantities that are commonly regarded as manifestations, or consequences, of either the $\mathbf{P}$ field itself or of its value changing. They include surface charge $\boldsymbol{\sigma}$ and bound charge $\boldsymbol{\varrho}^{\text{bound}}$ (Sec. X), polarization current $\mathbf{J}^{(p)}$(Sec. XI) and the macroscopic ($\vec{k} = 0$) electric field $\mathbf{E}$ (Sec. XIII.D).

Finnis's work (Finnis, 1998) and Sec. VII make it easy to write down an expression for $\boldsymbol{\sigma} = \boldsymbol{\sigma}[\rho]$, which is a linear functional of ρ. Its linearity means that, if ρ can be decomposed as $\rho = \sum_i \rho_i$, where each ρ_i is either nonnegative or nonpositive, this becomes $\boldsymbol{\sigma}[\rho] = \sum_i \boldsymbol{\sigma}[\rho_i]$. It follows immediately that, when ρ changes continuously in response to a slowly varying stimulus, and if the set $\{\rho_i\}$ of charge packets is chosen such that each one changes continuously as ρ changes, but its integral remains constant, then the polarization current can be expressed as the sum, $\dot{\boldsymbol{\sigma}} = \mathbf{J}^{(p)}[\dot{\rho}] = \sum_i \mathbf{J}^{(p)}[\dot{\rho}_i]$.

If the widths of the charge packets are microscopic, their shapes are irrelevant to macroscale observables because homogenization transforms each packet into a point charge. Therefore the contribution of each packet ρ_i to $\mathbf{J}^{(p)}$ can be calculated from the time derivatives of its integral, $q_i = \int \rho_i \, \mathrm{d}x$, and its center, $x_i = q_i^{-1} \int x \, \rho_i \, \mathrm{d}x$. If the packets can be chosen such that the integral of each one is time-invariant ($\dot{q}_i = 0$), we can use the MTOP to calculate $\mathbf{J}^{(p)}$ from the evolving *bulk* microstructure.

It follows immediately from the results stated in Sec. VII.H.3, and proved in Appendix E, that the macroscopic potential $\boldsymbol{\Phi}$ is zero in an isolated macroscopically-uniform material whose surfaces are not charged. It follows from this that a macroscopic $\mathbf{E}$ field cannot exist in such a material. Nevertheless, in Sec. XIII.D I prove this by expressing $\boldsymbol{\Phi}$ in terms of the microstructure ρ using the results of Sec. VII. In Sec. XIII.D I point out a fatal flaw in the cavity construction introduced by Lorentz to relate the macroscopic $\mathbf{E}$ field to $\mathbf{P}$, and in Sec. XIV I refute Bethe's derivation of his approximate expression for the mean inner potential.

I conclude that neither $\mathbf{P}$ nor the negative of its spatial derivative $\boldsymbol{\varrho}^{\text{bound}}$ are required elements of electromagnetic theory. I show that the quantization and multivaluedness of $\mathbf{P}$ found within the MTOP are consequences of requiring that $\mathbf{P}$ be a property of the bulk and of defining the excess charge at a surface as $\boldsymbol{\sigma}^{\text{bound}} = \mathbf{P} \cdot \hat{n}$. As Fig. 11 illustrates, the value of $\boldsymbol{\sigma}^{\text{bound}}$ depends on how the surface is terminated (e.g., on a plane of net positive charge or on a plane of net negative charge). It follows that both $\boldsymbol{\sigma}^{\text{bound}}$ and $\mathbf{P}$ must be multivalued unless the excess surface charge is defined as $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\text{bound}} + \boldsymbol{\sigma}^{\text{free}}$, where $\boldsymbol{\sigma}^{\text{free}}$ takes full account of the dependence of $\boldsymbol{\sigma}$ on surface termination.

If it is accepted that $\mathbf{P}$ is an unnecessary element of the theory, the importance of scrapping it should be obvious from its history: It has been interpreted in at least three different ways: as a property of the ether, as a dipole moment density, and as a property of the phase of a material's wavefunction. It can be misleading with regard to physical mechanisms; for example, expressing the potential energy per unit volume as $U = -\mathbf{P} \cdot \mathbf{E}$ suggests that $\mathbf{E}$ couples to the bulk, whereas expressing it as $U = -\boldsymbol{\sigma}\mathbf{E}$ makes clear that it only couples to charges at the surface, initially, and couples to the bulk indirectly by driving charge through it. It can also lead to false conclusions, such as that lack of inversion symmetry implies the existence of a uniform ($\vec{k} = 0 \;\not\Leftrightarrow\; \vec{k} \to 0$) macroscopic $\mathbf{E}$ field in the bulk of a crystal.

X. SURFACE CHARGE (σ)

A charged surface or interface is not stable unless the electric potential from it is compensated by, for example, an oppositely charged surface or interface. The instability of isolated charged surfaces is due to the divergence of the electric potential (see Sec. XIII). If a pristine isolated crystal surface is charged, and therefore unstable unless neutralized by a change in its composition with respect to the bulk, it is classified as *polar*. An important question, about which a great deal has been written (Bristowe *et al.*, 2014, 2011; Finnis, 1998; Goniakowski *et al.*, 2008; Goniakowski and Noguera, 2011, 2014, 2016; Noguera, 2000; Noguera and Goniakowski, 2013; Stengel, 2011; Stengel and Vanderbilt, 2009; Tasker, 1979) is how to determine whether a particular surface is polar or non-polar and to quantify its instability by calculating its surface charge.

A. Calculating surface charge from ρ

It does not seem difficult to intuit the meaning of the surface areal charge density $\boldsymbol{\sigma}$ when one first encounters the concept. However, as soon as one tries to define it, in order to calculate it, difficulties become apparent.

The obvious way to calculate $\boldsymbol{\sigma}$ is simply to integrate the volumetric charge density $\rho(\vec{r})$ from a point above the surface to a point far beneath it. Assuming that the surface is perpendicular to the $x-$axis, and that $\varrho(x)$ is the average of $\rho(\vec{r})$ over the $yz-$plane at x, the obvious definition of the $yz-$averaged areal density of excess surface charge is

$$\sigma_s(x_b) \equiv \int_{x_L}^{x_b} \varrho(x)\ \mathrm{d}x \tag{48}$$

Fig. 11 illustrates why this definition fails. It depicts the 'surface' of a one-dimensional crystal whose microscale charge distribution is a semi-infinite periodic array of alternating positive and negative point charges of magnitude one. For this simple case the integral in Eq. 48 becomes the sum $\sigma_s(x_b) = 1 - 1 + 1 - 1 + \cdots$. Its value is either zero or one, depending on the choice of x_b, and it continues to vary between these values *ad infinitum* as x_b increases. Therefore $\sigma_s(x_b)$ is a microscopic function of x_b and, as a result, Eq. 48 fails as a definition of $\boldsymbol{\sigma}$.

Finnis presented an elegant solution to this problem in Finnis, 1998, and a generalization of his result to amorphous materials is derived by a different route in Sec. VII and Appendix E. I quote and explain the more general result below. I then quote Finnis's result for crystals, which is simpler and easier to relate to the example depicted in Fig. 11.

B. Macroscale surface charge

The problem of how to express $\boldsymbol{\sigma}$ in terms of ρ is easy to resolve once it is realised that $\boldsymbol{\sigma}$ only has meaning at the macroscale. Microscopically, surfaces and interfaces are ill-defined entities because their widths are indeterminate: in the vicinity of a surface, both structure and composition differ from the bulk, in general, but they gradually become more bulk-like with depth. This gradual relaxation means that there is no clear boundary separating surface-like material from bulk-like material.

However, as Fig. 11 illustrates, even if a material could be terminated abruptly at a plane and prevented from changing its local structure (bond lengths and angles) or composition, such that surface structure and composition were identical to the bulk, the concept of a surface charge density simply does not apply at the microscale: The microstructure is defined on a simply connected subset of $\mathbb{R}^3$. One can define an areal charge density on any plane (e.g., $\sigma(y,z;x) \equiv \rho(x,y,z)\,\mathrm{d}x$), but no special surface plane exists.

However, as explained in Sec. VI, the spatial resolution ε_x at the macroscale is unavoidably finite and all points separated by microscopic distances coincide at the macroscale. As a result, the surface region of indeterminate width is contracted to zero width by the homogenization transformation. It becomes a two dimensional manifold.

The mesoscale average $\bar{\rho}$ of ρ at any microscale point whose macroscale image is on this manifold differs significantly (by more than $\varepsilon_\varrho/2$), in general, from its value elsewhere. To understand why, consider the material depicted in Fig. 12. There exist planes (e.g., Plane 4) on which the planar average of ρ does not vanish. It follows that the three dimensional mesoscale average $\bar{\rho}$ at any point (not just points on the charged planes) only vanishes as a result of cancellation of positive and negative contributions whose displacements from the point have components normal to those planes. If all material from one side of such a plane is removed to create a surface, this balance is disrupted and $\bar{\rho}$ becomes finite, in general, at any point within a distance $\varepsilon_x/2$ of the plane. The areal charge density $\boldsymbol{\sigma}$ at a point on the surface manifold is simply the integral of $\bar{\rho}$ over the point's preimage under the homogenization transformation. Therefore it is the integral of $\bar{\rho}$ along on an interval of width ε_x on an axis normal to the surface. The macroscopic volumetric charge density $\boldsymbol{\varrho}$ is simply the average of $\bar{\rho}$ on this interval.

It is logical, then, to define the surface charge as

$$\boldsymbol{\sigma} = \int_{\mathbf{x}_L}^{\mathbf{x}_b} \boldsymbol{\varrho}(\mathbf{x})\,\mathbf{dx}. \tag{49}$$

This integral converges with respect to both of its limits because $\boldsymbol{\varrho} = 0$ in the bulk and in the vacuum above the surface. By substituting the definition of $\boldsymbol{\varrho}$ as the mesoscale average $\bar{\rho}$ of ρ (Eq. 45), it is shown in Sec. VII that

$$\boldsymbol{\sigma} \equiv \int_{x_L}^{x_b} \rho(x)\ \mathrm{d}x - \bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b) \equiv \sigma_s(x_b) + \sigma_b(x_b), \tag{50}$$

where x_b is *any* point deep below the surface and $\bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b) \equiv -\sigma_b(x_b)$ is defined as follows: A mesoscopic neighbourhood of x_b is partitioned into a set of contiguous microscopic intervals $\mathfrak{I}(\bar{x}_m, \Delta_m)$ such that x_b is at a boundary between two of these *microintervals*, and such that the integral of ρ on each microinterval is zero. The second condition is always possible because ρ fluctuates microscopically about zero. The dipole moment density of the m^{th} interval is $\mathcal{M}_\rho^{\langle 1\rangle}(\bar{x}_m,\Delta_m) \equiv \Delta_m^{-1}\int_{-\Delta_m/2}^{\Delta_m/2}\rho(\bar{x}_m+u)\,u\;\mathrm{d}u$ and $\bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b)$ is defined as the average of this quantity over all microintervals in the discrete and finite set that partitions the mesoscale neighbourhood of x_b. It is shown in Appendix E that the value of $\bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b)$ is the same for all sets of microintervals that satisfy the conditions stated above.

Although both $\sigma_s(x_b)$ and $\sigma_b(x_b)$ depend sensitively on x_b, their sum is independent of it. This is easy to see in the special case of a periodic bulk charge density, $\rho(x+a)=\rho(x)$. The points x_b+ma, where $m\in\mathbb{Z}$, can be chosen as the microinterval boundary points. All microintervals are identical, in this case, and Eq. 50 simplifies to Finnis's result:

$$\begin{aligned}\boldsymbol{\sigma} &= \int_{x_L}^{x_b}\rho(x)\;\mathrm{d}x - \frac{1}{a}\int_0^a \rho(x_b+u)\,u\;\mathrm{d}u \\ &= \sigma_s(x_b) - \mathcal{P}(x_b+a/2;a). \qquad (51)\end{aligned}$$

Note that the definitions of $\mathcal{M}_\rho^{\langle 1\rangle}$ and $\mathcal{P}$ are identical. I use $\mathcal{P}$ when it is useful to make clear that it is a dipole moment density. I use $\mathcal{M}_\rho^{\langle 1\rangle}$ when I favour consistency with Sec. VII and with related quantities that will be introduced in Sec. XIII.

Referring again to Fig. 11, and comparing the choices $x_b=x_1$ and $x_b=x_2$, we find that $\sigma_s(x_1)=0$ and $\sigma_s(x_2)=1$. If the minimum distance between positive and negative charges is denoted by b, then $\sigma_b(x_1)=b/a$ and $\sigma_b(x_2)=-(a-b)/a$. Therefore, $\sigma_s(x_1)+\sigma_b(x_1)=\sigma_s(x_2)+\sigma_b(x_2)=b/a$.

It may be illuminating to note that applying the homogenization transformation is a lot like taking a thermodynamic limit; and we can think of macroscopic quantities as thermodynamic quantities which, in a non-ergodic system, can only be defined on macroscopic length scales. Indeed, Finnis's reasoning when deriving Eq. 51 differed slightly from the reasoning outlined above. He reasoned that one should average over an ensemble of terminating planes (x_b), in order to "*reconcile the atomistic picture, in which* [surface] *excesses appear to oscillate on the atomic length scale as a function of the* [surface] *region size, with the thermodynamic picture.*" In the language that I have chosen to use here and in Sec. VI, he found the mesoscale average of the microscopic function $\sigma_s(x_b)$. The same result is found by substituting $\boldsymbol{\varrho}=\bar{\rho}$ into Eq. 49 because, by changing the order of integration, the integral of a mesoscale average becomes the mesoscale average of an integral.

The subscripts of σ_s and σ_b are abbreviations of 'surface' and 'bulk', respectively. σ_s includes all contributions from compositional differences between the surface and the bulk, including charged adsorbates, surpluses or deficits of electrons, charged impurities, and non-stoichiometry associated with reconstructions or coordination defects. On the other hand, σ_b, depends only on the charge density in the bulk and is independent of the surface composition. However, it is simplistic and wrong to view σ_s as the contribution from extrinsic surface charges and σ_b as the contribution from the bulk charge distribution. For example, it is always possible to choose x_b such that $\sigma_b(x_b)=0$ and $\boldsymbol{\sigma}=\sigma_s(x_b)$. Therefore, as well as containing all extrinsic contributions to $\boldsymbol{\sigma}$, σ_s can contain some, or all, of the contribution from the bulk. Choosing $\sigma_b(x_b)=0$ is equivalent to the "dipole-free unit cell" strategy used by Goniakowski *et al.* to deduce surface charge and stability (Goniakowski *et al.*, 2008).

C. Surface Stability

A pristine crystal surface is specified by the structure of the bulk crystal, the surface-plane orientation, and the surface termination. Consider the crystal depicted in Fig. 12 and the two examples given of surfaces of that crystal, which I'll refer to as Surface 1 and Surface 2. These surfaces are defined by the planes (Plane 1 and Plane 2, respectively) at which one could imagine cleaving the perfect crystal. Each plane is defined by an orientation, which can be specified by the outward surface normal $\hat{n}$, and by a position along an axis parallel to $\hat{n}$. The importance of the relative displacements of parallel surface planes is illustrated by the fact that Surface 2 is negatively charged, whereas the surface created by cleaving at Plane 3 would be positively charged. In Fig. 12 I am considering an unphysical frozen surface where I have not allowed any relaxation from the bulk charge density Nevertheless, this unphysical surface suffices to allow the surface polarity to be quantified using Eq. 51.

If we could prepare the frozen surfaces depicted in Fig. 12, when we allowed them to relax in a vacuum they would relax and/or reconstruct. The surface might even melt. However, as long as the bulk crystal did not melt, or the surface didn't banish ions or electrons from it (in the case of Surface 2 we would probably need to apply an electric field to prevent this), none of the structural change near the surface would have any impact on the surface charge, $\boldsymbol{\sigma}$, because the integral of the charge density between the surface and the crystalline bulk, which is the first term in Eq. 51, would be unchanged.

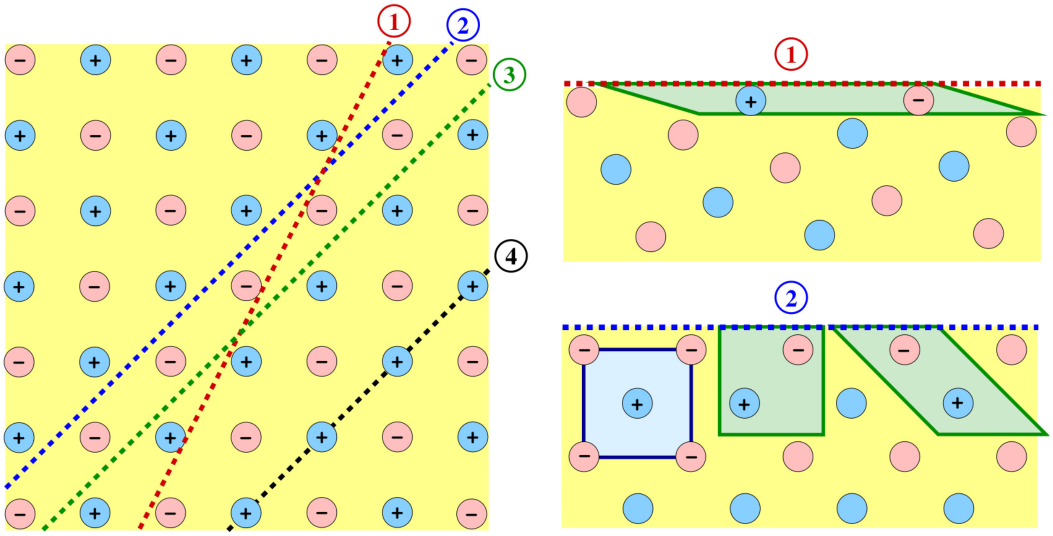

FIG. 12 A lattice of point charges can notionally be cleaved along an infinite number of lattice planes. Here, four lattice planes are shown in cross section as dashed lines. Plane 4 intersects a lattice of positive point charges and is included to illustrate a point made in Sec. X.B. Cleaving along Planes 1, 2, and 3, produce Surfaces 1, 2, and 3, respectively; Surfaces 1 and 2 are shown on the right. Deducing whether a surface is polar is as simple as constructing a unit cell of the crystal (shown in green), which has two lattice vectors that are parallel to the surface plane and one face at the surface, and calculating the dipole moment of this cell. The surface is polar if and only if the projection of the dipole moment onto the surface normal is non-zero. Surface 1 is non-polar but Surface 2 is polar. Surface 3 would be positively charged, despite Planes 2 and 3 being parallel and Surface 2 being negatively charged. Therefore, the orientation of the plane determines whether or not it is polar, but not the value of the surface charge density. One could also construct the dipole-free unit cell (navy boundary) and, if this cannot be constructed with one of its faces at the surface, the integral of the charge density from its uppermost face to the surface is the surface charge density, $\boldsymbol{\sigma}$.

D. Interface charge

By treating an interface between two materials as a pair of adjoined surfaces it is straightforward to show that the interface charge is

$$\boldsymbol{\sigma} = \int_{x_1}^{x_2} \rho(x)\,\mathrm{d}x + \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_1) - \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_2) \tag{52}$$

where x_1 is an arbitrary point in the bulk of the material on one side of the interface and $x_2 > x_1$ is an arbitrary point in the bulk of the material on the other side.

As a sanity check, let us imagine a plane perpendicular to $\hat{x}$ at position x_b in the bulk of a mesocopically-uniform material. This plane can be viewed as an interface between two perfectly-aligned identical materials. The charge density at this imaginary interface is

$$\boldsymbol{\sigma} = \int_{x_b-u}^{x_b+u} \rho(x)\,\mathrm{d}x + \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b-u) - \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b+u)$$

where u can have any value since all points on either side of x_b are in the bulk. For the sake of brevity, I choose the limit $u \to 0^+$, which leads immediately to $\boldsymbol{\sigma} = 0$, as should be the case. It is not too difficult to prove that $\boldsymbol{\sigma} = 0$ for an arbitrary value of u in a non-periodic system.

E. Consistency with the standard model of macroscale electrostatics

Standard treatments of macroscale electrostatics tend to distinguish between a *free* charge density $\varrho^{\text{free}} = \boldsymbol{\nabla}\cdot\mathbf{D}$ and a *bound* charge density $\varrho^{\text{bound}} = -\boldsymbol{\nabla}\cdot\mathbf{P}$, where $\varrho = \epsilon_0\boldsymbol{\nabla}\cdot\mathbf{E} = \boldsymbol{\nabla}\cdot\mathbf{D} - \boldsymbol{\nabla}\cdot\mathbf{P} = \varrho^{\text{bound}} + \varrho^{\text{free}}$. Substituting into Eq. 49 gives

$$\begin{aligned}\boldsymbol{\sigma} &= \overbrace{\int_{\mathbf{x}_L}^{\mathbf{x}_b} \varrho^{\text{free}}(\mathbf{x})\,\mathbf{dx}}^{\boldsymbol{\sigma}^{\text{free}}} \overbrace{- \int_{\mathbf{x}_L}^{\mathbf{x}_b} \boldsymbol{\nabla}\cdot\mathbf{P}\,\mathbf{dx}}^{\boldsymbol{\sigma}^{\text{bound}}} \\ &= \int_{\mathbf{x}_L}^{\mathbf{x}_b} \varrho^{\text{free}}(\mathbf{x})\,\mathbf{dx} + \mathbf{P}\cdot\hat{\mathbf{n}}\end{aligned} \tag{53}$$

and consistency with Eq. 51 requires that

$$\begin{aligned}&\int_{\mathbf{x}_L}^{\mathbf{x}_b} \varrho^{\text{free}}(\mathbf{x})\mathbf{dx} + \mathbf{P}\cdot\hat{\mathbf{n}} \\ &\qquad = \int_{x_L}^{x_b} \rho(x)\,\mathrm{d}x + \boldsymbol{\mathcal{P}}\left(x_b + a/2, a\right)\cdot\hat{\mathbf{n}}.\end{aligned} \tag{54}$$

Both terms on the right hand side depend sensitively on x_b. However, because x_b is arbitrary, both terms on the left hand side must be independent of it if physical meaning can be attributed to them independently of one other.

One way to resolve this is by defining $\boldsymbol{\varrho}^{\text{free}}$ and $\mathbf{P}$ to be the mesoscale averages of ρ and $\boldsymbol{\mathcal{P}}$, respectively. That is, $\boldsymbol{\varrho}^{\text{free}} \equiv \bar{\rho} = \boldsymbol{\varrho}$ and $\mathbf{P} \equiv \bar{\boldsymbol{\mathcal{P}}} = 0$ (see Eq. 47), which implies that $\boldsymbol{\varrho}^{\text{bound}} = 0$ and $\boldsymbol{\sigma}^{\text{bound}} \equiv \mathbf{P} \cdot \hat{\mathbf{n}} = 0$. These definitions preserve consistency between the standard model of electrostatics in dielectrics and the apparently-reasonable definitions of $\boldsymbol{\varrho}$ and $\boldsymbol{\sigma}$ presented herein; namely, $\boldsymbol{\varrho}$ is the mesoscale average of ρ and $\boldsymbol{\sigma}$ is its integral across a surface. However, achieving consistency in this way entails discarding several quantities from the standard model of electrostatics: $\mathbf{D}$, $\mathbf{P}$, and $\boldsymbol{\varrho}^{\text{bound}}$ all vanish and $\boldsymbol{\varrho}^{\text{free}}$ is simply $\boldsymbol{\varrho}$. The macroscopic Maxwell equations are now identical in form to their microscopic counterparts because averaging commutes with differentiation; for example,

$$\epsilon_0 \boldsymbol{\nabla} \cdot \mathcal{E}(x) = \rho(x) \Rightarrow \epsilon_0 \boldsymbol{\nabla} \cdot \mathbf{E}(\mathbf{x}) = \boldsymbol{\varrho}(\mathbf{x}). \tag{55}$$

1. Quantization of $\mathbf{P}$

It is important to consider carefully whether or not the less drastic option of keeping quantities $\mathbf{P}$, $\mathbf{D}$, $\boldsymbol{\varrho}^{\text{bound}}$, and $\boldsymbol{\varrho}^{\text{free}}$ within the macroscale theory is logical or viable. In this section I assume that $\mathbf{P}$ remains an element of the theory and I show that consistency with the definitions of $\boldsymbol{\varrho}$ (Eq. 45) and $\boldsymbol{\sigma}$ (Eq. 49) requires it to be quantized.

The fact that $\mathbf{P}$ would be quantized for a classical crystal in the same way as it is quantized in the MTOP appears to have been appreciated from the beginning (Vanderbilt, 2018; Vanderbilt and King-Smith, 1993). I explain it here for completeness, and also to emphasize it, because it is easy to misinterpret the term *quantum of polarization* as referring to something quantum mechanical.

As before I consider the surface of a pristine perfect crystal whose outward unit normal is $\hat{n}$ and whose structure and composition have not been allowed to change after all the material on one side of the surface plane was removed. I choose the crystal's primitive lattice vectors $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ such that $\mathbf{a}_1 \cdot \hat{n} > 0$, $\mathbf{a}_2 \cdot \hat{n} = \mathbf{a}_3 \cdot \hat{n} = 0$, and $\hat{n} \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = A_\Omega > 0$. The volume of the bulk crystal's unit cell is $\Omega \equiv \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = |\mathbf{a}_1| A_\Omega$. The surface, which is a two dimensional lattice, has primitive lattice vectors $(\mathbf{a}_2, \mathbf{a}_3)$ and the area of its primitive unit cell is A_Ω. Given the surface normal $\hat{n}$, this choice of primitive unit cell of the bulk crystal allows all surface terminations to be identified by a single parameter α, which is the position along $\hat{n}$ at which the uppermost primitive cell is sliced to form the surface. For example, surfaces formed by cleaving at Planes 2 and 3 of Fig. 12 differ only by their values of α. In this simple case the value of α determines only whether the uppermost plane is a plane of cations or a plane of anions. In more general cases the electron density would also be divided; however, it would be unphysical to remove fractions of electrons by removing all density above the termination plane, so I assume that the integral of the density that remains in the uppermost cell is rounded up to a whole number. How this density is distributed has no bearing on the arguments to follow.

The excess bound charge at the surface of the crystal is $\boldsymbol{\sigma}^{\text{bound}} = \mathbf{P}_\perp \equiv \mathbf{P} \cdot \hat{n}$. Within the standard model of electrostatics $\mathbf{P}$ is a bulk quantity and so it must be independent of surface termination α. Therefore $\boldsymbol{\sigma}^{\text{bound}}$ must be the same for all surfaces whose outward normal is $\hat{n}$. However, as discussed in Sec. X.C, and as Fig. 12 illustrates, $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\text{bound}} + \boldsymbol{\sigma}^{\text{free}}$ is not the same for all values of α. One could choose to include all of the α-dependence of $\boldsymbol{\sigma}$ in $\boldsymbol{\sigma}^{\text{free}}$, which would leave $\boldsymbol{\sigma}^{\text{bound}}$ independent of surface termination. However, this is not the approach taken within the MTOP (Stengel, 2011; Vanderbilt and King-Smith, 1993). The MTOP assumes the standard convention that $\boldsymbol{\varrho}^{\text{free}}$ and $\boldsymbol{\sigma}^{\text{free}}$ only contain contributions from charges that are not intrinsic to the material (Ashcroft and Mermin, 1976; Jackson, 1998). As a consequence of preserving this old convention, $\mathbf{P}$ must be quantized (Vanderbilt, 2018; Vanderbilt and King-Smith, 1993). I now prove this.

There are no extrinsic charges in the idealized surfaces constructed; therefore $\boldsymbol{\sigma}^{\text{free}} = 0$ and $\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\text{bound}} = \mathbf{P}_\perp$. Now, because $\boldsymbol{\sigma}$ is known and single-valued, and because it can be changed to the value it would have for any other value of α by adding/removing the same numbers and types of particles (nuclei and electrons) to/from each unit cell of the surface lattice, $\boldsymbol{\sigma}^{\text{bound}}$ must be multivalued. Its set of values must be the set of values of $\boldsymbol{\sigma}$ for every possible choice of surface termination, α. These values differ by integer multiples of e/A_Ω. Therefore, $\mathbf{P}_\perp$ is quantized such that if $\Delta\mathbf{P}$ is the difference between two values of $\mathbf{P}$ that are consistent with Eqs. 45 and 49, then

$$\begin{aligned}\Delta\mathbf{P}_\perp \equiv \Delta\mathbf{P} \cdot \hat{n} = \frac{m_1\, e}{|\mathbf{a}_2 \times \mathbf{a}_3|} = \frac{m_1\, e}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}\, \mathbf{a}_1 \cdot \hat{n} \\ \Rightarrow \Delta\mathbf{P} = \frac{m_1\, e}{\Omega}\, \mathbf{a}_1 + \Delta\mathbf{P}_{\parallel,2}\, \mathbf{a}_2 + \Delta\mathbf{P}_{\parallel,3}\, \mathbf{a}_3\end{aligned}$$

where $\Delta\mathbf{P}_\parallel \equiv \Delta\mathbf{P}_{\parallel,2}\, \mathbf{a}_2 + \Delta\mathbf{P}_{\parallel,3}\, \mathbf{a}_3 = \Delta\mathbf{P} - \Delta\mathbf{P}_\perp\, \hat{n}$, and m_1 is an integer. By considering surfaces perpendicular to $\mathbf{a}_3 \times \mathbf{a}_1$ and $\mathbf{a}_1 \times \mathbf{a}_2$, the same logic would lead me to the following general expression for the difference between any two values of polarization that are consistent with Eqs. 45 and 49.

$$\Delta\mathbf{P} = \frac{e\mathbf{A}}{\Omega} \tag{56}$$

where $\mathbf{A} = m_1\, \mathbf{a}_1 + m_2\, \mathbf{a}_2 + m_3\, \mathbf{a}_3$ is an arbitrary lattice vector and $m_1, m_2, m_3 \in \mathbb{Z}$. This is identical to the quantization of $\mathbf{P}$ deduced within the MTOP, but this derivation makes clear that the quantization of $\mathbf{P}$ is not a quantum mechanical quantization because I have not invoked quantum mechanics to deduce it. It is a consequence of adopting the apparently arbitrary and unnecessary convention that, in the absence of extrinsic charges, $\boldsymbol{\sigma}^{\text{bound}}$ is the total surface charge density, and of preserving the macroscale theory's internal consistency while shoehorning $\mathbf{P}$ into it.

F. Mapping to a set of localized charge packets

In this section I present a result that is pivotal for understanding the relationship between this work, which is founded on a systematic approach to structure homogenization, and the MTOP's definition of polarization current, which is founded on quantum mechanical perturbation theory.

Let us express the charge density as the sum, $\rho(x) = \sum_i \rho_i(x)$, of a set of localized charge packets, $\{\rho_i\}$, where each ρ_i is either nonpositive or nonnegative. The total charge in the i^{th} packet is q_i and its center of charge is x_i. That is,

$$q_i \equiv \int_{-\infty}^{\infty} \rho_i(x)\,\mathrm{d}x\,, \quad x_i \equiv \frac{1}{q_i}\int_{-\infty}^{\infty} \rho_i(x)\,x\;\mathrm{d}x\,.$$

I assume that each ρ_i can be chosen such that it is negligible outside interval $\mathfrak{I}(x_i, \varepsilon_x)$. For a system of nuclei and electrons whose charge density is given by Eq. 3, the nonnegative packets are $\rho_i(x) \equiv Z_i\,e\,\delta(x - X_i)$ and the nonpositive packets are $\rho_i(x) \equiv -e\,n_i(x)$, where $n(x) \equiv \sum_i n_i(x)$ is the electron number density partitioned into a set of packets, $\{n_i\}$.

The localization transformation $\rho_i(x) \to q_i\,\delta(x - x_i)$ conserves charge and preserves ρ_i's center of charge. Therefore, the transformation of ρ into the discrete distribution of point charges $\rho^q(x) \equiv \sum_i q_i\,\delta(x - x_i)$ is an isotropic spatial redistribution of charge. By 'isotropic' I mean that it does not change the center of charge of either ρ or $\tilde{\varrho} \equiv \sum_{i\in\mathcal{I}} \rho_i$, where $\mathcal{I}$ is *any* subset of the set of packet indices. The equitable movement of charge in both directions cannot change the macrostructure if charge is only moved across distances smaller than $\varepsilon_x = \mathbf{dx}$. Therefore, the macroscale counterpart $\boldsymbol{\varrho}^q$ of ρ^q cannot differ from $\boldsymbol{\varrho}$. From this fact, and from Eq. 50, it follows that

$$\boldsymbol{\sigma} = \int_{\mathbf{x}_L}^{\mathbf{x}_b} \boldsymbol{\varrho}^q(\mathbf{x})\,\mathbf{dx} = \sum_{i:x_i<x_b} q_i - \bar{\mathcal{M}}^{\langle 1\rangle}_{\rho^q}(x_b) \tag{57}$$

Note that the derivation of Eq. 39 in Sec. VII assumed that space could be partitioned into microintervals whose net charges were equal. This is always possible for a continuous charge density, but it is not possible for the ρ^q's resulting from every possible partitioning of ρ into packets because the integral of ρ^q on each microinterval is a sum of point charges. If the magnitudes of these charges are irregular it is not possible, in general, to partition space such that each interval's average charge density is precisely the same. It seems likely that a more general derivation, which does not require each microinterval to have the same average charge density, is discoverable.

In the bulk of a crystal with periodicity a, the charge packets can be chosen such that, for any packet ρ_i, whose center is x_i, there are identical packets with centers at $x_i + ma$ for all $m \in \mathbb{Z}$. In this case, the surface charge is

$$\boldsymbol{\sigma} = \sum_{i:x_i<x_b} q_i - \frac{1}{a}\sum_{i:x_i\in(x_b,x_b+a)} q_i\,x_i, \tag{58}$$

where I am assuming that $\sum_{i:x_i\in(x_b,x_b+a)} q_i = 0$. The time derivative of Eq. 58 is the current $\mathbf{J}$. If the conduction current vanishes, it is the polarization current $\mathbf{J}^{(p)}$.

XI. CURRENT (J)

As discussed in Sec. IV.H, unless it is prohibited by symmetry, a polarization current $\mathbf{J}^{(p)}$ flows through the bulk of a material in response to any stimulus $\zeta \to \zeta + \Delta\zeta$, where ζ might be temperature, an externally-applied electric, magnetic, or stress field, or anything else that changes the material's equilibrium or steady-state-nonequilibrium microstructure.

A. Polarization current as a rate of change of surface charge

When all charge that flows through the bulk, by any mechanism, or in response to any stimulus, accumulates at surfaces, the current density deduced from Eq. 50 is simply

$$\mathbf{J} = \dot{\boldsymbol{\sigma}} = \int_{x_L}^{x_b} \frac{\partial\rho}{\partial t}\bigg|_{(x,t)} \mathrm{d}x - \frac{\partial\bar{\mathcal{M}}^{\langle 1\rangle}_{\rho}}{\partial t}\bigg|_{(x_b,t)} \tag{59}$$

Therefore, the current can be calculated if the time-dependent charge density $\rho(x,t)$ (or $\boldsymbol{\varrho}$) is known everywhere. However, the amount of charge that can flow in an isolated material is limited and surface microstructures tend to be more difficult to calculate than microstructures in the bulk of crystalline materials. Therefore, we would like to be able to calculate $\mathbf{J}^{(p)}$ from the evolving equilibrium bulk microstructure.

If we can partition the electron density of a crystal into packets, we can use Eq. 58 to express the current density as

$$\mathbf{J} = \dot{\boldsymbol{\sigma}} = \sum_{i:x_i<x_b} \dot{q}_i - \frac{1}{a}\sum_{i:x_i\in(x_b,x_b+a)} (\dot{q}_i x_i + q_i \dot{x}_i)\,, \tag{60}$$

where each pair (x_i, q_i) is either the position and charge of a nucleus or the center and $-e$ times the integral of a packet n_i of electron density; and where x_b has been chosen to not coincide with any of the x_i's. Let us denote $(x_b, x_b + a)$, which is a primitive unit cell of the crystal, by Ω; let us denote the sum of all n_i whose centers are in Ω by $n_\Omega(x,t) = \sum_{i:x_i\in\Omega} n_i(x,t)$; and let us denote the charge distribution of the nuclei in Ω by $\rho^+_\Omega(x,t)$. Finally, let us assume that the packets have been chosen such that the integral of $\rho_\Omega = \rho^-_\Omega + \rho^+_\Omega$, where $\rho^-_\Omega(x,t) = -en_\Omega(x,t)$, is zero. Then, since the charge density in each primitive cell in the bulk is identical, the electron density in the bulk must be $n(x,t) = \sum_{m\in\mathbb{Z}} n_\Omega(x + ma, t)$, where $a = |\Omega|$ is the lattice constant and ma is a lattice vector. We have then realised the situation described in the discussion of Fig. 3 in Sec. V: we have partitioned the charge density of the crystal's bulk into a set of identical neutral charge densities that are displaced from one another by lattice vectors. Then, because the bulk must remain charge neutral, the integral (M_e) of n_Ω remains constant and the polarization current is given by

$$\begin{aligned}\mathbf{J}^{(p)} &= -\frac{1}{a}\sum_{i:x_i\in\Omega}\left(\dot{q}_i x_i + q_i \dot{x}_i\right) \qquad (61)\\ &= -\frac{M_e e}{a}\left(\dot{X}^+_\Omega - \dot{X}^-_\Omega\right) = -\frac{\dot{d}_\Omega}{a} = \dot{\boldsymbol{\mathcal{P}}}(x_b + a/2, a)\cdot\hat{\mathbf{n}}\end{aligned}$$

where X^+_Ω and X^-_Ω are the centers of charge of ρ^+_Ω and ρ^-_Ω, respectively; and I have used the fact that the sum of all the positive q_i's in Ω and the sum of all the negative q_i's in Ω are both time independent and equal to $M_e e$ and $-M_e e$, respectively. If each packet n_i has an integral that remains constant, the polarization current can also be expressed as

$$\mathbf{J}^{(p)} = -\frac{1}{a}\sum_{i:x_i\in\Omega} q_i \dot{x}_i \qquad (62)$$

and when the integral of each n_i is one (× spin degeneracy), this is equivalent to the MTOP definition of $\mathbf{J}^{(p)}$.

The generalization of Eq. 61 to amorphous materials is

$$\mathbf{J}^{(p)} = -\frac{1}{\ell}\sum_{i:x_i\in\mathfrak{I}(x_b,\ell)}\left(\dot{q}_i x_i + q_i \dot{x}_i\right). \qquad (63)$$

Although the result for crystalline systems appears to be exact and precise, there are variations ($\sim a/\ell$) in the value for amorphous systems with the choices of x_b and ℓ because of differences in the averages of $\dot{q}_i x_i + q_i \dot{x}_i$ on different intervals. Furthermore, because the net charge $Q = \sum_{i:x_i\in\mathfrak{I}(x_b,\ell)} q_i$ of interval $\mathfrak{I}(x_b,\ell)$ is not, in general, zero, the value of $\mathbf{J}^{(p)}$ calculated from this expression has an origin dependence unless Q is constant. In practice, it may be easier to find a set of packets n_i whose integrals are constant ($\Rightarrow \dot{q}_i = 0$) and to calculate $\mathbf{J}^{(p)} \equiv -(1/\ell)\sum_{i:x_i\in\mathfrak{I}(x_b,\ell)} q_i \dot{x}_i$.

If we define the conduction current as $\mathbf{J}^{(c)} \equiv \mathbf{J} - \mathbf{J}^{(p)}$, it is simply equal to $-e$ times the sum of the rates of change of the integrals of the packets n_i with centers in the surface region, i.e.,

$$\mathbf{J}^{(c)} = \sum_{i:x_i<x_b} \dot{q}_i. \qquad (64)$$

Regardless of the choice of x_b, this does not contain any contribution from bulk-like primitive cells, because the sum of the integrals of the packets in each bulk-like cell is constant and equal to M_e.

B. $\mathcal{H}$-representability and $\mathcal{H}(t)$-representability of n

I have established that $\mathbf{J}^{(p)}$ can be calculated from any evolving bulk microscopic charge density $\rho(x,t)$ that can can be expressed as a sum $\sum_i \rho_i$ of moving packets of fixed amounts charge, each of which is either non-positive or non-negative. When such a representation exists, it is not unique because, for example, one can always add to any given representation a co-moving pair of packets of equal and opposite charge, or combine multiple packets into a single packet.

Clearly the distribution of nuclear charge admits such a representation, so in this section I focus on electrons.

1. Electrons

I say that a number density $n(x)$ is *$\mathcal{H}$-representable* if there exists a projector $\hat{\mathrm{P}}$ onto a Hilbert space of dimension $N_e = \int n$ such that $n(x) = \langle x|\hat{\mathrm{P}}|x\rangle$. I say that a number density $n(x,t)$ is *$\mathcal{H}(t)$-representable* (*'Ht representable'*) if it is $\mathcal{H}$-representable at all relevant times t and if its time-dependent projector $\hat{\mathrm{P}}(t)$ evolves smoothly with t.

It is known that the ground state electron density of any material is either noninteracting v-representable or arbitrarily close to a noninteracting v-representable density (van Leeuwen, 2003). This means that it can be represented as a set of packets $n_i = |\varphi_i|^2$ of integral one (two for spin-degenerate packets), where the φ_i's are the lowest-eigenvalue eigenstates of a single electron Hamiltonian, $\hat{h}$.

I make an adiabatic approximation by assuming that the ground state density's time dependence can be expressed as a parametric dependence on a slowly- and smoothly-

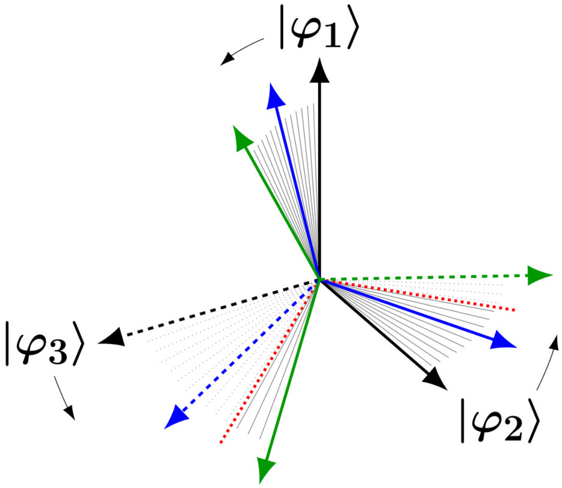

FIG. 13 Consider the ground state number density $n(x;\zeta)$ of two spin-zero electrons in a static confining potential $v^{\text{ext}}(x;\zeta) = -e\phi(x;\zeta)$, where ζ is some physical parameter (e.g., $\mathbf{E}_{\text{ext}}$). Since $n(\zeta): \mathbb{R} \to \mathbb{R}_+; x \mapsto n(x;\zeta)$ is arbitrarily close to a non-interacting v-representable density (van Leeuwen, 2003), it can be represented as a linear subspace $\mathcal{H}_\zeta \equiv \text{span}\{|\varphi_1(\zeta)\rangle, |\varphi_2(\zeta)\rangle\}$ of the Hilbert-Lebesgue space $\mathcal{L}^2(\mathbb{R}) = \text{span}\{|\varphi_i(\zeta)\rangle\}_{i=1}^{\infty}$, where $\{|\varphi_i(\zeta)\rangle\}_{i=1}^{\infty}$ are the orthonormal eigenvectors of operator $\sum_{i=1}^{\infty} \epsilon_i |\varphi_i(\zeta)\rangle\langle\varphi_i(\zeta)|$; $\epsilon_1, \epsilon_2 \in \mathbb{R}$ are its lowest eigenvalues; and $n(x;\zeta) = \langle x|\hat{\mathrm{P}}_\zeta|x\rangle$, where $\hat{\mathrm{P}}_\zeta = |\varphi_1(\zeta)\rangle\langle\phi_1(\zeta)| + |\phi_2(\zeta)\rangle\langle\phi_2(\zeta)|$ is the projector onto $\mathcal{H}_\zeta$. $|\varphi_1(\zeta)\rangle$ and $|\varphi_2(\zeta)\rangle$ are represented above, for some value of ζ, by solid black arrows and $|\varphi_3(\zeta)\rangle$, which is in the orthogonal complement $\mathcal{H}_\zeta^\perp$ of $\mathcal{H}_\zeta$, is the black dashed arrow. If ζ begins to vary continuously, and if $n(\zeta)$'s response to this variation is non-singular, $\mathcal{H}_\zeta$ changes continuously by rotating within $\mathcal{L}^2(\mathbb{R})$. The directions of $|\varphi_1(\zeta)\rangle$ and $|\varphi_2(\zeta)\rangle$ change continuously (from black to blue) by mixing with vectors from $\mathcal{H}_\zeta^\perp$. However there may exist a critical value ζ_c (dotted red lines) at which ϵ_3 becomes lower than ϵ_2, causing $\mathcal{H}_\zeta$ to change abruptly to $\text{span}\{|\varphi_1(\zeta)\rangle, \varphi_3(\zeta)\}$. At such a point there is a discontinuous redistribution of electron density in $\mathbb{R}$. Before the dotted red line is reached the current is $-e(\dot{x}_1 + \dot{x}_2)$, where $\dot{x}_i$ denotes the time derivative of the center of $|\varphi_i(x;\zeta)|^2 = |\langle x|\varphi_i(\zeta)\rangle|^2$. In a crystal the dimension of $\mathcal{H}_\zeta$ is very large ($\sim 10^{24}$), but its basis can be transformed to one for which there are the same number of centers in each primitive unit cell. Then, if the density's response is continuous, the current in the bulk can be calculated from the velocities of the centers in a single unit cell. This is the essence of the MTOP method of calculating $\mathbf{J}^{(p)}$. When the density's response is singular, multiple basis vectors can be exchanged between $\mathcal{H}_\zeta$ and $\mathcal{H}_\zeta^\perp$ in less time than it takes for electrons to respond. In that case the MTOP approach fails because, for example, if $|\varphi_i(\zeta)\rangle$ and $|\varphi_j(\zeta)\rangle$ are replaced in $\mathcal{H}_\zeta$'s basis by $|\varphi_k(\zeta)\rangle$ and $|\varphi_l(\zeta)\rangle$, the value of $\mathbf{J}$ calculated by assuming that the density at x_i was displaced by $x_k - x_i$ to x_k and the density at x_j was displaced by $x_l - x_j$ to x_l, would differ, in general, from the value calculated by assuming that the densities at x_i and x_j were displaced to x_l and x_k, respectively.

varying stimulus $\zeta(t)$. I express it as

$$n(x;\zeta) = \sum_{i \le N_e} |\varphi_i(x;\zeta)|^2 = \langle x|\hat{\mathrm{P}}_\zeta|x\rangle \tag{65}$$

where $\varphi_i(x;\zeta) \equiv \langle x|\varphi_i(\zeta)\rangle$ and

$$\hat{\mathrm{P}}_\zeta \equiv \sum_{i \le N_e} |\varphi_i(\zeta)\rangle\langle\varphi_i(\zeta)| \tag{66}$$

is a projector onto the Hilbert space $\mathcal{H}_\zeta$ spanned by the N_e eigenvectors $|\varphi_i(\zeta)\rangle$ of the single electron Hamiltonian $\hat{h}(\zeta)$ with the lowest eigenvalues. $\mathcal{H}_\zeta$ is an N_e-dimensional subspace of $\mathcal{L}^2(\mathbb{R})$, the infinite-dimensional Lebesgue space of square integrable functions on $\mathbb{R}$. It changes as ζ changes and the eigenstates of $\hat{h}(\zeta)$ change.

To understand the representability problem, it may be useful to visualize it as it is depicted in Fig. 13. In an insulator each vector in the basis $\{|\varphi_i(\zeta)\rangle\}_{i=1}^{N_e}$ of $\mathcal{H}_\zeta$ changes gradually with ζ as vectors from its orthogonal complement $\mathcal{H}_\zeta^\perp$ are mixed into them. Therefore $\mathcal{H}_\zeta$ rotates smoothly within $\mathcal{L}^2(\mathbb{R})$ as ζ changes. This is because there is a gap in the eigenspectrum of $\hat{h}(\zeta)$ between the N_e^{th} and the $(N_e+1)^{\text{th}}$ lowest eigenvalues, which never closes as ζ changes. In a metal, by contrast, the N_e^{th} eigenvalue is in a region of the spectrum where there is a quasicontinuum of eigenvalues. As ζ changes, the ordering of the eigenvalues is quasicontinuously changing, and each time the N_e^{th} eigenvalue and the $(N_e+1)^{\text{th}}$ eigenvalue cross, the N_e^{th} basis vector $|\varphi_{N_e}(\zeta)\rangle \in \mathcal{H}_\zeta$ is replaced with a vector $|\varphi_{N_e+1}(\zeta)\rangle \in \mathcal{H}_\zeta^\perp$ to form $\mathcal{H}_{\zeta+\mathrm{d}\zeta}$.

The serene rotation of the basis vectors in a insulator is illustrated by the rotation of the basis $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ in Fig. 13 *before* $|\varphi_2\rangle$ reaches the red dashed line, which indicates where the second and third eigenvalues become equal in this two-electron example. As soon as the variation of ζ rotates $|\varphi_2\rangle$ past the red line, $\mathcal{H}_\zeta$ changes abruptly from $\text{span}\{|\varphi_1\rangle, |\varphi_2\rangle\}$ to $\text{span}\{|\varphi_1\rangle, |\varphi_3\rangle\}$. The quasicontinuum of eigenvalues in a metal means that, instead of $\mathcal{H}_\zeta$ smoothly rotating, there is a rapid *click-clacking* of vectors in and out of its basis.

If ζ changes infinitely slowly, the electrons have time to reach, and settle at, each instantaneous $\mathcal{H}$-representation before it changes. The system can then be assumed to be close to equilibrium almost all of the time. However immediately after each exchange of basis vectors between $\mathcal{H}_\zeta$ and $\mathcal{H}_\zeta^\perp$, it could be far from equilibrium. This is likely to be the case if states are widely-separated spatially, such as when they are localized on opposite surfaces or on oppositely-charged electrodes attached to different parts of the material. At values of ζ at which the $\mathcal{H}$-representation changes, the response of $n(x;\zeta)$ to changes of ζ is singular and occurs via a nonequilibrium dynamical process involving many electrons, in general.

If the exchange of basis vectors between $\mathcal{H}_\zeta$ and $\mathcal{H}_\zeta^\perp$ occurs frequently, as is the case in a metal, the electron density does not have time to reach each new $\mathcal{H}$-representation before it changes again. Therefore the response of electrons in a metal to applied fields is singular and governed by nonequilibrium dynamics.

In an insulator $\mathcal{H}_\zeta$ rotates smoothly and its dimension is $\gtrsim 10^{24}$ for materials at the human scale. The set of vectors that span it can always be transformed unitarily among themselves to localize them or delocalize them. These transformations do not change $\mathcal{H}_\zeta$ or the projector $\hat{\mathrm{P}}_\zeta$. Therefore they do not change the density $n(x;\zeta) = \langle x|\hat{\mathrm{P}}_\zeta|x\rangle$ or the current $\mathbf{J}^{(p)}$. Therefore, although it is common to transform the eigenfunctions of $\hat{h}(\zeta)$ to a more localized set of basis functions by linearly combining them, in principle this is not necessary.

XII. SINGLE PARTICLE STATES

As mentioned above, one can change the minimal (N_e-fold) basis of $\mathcal{H}_\zeta$ by rotating it. As shown in Sec. V, one can also change from a position (x) representation, or basis, of each basis vector of $\mathcal{H}_\zeta$ to a wavevector (k) representation of each state, i.e.,

$$|\varphi_i(\zeta)\rangle = \int_{\mathbb{R}} \mathrm{d}x\, \varphi_i(x;\zeta)\, |x\rangle$$

where $\varphi_i(x;\zeta) \equiv \langle x|\varphi_i(\zeta)\rangle$ and $\langle x|x'\rangle = \delta(x-x')$, or

$$|\varphi_i(\zeta)\rangle = \int_{\mathbb{R}} \mathrm{d}k\, \tilde{\varphi}_i(k;\zeta)\, |k\rangle\,,$$

where $\tilde{\varphi}_i(k;\zeta) \equiv \langle k|\varphi_i\rangle$, $\langle k|k'\rangle = \delta(k-k')$, and $\langle x|k\rangle \equiv e^{ikx}/\sqrt{2\pi}$. One can also use a mixed position/wavevector representation, as we have already seen in Sec. V.B.1. This is a common way to exploit the translational symmetry of a crystal: Roughly speaking, real space is used to describe the electronic structure in a single primitive unit cell Ω and reciprocal space is used to describe variations of the structure between different primitive cells.

The polarization current can be calculated in any minimal basis $\{|\varphi_i(\zeta)\rangle\}_{i=1}^{N_e}$ of $\mathcal{H}_\zeta$ and for any choice of the basis in which each vector $|\varphi_i(\zeta)\rangle$ is represented as a function. All that is needed is the position operator for the chosen representation, which is simply x when working with $\varphi_i(x;\zeta) \equiv \langle x|\varphi_i(\zeta)\rangle$ and is $i\,\partial/\partial k$ when working with $\tilde{\varphi}_i(k;\zeta) \equiv \langle k|\varphi_i(\zeta)\rangle$.

A. Bloch and Wannier functions

As discussed in Sec. V.B.1, when describing the bulk of a material theoretically, or when simulating the bulk of a material, it is common to use Born-von Kármán periodic boundary conditions (Born and von Kármán, 1912). This is equivalent to representing the material's bulk in a torus $\mathbb{T}$ (or $\mathbb{T}^n$, in n-d), which obviates the need to deal with surfaces. If the material is a crystal, the absence of any surfaces means that, at thermal or mechanical equilibrium, the distributions of electrons and nuclei have the exact $|\Omega|$-periodicity of the crystal.

In solid state physics it is common to use $\mathbb{T}$ to study the electronic subsystem in the presence of a $|\Omega|$-periodic distribution of static nuclei. In the limit of heavy nuclei, the energy of interaction between electrons and nuclei is $(n, v^{\text{ext}}) \equiv \int_{\mathbb{T}} n(x) v^{\text{ext}}(x)\,\mathrm{d}x$, where v^{ext} is a $|\Omega|$-periodic *external potential.* It is also common to simplify the electronic structure of the crystal with the assumption that the electron density, $n(x)$, can be built from the eigenfunctions of an effective one-electron Hamiltonian, $\hat{h}(x)$, which inherits $|\Omega|$-periodicity from v^{ext}.

The eigenfunctions of a $|\Omega|$-periodic operator are known as *Bloch functions*, and they have the form $\varphi_{\alpha k}(x) \equiv \langle x|\varphi_{\alpha k}\rangle = e^{ikx}u_{\alpha k}(x)$, where the *Bloch state* $|\varphi_{\alpha k}\rangle$ is an eigenstate of $\int_{\mathbb{T}} \mathrm{d}x\, \hat{h}(x)\, |x\rangle\langle x|$ and $u_{\alpha k}(x)$ is known as a *periodic Bloch function* (PBF) because it has the crystal's $|\Omega|$-periodicity (see Sec. V.B.1). Both $\varphi_{\alpha k}(x)$ and $u_{\alpha k}(x)$ are delocalized over the entirety of $\mathbb{T}$, and it follows from the eigenvalue equation, $\hat{h}\varphi_{\alpha k} = \epsilon_{\alpha k}\varphi_{\alpha k}$, that $u_{\alpha k}$ is an eigenfunction of the k-dependent Hamiltonian $\hat{h}_k(x) \equiv e^{-ikx}\hat{h}(x)e^{ikx}$.

In the large-torus limit, $\hat{h}_k$ varies continuously with k. Therefore the eigenstates at different values of k can be labelled such that $u_{\alpha k}$ and its eigenvalue $\epsilon_{\alpha k}$ vary continuously with k. When the eigenvalues are plotted as functions of k, the set of points $\{(k, \epsilon_{\alpha k})\}_k$ forms a surface. Note that I use a subscript k on the parentheses to denote that different elements of the set correspond to different values of k; and the absence of a subscript α denotes that α is the same for all elements of the set. Therefore $\{(k, \epsilon_{\alpha k})\}_k$ and $\{(k, \epsilon_{\beta k})\}_k$ are different surfaces if $\alpha \neq \beta$. Each surface, or set of intersecting surfaces, is usually referred to as a *band* because its projection onto the energy (eigenvalue) axis is a band, or interval, of energies.

Let us assume that the set of all Bloch functions, $\{\varphi_{\alpha k}\}_{\alpha k}$, has been chosen to be orthonormal. Then, in an insulator, the electron density can be expressed as

$$n(x) = \sum_{\alpha k} |\varphi_{\alpha k}(x)|^2 = \langle x|\left(\sum_{\alpha k} |\varphi_{\alpha k}\rangle\langle\varphi_{\alpha k}|\right)|x\rangle\,,$$

where the sum over αk is a sum over a finite number of 'occupied' states, which for our purposes simply means states that contribute to n, and each state is either vacant or occupied by one electron. In Sec. V.B.1 it was shown how $\mathbf{J}^{(p)}$ could be calculated directly from the set of PBFs using Eq. 14. Now let us try to calculate it using

Eq. 62, and let us also try to relate it more directly to the depiction, in Fig. 13, of the mathematical process by which a polarization current arises.

The Bloch states are an orthonormal basis for the $\mathcal{H}$-representation of n and, by symmetry, there are the same number of Bloch state centers in each primitive unit cell of the crystal. Therefore, we could use Eq. 62 to calculate $\mathbf{J}^{(p)}$ from the velocities of their centers in one particular primitive cell. However, although this would yield the correct value of $\mathbf{J}^{(p)}$, because the Bloch functions are delocalized, this mapping of the density onto a set of point charges is conceptually inconsistent with the mapping envisaged in Sec. X.F. There we assumed a mapping of the charge density, $\rho^- = -en$, onto charge packets of microscopic widths, rather than delocalized distributions. Therefore, let us seek a localized basis for the $\mathcal{H}(t)$-representation of $n(x;\zeta(t))$.

First, let us express the density as $n(x;\zeta) = \sum_\alpha n_\alpha(x;\zeta)$, where $n_\alpha(x;\zeta) \equiv \sum_k |\varphi_{\alpha k}(x;\zeta)|^2$ is the density from all Bloch states that contribute to band α. We know that each n_α is $\mathcal{H}(t)$-representable because $\hat{\mathrm{P}}_\alpha(\zeta) \equiv \sum_k |\varphi_{\alpha k}(\zeta)\rangle\langle\varphi_{\alpha k}(\zeta)|$ is the projector onto its $\mathcal{H}$-representation. Therefore the polarization current can be calculated as the sum $\mathbf{J}^{(p)} \equiv \sum_\alpha \mathbf{J}^{(p)}_\alpha$, where $\mathbf{J}^{(p)}_\alpha$ is the polarization current from the variation of $n_\alpha(x;\zeta)$ with ζ. This becomes more complicated when bands intersect one another (Souza *et al.*, 2001), but I will assume that they do not.

Let us focus on the contribution $\mathbf{J}^{(p)}_\alpha$ of band α to $\mathbf{J}^{(p)}$. We can transform the Bloch functions to a more localized set with the generalized Fourier transform

$$w_{\alpha X}(x;\zeta) \equiv \frac{1}{\sqrt{N_\Omega}} \sum_k e^{-ikX} e^{i\theta_\alpha(k)} \varphi_{\alpha k}(x;\zeta), \qquad (67)$$

where N_Ω is the number of primitive unit cells in $\mathbb{T}$; X identifies a particular position (x) within the torus; and $\theta_\alpha(k)$ is any x-independent constant or function of k. The function $w_{\alpha X}$, which is localized in real space, is known as a *Wannier function* (Blount, 1962; Ferreira and Parada, 1970; Kohn, 1973; Wannier, 1937). It can be shown that, for any choice of X_0, the set $\{w_{\alpha X} : X \in \mathcal{X}(X_0)\}$, where $\mathcal{X}(X_0) \equiv \{X_0 + m|\Omega| : 0 \le m \le N_\Omega - 1\}$, is both orthonormal and satisfies

$$\sum_k |\varphi_{\alpha k}(x;\zeta)|^2 = \sum_X |w_{\alpha X}(x;\zeta)|^2 = n_\alpha(x;\zeta).$$

Therefore, the set $\{|w_{\alpha X}(\zeta)\rangle : X \in \mathcal{X}(X_0)\}$ of N_Ω *Wannier states*, where $|w_{\alpha X}(\zeta)\rangle \equiv \int_{\mathbb{T}} \mathrm{d}x\, w_{\alpha X}(x;\zeta)\,|x\rangle$, is a minimal orthonormal basis of the $\mathcal{H}$-representation of n_α, and $\hat{\mathrm{P}}_\alpha(\zeta) = \sum_{X\in\mathcal{X}(X_0)} |w_{\alpha X}(\zeta)\rangle\langle w_{\alpha X}(\zeta)|$.

Because an integer multiple of $|\Omega|$ separates any two of the chosen points X, each primitive unit cell contains one of them. Furthermore, by substituting $\varphi_{\alpha k} = e^{ikx}u_{\alpha k}$ into Eq. 67 and using the periodicity of $u_{\alpha k}$, it can be shown that any given Wannier function in the set transforms into any other under a lattice translation $m|\Omega|$, where $m \in \mathbb{Z}$. Therefore, each primitive cell contains the center of exactly one of them and we have decomposed n_α into a periodic array of identical localized packets of electron density, $|w_{\alpha X}(x)|^2$. It follows from Eq. 62 that $\mathbf{J}^{(p)}_\alpha = e\dot{X}/|\Omega|$.

The degree to which the Wannier functions are localized depends on the choice of the function $\theta_\alpha(k)$, but the most localized set, which is commonly known as the set of *maximally-localized Wannier functions* (MLWF) (Marzari *et al.*, 2012; Marzari and Vanderbilt, 1997), is obtained when it is a k-independent constant (see Ferreira and Parada, 1970 and Appendix C).

1. Interpretation of Wannier functions

Wannier functions, whether maximally localized or not, are not specific to quantum mechanics and there is no obvious reason to attach any particular physical meaning to them.

If a density is $\mathcal{H}$-representable, its $\mathcal{H}$-representation has an infinite number of orthonormal minimal bases. Among those, there must exist a maximally localized basis and a maximally delocalized basis. This is a mathematical observation which does not imply that elements of these extreme sets have any further meanings or any physical meanings. Therefore claims that MLWFs have greater physical significances than elements of other minimal bases should be substantiated and the precise physical meanings attributed to them should be clarified.

The Wannier states of band α are eigenstates of any operator of the form $\hat{\mathcal{D}} \equiv \sum_X d_X\, |w_{\alpha X}\rangle\langle w_{\alpha X}|$, which means that the Wannier functions are eigenfunctions of the generally-nonlocal integral operator whose kernel is $\langle x'|\hat{\mathcal{D}}|x\rangle$. However, because electrons want to *delocalize*, rather than localize, Wannier functions are not, in general, either the eigenfunctions, or approximately equal to the eigenfunctions, of an operator that could reasonably be interpreted as the Hamiltonian of a real or idealized physical system. Therefore if, in a many-particle system, there existed single-particle states that could be regarded as 'physical', in the sense that they resembled states that individual particles would like to occupy in a system with simplified energetics (e.g., mean field interactions), they would not be localized, in general, and certainly not by their mutual repulsion. For example, changing a single-particle Hamiltonian, $\hat{h} \equiv \hat{t} + \hat{v}^{\mathrm{ext}}$, by adding a repulsive mean-field Coulomb potential from one or more localized clouds of negative charge to $\hat{v}^{\mathrm{ext}} < 0$, would not make its eigenfunctions more localized, in general.

Furthermore, there is no 'natural' or right way to partition the density into the same number of packets as there are electrons. Therefore $\hat{v}^{\text{ext}}$, which together with N_e determines the character of chemical bonds, and which is usually the *only* localizing influence on electrons, does not localize partitions of the density *individually.* It localizes the density as a whole. There is nothing within rigorous physical or chemical theory to suggest that it bestows a density with a substructure of localized partitions.

I emphasize this point because it has been claimed that Wannier functions, and MLWFs in particular, can provide insight into chemical bonding by elucidating the substructure of the electron density (Marzari *et al.*, 2012; Marzari and Vanderbilt, 1997). However, this claim has not been justified theoretically, but by references to the chemistry literature: It was claimed in Marzari and Vanderbilt, 1997 and Marzari *et al.*, 2012 that chemists use *localized molecular orbitals*, which are the analogs of MLWFs for molecules, for this purpose. However, the papers cited, namely Boys, 1960, Foster and Boys, 1960a, Foster and Boys, 1960b, and Edmiston and Ruedenberg, 1963, do not justify using localized orbitals to analyse bonds, and they did not introduce them to represent the parts of the electron density that are most important to bonding. They introduced them to deal more efficiently with those parts of the electron density that are *least* important to bonding, or to changes in bonding.

For example, when one is interested in a reaction that involves one reactive part of an otherwise-inert large molecule, it is unnecessary and computationally expensive to treat all parts of the molecule as reactive. One can freeze the electronic structure of the inert part and calculate its effects on the reactive part using methods that are much more computationally efficient than treating the whole molecule as reactive would be.

The same trick can be played when studying multiple large molecules, which are mostly the same, but have different functional groups in one relatively-small region. After calculating the electronic structure of one of the molecules, it should not be necessary to recalculate it from scratch for another molecule: it is more efficient to reuse parts of the density that are unaffected by the differences in functional groups. Localized orbitals were introduced to facilitate this partitioning of the electronic structure (Boys, 1960; Foster and Boys, 1960a,b). For example, in calculations based on density functional theory, a boundary can be chosen between the reactive part of the molecule, $\mathcal{R}$, and the unreactive part, $\mathcal{U}$, and the density can be partitioned using the centers $\bar{x}_\alpha$ of the localized functions w_α as $n(x) = \sum_{\bar{x}_\alpha \in \mathcal{R}} |w_\alpha(x)|^2 + \sum_{\bar{x}_\alpha \in \mathcal{U}} |w_\alpha(x)|^2$. The more localized the functions w_α are, the more well-defined the boundary is.

B. Localized orbital-based models of bonding

Some of the concepts that appear in introductory chemistry textbooks do not have theoretical justifications, and were introduced, or originated, when very little was known about electrons and bonding (Gillespie and Robinson, 2007; Lewis, 1916; Pauling, 1960). One example is the idea that the electron density possesses a substructure of integer-occupied localized orbitals. Another is the idea that covalency is a distinct type, or mechanism, of bonding, rather than just bonding that does not conform closely to either the ionic limit or the metallic limit. Both of these misconceptions have been perpetuated by the widespread use of approximations that are based on simplifying the mathematical form of the many-electron wavefunction (Ψ) to make calculations tractable.

For example, the *Hartree-Fock approximation*, which is probably the simplest useful mean-field approximation, is based on restricting Ψ to be a Slater determinant. When Ψ has this form, integer-occupied single particle states appear to have clear physical meanings. When the single particle states in the determinant are linear combinations of atom-centered basis functions, such as the orbitals in the Hartree-Fock wavefunctions of isolated atoms, each basis function 'belongs' to one of the atoms. Therefore, there appears to be a clear and meaningful qualitative distinction between a covalent bond and an ionic bond, but this is an artefact of Ψ's simplified form.

The primary reason for approximating Ψ as a determinant, or as a sum of few determinants, is not to simplify bonding *conceptually*, but to make *calculations* tractable. Therefore it seems valid to question whether localized atomic or molecular orbitals help to simplify bonding conceptually, or whether they complexify and obfuscate it. The artificial qualitative distinction between covalent bonding and ionic bonding illustrates that, at least when building the most basic understanding of bonding from the most computationally tractable form of wavefunction, they can be misleading.

Although most research scientists understand this, and also understand that the terms *ionic*, *covalent*, and *metallic* refer to varying degrees of localization of the electron density around nuclei, students are still being taught more traditional and misleading ideas, such as that a pair of covalently bonded atoms *share* electrons *between* them (Bacskay *et al.*, 1997; Grundmann, 2016; McQuarrie *et al.*, 2011; Zürcher, 2018). Therefore the purpose of this subsection is to emphasize that, in some ways and for some purposes, the essence of bonding is simpler than it appears from descriptions of it in terms of localized orbitals.

C. Summary of chemical bonding in terms of electron density

The electron density (n) is high where the microscopic electric potential from the nuclei (ϕ^{nuc}) is high, and it only has maxima at the positions of the nuclei. If two nuclei shared electrons between them, there would be a local maximum of the density between them, but there never is.

The shape of n is determined by the shape of ϕ^{nuc}; and ϕ^{nuc} is determined by the charges and positions of the nuclei. Localizing density where ϕ^{nuc} is high makes the potential energy of attraction between electrons and nuclei more negative, but delocalizing density makes both the electrons' kinetic energy and their mutual repulsion less positive. The ground state density is the lowest-energy compromise between these localizing and delocalizing influences.

Most of the density is localized around nuclei, and most of it is at points where its gradient $\boldsymbol{\nabla} n$ is directed towards the nearest nucleus.

1. Ionic bonding

The bonding for which a superposition of spherically symmetric electron densities most closely approximates the true electron density is referred to as *ionic* bonding; and the *ionic limit* of bonding is the limit in which an approximation of this form becomes exact.

If the time average of the net charge of each nucleus and its almost-spherical electron cloud was zero, the attraction between atoms would be very weak and arise from electron correlation, rather than electrostatics. Atoms bind chemically by becoming ions, thereby lowering the potential energy via their mutual attraction. However they do not become ions by donating or accepting electrons, but by donating or accepting electron *density.* When atoms are close enough to one another to bond chemically, and on the shortest time scales relevant to atomic motion ($\sim 10^{-15}$ seconds), there is no theoretical reason why ions' average charges are likely to be integers or close to integers.

2. Covalent and metallic bonding

If more of the density is in regions where $\boldsymbol{\nabla} n$ is not directed towards the nearest nucleus, we describe the bonding as either *covalent* or *metallic.* Bonding is metallic if the density in these regions is so delocalized that a significant fraction of the electrons are mobile. Otherwise we refer to it as covalent, for historical reasons. There is not a clear boundary between ionic and covalent bonding; and, even in the canonically-ionic compound NaCl, the magnitudes of the ions' charges are only $\sim 0.8\,e$ (Bao *et al.*, 2018; Jennison and Kunz, 1976; Kvashnin *et al.*, 2019; Li *et al.*, 2007).

Large atomic numbers, delocalized electrons, and high coordination numbers go hand in hand because, when an atom's radius is large, the interactions of its nucleus with electrons on the outskirts of its electron cloud are weak, and comparable in magnitude to its interactions with electrons on the outskirts of neighbouring atoms. Therefore energy is lowered by atoms arranging such that each one has many neighbours, whose electrons its nucleus interacts with. The energy is lowered further by the electron density on the outskirts of atoms delocalizing, so that more electrons have interactions of comparable strengths with multiple nuclei. This delocalization worsens the approximation of the density as a superposition of spherical densities.

The *metallic limit* is the limit in which the electron density becomes a superposition of spherically-symmetric densities *and* a uniform density.

D. The natural single particle substructure of the density

The electron density does possess a 'natural' substructure of single-particle states (φ_α) and their 'occupancies' (μ_α) (Coleman, 1963; Löwdin, 1955; McWeeny, 1960), which satisfy

$$n(\vec{r}) = \sum_\alpha \mu_\alpha |\varphi_\alpha(\vec{r})|^2, \quad \langle \varphi_\alpha | \varphi_\beta \rangle = \delta_{\alpha\beta},$$
$$\sum_\alpha \mu_\alpha = N_e, \quad \mu_\alpha \leq 1,\ \forall \alpha,$$

and we will assume that they are indexed in order of decreasing occupation number, such that $\alpha \leq \beta \iff \mu_\alpha \geq \mu_\beta$. These *natural orbitals* are the normalized eigenstates of the 1-particle reduced density matrix. Their properties, some of which are discussed in Appendix D.2, suggest that they are the only single-particle states to which physical meaning should be attached in a many-particle system.

The following exact expression for the energy, $E \equiv \langle \Psi | \hat{H} | \Psi \rangle$, of a normalized N_e-particle pure state, $|\Psi\rangle$, which is derived in Appendix D.2, helps to illustrate their physical meaning.

$$E = \sum_\alpha \mu_\alpha \left[\varepsilon_\alpha + \left(\sum_{\beta \neq \alpha} \Delta\varepsilon_{\alpha\beta} \right) + \frac{1}{N_e} W_\alpha^{(N_e-1)} \right]. \tag{68}$$

In this expression ε_α is equal to $\langle \varphi_\alpha | \hat{h}_\alpha | \varphi_\alpha \rangle$, where $\hat{h}_\alpha \equiv \hat{h} + \hat{v}_\alpha^{MF}$, $\hat{h} \equiv \hat{t} + \hat{v}^{\text{ext}}$, $\hat{t}$ is the single-particle kinetic energy operator, $\hat{v}^{\text{ext}}$ is the interaction of an electron with the nuclei, and $\hat{v}_\alpha^{MF}$ is a fraction $1/N_e$ (one electron's share) of the mean-field interaction between an electron and the $N_e - 1$ electrons occupying the dual state, $|\Theta_\alpha\rangle \equiv \langle\varphi_\alpha|\Psi\rangle$, of natural orbital $|\varphi_\alpha\rangle$. $\Delta\varepsilon_{\alpha\beta}$ is a *mediated* coupling between an electron in orbital $|\varphi_\alpha\rangle$ and an electron in orbital $|\varphi_\beta\rangle$. $W_\alpha^{(N_e-1)}$ is the energy of interaction between $N_e - 1$ electrons occupying state $|\Theta_\alpha\rangle$, which, by definition, is the state that the remaining electrons would be in when one of them occupied orbital $|\varphi_\alpha\rangle$.

The natural orbitals are not eigenstates of the Hamiltonians $\hat{h}_\alpha$ or $\hat{h}$, or of any Hamiltonian. This is to be expected of 'physical' single-particle states because Hamiltonian eigenstates are stationary states and the states that individual electrons occupy cannot be stationary if they interact with other electrons while occupying them. Interactions perturb the electron occupying state $|\varphi_\alpha\rangle$ and, sooner or later, displace it from that state to another one. Therefore each occupation number, μ_α, is less than one and can be interpreted as either the fraction of time for which the α^{th} natural orbital is occupied, or the probability that it is occupied at any given time.

The form of Eq. 68 suggests that it can be interpreted within a quasi-independent-electron picture as follows: When orbital $|\varphi_\alpha\rangle$ is occupied, the energy of the electron occupying it is the sum of the orbital energy, ε_α, and the energies of interaction with electrons occupying other natural orbitals. Some of this interaction energy has been written explicitly as $\sum_{\beta\neq\alpha}\Delta\varepsilon_{\alpha\beta}$ and the rest is contained in the terms $\left\{W_\gamma^{N_e-1}\right\}_{\gamma\neq\alpha}$. The total energy is an occupation-weighted sum, over all product states $|\varphi_\alpha\rangle \otimes |\Theta_\alpha\rangle$, of the energy of the electron in orbital $|\varphi_\alpha\rangle$ and the mutual repulsion of the remaining electrons.

In terms of the natural orbitals, the total kinetic energy $\langle\Psi|\hat{T}|\Psi\rangle$ is given exactly by $\sum_\alpha \mu_\alpha \langle\varphi_\alpha|\,\hat{t}\,|\varphi_\alpha\rangle$. If $|\Psi\rangle$ was expanded in terms of any other set of single-particle states, $\{|\psi_\alpha\rangle\}$, the kinetic energy would have the more complicated form $\sum_\alpha\sum_\beta \tilde{c}_{\alpha\beta}\langle\psi_\alpha|\,\hat{t}\,|\psi_\beta\rangle$. Therefore the natural orbitals are the only orbitals for which the kinetic energy of the interacting many-electron system can be expressed as a weighted sum of *single*-orbital contributions. Furthermore, if the magnitude of the electron-electron repulsion could be reduced gradually to zero, the natural orbitals of the ground state would become eigenstates of $\hat{h}$ in the noninteracting limit.

The ground state occupation numbers become equal to one when the interactions between electrons are turned off, or when they are replaced with mean-field interactions, as in the drastic Hartree-Fock approximation. In real systems the number of them that differ significantly from zero is larger, and possibly very much larger, than the total number of electrons (Cioslowski and Strasburger, 2021; Giesbertz and van Leeuwen, 2013). However, the rate at which they decay with increasing α is currently an open question because natural orbitals and their occupation numbers have only been calculated reliably for simple molecular systems and molecular systems may not be representative of condensed phases.

For example, in a small molecule the eigenstates of $\hat{h}$ must be localized on the molecule, whereas in crystals they are delocalized Bloch states. Furthermore, since N_e is very large in a crystal, $\hat{h}_\alpha$ is almost equal to $\hat{h}$. This is because $\hat{v}_\alpha^{MF}$ is the mean-field potential from an electron density with integral one, and this electron density, which is $1/N_e$ times the density of state $|\Theta_\alpha\rangle$, is delocalized unless state $|\varphi_\alpha\rangle \otimes |\Theta_\alpha\rangle$ has a very high energy. Therefore, the ground state natural orbitals in a crystal are delocalized in the noninteracting limit and it seems very unlikely that switching interactions on would localize them.

If the natural orbitals are all delocalized, each orbital overlaps significantly with a large number of others. Furthermore, the set of orbitals with which orbital $|\varphi_\alpha\rangle$ overlaps is likely to include many whose energies are very close to ε_α. This suggests that, in condensed phases, the natural orbitals are delocalized and that electrons move between different natural orbitals relatively freely and frequently. If the occupation numbers are much smaller than one, many of these transitions might be to vacant states, but if they are all close to one, most transitions would involve electrons swapping states with other electrons.

XIII. MACROSCOPIC POTENTIAL ($\mathbf{\Phi}$) AND FIELD ($\mathbf{E}$)

It is well known that the electric potential is a relative quantity and that, when its value at a point is quoted, this value is always the difference between the potential at the point and a reference potential. In theoretical work the reference point is often taken to be a notional point in the vacuum at infinity, in experimental work it may be a particular electrode, and in engineering it is common to reference potentials to the 'ground' or 'earth'. The difference in meaning between the terms *mean inner potential* and *macroscopic potential* is of little relevance to this work. The MIP is the average of ϕ over all points in a material; therefore it is a scalar constant. The macroscopic potential is a scalar field, defined at all points in a material, but defined only to a finite precision ε_Φ. To within this precision its value at each point in the bulk is equal to the MIP. I mostly refer to $\mathbf{\Phi}$ as the macroscopic potential in this section and as the MIP when discussing Bethe's approximate expression for it in Sec. XIV.

It is very important to be able to calculate changes in macroscopic potential. From a conceptual and theoretical viewpoint, our understanding of the relationship between macroscale and microscale electrostatics cannot be considered complete if we do not know how, in principle, to calculate the potential at the macroscale from the charge density or potential at the microscale. From a practical perspective, the MIP, which is believed to be positive, is a key quantity in several areas of experimental and computational science. For example, computer simulations can be used to calculate microscopic charge densities and, up to an unknown constant, microscopic potentials. To improve the designs of devices, such as batteries, fuel cells, chemical sensors, and solar cells, computational scientists simulate their constituent materials independently and, from those simulations, try to calculate the change in the average potential that an electron or ion would experience if it moved from one device component to another (Blumenthal *et al.*, 2017; Hörmann *et al.*, 2019). The MIP is also used in electron microscopy to analyse and interpret electron diffraction images (Yesibolati *et al.*, 2020). In both of these contexts, the distinction between the MIP and the *time average* of the potential *felt by* the charge carrier is rarely made, but it is the latter that is of interest and the MIP is used as an approximation to it.

In this section I show how to calculate the change $\Delta\mathbf{\Phi}$ in the macroscale electric potential $\mathbf{\Phi}$, between the vacuum above a material's surface and its bulk, from the microscopic charge density ρ. The results are easy to generalize to the change in potential across an interface between two materials by treating it as a pair of adjoined surfaces. They also allow the value of $\mathbf{\Phi}$ in an isolated material to be calculated relative to a distant point in the vacuum surrounding it.

Some of my results, such as the macroscale potential within a thin film whose surfaces are equally- and oppositely-charged, are well known and serve as a sanity check on my theory and reasoning. However, the main result, which I justified on symmetry grounds in Sec. IV, contradicts most of the literature on this subject over the past century. This result is that $\mathbf{\Phi} = 0$ in the bulk $\mathfrak{B}$ of an isolated material unless it has charged surfaces or unless its bulk contains charged macroscopic heterogeneities.

A. Change in potential across a surface ($\Delta\mathbf{\Phi}$)

To calculate $\Delta\mathbf{\Phi}$, I will consider the microscopic potential at an arbitrary position $\vec{r}_b \in \mathfrak{B}$ deep below the surface at $x = x_L$. I will calculate the potential within a finite chunk of material and then take the large size limit. It is important, when doing this, to order the limits appropriately. To illustrate the possible pitfalls, consider the well-known example of the macroscopic potential $\mathbf{\Phi}$ on the plane $\mathbf{x} = \mathbf{x}_b \approx (\mathbf{x}_L + \mathbf{x}_R)/2$ from equally- and oppositely-charged surfaces at $\mathbf{x}_L$ and $\mathbf{x}_R$. If I calculate $\mathbf{\Phi}$ for an isolated material that is finite in all directions and take the limit $|\mathbf{x}_L - \mathbf{x}_R| \to \infty$ before I take the limit of large size in the lateral (yz) directions, I find that $\mathbf{\Phi}$ vanishes. However, if I take the limit of large cross-sectional area first, I find that $\mathbf{\Phi}$ is linear in $\mathbf{x}$ and that $\mathbf{E}$ is constant. The appropriate order to choose for the limits depends on the aspect ratio of the material and on the position within the material at which the potential is being calculated.

I want to calculate the average potential in a bulk-like region of the material that is much closer to one surface than any other. The plane $x = x_L$ is parallel to this surface and in the vacuum just beyond it. Because all other surfaces are further away, it is appropriate to assume that $|x - x_L|$ is much smaller than the material's lateral dimensions. I will first calculate the microscopic potential due to the charge within a cylindrical region of the material, of radius R, whose axis is normal to the surface. I will use the cylindrical coordinates $\vec{r} = (x, s, \phi)$ or $\vec{r} = (x, \vec{s})$, where $\vec{s} \in \mathbb{R}^2$ is a vector in the plane parallel to the surface, $s = |\vec{s}|$, and ϕ is the azimuth.

The potential at $\vec{r}_b \equiv (x_b, s_b, \phi_b) = (x_b, \vec{s}_b)$ due to the charge density within a cylinder bounded by the surfaces $|\vec{s} - \vec{s}_b| = R$, $x = x_L$, and $x = x_b$ is

$$\Phi_L(x_b, \vec{s}_b; R)$$
$$= \kappa \int_{x_L}^{x_b} \left(\iint_{|\vec{s} - \vec{s}_b| < R} \frac{\rho(x, \vec{s})}{\sqrt{(x - x_b)^2 + |\vec{s} - \vec{s}_b|^2}} \, \mathrm{d}^2 s \right) \mathrm{d}x$$

where $\kappa \equiv (4\pi\epsilon_0)^{-1}$. I will first cast this expression into a more convenient form. Then I will calculate its mesoscale average over a range of positions $(x_b, \vec{s}_b)$ in the limit of large R. Finally, I will add the mesoscale average of the potential, Φ_r, from charge deeper below the surface than $\vec{r}_b$. The right-hand boundary of the region whose charge density contributes to Φ_r is the plane $x = x_r$, where $x_b < x_r < x_R$; I use a lower case subscript for Φ_r to distinguish the plane $x = x_r$ from the right hand surface plane $x = x_R$. I will be considering the cases $x_r = x_R$ and $|x_r - x_b| \ll |x_R - x_r|$ separately.

The average volumetric charge density on a disc of radius R, which is parallel to the surface and centered at $(x, \vec{s})$, is

$$\bar{\rho}(x, \vec{s}; R) \equiv \frac{1}{\pi R^2} \iint_{|\vec{u}| < R} \rho(x, \vec{s} + \vec{u}) \, \mathrm{d}^2 u$$

Defining $\Delta\rho(\vec{u}; x, \vec{s}, R) \equiv \rho(x, \vec{s} + \vec{u}) - \bar{\rho}(x, \vec{s}; R)$ allows Φ_L to be split into two terms:

$$\Phi_L(x_b,\vec{s}_b;R)=\overbrace{\frac{1}{2\epsilon_0}\int_{x_L}^{x_b}\bar{\rho}(x,\vec{s}_b;R)\left(\int_0^R\frac{u}{\sqrt{(x-x_b)^2+u^2}}\,\mathrm{d}u\right)\mathrm{d}x}^{\Phi_L^{[\bar{\rho}]}(x_b,\vec{s}_b;R)}+\overbrace{\kappa\int_{x_L}^{x_b}\iint_{|\vec{u}|<R}\frac{\Delta\rho(\vec{u};x,\vec{s}_b,R)}{\sqrt{(x-x_b)^2+|\vec{u}|^2}}\,\mathrm{d}^2u\,\mathrm{d}x}^{\Phi_L^{[\Delta\rho]}(x_b,\vec{s}_b;R)} \tag{69}$$

and similarly for Φ_r. It is easy to see that the surface excess of $\Delta\rho(\vec{u};x,\vec{s}_b,R)$ vanishes when averaged over the plane parallel to the surface. It is shown in Appendix F that the planar averages of $\Phi_L^{[\Delta\rho]}$ and $\Phi_r^{[\Delta\rho]}$ also vanish in the large R limit, i.e., that

$$\lim_{\xi_b/R\to 0}\left\{\frac{1}{\pi R^2}\iint_{|\vec{u}|<R}\Phi_L^{[\Delta\rho]}(x_b,\vec{s}_b+\vec{u};R)\,\mathrm{d}^2u\right\}=0,$$
$$\lim_{\xi_b/R\to 0}\left\{\frac{1}{\pi R^2}\iint_{|\vec{u}|<R}\Phi_r^{[\Delta\rho]}(x_b,\vec{s}_b+\vec{u};R)\,\mathrm{d}^2u\right\}=0,$$

where $\xi_b\equiv|x_L-x_b|$ is the depth of $\vec{r}_b$ below the surface. Therefore, the only contributors to the bulk average of the microscopic potential are $\Phi_L^{[\bar{\rho}]}$ and $\Phi_r^{[\bar{\rho}]}$.

B. Mesoscale average of $\Phi_L^{[\bar{\rho}]}$

Integrating over u in the expression for $\Phi_L^{[\bar{\rho}]}$, choosing $R>\xi_b$, and using a Taylor expansion gives

$$\begin{aligned}\Phi_L^{[\bar{\rho}]}(x_b,\vec{s}_b;R)&=\frac{1}{2\epsilon_0}\int_{x_L}^{x_b}\bar{\rho}(x,\vec{s}_b;R)\left(\sqrt{(x-x_b)^2+R^2}+(x-x_b)\right)\mathrm{d}x\\&=\frac{1}{2\epsilon_0}\int_{x_L}^{x_b}\bar{\rho}(x,\vec{s}_b;R)(x-x_b)\,\mathrm{d}x+\frac{R}{2\epsilon_0}\int_{x_L}^{x_b}\bar{\rho}(x,\vec{s}_b;R)\left[1+\frac{1}{2}\left(\frac{x-x_b}{R}\right)^2+\mathcal{O}\left(\left(\frac{x-x_b}{R}\right)^4\right)\right]\mathrm{d}x\end{aligned} \tag{70}$$

I assume that there exists a well-defined macroscopic average of the volumetric charge density on every plane parallel to the surface. By this I mean that, although $\bar{\rho}(x,\vec{s}_b;R)$ may exhibit microscopic fluctuations as R increases, it converges to a well-defined value rather than systematically growing or shrinking. Furthermore, as $\vec{s}_b$ is varied at fixed R, $\bar{\rho}(x,\vec{s}_b;R)$ fluctuates microscopically about the value to which it converges in the large R limit. If I also assume that the bulk of the material is charge-neutral, only the first term of the series expansion in Eq. 70 can survive the large R limit. In anticipation of this limit, and with the understanding that '=' means '≈' until the limit is taken, I write

$$\Phi_L(x_b,\vec{s}_b;R)=\frac{1}{2\epsilon_0}\left[(R-x_b)\int_{x_L}^{x_b}\bar{\rho}(x,\vec{s}_b;R)\,\mathrm{d}x+\int_{x_L}^{x_b}x\,\bar{\rho}(x,\vec{s}_b;R)\,\mathrm{d}x\right] \tag{71}$$

$\Phi_L^{[\bar{\rho}]}(x_b,\vec{s}_b;R)$ depends sensitively on x_b and so do both of its constituent terms on the right hand side. Therefore, I will average over x_b. Before doing so, let us consider the average over $\vec{s}_b$, which we should perform to calculate the three dimensional macroscopic bulk average. Because we will be taking the limit of large R, we can assume that at every value of $\vec{s}_b$ the planar averages of $\bar{\rho}(x, \vec{s}_b; R)$ and $\Phi_L(x_b, \vec{s}_b; R)$ are converged, to our desired precisions, on an area much smaller than πR^2. Because they are now insensitive to $\vec{s}_b$, I will drop it from their list of arguments. The mesoscale average over x_b of $\int_{x_L}^{x_b} \bar{\rho}(x\,; R)\,\mathrm{d}x$ is the average areal surface charge density, $\boldsymbol{\sigma}_L(R)$.

The mesoscale average over x_b of $\int_{x_L}^{x_b} x\,\bar{\rho}(x\,; R)\,\mathrm{d}x$ is $X_{\boldsymbol{\sigma}}^L\,\boldsymbol{\sigma}_L(R)$, where

$$X_{\boldsymbol{\sigma}}^L \equiv \frac{\left\langle \int_{x_L}^{x_b} x\,\bar{\rho}(x; R)\,\mathrm{d}x \right\rangle_{\varepsilon_x}}{\left\langle \int_{x_L}^{x_b} \bar{\rho}(x; R)\,\mathrm{d}x \right\rangle_{\varepsilon_x}},$$

is the center of the microscale distribution of excess surface charge. The microscale charge density is bulk-like at mesoscopic depths and so $X_{\boldsymbol{\sigma}}^L$ is within $\varepsilon_x/2$ of x_L, which means that $X_{\boldsymbol{\sigma}}^L \in \mathbf{x}_L \equiv \mathfrak{I}(x_L, \varepsilon_x)$. When working at the macroscale, $\mathbf{x}_L$ and $\mathbf{x}_b$ are treated as having precise values. Therefore, taking the mesoscale average of both sides of Eq. 71 gives

$$\boldsymbol{\Phi}_L(R) = \frac{1}{2\epsilon_0}\boldsymbol{\sigma}_L(R)\,(R - |\mathbf{x}_b - \mathbf{x}_L|) \tag{72}$$

C. Mesoscale average of Φ_r

By the same approach that led to Eq. 71, we can find the potential at $(x_b, \mathbf{s}_b)$ from charge within the cylinder bounded by the surfaces $|\vec{s} - \vec{s}_b| = R$, $x = x_b$, and $x = x_r$, where $x_r > x_b$ is any position to the right of x_b such that $|x_r - x_b| \ll R$. This potential is

$$\Phi_r(x_b, x_r; R) = \frac{1}{2\epsilon_0}\left[(R + x_b)\int_{x_b}^{x_r} \bar{\rho}(x; R)\,\mathrm{d}x - \int_{x_b}^{x_r} x\,\bar{\rho}(x; R)\,\mathrm{d}x\right] \tag{73}$$

Except in the case of thin material films, if $|x_L - x_b| \ll R$ then $|x_R - x_b| \not\ll R$, which invalidates the derivation of Eq. 73 for $x_r = x_R$. Therefore, I will separately treat two cases. First, I will consider the case of a thin film, for which $|x_L - x_R| \ll R$. I will set $x_r = x_R$, and calculate the mesoscale average of the potential in the center of the film from the charge density of the entire film. In the second case, I will assume that the surface at x_R is far away and that its macroscopic surface charge density $\boldsymbol{\sigma}_R$ is zero. In this case I will add to $\boldsymbol{\Phi}_L$ the contribution to the mesoscale average of the potential from charge density at positions $x > x_b$ which are still within the bulk of the material.

1. Case I: Thin film, $|x_L - x_R| \ll R$

Defining $\boldsymbol{\sigma}_R(R)$ as the average areal charge density of the surface at $\mathbf{x}_R$, and $X_{\boldsymbol{\sigma}}^R \in \mathbf{x}_R$ as the center of the microscale distribution of excess charge at the right-hand surface, allows me to express the mesoscale average of Eq. 73, when $x_r = x_R$, as

$$\boldsymbol{\Phi}_R(R) = \frac{1}{2\epsilon_0}\boldsymbol{\sigma}_R(R)\,(R - |\mathbf{x}_R - \mathbf{x}_b|) \tag{74}$$

The total potential at $\mathbf{x}_b$ is the sum of $\boldsymbol{\Phi}_L$ and $\boldsymbol{\Phi}_R$ in the large R limit. That is,

$$\boldsymbol{\Phi}(x_b) = \lim_{R\to\infty}\left\{\frac{R}{\epsilon_0}\left(\frac{\boldsymbol{\sigma}_L(R) + \boldsymbol{\sigma}_R(R)}{2}\right) - \frac{1}{2\epsilon_0}\left[\sigma_L(R)\,|\mathbf{x}_b - \mathbf{x}_L| + \boldsymbol{\sigma}_R(R)|\mathbf{x}_b - \mathbf{x}_R|\right]\right\} \tag{75}$$

The term proportional to R becomes infinite in this limit unless $\boldsymbol{\sigma}_L + \boldsymbol{\sigma}_R = 0$. Therefore, I assume that the surfaces have equal and opposite average areal charge densities and I define $\boldsymbol{\sigma} \equiv \boldsymbol{\sigma}_L = -\boldsymbol{\sigma}_R$. Then Eq. 75 becomes

$$\boldsymbol{\Phi}(\mathbf{x}_b) = \frac{\boldsymbol{\sigma}}{\epsilon_0}\left[\frac{1}{2}(\mathbf{x}_L + \mathbf{x}_R) - \mathbf{x}_b\right] \tag{76}$$

From Eq. 76, we can immediately calculate the change in the macroscopic potential between the charged surfaces at $\mathbf{x}_L$ and $\mathbf{x}_R$ as

$$\Delta\boldsymbol{\Phi} \equiv \boldsymbol{\Phi}(\mathbf{x}_L) - \boldsymbol{\Phi}(\mathbf{x}_R) = \frac{\boldsymbol{\sigma}}{\epsilon_0}(\mathbf{x}_R - \mathbf{x}_L) \tag{77}$$

For very large finite values of R we can write this as $\Delta\boldsymbol{\Phi} = Q/(\epsilon_0 A)$, where $Q \equiv A\,\boldsymbol{\sigma}\,(\mathbf{x}_R - \mathbf{x}_L)$ and $A \equiv \pi R^2$. This is the familiar formula for the magnitude of the potential difference between parallel plates carrying equal and opposite charges. Therefore an important sanity check on the theory has been passed.

2. Case II: Macroscopic sample, $|x_L - x_R| \not\ll R$

When the surface at x_R is not charged and is sufficiently far away that it does not contribute to the potential at x_b, the total microscopic potential at x_b from all charge at greater depths ($x > x_b$) can be assumed to emanate from bulk regions where the charge density is macroscopically uniform and neutral.

If the bulk of the material is charge neutral, there is no contribution to the potential in the vicinity of x_b from points $x_r > x_b$ sufficiently far from it. However, the potential at x_b will depend on the precise choice of x_r because the integrated charge between x_b and x_r depends sensitively on its value. Therefore, as before, I will take the mesoscale average over x_r of the microscopic potential, $\Phi_r(x_b, x_r; R)$, at x_b. Eqs. A1 and A3, which are derived in Appendix VII, can be used in Eq. 73 to write the following expression for the mesoscale average of $\Phi_r(x_b, x_r; R)$ over x_r.

$$\begin{aligned}\overline{\Phi_r(x_b;R)} = \frac{1}{2\epsilon_0}\Bigg[&(R+x_b)\left(\int_{x_b}^{x_r}\bar{\rho}(x;R)\,\mathrm{d}x - \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_r)\right) \\ &-\left(\int_{x_b}^{x_r} x\,\bar{\rho}(x;R)\,\mathrm{d}x - x_r\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_r) - \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 2\rangle}(x_r)\right)\Bigg]\end{aligned}$$

This is independent of the choice of x_r, and so x_r may be chosen such that $\int_{x_b}^{x_r}\bar{\rho}(x;R)\,\mathrm{d}x = 0$. The reason for making this choice is that it implies the following relationships.

$$\begin{aligned}\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_r) &= \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b) \\ \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 2\rangle}(x_r) &= \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 2\rangle}(x_b) \\ \int_{x_b}^{x_r} x\,\bar{\rho}(x;R)\,\mathrm{d}x &= (x_r - x_b)\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b)\end{aligned}$$

Using these formulae, Eq. 78 simplifies to

$$\overline{\Phi_r(x_b;R)} = \frac{1}{2\epsilon_0}\left(\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 2\rangle}(x_b) - R\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b)\right) \tag{78}$$

and so the total microscopic potential at x_b is

$$\begin{aligned}\Phi(x_b;R) &= \Phi_L(x_b;R) + \overline{\Phi_r(x_b;R)} \\ &= \frac{1}{2\epsilon_0}\Bigg[(R-x_b)\int_{x_L}^{x_b}\bar{\rho}(x_b;R)\,\mathrm{d}x \\ &\quad + \int_{x_L}^{x_b} x\,\bar{\rho}(x_b;R)\,\mathrm{d}x + \bar{\mathcal{M}}_{\bar{\rho}}^{\langle 2\rangle}(x_b) - R\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b)\Bigg]\end{aligned} \tag{79}$$

It is straightforward to show that the average of $\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 1\rangle}(x_b)$ over x_b is zero for any macroscopically-uniform charge density (see Eq. 47 of Sec. VIII). Making use of Eq. 72 we can write down the mesoscale average of $\Phi(x_b;R)$, which is

$$\begin{aligned}\boldsymbol{\Phi}(R) &= \frac{1}{2\epsilon_0}\left[\boldsymbol{\sigma}_L(R)\,(R - |\mathbf{x}_b - \mathbf{x}_L|) + \boldsymbol{\mathcal{M}}_{\boldsymbol{\bar{\rho}}}^{\boldsymbol{\langle 2\rangle}}(R)\right] \\ &= \frac{\boldsymbol{\sigma}_L(R)}{2\epsilon_0}\,(R - |\mathbf{x}_b - \mathbf{x}_L|)\end{aligned} \tag{80}$$

where $\boldsymbol{\mathcal{M}}_{\boldsymbol{\bar{\rho}}}^{\boldsymbol{\langle 2\rangle}}(R)$ is the mesoscale average over x_b of $\bar{\mathcal{M}}_{\bar{\rho}}^{\langle 2\rangle}(x_b)$, which is shown in Appendix E.3 to be zero. The remaining term diverges in the limit $R \to \infty$, which demonstrates that charged surfaces are not stable.

Now let us assume that $\boldsymbol{\sigma}_L = 0$. Eq. 80 becomes $\boldsymbol{\Phi} = 0$, which implies that the macroscopic potential, which is the mesoscale average of the microscopic potential, is zero in the bulk of any material that is mesoscopically charge-neutral in the bulk and which does not have charged surfaces or contain any charged macroscale heterogeneities. There are two very important points to note about this result.

The first is that it contradicts the prevailing view that the MIP is finite and positive (Bethe, 1928; Yesibolati *et al.*, 2020). It also contradicts a view commonly expressed or implied in textbooks on electromagnetism (Jackson, 1998) and solid state physics (Ashcroft and Mermin, 1976; Kittel, 2004), namely, that it is possible for the symmetry of a crystalline microstructure to endow a material with a macroscale electric field.

The second important point is that, because $\Phi[\rho]$ is a linear functional of ρ, this result is the *only* result that could emerge from an internally-consistent theory of structure homogenization. Mathematically, the spatial average $\boldsymbol{\varrho}$ of ρ is simply the weighted sum (integral) of the infinite set of charge densities $\{\rho_u : \rho_u(x+u) \equiv \rho(x), \forall u \in \mathbb{R} \text{ and } \forall x \in \mathbb{R}\}$. It follows from linearity that $\boldsymbol{\Phi} = \langle\Phi[\rho]\rangle_{\varepsilon_x} = \Phi[\langle\rho\rangle_{\varepsilon_x}] = \Phi[\boldsymbol{\varrho}]$. Therefore, if $\boldsymbol{\varrho} = 0$ everywhere, as is the case for an isolated uniform material whose surfaces are uncharged, then $\boldsymbol{\Phi} = 0$ everywhere.

D. Lorentz's fallacy: the macroscopic local field

As discussed in Sec. IV.E, unless there exist sources of macroscopic fields that are external to the material's bulk (e.g., an applied field $\mathbf{E}_{\text{ext}}$ or a net charge at one or more of its surfaces) the isotropy and homogeneity of its macrostructure $\boldsymbol{\varrho}$, which vanishes everywhere in the bulk, preclude the existence of a non-vanishing $\mathbf{E}$ field. Isotropy is incompatible with the existence of a vector field. As discussed in Sec. IV.F and Sec. XIII, if the macroscopic charge density is zero in the bulk of an isolated material whose surfaces are uncharged, there is no source of macroscopic potential $\boldsymbol{\Phi}$. If $\boldsymbol{\Phi} = 0$ throughout the bulk, $\mathbf{E} = 0$ throughout the bulk.

Nevertheless, most textbooks posit the existence of a non-vanishing $\mathbf{E}$ emanating from the bulks of crystals that lack inversion symmetry at the microscale. Furthermore, it is commonly believed that the net field acting at each point x in a material's bulk is $\mathbf{E} + \mathbf{P}/3\epsilon_0 + \mathcal{E}(x)$, where $\mathcal{E}$ is a microscopic field and $\mathbf{P}/3\epsilon_0$ is another macroscopic ($\vec{k} = 0$) contribution. The purpose of this section is to critically examine the reasoning used to infer the existence of the contribution $\mathbf{P}/3\epsilon_0$. Almost all derivations of this term are based on a construction and line of reasoning first presented by Hendrik A. Lorentz in a series of lectures given at Columbia University in 1906, which were subsequently published in book form (Lorentz, 1916) (p137). His construction, which is illustrated in Fig. 15, is sometimes known as the *Lorentz cavity*. This construction has been used in many textbooks (Born and Huang, 1954; Griffiths, 1999; Jackson, 1998; Kirkwood, 1936, 1940; van Vleck, 1937), but as I will now explain, both Lorentz's original argument and all of its descendents that I am aware of are fatally flawed.

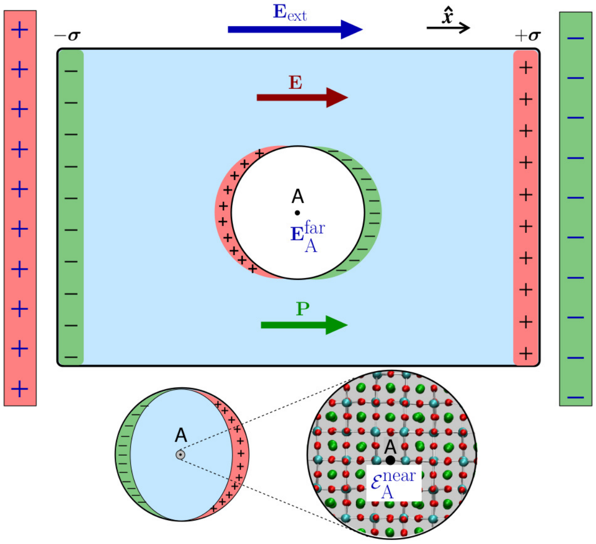

FIG. 14 The *Lorentz cavity* (Lorentz, 1916). See Sec. XIII.D.

Lorentz set out to calculate the average force or electric field acting on a microscopic particle A in the bulk of a dielectric in which there exists a macroscopic electric field $\mathbf{E}$. The particle could be a molecule, an atom, or an electron. He expressed the electric field at A as $\mathcal{E}_A = \mathbf{E}^{\rm far}_{\rm A} + \mathcal{E}^{\rm near}_{\rm A}$, where $\mathbf{E}^{\rm far}_{\rm A}$ is the field emanating from all material beyond a mesoscopic spherical region of radius R centered at A and $\mathcal{E}^{\rm near}_{\rm A}$ is the field emanating from all charges within the region, except A itself. As the boldface notation suggests, $\mathbf{E}^{\rm far}_{\rm A}$ is calculated by treating the material as a continuum; therefore it is a macroscopic quantity. This is reasonable because the length scale on which fluctuations of the microscopic charge density occur is much smaller than the distance ($> R \sim l$) to A. On the other hand $\mathcal{E}^{\rm near}_{\rm A}$ can be expressed as $\mathcal{E}^{\rm near}_{\rm A} = \mathbf{E}^{\rm near}_{\rm A} + \Delta\mathcal{E}^{\rm near}_{\rm A}$, where $\mathbf{E}^{\rm near}_{\rm A}$ and $\Delta\mathcal{E}^{\rm near}_{\rm A}$ are macroscopic and microscopic contributions, respectively. It is a microscopic quantity.

In some presentations of this approach, and as illustrated in Fig. 15, these two separate contributions are imagined in different and separated material systems. To calculate $\mathbf{E}^{\rm far}_{\rm A}$ a continuous material with a cavity in its bulk is imagined, with A at the center of the cavity. To calculate $\mathcal{E}^{\rm near}_{\rm A}$ a microscopically-varying spherical charge distribution is imagined, with A at its center. This is the charge that was evacuated to form the cavity and it is frozen in the arrangement it had prior to being evacuated.

From here, different authors have derived the term $\mathbf{P}/3\epsilon_0$ in different ways, but it tends to be thought of as arising either from the charge on the cavity's surfaces or from the dipole moment of the material evacuated from it. It is interesting that the version of the argument that appears in the first edition of Jackson's book (Jackson, 1962) differs substantially from the one appearing in its 1975 second edition (Jackson, 1975) and that the latter is very similar to the one in Ashcroft and Mermin's 1976 book Ashcroft and Mermin, 1976. However, it is not necessary to go in detail into these differences because we have already introduced the fatal flaw in Lorentz's reasoning and, to my knowledge, all variants of his derivation suffer from it.

Just as the charges at the surfaces of the materials depicted in Figs. 11 and 12 depend sensitively, in magnitude and sign, on how the surfaces are terminated, so too does the charge on the surface of the cavity and the dipole moment of the evacuated material. They both depend sensitively and microscopically on the cavity radius R and vanish when averaged over a continuous mesoscopic range of radii. Therefore the true value of the macroscopic field at A is the sum of only two contributions: the applied field and the field from charge at the material's surfaces.

E. LO-TO splitting

I have argued that inversion asymmetry of a crystal's microstructure does not endow it with a macroscopic $\mathbf{E}$ field. This implies that a macroscopic field is not created when the sublattices of an inversion symmetric crystal are relatively displaced. An oscillating rigid relative displacement of a crystal's sublattices can be regarded as a $\vec{k} = 0$ phonon, so another way of saying that $\mathbf{E}$ vanishes is to say that a $\vec{k} = 0$ phonon does not have an intrinsic electric field.

However, it is well known that the frequency of a $\vec{k} \to 0$ longitudinal optical (LO) phonon is increased by the electric field that is intrinsic to it, and which opposes its motion (Ashcroft and Mermin, 1976; Born and Huang, 1954; Coiana *et al.*, 2024; Jones and March, 1973; Lyddane *et al.*, 1941). Were it not for this field, the frequencies of some crystals' LO and TO phonons would be equal, by symmetry, in the long wavelength limit ($\vec{k} \to 0$). The breaking of this degeneracy by the LO phonon's intrinsic field is commonly referred to as *LO-TO splitting* (Ashcroft and Mermin, 1976; Born and Huang, 1954; Jones and March, 1973).

Therefore I am claiming that, in the $\vec{k} \to 0$ limit, an LO phonon of wavevector $\vec{k}$ creates an electric field of wavevector $\vec{k}$, but that a $\vec{k} = 0$ phonon does not create an electric field in a crystal whose surfaces are earthed.

A $\vec{k} = 0$ phonon does create a uniform (i.e., $\vec{k} = 0$) electric field if the surfaces are not earthed, because the polarization current that flows during the rigid relative motion of sublattices changes the areal charge densities on parallel opposing surfaces at equal and opposite rates. If this charge accumulates, a field emanates from it.

1. Why finite-wavevector LO phonons create electric fields

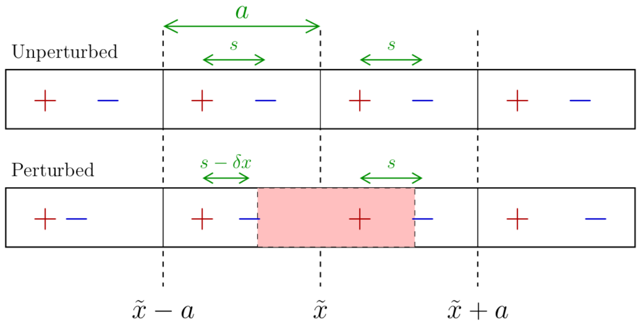

FIG. 15 See text of Sec. XIII.E.1. The net charge within the pink-shaded interval of width a is $+q$. In the unperturbed crystal, the net charge in every interval of width a is zero.

Suppose that the crystal's microstructure (ρ) is modulated along $\hat{x}$ by an LO phonon, of *finite* wavelength λ, propagating along $\hat{x}$. At each instant, this modulation creates regions of excess positive charge and regions of excess negative charge, which alternate along $\hat{x}$ with a wavelength of λ. This excess charge density wave, in turn, creates an electric field of the same wavelength, λ, which opposes the LO mode's motion.

To help understand *why* excesses of charge are created, it is instructive to consider the perfect crystal, without any perturbation, and to calculate the excess charge, $\boldsymbol{\sigma}(\tilde{x})$, on plane $\mathfrak{P}(\tilde{x})$, which is perpendicular to $\hat{x}$ at $\tilde{x}\in\mathfrak{B}$. To do so, we can treat the plane as a pair of adjoined surfaces, use Eq. 51 to calculate the excess charge on each one, and add them to give $\boldsymbol{\sigma}(\tilde{x})$. Because $\tilde{x}$ is in the bulk, and the bulk of an unperturbed crystal is uniform, we can choose $x_L = x_b = \tilde{x}$. We find that the excess charge at $\tilde{x}$ is

$$\begin{aligned}\boldsymbol{\sigma}(\tilde{x}) = \boldsymbol{\sigma}_+(\tilde{x}) + \boldsymbol{\sigma}_-(\tilde{x}) = \frac{1}{a}\int_{-a}^{0}\rho(\tilde{x}+u)\,u\;\mathrm{d}u \\ -\frac{1}{a}\int_{0}^{a}\rho(\tilde{x}+u)\,u\;\mathrm{d}u = 0, \quad (81)\end{aligned}$$

where $\boldsymbol{\sigma}_+$ and $\boldsymbol{\sigma}_-$ are the areal charge densities on the 'surfaces' at $\tilde{x}$ whose outward normals are $\hat{x}$ and $-\hat{x}$, respectively, and their cancellation follows from the periodicity of the crystal.

Now let us consider what happens when an LO phonon breaks periodicity by modulating the structure along the x axis. When this happens, $\boldsymbol{\sigma}_+(\tilde{x})$ and $\boldsymbol{\sigma}_-(\tilde{x})$ are no longer exactly equal in magnitude, in general, which means that $\boldsymbol{\sigma}(\tilde{x})$ does not vanish. Calculating its value is more complicated than in the periodic case because the crystal is no longer uniform. Therefore it is no longer valid to regard the point $\tilde{x}$ as both defining the position of our imaginary surfaces and as points in the 'bulk' beneath them. However, as an illustration, let us calculate the net charge in the interval $\mathfrak{I}(\tilde{x}+u,a)$ averaged over all u between $-a/2$ and $a/2$.

For the purpose of this illustration, let us suppose that each unit cell contains a single anion-cation pair and that the distance between the pair in interval $[\tilde{x}-a,\tilde{x}]$ is smaller by δx than the distance between the pair in interval $[\tilde{x},\tilde{x}+a]$, such that the difference between the dipole moments of these unit cells is $\Delta d = q\delta x$, where q is the cation's charge. Then, the average over $u\in(-a/2,a/2)$ of the net charge in interval $\mathfrak{I}(\tilde{x}+u,a)$ is $q\delta x/a = \Delta d/a$.

Now suppose that we have calculated the same quantity for every pair of adjacent unit cells in an interval $\mathfrak{I}(\tilde{x},\ell)$, where $a \ll \ell \ll \lambda$, and then repeated this calculation for a continuous range of values of $\tilde{x}$. Let us denote $1/\ell$ times the sum of all net charges in interval $\mathfrak{I}(x,\ell)$ by $\langle\rho\rangle^*(x)$ and $1/\ell$ times the sum of the cells' dipole moments by $\langle\mathcal{P}\rangle^*(x)$. Then, it can be shown that $\langle\rho\rangle^*(x) = -\boldsymbol{\nabla}\cdot\langle\mathcal{P}\rangle^*(x)$. The similarity of this expression to the relation $\rho = -\boldsymbol{\nabla}\cdot\mathbf{P}$ is not coincidental: Maxwell used a similar line of reasoning to deduce it, albeit with displacements of charges replaced by displacements of the ether.

This example illustrates that $\boldsymbol{\sigma}(\tilde{x})$ does not vanish in the presence of an LO perturbation because the symmetry reason for it vanishing no longer exists. Furthermore, because a more realistic charge density ρ would be a smooth function of position, $\boldsymbol{\sigma}(\tilde{x})$ would be a smooth function of $\tilde{x}$, with the same periodicity λ as the LO perturbation that created it. Therefore, there would be an excess charge density wave of periodicity λ, from which would emanate an electric field of periodicity λ.

2. Zero-wavevector LO phonons

If $\mathcal{E}_{\text{LO}}(\vec{k},u)$ denotes the electric field created by displacing a crystal by u along the eigenvector of an LO phonon of wavevector $\vec{k}$, my claim about the difference between the point $\vec{k}=0$ in reciprocal space and the $\vec{k}\to 0$ limit can be stated as follows:

$$0 \neq \lim_{\vec{k}\to 0}\mathcal{E}_{\text{LO}}(\vec{k},u) \neq \mathcal{E}_{\text{LO}}(0,u) = 0.$$

In other words, the limit $\vec{k}\to 0$ is a *singular limit* of $\mathcal{E}_{\text{LO}}(\vec{k},u)$.

The fact that $\mathcal{E}_{\text{LO}}(\vec{k},u)$ is discontinuous at $\vec{k}=0$ is well known when expressed in a different way: Squared phonon frequencies are eigenvalues of a crystal's dynamical matrix. Therefore, if $\mathcal{E}_{\text{LO}}(\vec{k},u)$ vanishes suddenly at

$\vec{k} = 0$, causing LO phonon frequencies at $\vec{k} = 0$ to be smaller than their values in the $\vec{k} \to 0$ limit, the dynamical matrix must be discontinuous at $\vec{k} = 0$. It is very well known that it is discontinuous, and *non-analytic corrections* are commonly applied to the $\vec{k} = 0$ dynamical matrix to calculate the $\vec{k} \to 0$ dynamical matrix (Baroni *et al.*, 2001; Born and Huang, 1954; Cochran, 1960; Cochran and Cowley, 1962; Giannozzi *et al.*, 1991; Gonze and Lee, 1997; Jones and March, 1973; Pick *et al.*, 1970). The $\vec{k} \to 0$ dynamical matrix is not one matrix, in general, because both it, and the LO phonon frequency, depend on the direction in reciprocal space from which the point $\vec{k} = 0$ is approached.

To understand why the $\vec{k} \to 0$ limit is singular, it is easier to think about the LO phonon's wavelength in the $\lambda \to \infty$ limit than its wavevector in the $\vec{k} \to 0$ limit: Imagine a microscopic or mesoscopic neighbourhood of a point in the bulk of an arbitrarily large perfect crystal, and then imagine that the crystal is perturbed by displacing it from equilibrium along the eigenvector of an LO phonon of wavelength λ. Now imagine increasing λ.

As λ becomes much larger than the size of the neighbourhood, and continues to increase, the microstructure within the neighbourhood looks more and more like it would look if the crystal's sublattices had been displaced rigidly relative to one another. Therefore it looks more and more like a crystal that has been perturbed by displacing it along the eigenvector of a $\vec{k} = 0$ phonon. Nevertheless, no matter how large λ becomes, if one moves a distance $\lambda/2$ in the direction of $\vec{k}$, the relative displacements of the atoms in the direction of $\vec{k}$ are reversed.

In other words, *microscopically*, increasing λ brings the structure closer to the $\vec{k} = 0$ structure, but *macroscopically* it does not; and it is the macroscopic structure that determines whether or not there is a macroscopic $\mathbf{E}$ field.

XIV. A POTENTIAL PARADOX

In this section I highlight some subtleties in the meaning and definition of the microscopic potential ϕ and its relationships with its macroscopic counterpart $\mathbf{\Phi}$ and the microscopic charge density ρ. As mentioned at the beginning of Sec. XIII, the value of the MIP is believed to be positive (Blumenthal *et al.*, 2017; Cendagorta and Ichiye, 2015; Hörmann *et al.*, 2019; Kathmann, 2021; Kathmann *et al.*, 2011; Leung, 2010; Madsen *et al.*, 2021; Pratt, 1992; Sanchez and Ochando, 1985; Sokhan and Tildesley, 1997a; Wilson *et al.*, 1987, 1988, 1989; Yesibolati *et al.*, 2020). This contradicts my finding that it is zero. Therefore, to illustrate the subtleties, I use the example of Hans Bethe's 1928 derivation of an approximate expression for the MIP, which is sometimes known as the *Bethe potential* ($\mathbf{\Phi}^{\text{Bethe}}$), from several different perspectives. I begin, in Sec. XIV.A, by outlining his derivation and line of reasoning.

My focus is on the 'paradox' referred to in the section title and I do not address the question(s) of most relevance to those using the Bethe potential, or one of its descendants, as a parameter in the analysis and/or interpretation of their calculations (e.g., theoretical electrochemistry) or experiments (e.g., electron holography). In most of these applications $\mathbf{\Phi}^{\text{Bethe}}$ is used as an approximation to the average potential *experienced by an electron* as it passes through the material. This quantity is likely to depend heavily on the electron's energy as it enters the material and the time that it spends inside the material. Furthermore, one should not calculate it from the probability density $n(\vec{r})$ of an electron being at $\vec{r}$, but on the conditional probability density $n_c(\vec{r}_1|\vec{r}_2)$ of there being an electron at $\vec{r}_1$ *given* that the probe electron is at $\vec{r}_2$.

As an illustration of the importance of basing calculations on n_c rather than n, consider the example of a neutral atom meeting a stray electron in a vacuum. One might deduce from the atom's electron density $n(\vec{r})$ that they would not be attracted to one another; but by considering how the distribution of the atom's electrons is changed by their interaction with the stray electron, one can quickly deduce that they do attract one another and that *all* singly-charged anions are stable in vacuum.

A. Bethe's fallacy: the mean inner potential

Bethe assumed that the charge densities of materials are not too dissimilar from a superposition of atomic charge densities. For a crystal with one spherically-symmetric atom in its primitive unit cell Ω, the expression he derived is

$$\mathbf{\Phi}^{\text{Bethe}} = \frac{2\pi e}{3\epsilon_0|\Omega|}\int_0^\infty n(r)r^4\,\mathrm{d}r > 0, \tag{82}$$

where $|\Omega|$ is the volume of Ω; and $n(r)$ is the number of electrons per unit volume in each atom's electron cloud at a distance r from its nucleus. Bethe deduced from Eq. 82 that $\mathbf{\Phi}$ is positive, which contradicts my finding that it is zero. I will now rederive Eq. 82 via a more explicitly-careful mathematical route than Bethe chose to present, but using his starting point and physical reasoning. His starting point was the expression

$$\phi_r(r) \equiv \frac{1}{\epsilon_0 r}\int_0^r \rho(u)u^2\,\mathrm{d}u + \frac{1}{\epsilon_0}\int_r^\infty \rho(u)u\,\mathrm{d}u\,, \tag{83}$$

for the electric potential ϕ_r at a distance r from the center of an isolated spherically-symmetric charge distribution $\rho(r)$. Eq. 83 can be derived from Gauss's law by assuming that the electric field inherits spherical symmetry from ρ and by expressing the potential at a distance r from the nucleus, $\phi_r(r)$, as the integral of the spherically-symmetric field from an infinitely-distant point to one whose distance from the nucleus is r, along an axis passing through the nucleus.

Bethe reasoned that the average potential in the crystal is the potential emanating from one atom, integrated over all points in space, and divided by the volume per atom, i.e., the $R \to \infty$ limit of

$$\bar{\Phi}(R) = \frac{4\pi}{|\Omega|} \int_0^R \phi_r(r) r^2 \, \mathrm{d}r \,. \tag{84}$$

His reason for integrating ϕ_r *over all space*, but dividing by the volume of *only one unit cell*, was that the potential in each cell has contributions from atoms in all other cells and, either by symmetry, or when averaged over all other cells, the sum of the contributions of the atom in a given cell to the averages of the electrostatic potential in all other cells must equal the sum of the contributions of atoms in all other cells to the average electrostatic potential in the given cell.

Bethe expressed each atom's charge density (charge per unit volume at a distance r from its center) as $\rho(r) = \rho^+(r) + \rho^-(r)$, where ρ^+ is the density of nuclear charge and $\rho^-(r) = -e\,n(r)$ is the density of electron charge. He expressed ρ^+ as a delta distribution, i.e.,

$$\rho(r) = Z\,e\,\delta(r) - e\,n(r), \tag{85}$$

but I will keep it more general for now. Substituting Eq. 83 into Eq. 84 and simplifying leads to

$$\begin{aligned}\bar{\Phi}(R) = &-\frac{2\pi}{3\epsilon_0|\Omega|} \int_0^R \rho(r) r^4 \, \mathrm{d}r \\ &+ \frac{\frac{4}{3}\pi R^3}{|\Omega|} \left[\frac{3}{2} \frac{\kappa Q(R)}{R} + \frac{1}{\epsilon_0} \int_R^\infty \rho(r) r \, \mathrm{d}r \right]\end{aligned} \tag{86}$$

where $Q(R) \equiv 4\pi \int_0^R \rho(r) r^2 \, \mathrm{d}r$ is the net charge in a sphere of radius R. When R is chosen large enough that the total charge outside this sphere is negligible, only the first term on the right hand side remains, i.e.,

$$\bar{\Phi}(R) = -\frac{2\pi}{3\epsilon_0|\Omega|} \int_0^R \rho(r) r^4 \, \mathrm{d}r$$

Since $\rho(r) < 0$ when r exceeds the spatial extent of ρ_+, $\bar{\Phi}(R)$ is positive and we recover Eq. 82 in the limit of large R if ρ^+ is localized at a point.

If, instead, we assume that ρ^+ has a finite width and denote the total nuclear charge by Q_+, we can use the atom's overall charge neutrality, i.e., $Q_+ \equiv 4\pi \int_0^\infty u^2 \rho^+(u) \, \mathrm{d}u = -4\pi \int_0^\infty u^2 \rho^-(u) \, \mathrm{d}u$, to express $\mathbf{\Phi}^{\text{Bethe}}$ as

$$\mathbf{\Phi}^{\text{Bethe}} = \frac{Q_+}{6\epsilon_0|\Omega|} \left(s_-^2 - s_+^2 \right) \tag{87}$$

where s_+^2 and s_-^2 are the mean squared distances of positive and negative charges, respectively, from the atom's center, i.e.,

$$s_\pm^2 \equiv \pm \frac{4\pi}{Q_+} \int_0^\infty \rho^\pm(u) u^4 \, \mathrm{d}u$$

From Eq. 87 it seems clear that the MIP being positive is a consequence of the electrons being more delocalized than the nuclei ($s_-^2 > s_+^2$). For example, if we assume that ρ^+ and ρ^- are constant within concentric spheres of radii r_+ and r_-, respectively, and zero outside them, then Bethe's derivation would lead to

$$\mathbf{\Phi}^{\text{Bethe}} = \frac{Q_+}{10\epsilon_0|\Omega|} \left(r_-^2 - r_+^2 \right) > 0.$$

It seems reasonable to interpret this expression as follows: Gauss' Law implies that $\phi_r(r)$ vanishes if $r > r_-$ because the net charge within a sphere of radius r_- centered at the atom's center is zero. The potential is positive at the atom's center and it decreases monotonically to its value of zero at $r = r_-$. Therefore the potential is positive in a sphere of radius r_- and zero everywhere else. It seems obvious, then, that $\mathbf{\Phi}$ must be positive.

I will now show, by illustration, why this obviously-right result must be wrong. Then I will explain the flaw in Bethe's reasoning and show that a more careful treatment of the problem leads to the conclusion that $\mathbf{\Phi}$ is zero. I illustrate the flaw from several different perspectives to highlight some of the many pitfalls that exist when working with the electric potential. Readers who have already spotted the flaw, or who don't like whodunnits, might want to skip the illustrations and proceed directly to Sec. XIV.B.

1. Existence of a flaw - Illustration 1

Bethe chose to build his material from a spherically-symmetric charge density, with a localized distribution of positive charge at its center and relatively delocalized distribution of negative charge surrounding it. However, just as there is no 'right' way to partition the charge density of a material for the purpose of defining its average dipole moment density (see Sec. III..1), there is no right way to partition it for the purpose of calculating its average potential. It is no less justified to build a crystal from a superposition of charge densities of the form

$$\rho(\vec{r}) = -e\,n(r) + \frac{Q_+}{N_A} \sum_{i:|\vec{R}_i|=A} \delta(\vec{r} - \vec{R}_i) \tag{88}$$

where, as before, the origin is chosen to coincide with a nucleus; $r = |\vec{r}|$; each $\vec{R}_i$ is a lattice vector and therefore a relative displacement of two nuclei (for simplicity

I assume that $T \to 0$); and the sum is over all $\vec{R}_i$'s of a given finite magnitude A, of which there are a total of $N_A = \sum_{i:|\vec{R}_i|=A} 1$.

An example of a building block of this form is shown schematically for a 2-d crystal in Fig. 16. It has negative charge at its center, positive charge further away, and it is charge neutral overall, meaning that the net flux of the electric field through any surface that encloses it is zero, as it is for an atom. The flux from an atom is zero at all points on a surface enclosing it, whereas the flux from the charge density of Eq. 88 is finite almost everywhere on a surface enclosing it, but with regions of the surface where it is positive and regions where it is negative. Nevertheless, Gauss's law implies that the net potential outside the surface from charge within it is zero, as it is for an atom.

Using the same physical reasoning with which Bethe deduced that $\boldsymbol{\Phi}^{\text{Bethe}} > 0$ for a material built from atoms, it can be shown that $\boldsymbol{\Phi}^{\text{Bethe}} < 0$ in a material built from this charge distribution, because $s_+ = A > s_-$. Furthermore, the magnitude of $\boldsymbol{\Phi}^{\text{Bethe}}$ depends on the value of A. For example, consider a simple cubic crystal with lattice spacing a and let us build it from Eq. 88 with the choice $A = a$ ($\Rightarrow N_A = 6$). I will refer to the crystal built in this way as *Crystal 2*, I will refer to the crystal built by Bethe from atoms as *Crystal 1*, and I will denote their MIPs, as derived using Bethe's approach, by $\boldsymbol{\Phi}_2^{\text{Bethe}}$ and $\boldsymbol{\Phi}_1^{\text{Bethe}}$, respectively. Then if, following Bethe, we assume the nuclei to be localized at a point, we find that $\boldsymbol{\Phi}_2^{\text{Bethe}} = \boldsymbol{\Phi}_1^{\text{Bethe}} - Q_+a^2/6\epsilon_0|\Omega|$.

Crystal 1 and Crystal 2 are identical in the bulk; they differ only near surfaces. However, because all surfaces of each crystal are charge-neutral ($\boldsymbol{\sigma} = 0$), and because $\boldsymbol{\varrho} = 0$ in the bulk in each case, the macroscale theory would not be internally consistent if $\boldsymbol{\Phi} > 0$ in one case and $\boldsymbol{\Phi} < 0$ in the other. If the value of $\boldsymbol{\Phi}$ is defined it must be the same in each case because the macrostructures of the two crystals are identical: electrostatically, each one is indistinguishable from empty space.

Crystal 1 and Crystal 2 could be made identical by adding a pair of charges of opposite signs, of magnitudes $Q_+/6$, and separated by a distance a, to each surface unit cell of one of the crystals. For example, to make Crystal 2 into Crystal 1 we would have to add a charge of $Q_+/6$ to the center of each of the electron clouds closest to its surface and a charge of $-Q_+/6$ at a displacement $a\,\hat{n}$ from the first charge, where $\hat{n}$ is the surface's outward unit normal. Let us temporarily assume that, from the perspective of a point whose depth below the surface is much greater than a, this is equivalent to adding an approximately-uniform areal density $\sigma_{\mathcal{P}} = -\left(Q_+a/6a^2\right)\hat{n} = -\left(Q_+a^2/6|\Omega|\right)\hat{n}$ of dipole moments to the surface. Then the result would be an upward shift of the potential in the crystal relative to the vacuum

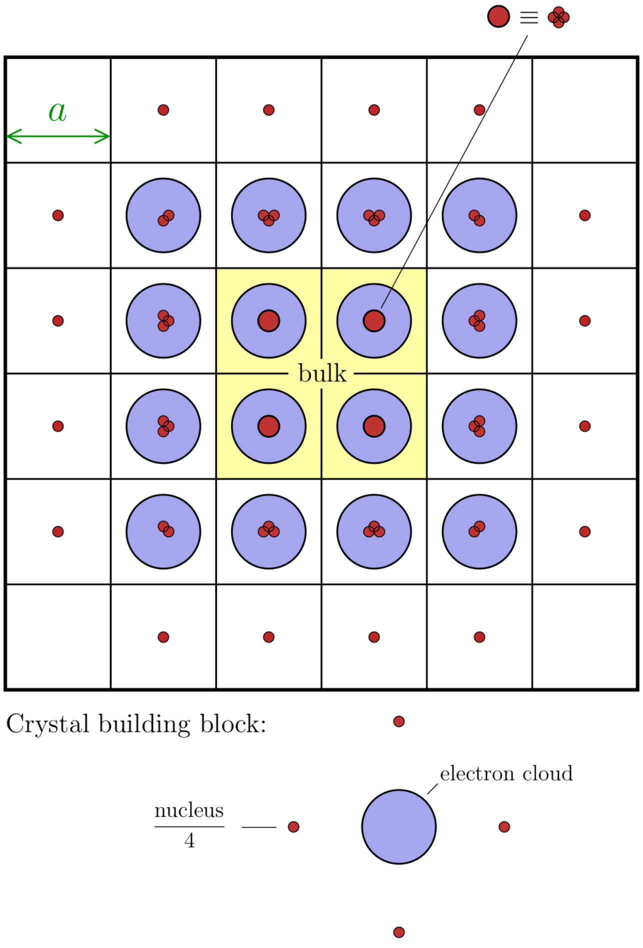

FIG. 16 If a crystal is built from a superposition of atoms (Crystal 1), Bethe's derivation leads to the result $\boldsymbol{\Phi}_1^{\text{Bethe}} > 0$. Crystal 2, shown schematically above, is built from a charge-neutral building block comprised of an electron cloud surrounded by fractions of nuclei (bottom). Bethe's derivation leads to the result $\boldsymbol{\Phi}_2^{\text{Bethe}} < 0$ for Crystal 2. Both crystals are identical in the bulk (yellow region), but differ at their surfaces; however the surfaces of both crystals are charge-neutral ($\boldsymbol{\sigma} = 0$). Crystal 2 can be transformed into Crystal 1 by cancelling each of the outermost $+Ze/4$ charges, where Z is the atomic number, with a charge of $-Ze/4$ at the same position and adding a charge of $+Ze/4$ to the centers of the outermost electron clouds. If we assume this to be equivalent to adding a dipole moment density (dipole moment per unit length, in this 2-d example) of $-(Ze/4)\,\hat{n}$ to each surface, where $\hat{n}$ is the surface's outward unit normal vector, it is easy to calculate the effect of such a layer on the MIP for the 3-d analogue of the crystal shown. It would shift the average potential below the surface relative to its value in the vacuum above the surface by exactly $\boldsymbol{\Phi}_1^{\text{Bethe}} - \boldsymbol{\Phi}_2^{\text{Bethe}}$, thereby making the values of the MIP calculated for Crystal 1 by two different routes equal. However, at the macroscale, adding this dipole moment density should have no effect on the value of $\boldsymbol{\Phi}$ because a microscopic distance is indistinguishable from a distance of zero at the macroscale (see Sec. VI). Therefore an isolated dipole moment shrinks to a point of no net charge under the homogenization transformation and the addition of a plane of such dipoles to a surface does not change either the net charge $\boldsymbol{\sigma}$ at the surface or the macroscopic charge density $\boldsymbol{\varrho}$ at any point inside or outside the material.

near the surface of $\Delta\phi = \sigma_{\mathcal{P}}/\epsilon_0 = Q_+ a^2/6\epsilon_0|\Omega|$. This cancels the difference between $\boldsymbol{\Phi}_2^{\text{Bethe}}$ and $\boldsymbol{\Phi}_1^{\text{Bethe}}$. Therefore, although the values of $\boldsymbol{\Phi}^{\text{Bethe}}$ derived by Bethe's method differ, it appears possible to correct the difference between them by changing the surface structure of either crystal to make the two crystals identical.

Unfortunately, although we have corrected the difference between the two values of $\boldsymbol{\Phi}^{\text{Bethe}}$, this does not solve our problem. We still do not have any reason to prefer one building block over another; therefore we do not have any reason to prefer one of the resulting crystal surface structures over the other. We have two derivations, which appear equally valid, and from which we deduce two different values of the MIP. This appears to imply that there is a flaw in the construction that Bethe used for his derivation.

Bethe did not involve surfaces in his derivation because, when $\boldsymbol{\sigma} = 0$, he regarded the MIP as a property of the bulk. However, the fact that $\boldsymbol{\Phi}_1^{\text{Bethe}}$ and $\boldsymbol{\Phi}_2^{\text{Bethe}}$ differ suggests that the MIP is a surface property, which can be changed by adding or removing equal amounts of positive and negative charge at each surface. This is problematic if we wish to identify $\boldsymbol{\Phi}^{\text{Bethe}}$ as the macroscopic potential $\boldsymbol{\Phi}$ because, as mentioned above, the addition of a layer of *microscopic* dipoles to a surface should not change $\boldsymbol{\Phi}$ in an internally-consistent linear macroscale theory. This is because, at the macroscale, a microscopic distance is equivalent to ($\overset{L}{\sim}$) a distance of zero (see Sec. VI and Sec. VII), and because a layer of microscopic dipoles, $qa\hat{n}$, is equivalent to two layers with equal and opposite charges per unit area that are separated by a distance a. Since $a \overset{L}{\sim} 0$ these two layers are equivalent, at the macroscale, to a single charge-neutral layer, which would not change $\boldsymbol{\Phi}$ inside the crystal. Therefore if Bethe's derivation is right, and if $\boldsymbol{\varrho}$ is a linear spatial average of ρ, the MIP cannot be identified as the macroscopic potential because that would be tantamount to saying that the same macroscale distribution of charge can give rise to different values of $\boldsymbol{\Phi}$. This would imply that, even when $\boldsymbol{\varrho}$ and the macroscale boundary conditions are known, the value of $\boldsymbol{\Phi}$ cannot be calculated; its value depends, in some way, on certain microscopic details of ρ that are lost by the $\rho \mapsto \boldsymbol{\varrho}$ homogenization transformation.

If we can assume that ϕ is a linear functional $\phi[\rho]$ of ρ, and that $\boldsymbol{\varrho}$ is a linear spatial average $\langle\rho\rangle$ of ρ, then the linearity of both operations implies that $\phi[\boldsymbol{\varrho}] = \phi[\langle\rho\rangle] = \langle\phi[\rho]\rangle$. Therefore, if there exist two microscopic charge densities, ρ and $\rho + \Delta\rho$, with the same macroscopic charge density $\boldsymbol{\varrho}$ ($\Rightarrow \langle\Delta\rho\rangle = 0$) and different macroscopic potentials, $\boldsymbol{\Phi}$ and $\boldsymbol{\Phi} + \boldsymbol{\Delta\Phi}$, then $\boldsymbol{\Phi}$ does not equal $\phi[\boldsymbol{\varrho}]$ and is a nonlinear functional $\boldsymbol{\Phi}[\rho]$ of ρ. Linearity would imply that $\boldsymbol{\Delta\Phi} \equiv \boldsymbol{\Phi}[\rho + \Delta\rho] - \boldsymbol{\Phi}[\rho] = \boldsymbol{\Phi}[\Delta\rho] \neq 0$; and it would also imply that the mesoscale spatial average $\langle\boldsymbol{\Delta\Phi}\rangle = \langle\boldsymbol{\Phi}[\Delta\rho]\rangle = \boldsymbol{\Phi}[\langle\Delta\rho\rangle]$ is zero because $\langle\Delta\rho\rangle = 0$; therefore $\boldsymbol{\Delta\Phi}$ would be a harmonic function ($\boldsymbol{\nabla^2}\,\boldsymbol{\Delta\Phi} = 0$) that fluctuates microscopically about zero. It would follow that $\boldsymbol{\Delta\Phi}$ and $\boldsymbol{\Phi}$ are microscopic quantities, not a macroscopic ones. On the other hand, nonlinearity of $\boldsymbol{\Phi}[\rho]$ implies that a material's macroscopic potential can depend on its history; for example, if the material's microstructure $\rho = \rho_1 + \Delta\rho_1 = \rho_2 + \Delta\rho_2$ is built by superimposing the charge densities ρ_1 and $\Delta\rho_1$, its macroscopic potential would differ, in general, from its value if it was built by superimposing the two different charge densities, ρ_2 and $\Delta\rho_2$. There are many problems that arise if we are tempted to assume that $\boldsymbol{\Phi}$ depends on microscopic details of ρ that are washed away by the homogenization transformation; I have only mentioned a few of them.

Returning to the example of Crystal 1 and Crystal 2: if we rigidly shift the MIP of Crystal 2 by $\Delta\phi > 0$ by coating its surfaces with a layer of dipoles, the same layer would shift the average potential in the vacuum just outside the crystal by $-\Delta\phi$. Outside Crystal 1, ϕ appears to be zero because the field from each atom is zero. Outside Crystal 2, the *average* of ϕ appears to be zero because the average electric field emanating from each building block is zero. Adding a dipole layer to Crystal 2 to turn it into Crystal 1 appears to shift the *mean vacuum potential* (MVP) in a layer surrounding the crystal up, while shifting $\boldsymbol{\Phi}^{\text{Bethe}}$ down by the same amount. In the limit of large distance from the crystal the potential vanishes, because the crystal is charge neutral overall, but it does not begin to decrease in magnitude significantly until the distance to the closest surface is comparable to one or more of the surface's linear dimensions; therefore the MVP is shifted by $-\Delta\phi$ in a *macroscopic* layer of vacuum surrounding the crystal. So if Bethe's derivation was correct, and if a macroscopic layer of microscopic dipoles could shift the average potential in a macroscopic region, the MVP would be zero in a macroscopic layer of vacuum surrounding the crystal both before and after it had been shifted by a finite amount $-\Delta\phi$! Clearly, this is absurd.

Now consider Fig. 17, which uses the concept of a dipole layer to illustrate one argument for why the MIP is positive. The crystal in question is identical to Crystal 1, so this construction appears to validate Bethe's result. However, there is no justification for dividing the material's microstructure into the blue and pink layers. If, for example, we combined each adjacent pair of pink and blue layers into a single charge-neutral layer, we would find that the MIP vanishes. Therefore, as with the construction Bethe used in his derivation, two equally-justified ways to partition and spatially-average the microstructure leads to two different values of the MIP.

The superposition principle, on which Bethe's derivation and much of electromagnetic theory are based, allows us to do the following: let us partition the space Ω occupied by an electron cloud into M partitions of volume $|\Omega|/M$ and let us divide the nucleus's charge into M 'pieces', such that for each partition there is a piece of nucleus with the same magnitude of charge. Now, after taking the large-M limit, let us displace the pieces to the partitions so that each one becomes charge neutral. The atom's spherical symmetry initially ensures that, after displacing all of the pieces, it is again spherically symmetric. After this redistribution of charge, the MIP must be zero because the nucleus and every partition have become charge neutral. If we view the displacement of nuclear charge as the superposition of the negative of an atom's charge density on each atom, this makes sense. We have simply superimposed a crystal's charge density and its negative, so of course the MIP of the superposition vanishes. However, we could also view the displacement of each piece of nucleus as the placement of a negative charge at the nucleus and a positive charge in the partition. Placing a dipole at a point in space changes the potential everywhere but, by symmetry, it cannot change the spatial average of the potential. Therefore placing all of these dipoles inside the crystal should not change the MIP. It appears that the superposition principle does not apply.

2. Existence of a flaw - Illustration 2

Another way to see that there must be problem with Bethe's result is to treat electrons as point particles instead of expressing ρ^- as a smooth and delocalized density. Using a line of reasoning similar to Bethe's we could say that the total potential in each unit cell from all electrons and nuclei outside the cell is approximately equal to the sum of the total potentials emanating from the particles inside the cell. Then we could calculate the total potential from each point particle at all points within a distance R of it, add together the total potentials from all particles within each unit cell of the crystal, and take the limit of large R to get the total potential emanating from each unit cell. This total would vanish because the potential emanating from a point charge Ze is the negative of Z times the total potential emanating from a point charge $-e$. Cancellation is obvious when $Z = 1$ (hydrogen), but Bethe's construction does not exclude this case, either explicitly or implicitly. Therefore his derivation leads to the conclusion that the magnitude of the spatial average of the potential from a proton is greater than the magnitude of the spatial average of the potential from an electron.

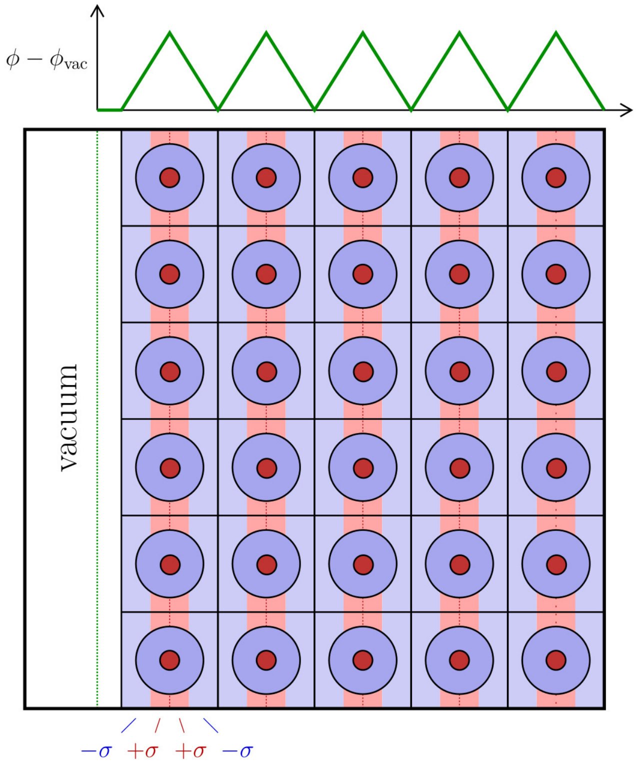

FIG. 17 One can think of each of the coloured layers parallel to the surface as planes of charge, with the layers coloured pink carrying charge densities of $\sigma > 0$ and those coloured blue carrying charge densities of $-\sigma$. Alternatively, one can think of each blue layer and the red layer next to it as a plane of dipoles. The plane of dipoles closest to the surface is pointing into the surface, causing an upward step in ϕ, which in this example denotes the planar average of the microscopic potential, relative to its value of ϕ_{vac} in the vacuum close to the surface (dotted green line). The next layer causes a downward step of ϕ, and so on. The average potential is clearly positive if ϕ_{vac} is regarded as the zero of potential. It would be negative if the layer closest to the vacuum was positive instead of negative, as in Fig. 16, and if the value of ϕ_{vac} was again set equal to zero, despite its value differing from its value above the surface of the crystal pictured above.

This suggests that the problem in Bethe's derivation might be related to his use of a continuous charge density for electrons and a (discrete) delta distribution of charge for nuclei. Usually this form of ρ is regarded as a time average of the true charge distribution: electrons are whizzing around the more massive nuclei so fast that their charge, when observed on a timescale of about $10^{-16} - 10^{-15}$ seconds, appears to be smeared into a continuous charge density. This timescale is too short for nuclei to move significantly, but long enough for each electron to trace out a very long trajectory. Nevertheless, because the integral of

$$\phi\left(\vec{u}\right) \equiv \overbrace{\sum_{i\in\text{nuclei}} \frac{\kappa Z e}{\left|\vec{u} - \vec{R}_i\right|}}^{\phi_n(\vec{u})} - \overbrace{\sum_{j\in\text{electrons}} \frac{\kappa e}{\left|\vec{u} - \vec{r}_j\right|}}^{\phi_e(\vec{u})}$$

over all space vanishes, which implies that its average over all space, $\langle\phi(\vec{u})\rangle_{\vec{u}}$, vanishes, one might expect the spatial average $\langle\langle\phi(\vec{u})\rangle\rangle_{\vec{u}}$ of its expectation value $\langle\phi(\vec{u})\rangle \equiv \langle\Psi|\phi(\vec{u})|\Psi\rangle \,/\, \langle\Psi|\Psi\rangle$ with respect to the material's wavefunction Ψ to also vanish. However, it is easy to reduce this expectation value and its average over all $\vec{u}$ to the forms

$$\begin{aligned}\langle\phi(\vec{u})\rangle =& \frac{\langle\Psi|\phi(\vec{u})|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \phi_n(\vec{u}) - \kappa\, e \int \frac{n(\vec{r})}{|\vec{u}-\vec{r}|}\,\mathrm{d}^3r\,, \\ \langle\langle\phi(\vec{u})\rangle\rangle_{\vec{u}} &= \langle\phi_n\rangle_{\vec{u}} + \left\langle -\kappa\, e \int \frac{n(\vec{r})}{|\vec{u}-\vec{r}|}\,\mathrm{d}^3r \right\rangle_{\vec{u}}, \end{aligned} \quad (89)$$

where $n(\vec{r})$ is the probability density that there is an electron at $\vec{r}$. At first glance, Eq. 89 might appear to validate Bethe's approach, because it appears to be the spatial average of the potential from a charge distribution whose form becomes equivalent to the one he used (Eq. 85) when there is spherical symmetry. It would be very strange if it were equivalent: it would mean that $\langle\langle\phi(\vec{u})\rangle\rangle_{\vec{u}}$ is finite but that the expectation value $\langle\langle\phi(\vec{u})\rangle_{\vec{u}}\rangle$ vanishes. Therefore, it would mean that its value is changed simply by changing the order of integration such that the integral with respect to $\vec{u}$ is performed first. We need to understand this better - both physically and mathematically.

B. The flaw

The flaw in Bethe's derivation is that he calculated the electrons' contribution to the potential from a volumetric density of negative charge, $\rho^-(\vec{r}) = -e\, n(\vec{r})$, which is defined at all points in space. Then, because the electron density is spread over a greater volume than the nuclear density on femtosecond time scales, the spatial average of the potential does not vanish.

To understand why it does not vanish, let us again consider the potential $\phi_r(r;\eta)$ at a distance r from the center of a spherically-symmetric nonpositive or nonnegative charge density $\rho(u;\eta)$, where the width of ρ is proportional to the value of η, which is a parameter. Let us denote the integral of ρ within a sphere of radius r by $Q(r;\eta)$ and its integral over all space by $Q_\infty \equiv \lim_{r\to\infty} Q(r;\eta) = \lim_{\eta\to 0} Q(r;\eta)$. Using Eq. 83 we can express the magnitude of the potential at a distance r from the center of ρ as

$$\begin{aligned}|\phi_r(r;\eta)| &= \left|\frac{\kappa Q(r;\eta)}{r} + \frac{1}{\epsilon_0}\int_r^\infty \rho(u;\eta)u\,\mathrm{d}u\right| \\ &< \left|\frac{\kappa Q(r;\eta)}{r} + \frac{1}{\epsilon_0}\int_r^\infty \rho(u;\eta)u\left(\frac{u}{r}\right)\mathrm{d}u\right| \\ \therefore |\phi_r(r;\eta)| &< \frac{\kappa|Q_\infty|}{r}.\end{aligned}$$

At very large distances ($r \gg \eta$) the potential is approximately equal to $\kappa Q_\infty/r$, and it gets closer to this value as r increases. At short distances ($r \sim \eta$) the magnitude of the potential is significantly smaller than $\kappa|Q_\infty|/r$ and the ratio $\phi_r(r;\eta)/\left(\kappa Q_\infty/r\right)$ gets smaller as r decreases.

Now let us consider the average of ϕ_r over all points within a fixed distance R of the center of ρ as the value of η changes. When η/R is very small, the magnitude of ϕ_r at almost all points is approximately $\kappa|Q_\infty|/r$ and it is only significantly smaller than that value in a volume fraction $\sim (\eta/R)^3$ of the sphere. In the limit $\eta/R \to 0$, the average potential is the same as it would be if ρ was the delta distribution of a point charge. However, as η increases, the magnitude of the average potential in the sphere of radius R decreases because the fraction of the volume occupied by points at which $|\phi_r|$ is significantly less than $\kappa|Q_\infty|/r$ increases. Therefore, the average potential in the sphere reduces in magnitude as η increases.

This is why the potential from the electrons does not cancel the potential from the nuclei in Bethe's derivation: the value of η is finite for the electrons, but vanishingly small for nuclei, which makes the magnitude of the average potential from the nuclei greater than that from the electron cloud.

I will now explain, from three different perspectives, what is wrong with Bethe's derivation and with his use of the charge density in Eq. 85.

1. Perspective 1

Bethe's use of a continuous electron charge density $\rho^-(\vec{r}) = -e\, n(\vec{r})$ suggests that he interpreted it as the time average of the electrons' instantaneous delta distribution of charge. However $n(\vec{r})$ is not a time average of the electrons' positions, it is a probability density that an electron (any one of them) is at position $\vec{r}$ at any precisely-specified time. The average, over a time interval $\mathfrak{I}(t,\Delta t)$, of the delta distribution of a set of moving particles is not a volumetric charge density, but a set of linear charge densities defined only along the segments of the trajectories followed during $\mathfrak{I}(t,\Delta t)$. Therefore the time average of the electrons' delta distribution is a set of linear charge densities defined on a set of curves and there is no charge at points that do not lie along these curves.

This means that the set of points at which the time-average of the electrons' charge distribution is nonzero is a set whose measure in $\mathbb{R}^3$ is zero, regardless of the magnitude of Δt; therefore electrons are not more delocalized than nuclei because both occupy zero volume. Using the fact that Ze/r is cancelled by $Z \times -e/r$, it is easy to show that the potential from the true time average of the electrons' charge distribution exactly cancels the potential from the nuclei.

2. Perspective 2

Although Eq. 89 appears to be the spatial average of the potential from a set of point nuclei and a continuous density of negative charge $\rho_-(\vec{r}) = -e\,n(\vec{r})$, it is not. The quantity

$$\langle \phi_e(\vec{u}) \rangle = -\kappa e \int \frac{n(\vec{r})}{|\vec{u}-\vec{r}|}\,\mathrm{d}^3 r \qquad (90)$$

is not the potential at $\vec{u}$ from $\rho^-(\vec{r}) = -e\,n(\vec{r})$ because Coulomb's law does not hold for the charge from a continuous charge density within a region of infinitesimal volume. To see that it does not hold, consider a spherical Gaussian surface of radius Δr centered at the point $\vec{r}$. The charge within the sphere is $\rho(\vec{r}) \times \frac{4}{3}\pi(\Delta r)^3$, but the electric field, $\mathcal{E}$, on its surface is not directed radially outward from $\vec{r}$ because it has a contribution from the nucleus. This would not present a problem for point charges because one could choose Δr to be arbitrarily small without changing the amount of charge enclosed by it; and, as Δr got smaller, the direction of the field passing through the surface would become arbitrarily close to radially outward.

However this reasoning does not apply to a continuous charge density because the magnitude of the charge enclosed by the surface scales like $(\Delta r)^3$ in the small Δr limit. A correct application of Gauss's law for a spherically-symmetric charge density leads to Eq. 83, which does not give the same result as Eq. 90. Equation 90 is the correct expression for the expectation value of the potential at $\vec{u}$ from the electrons, but it is not the correct expression for the potential from charge density ρ^-; when there is spherical symmetry, the latter is Eq. 83.

Equation 89 can be expressed in the slightly more general form

$$\langle\langle \phi(\vec{u}) \rangle\rangle_{\vec{u}} = \kappa \left\langle \int \frac{\rho(\vec{r})}{|\vec{u}-\vec{r}|}\,\mathrm{d}^3 r \right\rangle_{\vec{u}}$$

where $\rho = \rho^+ + \rho^-$ and ρ^+ is the distribution of the nuclei. Assuming that ρ has spherical symmetry, and choosing its center as the origin, the integral of $\langle \phi(\vec{u}) \rangle$ over all points within a distance R of its center is

$$\begin{aligned}\int_{|\vec{u}|<R} \langle \phi(\vec{u}) \rangle\,\mathrm{d}^3\vec{u} &= \kappa \int_{|\vec{u}|<R} \mathrm{d}^3 u \int_{\mathbb{R}^3} \mathrm{d}^3 r \left(\frac{\rho(\vec{r})}{|\vec{u}-\vec{r}|} \right) \\ &= \kappa \int_{\mathbb{R}^3} \mathrm{d}^3 r\,\rho(\vec{r}) \int_{|\vec{u}|<R} \frac{1}{|\vec{u}-\vec{r}|}\,\mathrm{d}^3 u\,.\end{aligned}$$

In the limit of large R the right hand side becomes arbitrarily close to

$$\kappa \left(\int_{\mathbb{R}^3} \rho(\vec{r})\,\mathrm{d}^3 r \right) \left(\int_0^R 4\pi u^2 \times \frac{1}{u}\,\mathrm{d}u \right),$$

which vanishes because $\int_{\mathbb{R}^3} \rho(\vec{r})\,\mathrm{d}^3 r$ vanishes. Therefore the total potential from the atom, divided by the volume of a primitive unit cell, which is the quantity calculated by Bethe, is zero. Furthermore, to deduce that it is zero I have not assumed anything about the degrees to which the nuclear density and the electron density $n(\vec{r})$ are localized.

3. Perspective 3

The relations $\mathcal{E} = -\boldsymbol{\nabla}\phi$ and $\rho/\epsilon_0 = -\nabla^2\phi$ are preserved by the homogenization transformation because $\boldsymbol{\nabla}$, ∇^2, and the spatial average are all linear operations, which commute when they are applied in a mutually-consistent manner. For example, if $\langle\,\rangle_x$ and $\langle\,\rangle_y$ denote averages along the x-axis and y-axis, respectively, then $\langle\rho\rangle_x/\epsilon_0 = -\left\langle \nabla^2\phi \right\rangle_x = -\nabla^2 \langle\phi\rangle_x$. However, it cannot generally be true that $\langle\rho\rangle_x/\epsilon_0 = -\nabla^2 \langle\phi\rangle_y$. Calculating the average potential, along an axis normal to a surface, from the average of the charge density on planes parallel to the surface does not, in general, lead to a meaningful result.

This is why it is not physically reasonable to calculate the MIP from the average of the green curve in Fig. 17 and it is why the value of the MIP deduced by averaging the charge distribution in layers parallel to the surface depends on the choice of the layers' positions and thicknesses. For example, if, instead of the division into the pink (P) and blue (B) layers depicted in Fig. 17, each layer was chosen to be a layer of atoms, the green curve would be flat because each layer would be charge neutral. A layer of atoms comprises two pink layers and two blue layers in the order BPPB, so I will call it a BPPB-layer. One could choose the first layer at the surface to be a negatively charged B-layer and all others to be charge-neutral PPBB-layers. In that case the green curve would be a straight line with a negative slope; therefore, not only would the MIP appear to be negative, there would be a macroscopic electric field in the material emanating from the plane of negative charge. If the first layer is a BPP-layer and subsequent layers alternate between B- and BPP-layers, the MIP appears to be negative, with the potential as a function of depth resembling a skewed version of the negative of the green curve in Fig. 17.

The electron charge density used by Bethe in his derivation can be regarded as the result of performing the following sequence of temporal and spatial averages: first, the time average of the electron charge distribution is calculated to give a set of curves carrying linear charge densities; next, this set of linear charge densities are turned into a volumetric charge density by averaging in the radial direction over a small width dr; finally, the resulting distribution is given spherical symmetry by setting the charge density at distance r equal to its spatial average on the spherical surface of radius r. The resulting charge density is then used to calculate the potential as a function of position along the radial direction, i.e., along an axis that, at its point of intersection with the surface on which the final spatial average is taken, is perpendicular to this surface. There is no reason to expect this procedure to produce results that are any more meaningful than those derived by partitioning the surface in Fig. 17 into artificial layers using an arbitrary, unjustified, and mutually-inconsistent sequence of partial spatiotemporal averages.

I avoided the problems with Bethe's derivation in Sec. XIII by not making any assumptions about the microscopic charge density, except that it is mathematically smooth; and by only averaging along the x-axis; and by using the same mesoscopic interval width for all spatial averages. Because I used a general form of ρ, the derivations of Sec. XIII apply to charge distributions that are arbitrarily close to delta distributions; and Coulomb's law can be applied to a smooth charge distribution in this limit.

XV. SUMMARY

Structure homogenization

This work lays some foundations of a theory of the relationship between a *microstructure* and its *macrostructure*. The microstructure is assumed to consist of one or more differentiable fields (e.g., $\nu : \mathbb{R}^n \to \mathbb{R}$) which fluctuate on the *microscale* a. The microstructure's macrostructure is the observable manifestations of these fields on the *macroscale*, $L \ggg a$.

Structure homogenization theory, in its most basic form, is founded on two premises: The first is that the Fourier transform $\tilde{\nu} \equiv \mathfrak{F}[\nu]$ of each field ν contributing to the microstructure satisfies

$$\int_{|k|<k_L} |\tilde{\nu}(k)|^2 \,\mathrm{d}k \gg \int_{|k|\in(k_L,k_a)} |\tilde{\nu}(k)|^2 \,\mathrm{d}k \ll \int_{|k|>k_a} |\tilde{\nu}(k)|^2 \,\mathrm{d}k\,, \qquad (91)$$

where $k_a \equiv 2\pi/\mathfrak{a}$ and $k_L \equiv 2\pi/\mathfrak{L}$; the microscale a and the macroscale L are defined by $\epsilon \sim a \iff \epsilon < \mathfrak{a}$ and $\epsilon \sim L \iff \epsilon > \mathfrak{L}$; $\mathfrak{a}$ is a property of the microstructure and $\mathfrak{L}$ is determined by both the microstructure and the scale on which the microstructure is observed. The second premise is that when ν is observed or measured with a probe of macroscopic dimensions (e.g., the pupil of an eye), what is observed is the weighted spatial average of ν on a mesoscopic domain. Roughly-speaking, Eq. 91 means that the microstructure fluctuates on the microscale a and the macroscale L, but there exists an intermediate *mesoscale* l, where $a \ll l \ll L$, on which its fluctuations are negligible. Therefore, on any mesoscopic domain, the average of the microscopic fluctuations of ν almost vanish and nonlinear contributions to its macroscopic variations are negligible.

It turns out that the homogenization transformation creates observable artefacts, which are consequences of the fact that perfect homogenization, meaning a total elimination of microscopic fluctuations, is only possible in the limit in which the macrostructure is an average of the microstructure over all points in its domain. This is the limit in which $\boldsymbol{\mathcal{V}}$, the counterpart of ν at the macroscale, is flat and featureless. For example, the Earth's surface macrostructure is close to this limit in Voyager 1's famous '*pale blue dot*' photograph.

Away from this limit, homogenization is imperfect, and the macroscopic field $\boldsymbol{\mathcal{V}}$ can only be defined to a precision, $\varepsilon_{\mathcal{V}}$, that is finite. The finite precision of $\boldsymbol{\mathcal{V}}$ means that if the only way to distinguish between two points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ is to observe the difference $\boldsymbol{\mathcal{V}}(\mathbf{x}_1) - \boldsymbol{\mathcal{V}}(\mathbf{x}_2)$ in its value, there is an unavoidable limit, ε_x, to the precisions with which positions and displacements can be measured or observed. An approximate relationship between $\varepsilon_{\mathcal{V}}$, ε_x, and the uncertainty ε_p in the gradient of $\boldsymbol{\mathcal{V}}$ is $\varepsilon_x \varepsilon_p \propto \varepsilon_{\mathcal{V}}$.

The uncertainty in positions and displacements implies a one to many relationship between points $\mathbf{x}$ at the macroscale and points x at the microscale. Effectively, microstructure homogenization is a compression of space which causes all microscopic distances to vanish. This spatial compression causes surfaces and interfaces, which are ill-defined at the microscale because their widths are indeterminate, to become well-defined and locally planar (zero width) at the macroscale: If the domain of ν is $\mathbb{R}^3$, they become two dimensional manifolds which carry excess fields, in general. For example, homogenizing a material's microscopic volumetric charge density ρ not only defines a macroscopic analogue $\boldsymbol{\varrho}$ of ρ within the material, it also turns the material's boundary into a two dimensional manifold (the surface), which carries an areal charge density, $\boldsymbol{\sigma}$.

I have derived expressions that relate boundary excess fields to the microscopic fields whose homogenization created them. For example, I have derived an expression $\boldsymbol{\sigma}[\rho]$ relating the areal charge density at a surface to the microscopic volumetric charge density, ρ. This expression generalizes Finnis's expression (Finnis, 1998) for the surface charge density of a crystal to amorphous microstructures.

Electrical macrostructure

I used the basic elements of the theory of structure homogenization to deduce how the microscopic fields ρ, $\mathcal{E}$, and ϕ, that appear in Maxwell's vacuum theory of electricity manifest as macroscopic fields, and to deduce the relationships between those macroscopic fields. The set $\{\rho, \mathcal{E}, \phi\}$ does not only define a set $\{\boldsymbol{\varrho}, \mathbf{E}, \boldsymbol{\Phi}\}$ of macroscopic-counterpart fields. It also defines macroscopic excess fields on lower-dimensional manifolds, such as surfaces, interfaces, edges, line defects, and point defects. These manifolds and fields are created by the spatial compression that is intrinsic to structure homogenization.

The linearity of the spatial averaging operation that turns microstructure into macrostructure means that the relationships between ρ, ϕ, and $\mathcal{E}$ are preserved by the homogenization transformation. Therefore $\boldsymbol{\varrho} = -\nabla^2 \boldsymbol{\Phi}$ and $\mathbf{E} = -\boldsymbol{\nabla}\boldsymbol{\Phi}$.

It is a well-known and obvious stability requirement that $\boldsymbol{\varrho} = 0$ in the bulk of every material. It follows that, in the bulk of a macroscopically-uniform material whose surfaces are charge-neutral, either $\boldsymbol{\Phi}$ is constant and $\mathbf{E} = 0$ or $\boldsymbol{\Phi}$ is a linear function of position and $\mathbf{E}$ is constant.

Both the $\mathbf{P}$ and $\mathbf{D}$ fields that appear in macroscopic electromagnetic theory have been interpreted, and their existences justified, in multiple mutually-inconsistent ways since Maxwell introduced them in the 19th century. I have pointed out that none of these interpretations or justifications are valid and that $\mathbf{P}$ and $\mathbf{D}$ appear within physical theory for historical reasons only: $\mathbf{P}$ is not observable, is not a necessary element of electromagnetic theory, cannot be defined uniquely, and its existence is prohibited by macroscale symmetry. Furthermore, it continues to cause a great deal of confusion, without adding to the utility of electromagnetic theory. Scrapping it removes the distinction between $\mathbf{E}$ and the electric displacement $\mathbf{D}$, so $\mathbf{D}$ should also be scrapped.

The only volumetric fields that are required at the macroscale are $\boldsymbol{\Phi}$ and its derivatives $\mathbf{E}$ and $\boldsymbol{\varrho}$; but the linearity of their interrelationships facilitates the decomposition of each one into components with distinct origins and effects. For example, when studying dielectric response it might be useful to write $\mathbf{E} = \mathbf{E}_{\text{ext}} + \boldsymbol{\Delta}\mathbf{E}$ and $\boldsymbol{\varrho} = \boldsymbol{\varrho}_0 + \boldsymbol{\Delta}\boldsymbol{\varrho}$, where $\mathbf{E}_{\text{ext}}$ is an externally-applied electric field, $\boldsymbol{\Delta}\boldsymbol{\varrho}$ is the change that it induces in the charge density, and $\boldsymbol{\Delta}\mathbf{E}$ is the field emanating from $\boldsymbol{\Delta}\boldsymbol{\varrho}$.

When studying the long wavelength electric fields that emanate from modulations of the structure by optical lattice vibrations, it makes more sense to express these modulations directly as changes in charge density ($\Delta\rho$ and/or $\Delta\boldsymbol{\varrho}$) than as a diverging polarization field. The electric field can be calculated directly from the former, whereas the latter must be translated into a charge density to deduce its field. Furthermore, expressing these modulations as a charge density makes the qualitative difference between the long wavelength limit ($\vec{k} \to 0$) and a rigid relative displacement of sublattices ($\vec{k} = 0$) clearer: a longitudinal optical phonon of wavelength $\lambda_L \sim L$ creates an electric field of wavelength λ_L; but if the material's surfaces are earthed, a rigid relative displacement of sublattices does not create any macroscopic field. Therefore if a material is at equilibrium, $\boldsymbol{\varrho} = 0$ implies that there cannot be any macroscopic electric field emanating from its bulk. The electric potential also vanishes unless it has a source. Therefore, if all surfaces of an electromagnetically-isolated ($\mathbf{E}^{\text{ext}} = 0$) material are neutral, $\mathbf{E}$ and $\boldsymbol{\Phi}$ both vanish in its bulk. This has important implications for materials physics.

The absence of a macroscopic field can also be understood as a demand of symmetry: symmetry is scale-dependent and the bulks of all compositionally- and structurally-uniform materials are isotropic at the macroscale, regardless of their microstructures. A vector field that has a linear relationship with $\boldsymbol{\varrho}$ cannot exist if $\boldsymbol{\varrho}$ is uniform because all directions are equivalent. Therefore if $\mathbf{E}$ does not vanish in the bulk of a homogeneous material, it is either externally applied or it emanates from an accumulation of charge at surfaces, interfaces, or other *macroscopic* heterogeneities.

On the macroscale, a material's response to an external field $\mathbf{E}_{\text{ext}}$ is the changing of the areal densities of charge at all points on surfaces and interfaces whose tangent planes are not parallel to $\mathbf{E}_{\text{ext}}$. When $\mathbf{E}_{\text{ext}}$ is perpendicular to two opposing surfaces, and parallel to all others, the net change in the macroscopic field in the material at equilibrium is $\mathbf{E}_{\text{ext}} - \Delta\boldsymbol{\sigma}/\epsilon_0$ where $\Delta\boldsymbol{\sigma} > 0$ is the magnitude of the changes in the surface charges induced by $\mathbf{E}_{\text{ext}}$.

When a crystal possesses a spontaneous polarization field, by which I mean *only* that its microstructure lacks inversion symmetry, any surface perpendicular to an axis of anisotropy would carry an areal surface charge density $\boldsymbol{\sigma}$ unless neutralized by extrinsic charges. A charged surface is unstable unless stabilized by another source of potential, such as an oppositely charged surface.

The definition of surface charge density $\boldsymbol{\sigma}$ as the integral of $\boldsymbol{\varrho}$ across the surface is equivalent to Finnis's definition, which I have generalized to non-crystalline microstructures. By relating currents to changes of surface charge, Finnis's result can be used to calculate the normal component of the current density $\mathbf{J}$ at any interface if the time dependence of ρ at the interface is known.

The current density in an insulator can also be calculated using the main practical result of the *Modern Theory of Polarization* (MTOP), which is a definition of the polarization current density $\mathbf{J}^{(p)}$ in terms of the time-dependent microstructure of a material's bulk. I have shown that this result follows from Finnis's result and that quantum mechanics is not required to derive it. My derivations make clear that the original MTOP definition of polarization current (King-Smith and Vanderbilt, 1993; Resta, 1993; Vanderbilt and King-Smith, 1993) is exact: The fact that it is expressed in terms of single particle states does not constitute an approximation.

Mathematical representations of classical microstructures

I have shown that many aspects of the mathematical structure of the quantum mechanical theory of electricity in materials are compatible with, or required features of, an internally-consistent statistical theory of a *deterministic* dynamical system of charged particles.

If a classical material comprised a large number of particles whose charges and masses were comparable in magnitude to those of electrons and nuclei, the particles' positions would change rapidly and the particles would respond sensitively to the act of observing them. Therefore it would be impossible to observe the particles' instantaneous positions, or even short segments of their trajectories. Therefore, as in quantum mechanics, the *observable* microstructure would not be the particles' positions, but their joint position probability distribution, p.

If a subset of the classical particles comprised N identical particles, it would be impossible to track those particles individually due to their high speeds of motion. Therefore they would be indistinguishable and, if they were massive and mutually-repulsive, the symmetry of their joint position probability distribution, p, and the impossibility of two of them having precisely the same position, would make p non-differentiable (see Appendix B). However the information possessed by p could be expressed as a differentiable anti-symmetric function $\Psi \in \mathcal{L}^2(\mathbb{R}^{3N}, \mathbb{C})$, for which $p = \Psi^*\Psi$.

Any set of 'single particle functions', $\{\varphi_i\}$, which is a complete orthonormal basis of $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$, can be expressed as the union of an infinite number of different, but not all mutually disjoint, sets that each contains exactly N of the single particle functions. From each of those subsets, a different N-particle Slater determinant can be formed, and the set of all those Slater determinants is a complete orthonormal basis of the subspace of $\mathcal{L}^2(\mathbb{R}^{3N}, \mathbb{C})$ containing only its anti-symmetric elements. Therefore, the statistical microstate of a set of mutually-repulsive and indistinguishable classical particles can be expressed exactly as a weighted sum of an infinite number of Slater determinants, or approximated by a weighted sum of a finite number of Slater determinants.

If it is approximated by a single Slater determinant, the particle density is approximated as $n(\vec{r}) \approx \sum_{i=1}^N |\varphi_i(\vec{r})|^2$, and the *Pauli exclusion principle* applies. The exclusion principle is how anti-symmetry (i.e., the requirement that no two particles can occupy the same position) manifests mathematically when a statistical microstate is approximated by a single determinant: if two of the single particle functions contributing to the determinant are the same, the determinant vanishes. When a statistical microstate is approximated by a sum of Slater determinants, the exclusion principle applies to each determinant individually.

If the material's statistical microstructure changes as some stimulus ζ varies, it may be possible to express the particles' number density exactly as $n(\vec{r};\zeta) = \sum_{i=1}^N n_i(\vec{r};\zeta)$, where $n_i(\vec{r};\zeta) \equiv |\varphi(\vec{r};\zeta)|^2$, and the set $\{\varphi_i(\zeta) \in \mathcal{L}^2(\mathbb{R}^3, \mathbb{C})\}_{i=1}^N$ is orthonormal and with elements that vary continuously as ζ changes. When $n(\vec{r};\zeta)$ admits such a representation, the polarization current that flows as ζ changes continuously is exactly $\mathbf{J}^{(p)} = q\dot{\zeta}\sum_{i=1}^N \mathrm{d}\vec{r}_i(\zeta)/\mathrm{d}\zeta$, where q is the particles' charge, and $\vec{r}_i \equiv \int_{\mathbb{R}^3} \mathrm{d}^3r\, \vec{r} n_i(\vec{r};\zeta)$. If there exists one such set of N single-particle states, there exists an infinite number of such sets, which are related to one another by rotations within the N-dimensional subspace of $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ that they span.

When the bulk of a crystal is represented in a torus, delocalized sets whose elements have the crystal's periodicity are known as *Bloch functions*, while localized non-periodic sets are known as *Wannier functions*. The most localized set is known as the set of *maximally localized Wannier functions* (MLWFs) (Ferreira and Parada, 1970; Marzari *et al.*, 2012; Marzari and Vanderbilt, 1997; Souza *et al.*, 2001).

MLWFs do not have an obvious physical interpretation, but there exist several reasons to attach physical meaning to so-called *natural states* and *natural orbitals*. Many of these were summarized by Coleman (Coleman, 1963), while others are illustrated by results derived in Appendix D.

Appendices

Appendix A: Areal charge densities and other integrals of macroscopic fields across interfaces

Here I quote expressions for the excesses of microscopic scalar fields at macroscale surfaces and interfaces. Derivations of these expressions can be found in Sec. VII and a derivation of a special case of one of them can be found in Finnis, 1998.

I use the microscopic charge density ρ as my example, but the expressions are applicable to any scalar field ν, including scalar components of vector or tensor fields. They are valid when there exists a mesoscale l on which the statistical characteristics of the microscopic ($\sim a$) fluctuations in ν do not vary appreciably and the macroscale counterpart $\boldsymbol{\mathcal{V}}$ of ν varies at most linearly. When that is the case it is possible to calculate a mesoscale average $\bar{\nu}$ of ν at each point, such that $\Delta\nu \equiv \nu - \bar{\nu}$ fluctuates microscopically about zero, and such that $\bar{\nu}$ equals $\boldsymbol{\mathcal{V}}$ to within the finite precision $\varepsilon_{\mathcal{V}}$ to which $\boldsymbol{\mathcal{V}}$ is defined.

Since ρ fluctuates microscopically about zero in a material's bulk, $\bar{\rho} = 0$ is among the values of its mesoscale average at each point. Therefore, I define $\Delta\rho \equiv \rho$ and I will express the quoted formulae in terms of ρ, rather than $\Delta\rho$. Before quoting them, it is necessary to explain the construction used to define the quantities appearing within them.

Consider a material with two surfaces which, on the mesoscale, are locally planar where they intersect the Cartesian x axis. Let x_L and x_R be points on the x axis in the vacuum immediately outside the material at the surfaces with outward unit normals $-\hat{x}$ and $+\hat{x}$, respectively. Let $x_b \in (x_L, x_R)$ be a point on the x axis that is arbitrary apart from the requirement that it is far enough away from both surfaces that it can be regarded as being in the bulk of the material. In each of the expressions quoted below, x_b should be regarded as a point in the bulk below whichever surface (at x_L or x_R) the expression pertains to.

Now consider a mesoscopic neighbourhood $\mathfrak{I}(x_b + u, \ell)$ of x_b, where $u \sim a$ and $\ell \sim l$. Let us assume that it is partitioned into contiguous microscopic ($\sim a$) intervals $\mathfrak{I}(\bar{x}_m, \Delta_m)$, such that x_b is at the boundary point shared by two of them, and such that the average of ρ on each interval vanishes. Let $\mathcal{M}_\rho^{\langle n\rangle}(\bar{x}_m, \Delta_m) \equiv \Delta_m^{-1}\int_{-\Delta_m/2}^{\Delta_m/2} \rho(\bar{x}_m + u)\, u^n \,\mathrm{d}u$ be the n^{th} moment of the m^{th} interval divided by its width $\Delta_m \sim a$, and let $\bar{\mathcal{M}}_\rho^{\langle n\rangle}(x_b)$ denote the average of $\mathcal{M}_\rho^{\langle n\rangle}(\bar{x}_m, \Delta_m)$ over all intervals in the set that partitions $\mathfrak{I}(x_b + u, \ell)$.

Then, if $\boldsymbol{\varrho}$ denotes the macroscopic counterpart of ρ, and $\mathbf{x}_L$ and $\mathbf{x}_R$ denote macroscale points in the vacuum beyond the surfaces normal to $-\hat{x}$ and $\hat{x}$, respectively, the areal densities of charge at these surfaces can be expressed as

$$\begin{aligned}\boldsymbol{\sigma}_L &= \int_{\mathbf{x}_L}^{\mathbf{x}_b} \boldsymbol{\varrho}(\mathbf{x})\mathbf{dx} \\ &= \int_{x_L}^{x_b} \rho(x)\,\mathrm{d}x - \bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b), \qquad \text{(A1)}\\ \boldsymbol{\sigma}_R &= \int_{\mathbf{x}_b}^{\mathbf{x}_R} \boldsymbol{\varrho}(\mathbf{x})\mathbf{dx} \\ &= -\int_{x_b}^{x_R} \rho(x)\,\mathrm{d}x + \bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b), \qquad \text{(A2)}\end{aligned}$$

and the integrals of $\mathbf{x}\boldsymbol{\varrho}(\mathbf{x})$ across the surfaces can be expressed as

$$\begin{aligned}\int_{\mathbf{x}_L}^{\mathbf{x}_b} \mathbf{x}\boldsymbol{\varrho}(\mathbf{x})\mathbf{dx} &= \int_{x_L}^{x_b} x\rho(x)\,\mathrm{d}x - x_b\bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b) \\ &\quad - \bar{\mathcal{M}}_\rho^{\langle 2\rangle}(x_b) \qquad \text{(A3)}\\ \int_{\mathbf{x}_b}^{\mathbf{x}_R} \mathbf{x}\boldsymbol{\varrho}(\mathbf{x})\mathbf{dx} &= \int_{x_b}^{x_r} x\rho(x)\,\mathrm{d}x + x_b\bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_b) \\ &\quad + \bar{\mathcal{M}}_\rho^{\langle 2\rangle}(x_b) \qquad \text{(A4)}\end{aligned}$$

The expression for the charge $\boldsymbol{\sigma}$ at an interface between two materials can be deduced from these expressions by assuming that $\mathbf{x}_{bL}$ and x_{bL} are macroscale and microscale points, respectively, in the bulk of one of the materials, and that $\mathbf{x}_{bR} > \mathbf{x}_{bL}$ and $x_{bR} > x_{bL}$ are points in the bulk of the other.

$$\begin{aligned}\boldsymbol{\sigma} &= \int_{\mathbf{x}_{bL}}^{\mathbf{x}_{bR}} \boldsymbol{\varrho}(\mathbf{x})\mathbf{dx} \\ &= \int_{x_{bL}}^{x_{bR}} \rho(x)\,\mathrm{d}x + \bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_{bL}) - \bar{\mathcal{M}}_\rho^{\langle 1\rangle}(x_{bR}) \qquad \text{(A5)}\end{aligned}$$

It follows from this expression that the excess of charge at any plane in the bulk of a macroscopically-uniform material vanishes. The plane can be treated as an interface and, since the plane itself is in the bulk, the values of x_{bL} and x_{bR} can both be chosen to be the point at which the x axis intersects the plane. Then the right hand side of Eq. A5 vanishes.

Appendix B: Pauli exclusion principle

Consider an observer's probability density function $p(\vec{r}_1, \vec{r}_2)$ for the positions, $\vec{r}_1$ and $\vec{r}_2$, of two identical classical particles, or two electrons with the same spin. Because the observer cannot distinguish one particle from the other, p must be symmetric with respect to exchange of their positions, i.e., $p(\vec{r}_1, \vec{r}_2) = p(\vec{r}_2, \vec{r}_1)$.

Let us denote the position half way between them by $\vec{r}$ and their relative displacement by $\vec{u}$, so that $p(\vec{r}_1, \vec{r}_2) = p(\vec{r} + \vec{u}/2, \vec{r} - \vec{u}/2)$. Let $p_r : \mathbb{R}^3 \to \mathbb{R}_+; \vec{r} \mapsto p_r(\vec{r})$ be the probability density function for their midpoint $\vec{r}$, and and let us define the conditional probability density function $p_{u|r} : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}_+; (\vec{u}, \vec{r}) \mapsto p_{u|r}(\vec{u}|\vec{r})$, where

$$p_{u|r}(\vec{u}|\vec{r}) \equiv \frac{p(\vec{r} + \vec{u}/2, \vec{r} - \vec{u}/2)}{p_r(\vec{r})}.$$

$p_{u|r}$ is the probability density function for their relative displacement being $\vec{u}$ when it is known that the point half way between them is $\vec{r}$.

Now let us assume that the particles interact with one another, and that this interaction strengthens as the distance between them, $|\vec{u}|$, reduces, such that in the limit $|\vec{u}| \to 0$ their mutual interaction overwhelms any other interactions they may have with their surroundings. Then, because space is isotropic, $p_{u|r}(\vec{u}|\vec{r})$ must become independent of the direction $\hat{u} \equiv \vec{u}/|\vec{u}|$ in the $|\vec{u}| \to 0$ limit. Therefore, in the limit $U \to 0^+$, we can express the probability that $|\vec{u}| < U$ as

$$\Pr(|\vec{u}| < U|\vec{r}) = \int_0^U 4\pi u^2 p_{u|r}(u\hat{u}|\vec{r})\,\mathrm{d}u\,,$$

where $\hat{u}$ is arbitrary and irrelevant and $U > 0$.

If the pair of particles were isolated in vacuum, and if they did not interact with one another, then $\Pr(|\vec{u}| < U|\vec{r})$ would be proportional to $\frac{4}{3}\pi U^3$. In other words, the probability of both particles being within a spherical neighbourhood of $\vec{r}$ would be proportional to the volume of that neighbourhood. It follows that if the particles repel one another, $\Pr(|\vec{u}| < U|\vec{r})$ must decay more rapidly than U^3 in the limit $U \to 0^+$. Let us assume that it decays as U^{3+m}, where $m \geq 0$. This implies that

$$p_{u|r}(U\hat{u}|\vec{r}) = \frac{1}{4\pi U^2}\frac{\mathrm{d}}{\mathrm{d}U}\Pr(|\vec{u}| < U|\vec{r}) \sim U^m,$$

as $U \to 0^+$. Therefore, as the distance $|\vec{u}|$ between the particles becomes very small, we have $p_{u|r}(\vec{u}|\vec{r}) \sim |\vec{u}|^m$, which implies that $p(\vec{r} + \vec{u}/2, \vec{r} - \vec{u}/2) \sim |\vec{u}|^m$.

Since both $\hat{u}$ and the position $\vec{r}$ of the particles' midpoint become irrelevant when the particles' mutual interacton overwhelms all other interactions, let us use the shorthand notation $p(u) \equiv p(\vec{r} + u\hat{u}/2, \vec{r} - u\hat{u}/2)$. Then we can express the limiting behaviour, and the constraint that exchange symmetry is preserved, as

$$p(u) = p(-u) \sim u^m$$
$$\Rightarrow p^{(1)}(u) = -p^{(1)}(-u) \equiv \frac{\mathrm{d}p}{\mathrm{d}u} \sim \begin{cases} 0, & \text{if } m = 0, \\ u^{m-1}, & \text{if } m > 0. \end{cases}$$

If $m = 0$, $p(u)$ is constant. Therefore it does not vanish as $u \to 0^+$ or $u \to 0^-$, but it does respect exchange symmetry in a neighbourhood of the point $u = 0$. However, since a non-constant function cannot be analytic if it is constant on a finite-measure subset of its domain, it is either non-analytic or it is a uniform distribution.

If m is not an integer, then $p(u)$ is non-differentiable at $u = 0$. Furthermore, if $0 < m < 1$, the magnitude of $p^{(1)}(u)$ diverges as $u \to 0$. A divergence of $\left|p^{(1)}\right|$ should be expected if there is a power law repulsion between the particles that increases in magnitude with decreasing separation, and if the rate of this increase does not change as their separation becomes very small.

The value of m cannot be an odd positive integer: If it were, p would take negative values.

If m is a positive even integer, or if $p(u) \sim |u|^m$ and $m > 1$, then $p(u)$ is differentiable, respects exchange symmetry, and only vanishes at $u = 0$. However, its derivative decreases smoothly in magnitude as $u \to 0^\pm$, which means that the repulsion between the particles must reduce in magnitude as they approach one another. In other words, when u is sufficiently small, the probability of both particles being within a sphere of radius u is greater than the probability of both particles being within a spherical shell of radius $U \gg u$, whose volume is the same, i.e., $4\pi U^2 \times \text{thickness} = \frac{4}{3}\pi u^3$.

Therefore, if the particles are mutually repulsive, and if their repulsion does not weaken and vanish as they approach one another, $p(\vec{r}_1, \vec{r}_2)$ cannot be differentiable at $(\vec{r}_1, \vec{r}_2) = (\vec{r}, \vec{r})$ if it respects exchange symmetry.

To overcome this nondifferentiability problem, the information contained in p can be expressed as a smooth antisymmetric function $\Psi = \sqrt{p}e^{i\theta}$, which changes sign wherever $\vec{r}_1 = \vec{r}_2$.

The *Pauli exclusion principle* follows from expressing $\Psi(\vec{r}_1, \vec{r}_2)$ in a basis of mutually-orthogonal functions of a single particle position. For example,

$$\Psi(\vec{r}_1, \vec{r}_2) = \sum_{\alpha\beta} C_{\alpha\beta}\phi_\alpha(\vec{r}_1)\phi_\beta(\vec{r}_2).$$

Since Ψ changes sign wherever $\vec{r}_1 = \vec{r}_2$, it must be antisymmetric with respect to exchange of $\vec{r}_1$ and $\vec{r}_2$, i.e.,

$$\begin{aligned}\Psi(\vec{r}_1, \vec{r}_2) &= \sum_{\alpha\beta} C_{\alpha\beta}\phi_\alpha(\vec{r}_1)\phi_\beta(\vec{r}_2) = -\Psi(\vec{r}_2, \vec{r}_1) \\ &= -\sum_{\alpha\beta} C_{\alpha\beta}\phi_\alpha(\vec{r}_2)\phi_\beta(\vec{r}_1) \\ &= -\sum_{\alpha\beta} C_{\beta\alpha}\phi_\beta(\vec{r}_2)\phi_\alpha(\vec{r}_1) \qquad \text{(B1)}\end{aligned}$$

From the orthogonality of the set of *orbitals* $\{\phi_\alpha\}$ we get

$$C_{\alpha\beta} = \int\int \Psi(\vec{r}_1, \vec{r}_2)\phi^*_\alpha(\vec{r}_1)\phi^*_\beta(\vec{r}_2)\,\mathrm{d}^3r_1\,\mathrm{d}^3r_2 = -C_{\beta\alpha},$$

which implies that $C_{\alpha\alpha} = 0$, $\forall\alpha$. Under the usually-misleading and usually-mistaken interpretation of $\phi_\alpha(\vec{r}_1)\phi_\beta(\vec{r}_2) - \phi_\beta(\vec{r}_1)\phi_\alpha(\vec{r}_2)$ as a state of the system in which one particle *occupies* orbital ϕ_α and the other particle occupies orbital ϕ_β, $C_{\alpha\alpha} = 0$ means that no two particles can occupy the same orbital at the same time.

The traditional idea of a covalent bond illustrates how misleading it can be to interpret single-position basis functions as states occupied by single electrons in a many electron system. When all basis functions are centered on nuclei, one cannot describe density between two nuclei without using basis functions centered on *both* nuclei, and this used to be misinterpreted as the nuclei 'sharing' the electrons that 'occupy' these basis functions.

Nowadays the term 'covalent' is usually used to mean only that bonding does not conform closely to either the metallic limit or the ionic limit. However, the idea that covalency is a qualitatively distinct type of chemical bonding endures to some degree (Bacskay *et al.*, 1997; Cohen, 1992, 1993; Grundmann, 2016; McQuarrie *et al.*, 2011; Mukhopadhyay *et al.*, 2016; Ouellette and Rawn, 2015; Zürcher, 2018).

Appendix C: Wannier functions of minimal width

Wannier function of minimal width is simply a different term for *maximally localized Wannier function* (Marzari and Vanderbilt, 1997). The title of this section acknowledges the work of Ferreira and Parada (Ferreira and Parada, 1970), on which it is based. My presentation differs from theirs in several ways, the most deliberate of which is my avoidance of quantum mechanical perturbation theory. This is to demonstrate that there is nothing specific to quantum mechanics in the theory of Wannier functions and their relationships with Bloch functions. All of this section would apply to the eigenfunctions $b_{k\alpha}(x) = b_\alpha(k, x) \equiv \langle x|b_\alpha, k\rangle$ of any periodic operator $\hat{h}(x) \equiv \langle x|\hat{h}|x\rangle = \hat{h}(x + m|\Omega|)$, $\forall m \in \mathbb{Z}$, where the circumference $S_{\mathfrak{B}}$ of $\mathbb{T}$ is an integer multiple of $|\Omega|$, and $\hat{h} : \mathcal{L}^2(\mathbb{T}) \to \mathcal{L}^2(\mathbb{T})$ is bounded and self-adjoint. For example, a one particle position probability density function $p(x)$ from any classically-modelled process in a crystal whose bulk is represented in $\mathbb{T}$ could be expressed as a smooth function $\psi(x) = \sqrt{p(x)}e^{i\theta(x)}$ and expanded in either a Bloch basis or a Wannier basis. Either basis could be used to build a basis of many-particle states to represent a function Ψ whose square modulus is a classical many-particle position probability density function.

Consider the Wannier transformation of Bloch function $b_\alpha(k, x) = e^{ikx}u_\alpha(k, x)$,

$$w_\alpha(x) = \int_{\Omega^*} f_\alpha(k)b_\alpha(k, x)\,\mathrm{d}k\,, \qquad \text{(C1)}$$

where, as in Sec. V.B.1, the integral $\int_{\Omega^*}\mathrm{d}k$ really means the sum $(2\pi/S_{\mathfrak{B}})\sum_{k\in\Omega^*}$ over the discrete and finite set $\Omega^* \equiv \{k \in 2\pi\mathbb{Z}/S_{\mathfrak{B}} : 0 \leq k < 2\pi/|\Omega|\}$, and f_α is a function that is periodic in reciprocal space and normalized on Ω^*, i.e., $\int_{\Omega^*} f^*_\alpha(k)f_\alpha(k)\,\mathrm{d}k = 1$ and $f_\alpha(k + G) = f_\alpha(k)$, $\forall G \in 2\pi\mathbb{Z}/|\Omega|$. As discussed in Sec. V.B.1, we are free to define each Bloch function $b_\alpha(k, x) = e^{ikx}u_\alpha(k, x)$ and its $|\Omega|$-periodic part, u_α, such that b_α is $|\Omega^*|$-periodic in reciprocal space. Therefore, let us make this choice and let us also normalize each Bloch function such that $\int_{\mathbb{T}} |b_\alpha(k, x)|^2\,\mathrm{d}x = 1$, which implies that $\int_{\mathbb{T}} |w_\alpha(x)|^2\,\mathrm{d}x = 1$.

We are interested in finding the most localized Wannier functions that can be constructed from the eigenfunctions of a Hermitian operator $\hat{h}(x)$ that varies smoothly with x. Therefore, let us assume that all required partial derivatives of each Bloch function $b_\alpha(k, x)$ with respect to both k and x exist, where, as in Sec. V.B.1, its derivative with respect to k is defined as

$$\partial_k b_\alpha(k, x) \equiv \lim_{S_{\mathfrak{B}}\to\infty}\left(\frac{S_{\mathfrak{B}}}{2\pi}\right)\left[b_\alpha\left(k + \frac{2\pi}{S_{\mathfrak{B}}}, x\right) - b_\alpha(k, x)\right]$$

Let us also assume that $f_\alpha(k)$ is a smooth function of k; if it were discontinuous, the Wannier function would not decay exponentially. Smoothness of b_α and f_α in reciprocal space requires that each of them respects the symmetry of the crystal's reciprocal lattice and the topology of the torus $\mathbb{T}$ in which the crystal is represented.

We will be assuming that we are in the limit $S_{\mathfrak{B}} \to \infty$, which means that the set of points in Ω^* is quasicontinuous. It also means that we can assume that the width of w_α is much smaller than $S_{\mathfrak{B}}$. Therefore, when we calculate the spread

$$W_\alpha(X) = \int_{\mathbb{T}} |w_\alpha(x)|^2 (x-X)^2 \, \mathrm{d}x \,,$$

of w_α about an arbitrary point $X \in \mathbb{T}$, or any integral whose integrand is localized in $\mathbb{T}$, $\int_{\mathbb{T}}$ should always be taken to mean $\int_{x_0}^{x_0+S_{\mathfrak{B}}}$ where x_0 is chosen such that the integrand is negligible at $x_0 = x_0 + S_{\mathfrak{B}}$.

Now let us define a generating function $G_\alpha(s,X)$, from which $W_\alpha(X)$ can be calculated, as follows

$$G_\alpha(s,X) \equiv \int_{\mathbb{T}} |w_\alpha(x)|^2 e^{-is(x-X)} \, \mathrm{d}x \tag{C2}$$

$$\begin{aligned} \Rightarrow W_\alpha(X) &= -\lim_{s\to 0} \partial_s^2 G_\alpha(s,X) \\ &= \lim_{s\to 0} \int_{\mathbb{T}} |w_\alpha(x)|^2 \,(x-X)^2\, e^{-is(x-X)} \, \mathrm{d}x \end{aligned} \tag{C3}$$

$G_\alpha(s,X)$ is a Fourier transform of $|w_\alpha(x)|^2$ after it has been displaced by $-X$; therefore, if X was the center of w_α, $G_\alpha(s,X)$ would be the Fourier transform of w_α after its center had been moved to the origin. Inserting Eq. C1 into Eq. C2 gives

$$\begin{aligned} G_\alpha(s,X) = &\int_{\Omega^*} \mathrm{d}k' \int_{\Omega^*} \mathrm{d}k \int_{\mathbb{T}} \mathrm{d}x \; \bar{f}_\alpha(k') \bar{b}_\alpha(k',x) \\ &\times f_\alpha(k+s) b_\alpha(k+s,x) e^{-is(x-X)}, \end{aligned} \tag{C4}$$

where the relationship $b_\alpha(k,x) = e^{ikx} u_\alpha(k,x)$ between Bloch state $b_\alpha(k,x)$ and the $|\Omega|$-periodic function $u_\alpha(k,x)$ has been used; and the $|\Omega^*|$-periodicities of b_α and f_α have been used to shift the domain of one of the integrations over Ω^* to the set $\Omega^* + s$, which contains the elements of Ω^* after they have been translated by s.

If we now multiply boths sides of the eigenvalue equation $\hat{h}(x) b_\alpha(k,x) = \epsilon_\alpha(k) b_\alpha(k,x)$ by e^{-ikx}, we find that $u_\alpha(k,x)$ is an eigenfunction of the k-dependent Hamiltonian $e^{-ikx}\hat{h}(x)e^{ikx}$, which is Hermitian if $\hat{h}(x)$ is Hermitian. Hermiticity means that, for each value of k, the set $\{u_\alpha(k,x)\}_{\alpha=1}^{\infty}$ of all eigenfunctions of $e^{-ikx}\hat{h}(x)e^{ikx}$ is complete and can be chosen to be orthonormal. Let us make this choice, which allows us to express $u_\alpha(k+s,x)$ as

$$\begin{aligned} u_\alpha(k+s,x) &= \sum_\beta C_{\alpha\beta}(k,s) u_\beta(k,x) \\ \Rightarrow b_\alpha(k+s,x) &= e^{isx} \sum_\beta C_{\alpha\beta}(k,s) b_\beta(k,x) \end{aligned} \tag{C5}$$

where

$$\begin{aligned} C_{\alpha\beta}(k,s) &\equiv \langle u_\beta, k | u_\alpha, k+s \rangle \\ &\equiv \int_{\mathbb{T}} \bar{u}_\beta(k,x) u_\alpha(k+s,x) \, \mathrm{d}x \end{aligned} \tag{C6}$$

When s is small, we can express $\bar{u}_\alpha(k+s,x)$ as

$$\bar{u}_\alpha(k+s,x) = \bar{u}_\alpha(k,x) + s\partial_k \bar{u}_\alpha(k,x) + \mathcal{O}\big(s^2\big) \tag{C7}$$

and if we multiply Eq. C7 by Eq. C5, integrate over $\mathbb{T}$, use the orthonormality of the u functions ($C_{\alpha\beta}(k,0) = \delta_{\alpha\beta}$), and rearrange, we get

$$C_{\alpha\alpha}(k,s) = 1 - s \sum_{\beta \neq \alpha} C_{\alpha\beta}(k,s) \,\langle \partial_k u_\alpha, k | u_\beta, k \rangle + \mathcal{O}\big(s^2\big).$$

If we now use this equation to calculate $\partial_s C_{\alpha\alpha}(k,0) = \lim_{s\to 0} (2s)^{-1} \left[C_{\alpha\alpha}(k,s) - C_{\alpha\alpha}(k,-s)\right]$, and use $\lim_{s\to 0} C_{\alpha\beta}(k,s) = \lim_{s\to 0} C_{\alpha\beta}(k,-s)$, which follows from the continuity of u_α in reciprocal space, we find that $\partial_s C_{\alpha\alpha}(k,0) = 0$. If we take the second partial derivative of Eq. C6 with respect to s, set $\beta = \alpha$, perform an integration by parts, and then take the $s \to 0$ limit, we find that

$$\begin{aligned} t_\alpha(k) \equiv \partial_s^2 C_{\alpha\alpha}(k,0) &\equiv -\lim_{s\to 0} \partial_s^2 C_{\alpha\alpha}(k,s) \\ &= \langle \partial_k u_\alpha, k | \partial_k u_\alpha, k \rangle \,. \end{aligned} \tag{C8}$$

Inserting Eq. C5 into Eq. C4 gives

$$\begin{aligned} G_\alpha(s,X) = \sum_\beta &\int_{\Omega^*} \mathrm{d}k' \int_{\Omega^*} \mathrm{d}k \, \bar{f}_\alpha(k') f_\alpha(k+s) \\ &\times C_{\alpha\beta}(k,s) e^{isX} \int_{\mathbb{T}} \bar{b}_\alpha(k',x) b_\beta(k,x) \, \mathrm{d}x \,, \end{aligned}$$

and the orthogonality of the set of Bloch functions means that only the $\beta = \alpha$ term in the sum is nonzero, and only the $k' = k$ term in the integral over k' is nonzero. Therefore,

$$G_\alpha(s,X) = \int_{\Omega^*} \bar{f}_\alpha(k) f_\alpha(k+s) C_{\alpha\alpha}(k,s) e^{isX} \, \mathrm{d}k \,. \tag{C9}$$

Using $C_{\alpha\alpha}(k,0) = 1$, $\partial_s C_{\alpha\alpha(k,0)} = 0$, and equations C9, C8, and C3, it follows that

$$W_\alpha(X) = \int_{\Omega^*} \bar{f}_\alpha(k) \left[-\partial_k^2 - 2iX\partial_k + t_\alpha(k) + X^2\right] f_\alpha(k) \, \mathrm{d}k \,,$$

and if we define $g_\alpha(k) \equiv e^{-ikX} f_\alpha(k)$, we can express this as

$$W_\alpha(X) = \int_{\Omega^*} \bar{g}_\alpha(k) \left[-\partial_k^2 + t_\alpha(k)\right] g_\alpha(k) \, \mathrm{d}k \,. \tag{C10}$$

This is stationary with respect to norm-preserving variations of g_α when g_α is any eigenfunction of

$\hat{\mathcal{q}}_\alpha(k) \equiv \langle k|\hat{\mathcal{q}}_\alpha|k\rangle \equiv -\partial_k^2 + t_\alpha(k)$, and the stationary values of $W_\alpha(X)$ are the eigenvalues of $\hat{\mathcal{q}}_\alpha(k)$. Note that the relationship in reciprocal space between u_α and $t_\alpha(k)$ mirrors the relationship in real space between the Bloch function b_α and what would be referred to within quantum mechanics as its kinetic energy. $t_\alpha(k)$ plays the role of a positive potential in the definition of $\hat{\mathcal{q}}_\alpha(k)$.

If we now set $g_\alpha(k) \equiv |\Omega^*|^{-\frac{1}{2}} e^{i\theta_\alpha(k)}$, f_α has the required normalization and Eq. C10 becomes

$$W_\alpha(X) = \frac{1}{|\Omega^*|}\int_{\Omega^*}\left[|\partial_k\theta_\alpha|^2 + t_\alpha(k)\right]\mathrm{d}k\,, \tag{C11}$$

which has its minimum value when θ_α is a constant, and this minimum value is

$$W_\alpha^{\min}(X) = \frac{1}{|\Omega^*|}\int_{\Omega^*}\langle\partial_k u_\alpha, k|\partial_k u_\alpha, k\rangle\,\mathrm{d}k\,. \tag{C12}$$

Finally, let us return to equations C2 and C9 to find the center $\langle x\rangle_\alpha$ of the Wannier function whose spread about point X is minimal:

$$\begin{aligned}\langle x\rangle_\alpha &= X + i\lim_{s\to 0}\partial_s G_\alpha(k,s)\\ &= X + i\int_{\Omega^*}\bar{f}_\alpha(k)\left[\partial_k + iX\right]f_\alpha(k)\,\mathrm{d}k\\ &= \int_{\Omega^*}\bar{f}_\alpha(k)\left(i\partial_k f_\alpha(k)\right)\mathrm{d}k = X\end{aligned} \tag{C13}$$

I emphasize that this section does not have any quantum mechanical content. I have presented a mathematical derivation, which is as applicable within classical statistical mechanics as it is within quantum mechanics.

Appendix D: Natural expressions for energy at the microscale

The information contained in the position probability density function (pdf), $p(x_1,\cdots,x_N)$, of a classical or quantum mechanical system of N identical particles is also contained in any function of the form $\Psi(x_1,\cdots,x_N) \equiv \sqrt{p(x_1,\cdots,x_N)}e^{i\theta}$, where θ is arbitrary. Not only is it possible to represent the pdf's information as an element of a Hilbert-Lebesgue space, $\mathcal{L}^2(\mathbb{R}^N)$, it appears mandatory to do so when the statistical state function is required to be differentiable, as discussed in Appendix B.

1. Natural states

A *natural p-state* $\mathcal{X}_\alpha(x_1,\cdots,x_p)$ of an isolated system of $N=p+q$ identical particles in a pure state $\Psi(x_1,\cdots,x_N)\in\mathcal{L}^2\left(\mathbb{R}^N\right)$ is an eigenstate of its p^{th}-order reduced density matrix (or simply *p-matrix*). That is,

$$\begin{aligned}\int\Gamma_p(x_1\cdots x_p;x_1'\cdots x_p')\mathcal{X}_\alpha(x_1',\cdots,x_p')\,\mathrm{d}x_1'\cdots\mathrm{d}x_p'\\ = \lambda_\alpha\mathcal{X}_\alpha(x_1,\cdots,x_p),\end{aligned}$$

where λ_α is a nonnegative real number and

$$\begin{aligned}\Gamma_p(x_1\cdots x_p;x_1'\cdots x_p') \equiv \int\Psi(x_1\cdots x_p,x_{p+1}\cdots x_N)\\ \times\bar{\Psi}(x_1'\cdots x_p',x_{p+1}\cdots x_N)\,\mathrm{d}x_{p+1}\cdots\mathrm{d}x_N\,.\end{aligned}$$

Natural states have many nice properties; for example, if $\{\tilde{\mathcal{X}}_\alpha\}$ and $\{\tilde{\mathcal{Y}}_\beta\}$ are not sets of natural states, but are any other complete orthonormal bases of the p-particle and q-particle Hilbert spaces, respectively, then Ψ can be expressed exactly as the double infinite sum

$$\begin{aligned}&\Psi(x_1\cdots x_N)\\ &\quad=\sum_{\alpha,\beta}\tilde{C}_{\alpha\beta}\tilde{\mathcal{X}}_\alpha(x_1\cdots x_p)\tilde{\mathcal{Y}}_\beta(x_{p+1}\cdots x_N),\end{aligned} \tag{D1}$$

for some set of constants $\tilde{C}_{\alpha\beta}\in\mathbb{C}$. However, if $\{\mathcal{X}_\alpha\}$ and $\{\mathcal{Y}_\beta\}$ are the sets of natural p- and q-states this expression simplifies to the single infinite sum

$$\begin{aligned}&\Psi(x_1\cdots x_N)\\ &\quad=\sum_\alpha C_\alpha\mathcal{X}_\alpha(x_1\cdots x_p)\mathcal{Y}_\alpha(x_{p+1}\cdots x_N),\end{aligned} \tag{D2}$$

where $C_\alpha\in\mathbb{C}$ and $\mathcal{X}_\alpha$ and $\mathcal{Y}_\alpha$ are eigenstates of the p-matrix and the q-matrix, respectively, with the same eigenvalue, $\lambda_\alpha\equiv|C_\alpha|^2$. Furthermore,

$$\begin{aligned}C_\alpha\mathcal{Y}_\alpha(x_{p+1}\cdots x_N) &= \left(\langle x_{p+1}\cdots x_N|\otimes\langle\mathcal{X}_\alpha|\right)|\Psi\rangle\\ &\equiv\int\bar{\mathcal{X}}_\alpha(x_1\cdots x_p)\Psi(x_1\cdots x_N)\,\mathrm{d}x_1\cdots\mathrm{d}x_p\,,\end{aligned}$$

which means that $\mathcal{Y}_\alpha$ is simply Ψ after it has been projected onto the Hilbert subspace that is orthogonal to $\mathcal{X}_\alpha$. It also means that both $\mathcal{X}_\alpha$ and $\mathcal{Y}_\alpha$ inherit antisymmetry with respect to interchange of coordinates from Ψ.

I refer the reader to Coleman's 1963 article (Coleman, 1963) for a clear explanation of many of the nice properties of natural states. These properties suggest that natural p-states are the only p-particle states to which physical meaning should be attached in a system comprised of more than p particles. I state only two of these properties here.

Property 1: It can be shown (see Coleman's Theorem 3.1) that if Φ is restricted to the mathematical form

$$\Phi(x_1 \cdots x_N) = \sum_{\alpha \le u,\, \beta \le v} \tilde{C}_{\alpha\beta} \tilde{\mathcal{X}}_\alpha(x_1 \cdots x_p) \tilde{\mathcal{Y}}_\beta(x_{p+1} \cdots x_N),$$

where $u \le v < \infty$, and if $\|\Psi - \Phi\|^2$ is minimized with respect to the set of coefficients $\{\tilde{C}_{\alpha\beta}\}$ and the sets of functions, $\{\tilde{\mathcal{X}}_\alpha\}_{\alpha \le u}$ and $\{\tilde{\mathcal{Y}}_\beta\}_{\beta \le v}$, the minimum is obtained by the following truncation of the sum in Eq. D2:

$$\Phi(x_1 \cdots x_N) = \sum_{\alpha \le u} C_\alpha \mathcal{X}_\alpha(x_1 \cdots x_p) \mathcal{Y}_\alpha(x_{p+1} \cdots x_N),$$

where the coefficients are indexed such that $\alpha < \beta \Rightarrow |C_\alpha| \ge |C_\beta|$.

Property 2: It can also be shown (see Coleman's Theorem 3.3) that if p is odd and $2p < N$, then

$$\int \bar{\mathcal{X}}_\alpha(x_1 \cdots x_p) \mathcal{Y}_\alpha(x_1 \cdots x_N)\, \mathrm{d}x_1 \cdots \mathrm{d}x_p = 0. \tag{D3}$$

In what follows I will often use j as shorthand for x_j and $\mathrm{d}\mu_{(1\cdots M)}$ as shorthand for $\mathrm{d}x_1 \cdots \mathrm{d}x_M$. Therefore,

$$\begin{aligned} |1 \cdots N\rangle &= |1 \cdots p\rangle \otimes |p+1 \cdots N\rangle \\ &\equiv |x_1 \cdots x_p\rangle \otimes |x_{p+1} \cdots x_N\rangle = |x_1 \cdots x_N\rangle . \end{aligned}$$

If $|f\rangle$ is a p-state, $|F\rangle$ is a q-state, and $q > p$, then $\langle f|F\rangle$ denotes the $(q-p)$-state,

$$\langle f|F\rangle \equiv \int \bar{f}(1 \cdots p) F(1 \cdots p \cdots q)\, |p+1 \cdots q\rangle\, \mathrm{d}\mu_{(1\cdots q)} .$$

2. Natural orbitals

In this section I focus on the *natural orbitals* (natural 1-states) of an isolated system with wavefunction Ψ and Hamiltonian $\hat{H}$. I derive an exact expression for the expectation value of $\hat{H}$, which might be useful, but is intended primarily as an illustration to strengthen the case for the natural orbitals being the most 'physical' single particle states.

Let us begin with the exact expression

$$\Psi(1 \cdots N) = \sum_\alpha c_\alpha \varphi_\alpha(1) \Theta_\alpha(2 \cdots N), \tag{D4}$$

where $\{\varphi_\alpha\}$ and $\{\Theta_\alpha\}$ are the sets of natural orbitals and natural $(N-1)$-states, respectively. The functions in each set are mutually orthogonal and normalized to one, i.e., $\langle\varphi_\alpha|\varphi_\beta\rangle = \delta_{\alpha\beta}$ and $\langle\Theta_\alpha|\Theta_\beta\rangle = \delta_{\alpha\beta}$. Ψ is also chosen to be normalized to one, which implies that $\sum_\alpha \lambda_\alpha = 1$, where $\lambda_\alpha \equiv |c_\alpha|^2$.

The N-particle Hamiltonian $\hat{H} = \hat{H}_1 + \hat{W}$ is the sum of a 1-particle term, $\hat{H}_1$, and a 2-particle term, $\hat{W}$. The 1-particle term is such that $\langle 1 \cdots N|\hat{H}_1|1 \cdots N\rangle = \sum_i \hat{h}(i)$, where $\hat{h}$ could be any 1-particle operator. However, assuming that we are dealing with a quantum mechanical system, and specifically a system of identical fermions, it has the form $\hat{h}(i) = \hat{t}(i) + \hat{v}^{\text{ext}}(i)$, where $\hat{t}$ is the single-particle kinetic energy operator and $\hat{v}^{\text{ext}}$ is the external potential operator for space/spin coordinate i. The 2-particle term is such that $\langle 1 \cdots N|\hat{W}|1 \cdots N\rangle = \sum_{i,j>i} \hat{w}(i,j)$, where $\hat{w}(i,j)$ is the interaction between particles with coordinates i and j. I will use $\hat{W}$ more generally to denote the 2-particle interaction operator of a system with M particles, where M can be deduced from its context: it is the number of particles of the state on which $\hat{W}$ acts.

The expectation value of the one-particle energy can be expressed as

$$\begin{aligned} \langle\Psi|\hat{H}_1|\Psi\rangle \equiv \sum_{\alpha,\beta} \bar{c}_\alpha c_\beta \int & \bar{\varphi}_\alpha(x_1) \bar{\Theta}_\alpha(x_2 \cdots x_N) \\ & \times \left(\sum_i \hat{h}(x_i) \right) \varphi_\beta(x_1) \Theta_\beta(x_2 \cdots x_N)\, \mathrm{d}\mu_{(1\cdots N)} , \end{aligned}$$

and if we use the orthonormality, $\langle\Theta_\alpha|\Theta_\beta\rangle = \delta_{\alpha\beta}$, of the natural $(N-1)$-states and the antisymmetry of Ψ, we can simplify it to the form

$$\langle\Psi|\hat{H}_1|\Psi\rangle = \sum_\alpha \mu_\alpha \left(t_\alpha + v_\alpha^{\text{ext}} \right) \tag{D5}$$

where $\mu_\alpha \equiv N\lambda_\alpha \equiv N|c_\alpha|^2$, $t_\alpha \equiv \langle\varphi_\alpha|\,\hat{t}\,|\varphi_\alpha\rangle$, and

$$v_\alpha^{\text{ext}} \equiv \langle\varphi_\alpha|\hat{v}^{\text{ext}}|\varphi_\alpha\rangle = \int v^{\text{ext}}(x) n_\alpha(x)\, \mathrm{d}x ,$$

where $n_\alpha(x) \equiv |\varphi_\alpha(x)|^2$ denotes the density of the α^{th} natural orbital.

The expectation value of the two-particle energy is

$$\begin{aligned}\langle\Psi|\hat{W}|\Psi\rangle &\equiv \sum_{\alpha,\beta}\bar{c}_\alpha c_\beta \int \bar{\varphi}_\alpha(1)\bar{\Theta}_\alpha(2\cdots N)\left(\sum_{i,j>i}\hat{w}(i,j)\right)\varphi_\beta(1)\Theta_\beta(2\cdots N)\,\mathrm{d}\mu_{(1\cdots N)}\\ &= \sum_{\alpha,\beta}\bar{c}_\alpha c_\beta \int \bar{\varphi}_\alpha(1)\bar{\Theta}_\alpha(2\cdots N)\left(\sum_{i>1,j>i}\hat{w}(i,j)+\sum_{j>1}\hat{w}(1,j)\right)\varphi_\beta(1)\Theta_\beta(2\cdots N)\,\mathrm{d}\mu_{(1\cdots N)}\end{aligned}$$

The first sum of interaction terms in the parentheses on the second line does not involve coordinate 1; therefore we can use $\langle\varphi_\alpha|\varphi_\beta\rangle = \delta_{\alpha\beta}$ to express this as

$$\begin{aligned}&\langle\Psi|\hat{W}|\Psi\rangle\\ &\quad=\sum_\alpha \lambda_\alpha W_\alpha^{(N-1)} + \sum_{\alpha,\beta}\bar{c}_\alpha c_\beta\,\langle\Theta_\alpha|\hat{\mathcal{U}}_{\alpha\beta}|\Theta_\beta\rangle \quad \text{(D6)}\\ &\quad=\sum_\alpha \lambda_\alpha W_\alpha^{(N-1)} + \sum_{\alpha,\beta}\bar{c}_\alpha c_\beta\,\langle\varphi_\alpha|\hat{\mathcal{V}}_{\alpha\beta}|\varphi_\beta\rangle\,, \quad \text{(D7)}\end{aligned}$$

where $W_\alpha^{(N-1)} \equiv \langle\Theta_\alpha|\hat{W}|\Theta_\alpha\rangle$ and I have introduced the $(N-1)$-particle operator,

$$\hat{\mathcal{U}}_{\alpha\beta} \equiv \sum_{j>1}\int \bar{\varphi}_\alpha(1)\hat{w}(1,j)\varphi_\beta(1)\,|2\cdots N\rangle\langle 2\cdots N|\,\mathrm{d}\mu_{(1\cdots N)}\,,$$

and the 1-particle operator,

$$\hat{\mathcal{V}}_{\alpha\beta} \equiv \sum_{j>1}\int \bar{\Theta}_\alpha(2\cdots N)\hat{w}(1,j)\Theta_\beta(2\cdots N)\,|1\rangle\langle 1|\,\mathrm{d}\mu_{(1\cdots N)}\,.$$

Combining Eqs. D5 and D7, we find that

$$\begin{aligned}E &\equiv \langle\Psi|\hat{H}|\Psi\rangle\\ &= \sum_\alpha \mu_\alpha\left(t_\alpha + v_\alpha^{\text{ext}} + \frac{1}{N}\sum_\beta \bar{c}_\alpha c_\beta\,\langle\varphi_\alpha|\hat{\mathcal{V}}_{\alpha\beta}|\varphi_\beta\rangle\right)\\ &\quad+\sum \lambda_\alpha W_\alpha^{(N-1)} \qquad \text{(D8)}\end{aligned}$$

Now let us express the second sum in Eq. D7 as $\sum_{\alpha,\beta} = \sum_{\alpha=\beta} + \sum_{\alpha,\beta\neq\alpha}$. Using the symmetry of $|\Theta_\alpha|^2$ with respect to exchange of coordinates, the sum of $\alpha=\beta$ terms reduces to a sum of mean-field interactions between the density, n_α, of orbital φ_α and the density $n_\alpha^{(N-1)}$ of state Θ_α. Therefore we get

$$\begin{aligned}\langle\Psi|\hat{W}|\Psi\rangle &= \sum_\alpha \lambda_\alpha\left[W_\alpha^{(N-1)} + N\,\langle\varphi_\alpha|\hat{v}_\alpha^{MF}|\varphi_\alpha\rangle\right]\\ &\quad+\sum_{\alpha,\beta\neq\alpha}\mu_\alpha\Delta\varepsilon_{\alpha\beta}, \qquad \text{(D9)}\end{aligned}$$

where I have introduced a mean field potential from density $n_\alpha^{(N-1)}$,

$$\begin{aligned}\hat{v}_\alpha^{MF} &\equiv \int v_\alpha^{MF}(1)\,|1\rangle\langle 1|\,\mathrm{d}\mu_{(1)}\,,\\ v_\alpha^{MF}(1) &\equiv \frac{1}{N}\int \hat{w}(1,2)n_\alpha^{(N-1)}(2)\,\mathrm{d}\mu_{(2)}\,,\end{aligned}$$

and where

$$\Delta\varepsilon_{\alpha\beta} \equiv \frac{1}{N}\left(\frac{c_\beta}{c_\alpha}\right)\langle\varphi_\alpha|\hat{\mathcal{U}}_{\alpha\beta}|\varphi_\beta\rangle = \frac{1}{N}\left(\frac{c_\beta}{c_\alpha}\right)\langle\Theta_\alpha|\hat{\mathcal{V}}_{\alpha\beta}|\Theta_\beta\rangle\,.$$

Combining Eqs. D7 and D9 gives

$$\begin{aligned}E &= \sum_\alpha \mu_\alpha\left[t_\alpha + v_\alpha^{\text{ext}} + v_\alpha^{MF} + \sum_{\beta\neq\alpha}\Delta\varepsilon_{\alpha\beta} + \frac{1}{N}W_\alpha^{(N-1)}\right].\\ &= \sum_\alpha \mu_\alpha\left(\varepsilon_\alpha + \sum_{\beta\neq\alpha}\Delta\varepsilon_{\alpha\beta}\right) + \sum_\alpha \lambda_\alpha W_\alpha^{(N-1)}, \quad \text{(D10)}\end{aligned}$$

where $\varepsilon_\alpha \equiv \langle\varphi_\alpha|\hat{t} + \hat{v}^{\text{ext}} + \hat{v}_\alpha^{MF}|\varphi_\alpha\rangle$.

It can be shown (see Coleman (Coleman, 1963)) that $\lambda_\alpha \leq 1/N \Rightarrow \mu_\alpha \leq 1$, with equality if and only if Ψ has the form $\Psi(x_1\cdots x_N) = \hat{\mathcal{A}}\{\varphi_\alpha(x_1)\Theta_\alpha(x_2\cdots x_N)\}$. Therefore, $\mu_\alpha = 1$ in the Hartree-Fock approximation and the Hartree-Fock energy is simply

$$E_{HF} = \sum_\alpha\left(\mu_\alpha\varepsilon_\alpha + \lambda_\alpha W_\alpha^{(N-1)}\right), \qquad \text{(D11)}$$

where $W_\alpha^{(N-1)}$ is a sum of mean-field and exchange interactions between orbitals in the set $\{\varphi_\beta\}_{\beta\neq\alpha}$. By expressing each $W_\beta^{(N-1)}$, where $\beta\neq\alpha$, in terms of orbitals, and rearranging, we could cast E_{HF} into a more recognizable and conventional form.

The form of Eq. D10 suggests the following interpretation: a system of N electrons can be regarded as being in a superposition of states, with each state being a product of a one electron state and its dual $N-1$ state. The *independent-electron energy*, ε_α, of state $|\varphi_\alpha\rangle$ is the sum of its kinetic energy and its interaction with an effective potential, which is the sum of the external potential from the nuclei and a mean field interaction with its dual state. The mean field interaction is $1/N$ times the interaction with charge density $-en_\alpha^{(N-1)}$. Each state $|\varphi_\alpha\rangle$ is coupled to every other state $|\varphi_\beta\rangle$ with which it overlaps and state $|\varphi_\alpha\rangle$'s share of the coupling energy is $\Delta\varepsilon_{\alpha\beta}$. This is a *dressed* coupling, meaning that it is mediated by the electrons in the dual states $|\Theta_\alpha\rangle$ and $|\Theta_\beta\rangle$. In addition to the sum of independent-electron energies and the sum of the coupling energies between natural orbitals,

E includes a weighted sum of the dual states' electron-electron repulsion.

In the thermodynamic limit ($N \to \infty$) the energy of direct coupling between $|\varphi_\alpha\rangle$ and all other natural orbitals vanishes unless φ_α is delocalized and the number of states with which it overlaps scales with N. If all states were localized the energy in the thermodynamic limit would be

$$E = \sum_\alpha \left(\mu_\alpha \varepsilon_\alpha + \lambda_\alpha W_\alpha^{(N-1)} \right),$$

which is a sum of one electron energies and $(N-1)$ electron energies.

a. Appendix to Appendix D.2

Note that $W_\beta^{(N-1)}$, where $\beta \neq \alpha$, depends indirectly on $|\varphi_\alpha\rangle$, and this dependence could be made explicit, but I will not do this. However, I will draw out the dependence of $\Delta\varepsilon_{\alpha\beta}$ on orbitals other than $|\varphi_\alpha\rangle$ and $|\varphi_\beta\rangle$. These dependences enter $\Delta\varepsilon_{\alpha\beta}$ via $\hat{\mathcal{V}}_{\alpha\beta}$, because $|\Theta_\alpha\rangle$ and $|\Theta_\beta\rangle$ both contain finite overlaps with *at least* $N-1$ natural 1-states.

The overlap of $|\Theta_\alpha\rangle$ with $|\varphi_\alpha\rangle$ vanishes by Eq. D3; however $\langle\varphi_\alpha|\Theta_\beta\rangle$ does not vanish, in general, if $\beta \neq \alpha$. Therefore, let us express $|\Theta_\beta\rangle$ as the sum of a state with finite overlap with $|\varphi_\alpha\rangle$ and a state $|\Theta_{\beta\perp\alpha}\rangle$ whose projection onto $|\varphi_\alpha\rangle$ vanishes. To facilitate this decomposition, let us define the *annihilation operator* $\hat{a}_\alpha$ and the *creation operator* $\hat{a}_\alpha^\dagger$ by their actions on an M-particle state χ_M and an $(M-1)$-particle state χ_{M-1}, respectively.

$$\begin{aligned}\left(\hat{a}_\alpha \chi_M\right)(1\cdots M-1) &\equiv M^{\frac{1}{2}} \int \chi_M(1\cdots M)\bar{\varphi}_\alpha(M)\,\mathrm{d}x_M \\ \left(\hat{a}_\alpha^\dagger \chi_{M-1}\right)(1\cdots M) &\equiv M^{-\frac{1}{2}} \hat{\mathcal{A}}\left\{\chi_{M-1}(1\cdots M-1)\varphi_\alpha(M)\right\}\end{aligned}$$

where $\hat{\mathcal{A}}$ is the antisymmetrization operator. With a bit of algebra it can be shown that $\hat{a}_\alpha\hat{a}_\alpha^\dagger + \hat{a}_\alpha^\dagger\hat{a}_\alpha = \hat{\mathcal{I}}$, where $\hat{\mathcal{I}}$ is the identity. Note that this notation is a bit sloppy and, as a result, this expression for the identity is misleading. We should really express it as $\hat{a}_{M+1,\alpha}\hat{a}_{M+1,\alpha}^\dagger + \hat{a}_{M,\alpha}^\dagger\hat{a}_{M,\alpha} = \hat{\mathcal{I}}_M$, where $\hat{a}_{M,\alpha}$ acts on M-particle states to produce $(M-1)$-particle states, $\hat{a}_{M,\alpha}^\dagger$ acts on $(M-1)$-particle states to produce M-particle states, and $\hat{\mathcal{I}}_M$ is the identity in the M-particle Hilbert space. With this in mind, let us proceed with the simpler sloppy notation. We can write

$$\begin{aligned}\Theta_\beta(2\cdots N) &= \hat{a}_\alpha^\dagger\hat{a}_\alpha\Theta_\beta(2\cdots N) + \hat{a}_\alpha\hat{a}_\alpha^\dagger\Theta_\beta(2\cdots N) \\ &= \hat{a}_\alpha^\dagger\Theta_{\beta-\alpha}(2\cdots N-1) + \Theta_{\beta\perp\alpha}(2\cdots N)\end{aligned}$$

where $\langle\varphi_\alpha|\Theta_{\beta-\alpha}\rangle$ and $\langle\varphi_\alpha|\Theta_{\beta\perp\alpha}\rangle$ both vanish. Then,

$$\begin{aligned}N\left(\frac{c_\alpha}{c_\beta}\right)\Delta\varepsilon_{\alpha\beta} &\equiv \langle\Theta_\alpha|\hat{\mathcal{U}}_{\alpha\beta}|\hat{a}_\alpha^\dagger\Theta_{\beta-\alpha}\rangle + \langle\Theta_\alpha|\hat{\mathcal{U}}_{\alpha\beta}|\Theta_{\beta\perp\alpha}\rangle \\ &= \int \bar{\varphi}_\alpha(1)\bar{\theta}_{\alpha\beta}(2)\hat{w}(1,2)\varphi_\alpha(2)\varphi_\beta(1)\,\mathrm{d}\mu(1,2) \\ &\quad + \int \bar{\varphi}_\alpha(1)\bar{\Theta}_\alpha(2\cdots N)\hat{w}(1,2) \\ &\qquad \times \Theta_{\beta\perp\alpha}(2\cdots N)\varphi_\beta(1)\,\mathrm{d}\mu(1\cdots N) \qquad \text{(D12)}\end{aligned}$$

where $|\theta_{\alpha\beta}\rangle \equiv \langle\Theta_{\beta-\alpha}|\Theta_\alpha\rangle$ is a 1-particle state that is orthogonal to $|\varphi_\alpha\rangle$ and, to reach the second equation from the first, I have used the orthogonality of Θ_α to φ_α, as follows: in the expression for $\hat{a}_\alpha^\dagger\Theta_{\beta-\alpha}$, I expanded the antisymmetrized product of $\Theta_{\beta-\alpha}$ and φ_α as a sum; then I used the fact that each integral for which the argument of φ_α is not 2 vanishes.

3. Non-overlapping bodies

In this section I present one way to understand the forces and torques exerted by two bodies on one another when they do not overlap spatially and when each body is charge neutral overall. Each body could be an atom, a molecule, a nanoparticle, or any other object composed of more than one charged particle. I will refer to the bodies as C-particles, where 'C' abbreviates *composite.* I will identify them individually as CP1 and CP2 and I will refer to the isolated system comprised *only* of of CP1 and CP2 as CP1+CP2. For simplicity I assume that each C-particle is composed of only two species of more elementary particle: electrons, and nuclei of atomic number Z. The set of all space/spin coordinates $\vec{r}_i$ of the p particles in CP1 will be denoted by $\vec{R} = \{\vec{r}_1, \vec{r}_2, \cdots \vec{r}_p\}$ and the set of all positions $\vec{s}_j$ of the q particles in CP2 will be denoted by $\vec{S} = \{\vec{s}_1, \vec{s}_2, \cdots \vec{r}_q\}$. As before, I will assume that CP1+CP2 is in a pure state whose wavefunction is $\Psi(\vec{R}, \vec{S})$; and because there is no overlap between CP1 and CP2, it does not change the energy, or the expectation value of any observable, if Ψ is chosen to not have the correct (anti-)symmetry with respect to interchange of coordinates *between* CP1 and CP2. Therefore Ψ can be expressed as (see above or Coleman, 1963)

$$\Psi(\vec{R}, \vec{S}) = \sum_\alpha C_\alpha \mathcal{X}_\alpha(\vec{R})\mathcal{Y}_\alpha(\vec{S}) \qquad \text{(D13)}$$

where each $\mathcal{X}_\alpha$ is an eigenfunction of the integral operator with kernel

$$D_1(\vec{R}; \vec{R}') \equiv \int_{\mathbb{R}^{3p}} \Psi(\vec{R}, \vec{S})\Psi^*(\vec{R}', \vec{S})\,\mathrm{d}^{3q}S$$

and each $\mathcal{Y}_\alpha$ is an eigenfunction of the integral operators with kernel

$$D_2(\vec{S}; \vec{S}') \equiv \int_{\mathbb{R}^{3q}} \Psi(\vec{R}, \vec{S})\Psi^*(\vec{R}, \vec{S}')\,\mathrm{d}^{3p}R\,.$$

That is,

$$\int_{\mathbb{R}^{3q}} D_1(\vec{R};\vec{R}')\mathcal{X}_\alpha(\vec{R}')\,\mathrm{d}^{3p}R' = \lambda_\alpha \mathcal{X}_\alpha(\vec{R}),$$
$$\int_{\mathbb{R}^{3p}} D_2(\vec{S};\vec{S}')\mathcal{Y}_\alpha(\vec{S}')\,\mathrm{d}^{3q}S' = \lambda_\alpha \mathcal{Y}_\alpha(\vec{S}),$$

where $\lambda_\alpha = |C_\alpha|^2$. The functions $\mathcal{X}_\alpha(\vec{R}) = \mathcal{X}_\alpha(\vec{r}_1\cdots\vec{r}_p)$ and $\mathcal{Y}_\alpha(\vec{S}) = \mathcal{X}_\alpha(\vec{s}_1\cdots\vec{s}_q)$ are mutually-orthogonal, normalised to one, and have the appropriate symmetry with respect to interchange of any two identical particles on the same C-particle. For example if $\vec{r}_i$ and $\vec{r}_j$ are the coordinates of electrons on CP1, then

$$\mathcal{X}_\alpha(\vec{r}_1\cdots\vec{r}_i\cdots\vec{r}_j\cdots\vec{r}_p) = -\mathcal{X}_\alpha(\vec{r}_1\cdots\vec{r}_j\cdots\vec{r}_i\cdots\vec{r}_p).$$

The full Hamiltonian $\hat{h}$ of CP1+CP2 can be expressed as $\hat{h} = \hat{h}_1 + \hat{h}_2 + \hat{v}_{12}$, where $\hat{h}_1$ and $\hat{h}_2$ are the Hamiltonians of CP1 and CP2, respectively, and $\hat{v}_{12}$ is the interaction between them, which I express as

$$\hat{v}_{12} = \sum_{\substack{i\in\mathrm{CP1}\\ j\in\mathrm{CP2}}} v_{12}(|\vec{r}_i - \vec{s}_j|)$$

Energy

The energy of the combined system (CP1+CP2) is

$$E = \langle\Psi|\hat{h}|\Psi\rangle = \sum_{\alpha\beta} C^*_\alpha C_\beta \int_{\mathbb{R}^{3p}}\int_{\mathbb{R}^{3q}} \mathcal{X}^*_\alpha(\vec{R})\mathcal{Y}^*_\alpha(\vec{S})\left(\hat{h}_1 + \hat{h}_2 + \hat{v}_{12}\right)\mathcal{X}_\beta(\vec{R})\mathcal{Y}_\beta(\vec{S})\,\mathrm{d}^{3q}S\,\mathrm{d}^{3p}R$$
$$= \sum_\alpha \lambda_\alpha\left(\mathcal{E}^\alpha_1 + \mathcal{E}^\alpha_2\right) + \sum_{\alpha\beta} C^*_\alpha C_\beta \int_{\mathbb{R}^{3p}}\int_{\mathbb{R}^{3q}} \mathcal{X}^*_\alpha(\vec{R})\mathcal{Y}^*_\alpha(\vec{S})\;\hat{v}_{12}\;\mathcal{X}_\beta(\vec{R})\mathcal{Y}_\beta(\vec{S})\,\mathrm{d}^{3q}S\,\mathrm{d}^{3p}R = \sum_\alpha \lambda_\alpha\left(\mathcal{E}^\alpha_1 + \mathcal{E}^\alpha_2\right) + \langle\Psi|\hat{v}_{12}|\Psi\rangle$$

where $\mathcal{E}^\alpha_1 \equiv \langle\mathcal{X}_\alpha|\hat{h}_1|\mathcal{X}_\alpha\rangle$ and $\mathcal{E}^\alpha_2 \equiv \langle\mathcal{Y}_\alpha|\hat{h}_2|\mathcal{Y}_\alpha\rangle$. The interaction energy between the C-particles is

$$\begin{aligned} E_{12} &= \langle\Psi|\hat{v}_{12}|\Psi\rangle \\ &= \int_{\mathbb{R}^{3p}}\int_{\mathbb{R}^{3q}} |\Psi(\vec{r}_1\cdots\vec{r}_p,\vec{s}_1\cdots\vec{s}_q)|^2 \\ &\times\left(\sum_{ij} v_{12}(|\vec{s}_j - \vec{r}_i|)\right)\mathrm{d}^3r_1\cdots\mathrm{d}^3r_p\,\mathrm{d}^3s_1\cdots\mathrm{d}^3s_q \end{aligned}$$

The sum over i and j can be split into separate sums for each different ordered pair of particle types, i.e.,

$$E_{12} = E^{ee}_{12} + E^{nn}_{12} + E^{en}_{12} + E^{ne}_{12}$$

where each different ordered pair is indicated by one of '*ee*', '*en*', '*ne*', and '*nn*', where '*e*' refers to electrons, '*n*' refers to nuclei, and the ordering of the subscripts indicates which C-particle each particle belongs to. For example, E^{ee}_{12} is the energy of interaction between the electrons on CP1 and the electrons on CP2, E^{ne}_{12} is the energy of interaction between the nuclei on CP1 and the electrons on CP2, etc..

Definitions of probability density functions

Using the symmetry of $|\Psi|^2$ with respect to interchange of two identical particles on the same C-particle, it is easy to show that

$$\begin{aligned} E_{12} = \int_{\mathbb{R}^3}\int_{\mathbb{R}^3} &\big[\rho_{ee}(\vec{r},\vec{s}) - Z\rho_{ne}(\vec{r},\vec{s}) - Z\rho_{en}(\vec{r},\vec{s}) \\ &+ Z^2\rho_{nn}(\vec{r},\vec{s})\big]\phi(|\vec{s}-\vec{r}|)\,\mathrm{d}^3r\,\mathrm{d}^3s \end{aligned} \tag{D14}$$

where $\phi(u) \equiv \kappa e^2/u$ and the probability density functions $\rho_{ee}(\vec{r},\vec{s})$, $\rho_{ne}(\vec{r},\vec{s})$, $\rho_{en}(\vec{r},\vec{s})$, and $\rho_{en}(\vec{r},\vec{s})$ are the probability densities there exist a particle of type indicated by the first subscript on CP1 at $\vec{r}$ and a particle of type indicated by the second subscript on CP2 at $\vec{s}$. For example, if $\vec{r}_i$ are the coordinats of an electron on CP1 and $\vec{s}_j$ are the coordinates of a nucleus on CP2, then

$$\begin{aligned} \rho_{en}(\vec{r},\vec{s}) \equiv p_e q_n &\int_{\mathbb{R}^{3p}}\int_{\mathbb{R}^{3q}} |\Psi(\vec{r}_1\cdots\vec{s}_q)|^2 \\ &\times\delta(\vec{r}-\vec{r}_i)\delta(\vec{s}-\vec{s}_j)\,\mathrm{d}^3r_1\cdots\mathrm{d}^3s_q \end{aligned}$$

where $p_e = Zp/(Z+1)$ is the number of electrons on CP1, $q_n = q/(Z+1)$ is the number of nuclei on CP2, and $i \in \{1,2,\cdots,p\}$ and $j \in \{1,2,\cdots,q\}$. The probability density of finding any one of the electrons belonging to CP1 at $\vec{r}$ is denoted by $\rho^{(1)}_e(\vec{r})$, the probability density of finding any one of the nuclei belonging to CP2 at $\vec{s}$ is denoted by $\rho_n^{(2)}(\vec{s})$, etc., i.e.,

$$\begin{aligned}
\rho_e^{(1)}(\vec{r}) &\equiv \frac{1}{q_e}\int_{\mathbb{R}^3} \rho_{ee}(\vec{r},\vec{s})\,\mathrm{d}^3 s = \frac{1}{q_n}\int_{\mathbb{R}^3} \rho_{en}(\vec{r},\vec{s})\,\mathrm{d}^3 s\,,\\
\rho_n^{(1)}(\vec{r}) &\equiv \frac{1}{q_e}\int_{\mathbb{R}^3} \rho_{ne}(\vec{r},\vec{s})\,\mathrm{d}^3 s = \frac{1}{q_n}\int_{\mathbb{R}^3} \rho_{nn}(\vec{r},\vec{s})\,\mathrm{d}^3 s\,,\\
\rho_e^{(2)}(\vec{s}) &\equiv \frac{1}{p_e}\int_{\mathbb{R}^3} \rho_{ee}(\vec{r},\vec{s})\,\mathrm{d}^3 r = \frac{1}{q_n}\int_{\mathbb{R}^3} \rho_{ne}(\vec{r},\vec{s})\,\mathrm{d}^3 r\,,\\
\rho_n^{(2)}(\vec{s}) &\equiv \frac{1}{p_e}\int_{\mathbb{R}^3} \rho_{ne}(\vec{r},\vec{s})\,\mathrm{d}^3 r = \frac{1}{p_n}\int_{\mathbb{R}^3} \rho_{nn}(\vec{r},\vec{s})\,\mathrm{d}^3 r\,,
\end{aligned}$$

The conditional probability density functions $\rho_{e|n}^{(1)}$, $\rho_{n|e}^{(1)}$, $\rho_{e|e}^{(1)}$, $\rho_{n|n}^{(1)}$, $\rho_{e|n}^{(2)}$, $\rho_{n|e}^{(2)}$, $\rho_{e|e}^{(2)}$, and $\rho_{n|n}^{(2)}$ are defined as follows.

$$\begin{aligned}
\rho_{ee}(\vec{r},\vec{s}) &= \rho_{e|e}^{(1)}(\vec{r}|\vec{s})\rho_e^{(2)}(\vec{s}) = \rho_{e|e}^{(2)}(\vec{s}|\vec{r})\rho_e^{(1)}(\vec{r}),\\
\rho_{nn}(\vec{r},\vec{s}) &= \rho_{n|n}^{(1)}(\vec{r}|\vec{s})\rho_n^{(2)}(\vec{s}) = \rho_{n|n}^{(2)}(\vec{s}|\vec{r})\rho_n^{(1)}(\vec{r}),\\
\rho_{en}(\vec{r},\vec{s}) &= \rho_{e|n}^{(1)}(\vec{r}|\vec{s})\rho_n^{(2)}(\vec{s}) = \rho_{n|e}^{(2)}(\vec{s}|\vec{r})\rho_e^{(1)}(\vec{r}),\\
\rho_{ne}(\vec{r},\vec{s}) &= \rho_{n|e}^{(1)}(\vec{r}|\vec{s})\rho_e^{(2)}(\vec{s}) = \rho_{e|n}^{(2)}(\vec{s}|\vec{r})\rho_n^{(1)}(\vec{r}).
\end{aligned}$$

Finally, the densities at $\vec{r}$ of electrons and nuclei for the combined system CP1+CP2 are $\rho_e(\vec{r}) = \rho_e^{(1)}(\vec{r}) + \rho_e^{(2)}(\vec{r})$ and $\rho_n(\vec{r}) = \rho_n^{(1)}(\vec{r}) + \rho_n^{(2)}(\vec{r})$, respectively, and the charge densities of CP1, CP2, and CP1+CP2 are, respectively, $-e\rho^{(1)} \equiv -e\left(\rho_e^{(1)} - Z\rho_n^{(1)}\right)$, $-e\rho^{(2)} \equiv -e\left(\rho_e^{(2)} - Z\rho_n^{(2)}\right)$, and $-e\rho \equiv -e\left(\rho_e - Z\rho_n\right)$.

Simplifying the interaction energy

Using the definitions in the previous section, Eq. D14 can be rewritten as

$$\begin{aligned}
E_{12} = {} & \frac{1}{2}\int_{\mathbb{R}^3} \rho_e^{(1)}(\vec{r})\left(\int_{\mathbb{R}^3}\left[\rho_{e|e}^{(2)}(\vec{s}|\vec{r}) - Z\rho_{n|e}^{(2)}(\vec{s}|\vec{r})\right]\phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s\right)\mathrm{d}^3 r\\
& + \frac{1}{2}\int_{\mathbb{R}^3} \rho_e^{(2)}(\vec{s})\left(\int_{\mathbb{R}^3}\left[\rho_{e|e}^{(1)}(\vec{r}|\vec{s}) - Z\rho_{n|e}^{(1)}(\vec{r}|\vec{s})\right]\phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r\right)\mathrm{d}^3 s\\
& + \frac{Z}{2}\int_{\mathbb{R}^3} \rho_n^{(1)}(\vec{r})\left(\int_{\mathbb{R}^3}\left[Z\rho_{n|n}^{(2)}(\vec{s}|\vec{r}) - \rho_{e|n}^{(2)}(\vec{s}|\vec{r})\right]\phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s\right)\mathrm{d}^3 r\\
& + \frac{Z}{2}\int_{\mathbb{R}^3} \rho_n^{(2)}(\vec{s})\left(\int_{\mathbb{R}^3}\left[Z\rho_{n|n}^{(1)}(\vec{r}|\vec{s}) - \rho_{e|n}^{(1)}(\vec{r}|\vec{s})\right]\phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r\right)\mathrm{d}^3 s
\end{aligned}$$

Defining

$$\begin{aligned}
\delta\rho_{e|e}^{(1)}(\vec{r}|\vec{s}) &\equiv \rho_{e|e}^{(1)}(\vec{r}|\vec{s}) - \rho_e^{(1)}(\vec{r}),\\
\delta\rho_{e|e}^{(2)}(\vec{s}|\vec{r}) &\equiv \rho_{e|e}^{(2)}(\vec{s}|\vec{r}) - \rho_e^{(2)}(\vec{s})\\
\delta\rho_{e|n}^{(1)}(\vec{r}|\vec{s}) &\equiv \rho_{e|n}^{(1)}(\vec{r}|\vec{s}) - \rho_e^{(1)}(\vec{r}),\\
\delta\rho_{e|n}^{(2)}(\vec{s}|\vec{r}) &\equiv \rho_{e|n}^{(2)}(\vec{s}|\vec{r}) - \rho_e^{(2)}(\vec{s}),
\end{aligned}$$

and also

$$\begin{aligned}
\delta\rho_{n|e}^{(1)}(\vec{r}|\vec{s}) &\equiv \rho_{n|e}^{(1)}(\vec{r}|\vec{s}) - \rho_n^{(1)}(\vec{r}),\\
\delta\rho_{n|e}^{(2)}(\vec{s}|\vec{r}) &\equiv \rho_{n|e}^{(2)}(\vec{s}|\vec{r}) - \rho_n^{(2)}(\vec{s})\\
\delta\rho_{n|n}^{(1)}(\vec{r}|\vec{s}) &\equiv \rho_{n|n}^{(1)}(\vec{r}|\vec{s}) - \rho_n^{(1)}(\vec{r}),\\
\delta\rho_{n|n}^{(2)}(\vec{s}|\vec{r}) &\equiv \rho_{n|n}^{(2)}(\vec{s}|\vec{r}) - \rho_n^{(2)}(\vec{s}),
\end{aligned}$$

allows E_{12} to be expressed as follows.

$$
\begin{aligned}
E_{12} = {} & \frac{1}{2}\int_{\mathbb{R}^3} \rho_e^{(1)}(\vec{r}) \left(\int_{\mathbb{R}^3} \left[\rho^{(2)}(\vec{s}) + \delta\rho_{e|e}^{(2)}(\vec{s}|\vec{r}) - Z\delta\rho_{n|e}^{(2)}(\vec{s}|\vec{r}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s \right) \mathrm{d}^3 r \\
& + \frac{1}{2}\int_{\mathbb{R}^3} \rho_e^{(2)}(\vec{s}) \left(\int_{\mathbb{R}^3} \left[\rho^{(1)}(\vec{r}) + \delta\rho_{e|e}^{(1)}(\vec{r}|\vec{s}) - Z\delta\rho_{n|e}^{(1)}(\vec{r}|\vec{s}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r \right) \mathrm{d}^3 s \\
& - \frac{Z}{2}\int_{\mathbb{R}^3} \rho_n^{(1)}(\vec{r}) \left(\int_{\mathbb{R}^3} \left[\rho^{(2)}(\vec{s}) + \delta\rho_{e|n}^{(2)}(\vec{s}|\vec{r}) - Z\delta\rho_{n|n}^{(2)}(\vec{s}|\vec{r}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s \right) \mathrm{d}^3 r \\
& - \frac{Z}{2}\int_{\mathbb{R}^3} \rho_n^{(2)}(\vec{s}) \left(\int_{\mathbb{R}^3} \left[\rho^{(1)}(\vec{r}) + Z\delta\rho_{e|n}^{(1)}(\vec{r}|\vec{s}) - \delta\rho_{n|n}^{(1)}(\vec{r}|\vec{s}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r \right) \mathrm{d}^3 s
\end{aligned}
$$

Now I define $\bar{v}^{(1)}$ and $\bar{v}^{(2)}$ to be the mean-field potentials felt by electrons from the charges on CP1 and CP2, respectively, i.e.,

$$
\begin{aligned}
\bar{v}^{(1)}(\vec{s}) &\equiv \int_{\mathbb{R}^3} \left[\rho_e^{(1)}(\vec{r}) - Z\rho_n^{(1)}(\vec{r}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r\,, \\
\bar{v}^{(2)}(\vec{r}) &\equiv \int_{\mathbb{R}^3} \left[\rho_e^{(2)}(\vec{s}) - Z\rho_n^{(2)}(\vec{s}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s\,,
\end{aligned}
$$

and I define the following *correlation potentials*:

$$
\begin{aligned}
\delta\bar{v}_e^{(1)}(\vec{s}) &\equiv \int_{\mathbb{R}^3} \left[\delta\rho_{e|e}^{(1)}(\vec{r}|\vec{s}) - Z\delta\rho_{n|e}^{(1)}(\vec{r}|\vec{s}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r\,, \\
\delta\bar{v}_e^{(2)}(\vec{r}) &\equiv \int_{\mathbb{R}^3} \left[\delta\rho_{e|e}^{(2)}(\vec{s}|\vec{r}) - Z\delta\rho_{n|e}^{(2)}(\vec{s}|\vec{r}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s \\
\delta\bar{v}_n^{(1)}(\vec{s}) &\equiv \int_{\mathbb{R}^3} \left[\delta\rho_{e|n}^{(1)}(\vec{r}|\vec{s}) - Z\delta\rho_{n|n}^{(1)}(\vec{r}|\vec{s}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r\,, \\
\delta\bar{v}_n^{(2)}(\vec{r}) &\equiv \int_{\mathbb{R}^3} \left[\delta\rho_{e|n}^{(2)}(\vec{s}|\vec{r}) - Z\delta\rho_{n|n}^{(2)}(\vec{s}|\vec{r}) \right] \phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 s
\end{aligned}
$$

where, for example, $\left\langle \delta\bar{v}_e^{(2)}, \rho_e^{(1)} \right\rangle \equiv \int_{\mathbb{R}^3} \delta\bar{v}_e^{(2)}(\vec{r})\rho_e^{(1)}(\vec{r})\,\mathrm{d}^3 r$ is a self-energy which corrects the mean field electron-electron interaction energy $\left\langle \bar{v}_e^{(2)}, \rho_e^{(1)} \right\rangle$ by adding in the mutual responses of electrons on CP1 and CP2 to one another. No assumptions have been made about the mechanism of this synchronous motion, and relativistic retardation effects can be accounted for fully within $\delta\bar{v}_e^{(2)}(\vec{r})$. Using these definitions, the interaction energy can now be written as

$$
\begin{aligned}
E_{12} = {} & \frac{1}{2}\int_{\mathbb{R}^3} \rho^{(1)}(\vec{r})\bar{v}^{(2)}(\vec{r})\,\mathrm{d}^3 r + \frac{1}{2}\int_{\mathbb{R}^3} \rho^{(2)}(\vec{s})\bar{v}^{(1)}(\vec{s})\,\mathrm{d}^3 s \\
& + \frac{1}{2}\int_{\mathbb{R}^3} \rho_e^{(1)}(\vec{r})\delta\bar{v}_e^{(2)}(\vec{r})\,\mathrm{d}^3 r - \frac{Z}{2}\int_{\mathbb{R}^3} \rho_n^{(1)}(\vec{r})\delta\bar{v}_n^{(2)}(\vec{r})\,\mathrm{d}^3 r + \frac{1}{2}\int_{\mathbb{R}^3} \rho_e^{(2)}(\vec{s})\delta\bar{v}_e^{(1)}(\vec{s})\,\mathrm{d}^3 s - \frac{Z}{2}\int_{\mathbb{R}^3} \rho_n^{(2)}(\vec{s})\delta\bar{v}_n^{(1)}(\vec{s})\,\mathrm{d}^3 s \\
= {} & \overbrace{\int_{\mathbb{R}^3}\int_{\mathbb{R}^3} \rho^{(1)}(\vec{r})\rho^{(1)}(\vec{s})\phi(|\vec{r}-\vec{s}|)\,\mathrm{d}^3 r\,\mathrm{d}^3 s}^{E_{12}^{MF}[\rho^{(1)},\rho^{(2)}]} + \frac{1}{2}\left[\left\langle \rho_e^{(1)}, \delta\bar{v}_e^{(2)} \right\rangle - Z\left\langle \rho_n^{(1)}, \delta\bar{v}_n^{(2)} \right\rangle + \left\langle \rho_e^{(2)}, \delta\bar{v}_e^{(1)} \right\rangle - Z\left\langle \rho_n^{(2)}, \delta\bar{v}_n^{(1)} \right\rangle \right] \\
\therefore E_{12} = {} & E_{12}^{\mathrm{MF}}[\rho^{(1)},\rho^{(2)}] + \frac{1}{2}\left[\left\langle \rho^{(1)}, \delta\bar{v}_n^{(2)} \right\rangle + \left\langle \rho_e^{(1)}, \delta\bar{v}_e^{(2)} - \delta\bar{v}_n^{(2)} \right\rangle + \left\langle \rho^{(2)}, \delta\bar{v}_n^{(1)} \right\rangle + \left\langle \rho_e^{(2)}, \delta\bar{v}_e^{(1)} - \delta\bar{v}_n^{(1)} \right\rangle \right]
\end{aligned}
$$

Separation of time scales

We have found the following expression for the total energy of CP1+CP2, which is exact in the limit of zero overlap between CP1 and CP2 if CP1+CP2 is in a pure state Ψ.

$$\begin{aligned} E[\Psi] = &\sum_\alpha \lambda_\alpha\left(\mathcal{E}_1^\alpha + \mathcal{E}_2^\alpha\right) + E_{12}^{\rm MF}[\rho^{(1)},\rho^{(2)}] \\ &+ \frac{1}{2}\left[\left\langle \rho_e^{(1)}, \delta\bar{v}_e^{(2)}\right\rangle + \left\langle \rho_e^{(2)}, \delta\bar{v}_e^{(1)}\right\rangle\right] \\ &- \frac{Z}{2}\left[\left\langle \rho_n^{(1)}, \delta\bar{v}_n^{(2)}\right\rangle + \left\langle \rho_n^{(2)}, \delta\bar{v}_n^{(1)}\right\rangle\right] \end{aligned} \tag{D15}$$

Let us assume that an isolated C-particle is approximately spherical but thermally disordered. When two C-particles approach one another the interaction between them can break their near-spherical symmetry. If they are observed on a time scale that is short relative to the time scale on which they rotate about an axis passing through their centers, and that is short relative to the time scale on which the internal structure of a C-particle can rearrange, it is reasonable to assume that they are observed in a pure state. This is because there are no relevant symmetries on such a time scale.

Let us now consider the different types of correlation described by the $\delta\bar{v}$ terms on the right hand side of Eq. D15. The terms $-Z\left\langle \rho_n^{(1)}, \delta\bar{v}_n^{(2)}\right\rangle$ and $-Z\left\langle \rho^{(2)}, \delta\bar{v}_n^{(1)}\right\rangle$ account for the energy associated with synchronicity between the motion of nuclei on one C-particle and the motion of nuclei and electrons on the other. If we assume that nuclei move much more slowly that electrons and that electrons are free to move so that, on the time scale of nuclear motion, they perfectly screen any fields from nuclei on the other C-particle, then $\delta\bar{v}_n(1) = \delta\bar{v}_n(2) = 0$ and only the synchronous motion of electrons on different C-particles is relevant. Our assumption that electrons move freely also implies that $E_{12}^{\rm MF}[\rho^{(1)},\rho^{(2)}] = 0$, since both C-particles are globally charge-neutral and since on nuclear time scales electrons move rapidly to ensure local charge-neutrality. Therefore, it is expected that a very good approximation to the energy of CP1+CP2 is provided by

$$E \approx \sum_\alpha \lambda_\alpha\left(\mathcal{E}_1^\alpha + \mathcal{E}_2^\alpha\right) + \frac{1}{2}\left\langle \rho_e^{(1)}, \delta\bar{v}_e^{(2)}\right\rangle + \frac{1}{2}\left\langle \rho_e^{(2)}, \delta\bar{v}_e^{(1)}\right\rangle,$$

or

$$\begin{aligned} E[\{\lambda_\alpha,\mathcal{X}_\alpha,\mathcal{Y}_\alpha\}] \approx &\sum_\alpha \lambda_\alpha\left(E_1[\mathcal{X}_\alpha] + E_2[\mathcal{Y}_\alpha]\right) \\ &+ \frac{1}{2}\left\langle \rho_e^{(1)}[\{\lambda_\alpha,\mathcal{X}_\alpha\}], \delta\bar{v}_e^{(2)}[\{\lambda_\alpha,\mathcal{X}_\alpha,\mathcal{Y}_\alpha\}]\right\rangle \\ &+ \frac{1}{2}\left\langle \rho_e^{(2)}[\{\lambda_\alpha,\mathcal{Y}_\alpha\}], \delta\bar{v}_e^{(1)}[\{\lambda_\alpha,\mathcal{X}_\alpha,\mathcal{Y}_\alpha\}]\right\rangle \end{aligned} \tag{D16}$$

Appendix E: Excess field invariance proofs

Equations 39, 42, 43 and 44, which are expressions for the macroscale interfacial excesses of $\Delta\nu(x)$ and $x\,\Delta\nu(x)$, are the most important results of Sec. VII and among the most important results of the homogenization theory presented in this work.

On first examination these expressions appear to depend on x_b and on how the mesoscale neighbourhood of x_b is partitioned into microscopic intervals. Since any such dependence would make them ill-defined quantities, it is crucial to the importance and generality of these expressions that all choices of x_b and $\Pi(x_b,\ell)$, which satisfy the conditions stated in Sec. VII.F, give the same values for $\boldsymbol{\mathcal{S}}_0^{\langle\Delta\nu\rangle}$ and $\boldsymbol{\mathcal{S}}_1^{\langle\Delta\nu\rangle}$. This section is devoted to proving that this is indeed the case.

1. Derivatives of $\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}$ and $\bar{\mathcal{M}}_{\Delta\nu}^{\langle 2\rangle}$ with respect to x_b

The derivatives of $\bar{\mathcal{M}}_{\Delta\nu}^{\langle n\rangle}(x_b)$ with respect to x_b. will be used to demonstrate that the surface excesses calculated in Sec. VII.H.1 and VII.H.2 are independent of x_b. This will demonstrate that x_b is a parameter that determines the values of each term on the right hand sides of Eqs. 39 and 42, but not their sums - $\boldsymbol{\mathcal{S}}_0^{[\boldsymbol{\Delta\nu}]}(\mathbf{x}_b)$ and $\boldsymbol{\mathcal{S}}_1^{[\boldsymbol{\Delta\nu}]}(\mathbf{x}_b)$, respectively. I will assume that $\mathfrak{a}/\varepsilon_x$ is sufficiently small that the kernel average $\left\langle \mathcal{M}_{\Delta\nu}^{\langle n\rangle};\mu\right\rangle_{\varepsilon_x}^*$ can be replaced with a simple average, i.e.,

$$\bar{\mathcal{M}}_{\Delta\nu}^{\langle n\rangle}(x_b) = \frac{1}{\ell}\sum_m \mathcal{M}_{\Delta\nu}^{\langle n\rangle}(\bar{x}_m,\Delta_m)\Delta_m \tag{E1}$$

where $\mathcal{M}_{\Delta\nu}^{\langle n\rangle}(\bar{x}_m,\Delta_m) \equiv \Delta_m^{-1}\int_{x_m^-}^{x_m^+}(x-\bar{x}_m)^n\,\Delta\nu(x)\,\mathrm{d}x$. For convenience I have denoted the left-hand and right-hand boundaries of $\mathfrak{I}_m$ by $x_m^+ \equiv \bar{x}_m + \frac{1}{2}\Delta_m \in \Pi(x_b,\ell)$ and $x_m^- \equiv \bar{x}_m - \frac{1}{2}\Delta_m \in \Pi(x_b,\ell)$, respectively.

The simple average, $\langle\Delta\nu\rangle_\ell(x_b)$, fluctuates microscopically and continuously within the range $\mathfrak{I}(0,\varepsilon_{\mathcal{V}})$ as x_b changes. However, the definition of $\langle\,\cdot\,\rangle_\ell^*$ stipulates that $x_b \in \Pi(x_b,\ell)$ and that the average of $\Delta\nu$ is the same on every microinterval, which means that, in general, $x_b \pm \ell/2 \neq x_{\pm M}$ and that $0 \leq \left|\langle\Delta\nu\rangle_{\varepsilon_x}(x_b) - \bar{\mathcal{M}}_{\Delta\nu}^{\langle 0\rangle}\right| < \varepsilon_{\mathcal{V}}$. I will preserve the constraint $\bar{\mathcal{M}}^{\langle 0\rangle}_{\Delta\nu}=0$ as x_b changes, which means that every element of $\Pi(x_b,\ell)$, including x_M and x_{-M}, changes with x_b. Therefore, $\ell = x_M - x_{-M}$ also changes. I will assume that $\nu(x_b)\neq\bar{\nu}(x_b)$, leaving the special case $\nu(x_b)=\bar{\nu}(x_b)$ to the interested reader. Using primes to denote total derivatives with respect to x_b, I can write

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x_b}\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b) &= \frac{1}{\ell}\sum_m \frac{\mathrm{d}}{\mathrm{d}x_b}\left[\Delta_m \mathcal{M}^{\langle n\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\right]\\ &\quad -\frac{\ell'}{\ell}\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b) \end{aligned}\tag{E2}$$

where

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}x_b}\left[\Delta_m \mathcal{M}^{\langle n\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\right]\\ &\quad=\left(\frac{\Delta_m}{2}\right)^n\left[x_m^{+\prime}\Delta\nu(x_m^+)+(-1)^{n+1}\,x_m^{-\prime}\Delta\nu(x_m^-)\right]\\ &\qquad - n\Delta_m\bar{x}_m'\mathcal{M}^{\langle n-1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\end{aligned}\tag{E3}$$

I will derive expressions for the derivatives of $\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}$ and $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}$ below. Before doing so, I will simplify this task by deducing a relationship between $x_m^{-\prime}$ and $x_m^{+\prime}$ from the constraints $\langle\Delta\nu\rangle_{\Delta_m}(\bar{x}_m)=\overline{\Delta\nu}(x_b)=0\Rightarrow\overline{\Delta\nu}'(x_b)=0$.

$$\begin{aligned}\overline{\Delta\nu}\,'(x_b) &= \frac{\mathrm{d}}{\mathrm{d}x_b}\left(\frac{1}{\Delta_m}\int_{x_m^-}^{x_m^+}\Delta\nu(x)\,\mathrm{d}x\right)=0\\ \Rightarrow -\frac{\Delta_m'}{\Delta_m}\overline{\Delta\nu}(x_b) &+ \frac{1}{\Delta_m}\left[x_m^{+\prime}\Delta\nu(x_m^+)-x_m^{-\prime}\Delta\nu(x_m^-)\right]=0\end{aligned}$$

$\overline{\Delta\nu}=0$ means that $x_m^{+\prime}\Delta\nu(x_m^+)=x_m^{-\prime}\Delta\nu(x_m^-)$. Recursively applying this relationship and using the fact that $x_0\equiv x_b\Rightarrow x_0'=1$ we find that

$$x_m'\Delta\nu(x_m)=\Delta\nu(x_b),\ \ \forall\, x_m\in\Pi(x_b,\ell)\tag{E4}$$

Because a microscopic change of x_b cannot change ℓ by more than $\mathfrak{a}$, the second term on the right hand side of Eq. E2 is negligible for our purposes. Therefore, by substituting Eq. E4 into Eq. E5 we find that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x_b}\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b) &= \frac{\Delta\nu(x_b)}{\ell}\left[1-(-1)^n\right]\sum_m\left(\frac{\Delta_m}{2}\right)^n\\ &\quad -\frac{n}{\ell}\sum_m\Delta_m\bar{x}_m'\mathcal{M}^{\langle n-1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\end{aligned}\tag{E5}$$

a. Case I: $\mathrm{d}\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}\Big/\mathrm{d}x_b$

For $n=1$, Eqs. E3 and E4 mean that

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}x_b}\left[\Delta_m\mathcal{M}^{\langle 1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)\right]\\ &\quad=\frac{\Delta_m}{2}\left[x_m^{+\prime}\Delta\nu(x_m^+)+x_m^{-\prime}\Delta\nu(x_m^-)\right]=\Delta_m\Delta\nu(x_b)\end{aligned}$$

Subsistuting this and $\bar{\mathcal{M}}^{\langle 0\rangle}_{\Delta\nu}(x_b)=0$ into Eq. E5 gives

$$\frac{\mathrm{d}}{\mathrm{d}x_b}\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b)=\Delta\nu(x_b)\tag{E6}$$

This result can be derived at greater length without requiring that $\overline{\Delta\nu}\,'(x_b)=0$ or $\overline{\Delta\nu}(x_b)=0$.

b. Case II: $\mathrm{d}\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}\Big/\mathrm{d}x_b$

Inserting $n=2$ in Eq. E5 and using Eqs. E3 and E4 gives

$$\frac{\mathrm{d}}{\mathrm{d}x_b}\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b)=-\frac{2}{\ell}\sum_m\bar{x}_m'd_m\tag{E7}$$

where $d_m\equiv\Delta_m\mathcal{M}^{\langle 1\rangle}_{\Delta\nu}(\bar{x}_m,\Delta_m)$ is the first moment of $\Delta\nu$ in $\mathfrak{I}_m$. I define X_d to be the first-moment-weighted average of the interval midpoints, $\bar{x}_m$.

$$X_d\equiv\frac{\sum_m\bar{x}_m d_m}{\sum_m d_m}\tag{E8}$$

In the limit $a/l\to 0$ in an infinite macroscopically-uniform material, X_d coincides both with x_b and with $\frac{1}{2}\left(x_{-M}+x_M\right)$. Rearranging Eq. E8 and taking the derivative with respect to x_b gives

$$\sum_m\bar{x}_m'd_m=X_d'\sum_m d_m+X_d\sum_m d_m'-\sum_m\bar{x}_m d_m'\tag{E9}$$

Using Eq. E4 and $\int_{x_m^-}^{x_m^+}\Delta\nu(x)\,\mathrm{d}x=0$, the derivative of d_m can be expressed as

$$\begin{aligned}d'_m&=\left(\frac{\Delta_m}{2}\right)\left[x_m^{+\prime}\Delta\nu(x_m^+)+x_m^{-\prime}\Delta\nu(x_m^-)\right]\\ &=\Delta_m\Delta\nu(x_b)\Rightarrow\sum_m d'_m=\ell\,\Delta\nu(x_b)\end{aligned}\tag{E10}$$

The last term on the right hand side of Eq. E9 is

$$\begin{aligned}\sum_m\bar{x}_m d'_m&=\Delta\nu(x_b)\sum_m\Delta_m\bar{x}_m\approx\Delta\nu(x_b)\int_{x_{-M}}^{x_M}x\ \mathrm{d}x\\ &=\Delta\nu(x_b)\,\frac{\ell}{2}\,\left(x_{-M}+x_M\right)\end{aligned}\tag{E11}$$

Eqs. E7, E8, E9, E10, and E11 can be combined with $\sum_m d_m=\ell\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b)$ to show that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x_b}\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b)&=2\,\Delta\nu(x_b)\left[\left(\frac{x_{-M}+x_M}{2}\right)-X_d\right]\\ &\quad-2\,X_d'\,\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b)\end{aligned}\tag{E12}$$

Now, because $\frac{1}{2}(x_{-M}+x_M)$ and X_d both get closer x_b as l increases, their difference vanishes and X_d' becomes one in the limit $\mathfrak{a}/l\to 0$. Therefore, in this limit,

$$\frac{\mathrm{d}}{\mathrm{d}x_b}\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b)=-2\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b)\tag{E13}$$

To derive Eqs. E6 and E13 we assumed that a/l could be brought to zero without straying into regions having different mesoscale averages $\bar{\nu}$. When this assumption is not valid, some of the terms that were discarded should be considered more carefully.

c. $\mathrm{d}\boldsymbol{\mathcal{S}}_0^{[\boldsymbol{\Delta\mathcal{V}}]}\Big/\mathrm{d}\mathbf{x}$ and $\mathrm{d}\boldsymbol{\mathcal{S}}_1^{[\boldsymbol{\Delta\mathcal{V}}]}\Big/\mathrm{d}\mathbf{x}$

It is straightforward to use Eqs. 39, 42, E6, and E13 to show that

$$\frac{\mathrm{d}\boldsymbol{\mathcal{S}}_0^{[\boldsymbol{\Delta\mathcal{V}}]}}{\mathrm{d}x_b} = 0 \tag{E14a}$$

$$\frac{\mathrm{d}\boldsymbol{\mathcal{S}}_1^{[\boldsymbol{\Delta\mathcal{V}}]}}{\mathrm{d}x_b} = \bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b) \tag{E14b}$$

2. Macroscopic moment densities are independent of the choice of microscopic intervals

It is important to demonstrate that, in the limit $a/l \to 0$, our results do not depend on the choice of the set of microintervals, $\Pi(x_b, \ell)$, that partition the space around x_b. In this section it is demonstrated that all sets of points which satisfy the requirements explained in Sec. VII.F and and Sec. VII.G give the same values of $\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b)$ and $\bar{\mathcal{M}}_{\Delta\nu}^{\langle 2\rangle}(x_b)$ and therefore the same values of $\boldsymbol{\mathcal{S}}_0^{[\boldsymbol{\Delta\mathcal{V}}]}$ and $\boldsymbol{\mathcal{S}}_1^{[\boldsymbol{\Delta\mathcal{V}}]}$.

a. $\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b)$

Let us assume that we are in the limit $a/l \to 0$ and that for a particular choice, $\Pi_1(x_b, \ell)$, of the set of microinterval boundary points, we find

$$\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_1) = \frac{1}{\ell}\sum_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x \tag{E15}$$

where $\langle \Delta\nu \rangle_{\Delta_m}(\bar{x}_m) = 0$ has allowed each integrand $(x - \bar{x}_m)\Delta\nu(x)$ to be simplified to $x\Delta\nu(x)$.

Now suppose that a new set $\Pi_2(x_b, \ell_s)$ of boundary points $s_m \equiv x_m + \delta x_m$ is formed by changing every point $x_m \in \Pi_1(x_b, \ell)$, *except* $x_0 = x_b$, by an amount δx_m, such that the average of $\Delta\nu(x)$ on each of the new microscopic intervals remains equal to $\overline{\Delta\nu}(x_b) = 0$ and such that the ordering of the points does not change ($s_{m+1} > s_m, \quad \forall\ m$). The new set of microintervals partitions the interval $[s_{-M}, s_M]$, where $s_M - s_{-M} = \ell_s$, $|s_M - x_b - \ell_s/2| < \mathfrak{a}$, and $\mathfrak{a}/\ell_s \sim \mathfrak{a}/l \to 0$. I denote the midpoint, width, left-hand boundary, and right-hand boundary of the new m^{th} interval by $\bar{s}_m$, Δ_m^s, s_m^-, and s_m^+, respectively. By construction, the average of $\Delta\nu(x)$ on each microinterval is zero. Therefore,

$$\begin{aligned}\int_{s_m^-}^{s_m^+}\Delta\nu(s)\,\mathrm{d}s &= \int_{x_m^-}^{x_m^+}\Delta\nu(x)\,\mathrm{d}x = 0\\ \Rightarrow \int_{x_m^+}^{s_m^+}\Delta\nu(x)\,\mathrm{d}x &= \int_{x_m^-}^{s_m^-}\Delta\nu(x)\,\mathrm{d}x\end{aligned} \tag{E16}$$

The new average moment density is

$$\begin{aligned}\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_2) = \frac{1}{\ell_s}\sum_m \Bigg[&\int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x\\ &+ \int_{x_m^+}^{s_m^+} x\,\Delta\nu(x)\,\mathrm{d}x - \int_{x_m^-}^{s_m^-} x\,\Delta\nu(x)\,\mathrm{d}x\Bigg]\end{aligned} \tag{E17}$$

After cancelling terms in the sum, this can be written as

$$\begin{aligned}\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_2) = &\frac{1}{\ell_s}\sum_{m=-M}^{M-1}\int_{x_m}^{x_{m+1}} x\,\Delta\nu(x)\,\mathrm{d}x\\ &+\frac{1}{\ell_s}\int_{x_M}^{s_M} x\,\Delta\nu(x)\,\mathrm{d}x + \frac{1}{\ell_s}\int_{s_{-M}}^{x_{-M}} x\,\Delta\nu(x)\,\mathrm{d}x\end{aligned} \tag{E18}$$

Although the widths of all microintervals defined by $\Pi_1(x_b, \ell)$ and $\Pi_2(x_b, \ell_s)$ are less than $\mathfrak{a}$, I have not assumed that $|\bar{s}_m - \bar{x}_m| < \mathfrak{a}$. If, for example, the width of each new interval was larger than each old interval, i.e., $0 < \Delta_m < \Delta_m^s < \mathfrak{a}$, this would imply that $0 < s_M - x_M \sim Ma \not\ll l$. Therefore, we cannot immediately dismiss the second and third terms on the right hand side as negligible. However, $\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_1)$ was assumed to be converged with respect to the magnitude of ℓ. Therefore, without changing its value significantly, I can expand the set Π_1 to encompass the ranges $[x_M, s_M]$ and $[s_{-M}, x_{-M}]$ by dividing these ranges into microintervals and adding their boundary points to $\Pi_1(x_b, \ell)$ to form a new set $\Pi_1^{\text{new}}(x_b, \ell^{\text{new}}) \supset \Pi_1(x_b, \ell)$ containing $M^{\text{new}} > M$ microinterval boundary points on each side of x_b, and such that $0 < x_{-M} - x_{-M^{\text{new}}} < \mathfrak{a}$, $0 < x_{M^{\text{new}}} - x_M < \mathfrak{a}$, and $0 < \ell^{\text{new}} - \ell_s \lesssim \mathfrak{a}$. Eq. E18 then becomes

$$\begin{aligned}\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_2) = &\frac{1}{\ell_s}\sum_{m=-M^{\text{new}}}^{M^{\text{new}}-1}\int_{x_m}^{x_{m+1}} x\,\Delta\nu(x)\,\mathrm{d}x\\ &-\frac{1}{\ell_s}\int_{s_M}^{x_{M^{\text{new}}}} x\,\Delta\nu(x)\,\mathrm{d}x - \frac{1}{\ell_s}\int_{x_{-M^{\text{new}}}}^{s_{-M}} x\,\Delta\nu(x)\,\mathrm{d}x\\ = &\frac{\ell^{\text{new}}}{\ell_s}\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_1^{\text{new}}) + \mathcal{O}\left(\frac{\mathfrak{a}}{l}\right)\end{aligned} \tag{E19}$$

Therefore, if $\mathcal{O}(\mathfrak{a}/l)$ terms are neglected,

$$\bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_2) = \bar{\mathcal{M}}_{\Delta\nu}^{\langle 1\rangle}(x_b, \Pi_1) \tag{E20}$$

I have assumed that $s_M > x_M$ and $s_{-M} < x_{-M}$, but a similar procedure can be followed to prove the same result for any other case, such as $s_M > x_M$ and $s_{-M} > x_{-M}$.

b. $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b)$

I will now demonstrate that $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b)$ is independent of how the region around x_b is partitioned. I will use the partitions $\Pi_1(x_b,\ell)$, $\Pi_2(x_b,\ell_s)$ and $\Pi_1^{\text{new}}(x_b,\ell_s)$, introduced in the previous section, where $\ell_s > \ell$, and I again assume the limit $\mathfrak{a}/l \to 0$, which implies that $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1) = \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1^{\text{new}})$. I will show that $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1) = \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1^{\text{new}}) = \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1)$.

$$\begin{aligned}
\ell_s\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_2) &= \sum_m \int_{s_m^-}^{s_m^+} (x-\bar{s}_m)^2\Delta\nu(x)\,\mathrm{d}x = \sum_m \left[\int_{s_m^-}^{s_m^+} (x-\bar{x}_m)^2\Delta\nu(x)\,\mathrm{d}x - 2\ (\bar{s}_m - \bar{x}_m)\ \int_{s_m^-}^{s_m^+} x\,\Delta\nu(x)\,\mathrm{d}x\right] \\
&= \sum_m \left[\int_{x_m^-}^{x_m^+} (x-\bar{x}_m)^2\Delta\nu(x)\,\mathrm{d}x + \int_{x_m^+}^{s_m^+} x^2\Delta\nu(x)\,\mathrm{d}x - \int_{x_m^-}^{s_m^-} x^2\Delta\nu(x)\,\mathrm{d}x\ - 2\,\bar{s}_m \int_{s_m^-}^{s_m^+} x\,\Delta\nu(x)\,\mathrm{d}x + 2\,\bar{x}_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x\right] \\
&= \ell\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1) + \sum_{m=M+1}^{M^{\text{new}}} \int_{x_m^-}^{x_m^+} (x-\bar{x}_m)^2\Delta\nu(x)\,\mathrm{d}x + \sum_{m=-M^{\text{new}}}^{-M-1} \int_{x_m^-}^{x_m^+} (x-\bar{x}_m)^2\Delta\nu(x)\,\mathrm{d}x + 2\sum_{m=M+1}^{M^{\text{new}}} \bar{x}_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x \\
&+ 2\sum_{m=-M^{\text{new}}}^{-M-1} \bar{x}_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x - 2\sum_{m=-M}^{M} \bar{s}_m \int_{s_m^-}^{s_m^+} x\,\Delta\nu(x)\,\mathrm{d}x + 2\sum_{m=-M}^{M} \bar{x}_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x \\
&= \ell^{\text{new}}\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1^{\text{new}}) + 2\sum_{m=-M^{\text{new}}}^{M^{\text{new}}} \bar{x}_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x - 2\sum_{m=-M}^{M} \bar{s}_m \int_{s_m^-}^{s_m^+} x\,\Delta\nu(x)\,\mathrm{d}x
\end{aligned} \tag{E21}$$

Denoting the dipole weighted mean positions (Eq. E8) of sets $\Pi_1^{\text{new}}(x_b,\ell_s)$ and $\Pi_2(x_b,\ell_s)$ by $X_d(\Pi_1^{\text{new}})$ and $X_d(\Pi_2)$, respectively, using Eq. E20, and neglecting $\mathcal{O}(\mathfrak{a}/l)$ terms, allows this to be written as

$$\begin{aligned}
&\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_2) = \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1^{\text{new}}) \\
&+ 2\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b,\Pi_1^{\text{new}})X_d(\Pi_1^{\text{new}}) - 2\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b,\Pi_2)X_d(\Pi_2) \\
&= \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1) + 2\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b,\Pi_1)\left[X_d(\Pi_1) - X_d(\Pi_2)\right]
\end{aligned}$$

In the limit $a/l \to 0$, both $X_d(\Pi_1)$ and $X_d(\Pi_2)$ tend to x_b and so

$$\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_2) = \bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b,\Pi_1) \tag{E22}$$

3. Mesoscale averages of $\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}(x_b)$

In general, both $\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b)$ and $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b)$ vary microscopically with x_b. It can be necessary to know their mesoscale averages over x_b, which are denoted by $\boldsymbol{\mathcal{M}}^{\langle 1\rangle}_{\boldsymbol{\Delta\nu}}$ and $\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}$, respectively. In Sec. VII.H.3 it was argued that idempotency of the mesoscale-averaging operation applied to surface integrals requires both $\boldsymbol{\mathcal{M}}^{\langle 1\rangle}_{\boldsymbol{\Delta\nu}}$ and $\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}$ to be zero in regions of mesoscale uniformity. In this section I prove that this requirement is satisfied in the $a/l \to 0$ limit.

I define $\ell_1 \sim l$ and $\ell_2 \sim l$ as the widths of the intervals on which the mesoscale averages of $\mathcal{M}^{\langle n\rangle}_{\Delta\nu}$ and $\bar{\mathcal{M}}^{\langle n\rangle}_{\Delta\nu}$, respectively, are calculated. I consider the average, over all $u \in \mathfrak{I}(-\ell_2/2,\ell_2/2)$, of $\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b+u)$. As usual, when $\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}$ is being evaluated at x_b+u, I partition an interval of width ℓ_2 centered at a microscopic distance from x_b+u into a set of microintervals. I denote the left-hand boundary, right-hand boundary, midpoint, and width of the m^{th} microinterval, $\mathfrak{I}_m(u)$, by $x_m^- = x_m^-(u)$, $x_m^+ = x_m^+(u)$, $\bar{x}_m = \bar{x}_m(u)$, and $\Delta_m = \Delta_m(u)$, respectively. For each value of u, the value of $\ell_1(u) = \ell_1(0) + \Delta\ell_1(u)$ is chosen such that $\overline{\Delta\nu}(x_b+u) = 0$. It is always possible to choose it such that $\Delta\ell_1(u) < \mathfrak{a}$. The values of $\bar{x}_m$ and Δ_m are chosen such that $\langle\Delta\nu\rangle_{\Delta_m(u)}(\bar{x}_m(u)) = \overline{\Delta\nu}(x_b+u) = 0$. To avoid clutter I will only make the dependences of x_m, x_m^-, x_m^+, Δ_m, $\bar{x}_m$, and ℓ_1 on u explicit in my notation when it is necessary for clarity.

The mesoscale average of $\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b)$ is

$$\begin{aligned}
&\left\langle\bar{\mathcal{M}}^{(1)}_{\nu}\right\rangle_{\ell_2}(x_b) = \frac{1}{\ell_2}\int_{-\ell_2/2}^{\ell_2/2}\left(\frac{1}{\ell_1}\sum_m \int_{x_m^-}^{x_m^+} x\,\Delta\nu(x)\,\mathrm{d}x\right)\mathrm{d}u \\
&= \frac{1}{\ell_2}\int_{-\ell_2/2}^{\ell_2/2}\left(\frac{1}{\ell_1}\int_{-\ell_1/2}^{\ell_1/2} v\,\Delta\nu(x_b+u+v)\,\mathrm{d}v\right)\mathrm{d}u
\end{aligned}$$

Because terms of order $\mathfrak{a}/\ell_1$ are negligible, we can ignore the dependence of ℓ_1 on u. This allows us to switch the order of the integrations over u and v.

$$\begin{aligned}&\left\langle \bar{\mathcal{M}}^{(1)}_{\nu}\right\rangle_{\ell_2}(x_b)\\ &= \frac{1}{\ell_2\,\ell_1}\int_{-\ell_1/2}^{\ell_1/2} u\left(\int_{\ell_2/2}^{-\ell_2/2}\Delta\nu(x_b+u+v)\,\mathrm{d}v\right)\mathrm{d}u\end{aligned}$$

The inner integral vanishes in the limit $\mathfrak{a}/\ell_2 \to 0$. Therefore,

$$\boldsymbol{\mathcal{M}}^{\langle 1\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b) \equiv \left\langle \bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}\right\rangle_{l}(x_b) = 0 \qquad \text{(E23)}$$

Because ν_B is constant on the microscale, it follows that $\boldsymbol{\mathcal{M}}^{\langle 1\rangle}_{\boldsymbol{\nu}}$ is also zero. This is simply a consequence of mesoscale uniformity implying local mesoscale isotropy. By *local mesoscale isotropy* I mean that, for all n,

$$\int_{-\ell/2}^{\ell/2} u^n\,\nu(x+u)\,\mathrm{d}u = \int_{-\ell/2}^{\ell/2} u^n\,\nu(x-u)\,\mathrm{d}u + \mathcal{O}(a/l)$$

I now want to prove that the mesoscale average $\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}$ of $\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}$ is zero, where

$$\begin{aligned}\bar{\mathcal{M}}^{\langle 2\rangle}_{\Delta\nu}(x_b+u) &= \frac{1}{\ell_1}\sum_m\int_{x_m^-}^{x_m^+}(x-\bar{x}_m)^2\,\Delta\nu(x)\,\mathrm{d}x\\ &= \frac{1}{\ell_1}\sum_m\left[\int_{x_m^-}^{x_m^+}(x-x_b-u)^2\,\Delta\nu(x)\,\mathrm{d}x\right.\\ &\quad\left.-\,2\left(\bar{x}_m-x_b-u\right)\int_{x_m^-}^{x_m^+}x\,\Delta\nu(x)\,\mathrm{d}x\right]\\ &= \frac{1}{\ell_1}\int_{x_{-M}}^{x_M}(x-x_b-u)^2\,\Delta\nu(x)\,\mathrm{d}x\\ &\quad-2\left[X_d-(x_b+u)\right]\bar{\mathcal{M}}^{\langle 1\rangle}_{\Delta\nu}(x_b+u) \qquad \text{(E24)}\end{aligned}$$

As defined in Sec. E.1.b, $X_d = X_d(u)$ is the first-moment-weighted average of the microinterval midpoints. Its value tends to x_b+u in the $a/\ell_1 \to 0$ limit; therefore, the second term on the right hand side vanishes in this limit. Since $\mathcal{O}(a/l)$ terms can be disregarded, we can assume that $x_{\pm M} = x_b + u \pm \ell_1/2$ and express the average over u as

$$\begin{aligned}&\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b)\\ &= \frac{1}{\ell_2}\int_{-\ell_2/2}^{\ell_2/2}\frac{1}{\ell_1}\left(\int_{-\ell_1/2}^{\ell_1/2}v^2\,\Delta\nu(x_b+u+v)\,\mathrm{d}v\right)\mathrm{d}u\\ &= \frac{1}{\ell_1}\int_{-\ell_1/2}^{\ell_1/2}v^2\left\langle\Delta\nu\right\rangle_{\ell_2}(x_b+v)\,\mathrm{d}v \qquad \text{(E25)}\end{aligned}$$

where, to reach the second line from the first, we reverse the order of integration, thereby neglecting the $\mathcal{O}(\mathfrak{a}/\ell_1)$ error made by ignoring the $u-$dependence of ℓ_1. The value of $\langle\Delta\nu\rangle_{\ell_2}(x_b+v)$ fluctuates microscopically about zero as v is varied, and its magnitude remains smaller than $\delta_A^{[\bar{\nu}]}/2$. Therefore, it is possible to choose a value of $\eta_1 < \mathfrak{a}$ for which $\int_{-\ell_1/2}^{\ell_1/2+\eta_1}\langle\Delta\nu\rangle(x_b+v)\,\mathrm{d}v = 0$. Therefore, we can add η_1 to the upper limit of the integral at the expense of a negligible $\mathcal{O}(a/l)$ error. Integrating by parts gives

$$\begin{aligned}&\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b)\\ &\qquad= -\frac{2}{\ell_1}\int_{-\ell_1/2}^{\ell_1/2+\eta_1}v\left(\int_{-\ell_1/2}^{v}\langle\Delta\nu\rangle_{\ell_2}(x_b+v')\,\mathrm{d}v'\right)\mathrm{d}v\end{aligned}$$

Let $\eta_2 > 0$ be the shortest distance for which $\int_{-\ell_1/2}^{v-\eta_2}\langle\Delta\nu\rangle_{\ell_2}(x_b+v')\,\mathrm{d}v' = 0$. Because $\langle\Delta\nu\rangle_{\ell_2}(x_b+v')$ fluctuates microscopically about zero, $\eta_2 < \mathfrak{a}$ and $\left|\int_{v-\eta_2}^{v}\langle\Delta\nu\rangle_{\ell_2}(x_b+v')\,\mathrm{d}v'\right| < \eta_2\,\delta_A^{[\bar{\nu}]}/2$. Therefore,

$$\left|\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b)\right| < \frac{\mathfrak{a}}{\ell_1}\,\delta_A^{[\bar{\nu}]}\left|\int_{-\ell_1/2}^{\ell_1/2+\eta_1}v\,\mathrm{d}v\right| \sim \frac{1}{2}\frac{\mathfrak{a}^3}{\ell_2}\delta_B^{[\nu]} \qquad \text{(E26)}$$

This vanishes in the $\mathfrak{a}/\ell_2 \to 0$ limit and so

$$\boldsymbol{\mathcal{M}}^{\langle 2\rangle}_{\boldsymbol{\Delta\nu}}(\mathbf{x}_b) = 0 + \mathcal{O}(a/l) \qquad \text{(E27)}$$

Appendix F: Proof that the mesoscale average of $\Phi_L^{[\Delta\rho]}$ vanishes

The charge density,

$$\Delta\rho(\vec{u};x,\vec{s}_b,R) \equiv \rho(x,\vec{s}_b+\vec{u}) - \bar{\rho}(x,\vec{s}_b;R),$$

is introduced in Sec. XIII.A, where $\bar{\rho}(x,\vec{s}_b;R)$ is the average of the volumetric charge density ρ on a disc of radius R, which is parallel to the surface and centered at $(x,\vec{s})$. $\Phi_L^{[\Delta\rho]}(x_b,\vec{s}_b;R)$, which is defined in Eq. 69, can be viewed as the microscopic electric potential at $(x_b,\vec{s}_b)$ from all of the *areal* charge densities $\Delta\rho(\vec{u};x,\vec{s}_b,R)\,\mathrm{d}x$ on all such discs with centers at $(x,\vec{s}_b)$ for x between x_L and x_b. I will now demonstrate that the average of $\Phi_L^{[\Delta\rho]}(x_b,\vec{s}_b;R)$ over all points on the disc of radius $\mathfrak{R}_1$ at $x = x_b$ vanishes as $\mathfrak{R}_1$ and the radius, $R \gg \mathfrak{R}_1$, of the discs on which $\Delta\rho$ is defined, get very large.

I will assume that R is sufficiently large that the mesoscale uniformity of microstructure allows us to neglect the dependence of the disc-averaged charge density $\bar{\rho}(x, \vec{s}; R)$ on the position $\vec{s}$ of the disc's center in the plane perpendicular to $\hat{x}$. The disc parallel to the surface of radius $\mathfrak{R}_1 \ll R$ and with its center at $(x_b, \vec{s}_b)$ is denoted by $\mathcal{D}_{\mathfrak{R}_1}(x_b, \vec{s}_b)$ and the disc of radius R centered at $(x, \vec{s}_b)$ is denoted by $\mathcal{D}_R(x, \vec{s}_b)$. I denote the potential at point $(x_b, \vec{s}_b + \vec{u}) \in \mathcal{D}_{\mathfrak{R}_1}(x_b, \vec{s}_b)$, from the areal charge density $\Delta\rho(\vec{v}; x, \vec{s}_b, R)\,\mathrm{d}x$ on $\mathcal{D}_R(x, \vec{s}_b)$ by $\mathrm{d}\Phi_L^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)$, i.e.,

$$\begin{aligned}
\mathrm{d}\Phi_L^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R) &= \kappa\,\mathrm{d}x \iint_{|\vec{v}|<R} \frac{\Delta\rho(\vec{v}; x, \vec{s}_b, R)}{\sqrt{(x_b - x)^2 + |\vec{v} - \vec{u}|^2}}\,\mathrm{d}^2v = \kappa\,\mathrm{d}x \iint_{|\vec{u}+\vec{v}|<R} \frac{\Delta\rho(\vec{v} + \vec{u}; x, \vec{s}_b, R)}{\sqrt{(x_b - x)^2 + v^2}}\,\mathrm{d}^2v \\
&= \underbrace{\kappa\,\mathrm{d}x \iint_{|\vec{v}|<\mathfrak{R}_2} \frac{\Delta\rho(\vec{u} + \vec{v}; x, \vec{s}_b, R)}{\sqrt{(x_b - x)^2 + v^2}}\,\mathrm{d}^2v}_{\mathrm{d}\Phi_{L,1}^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)} + \underbrace{\kappa\,\mathrm{d}x \iint_{\substack{|\vec{v}|>\mathfrak{R}_2 \\ |\vec{v}+\vec{u}|<R}} \frac{\Delta\rho(\vec{u} + \vec{v}; x, \vec{s}_b, R)}{\sqrt{(x_b - x)^2 + v^2}}\,\mathrm{d}^2v}_{\mathrm{d}\Phi_{L,2}^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)}, \qquad \text{(F1)}
\end{aligned}$$

where $\kappa = (4\pi\epsilon_0)^{-1}$, $v \equiv |\vec{v}|$, and I have split the integral over $\mathcal{D}_R(x, \vec{s}_b)$ into an integral over the disc $\mathcal{D}_{\mathfrak{R}_2}(x, \vec{s}_b + \vec{u})$, of radius $\mathfrak{R}_2 \ll R$ and center $(x, \vec{s}_b + \vec{u})$, and an integral over $\mathcal{D}_R(x, \vec{s}_b) \setminus \mathcal{D}_{\mathfrak{R}_2}(x, \vec{s}_b + \vec{u})$. We will be averaging $\mathrm{d}\Phi_L^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)$ over all $(x_b, \vec{u}) \in \mathcal{D}_{\mathfrak{R}_1}(x_b, \vec{s}_b)$; when we do so, the average of the first term, $\mathrm{d}\Phi_{L,1}^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)$, vanishes because the mesoscale uniformity of microstructure ensures that the average of $\Delta\rho$ on any disc of radius $\mathfrak{R}_1$ vanishes in the large-$\mathfrak{R}_1$ limit. Therefore let us focus our attention on $\mathrm{d}\Phi_{L,2}^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)$.

The average of $\mathrm{d}\Phi_L^{[\Delta\rho]}(x_b, \vec{u}; x, \vec{s}_b, R)$ over all $\vec{u}$ for which $|\vec{u}| < \mathfrak{R}_1$, is the average potential in $\mathcal{D}_{\mathfrak{R}_1}(x_b, \vec{s}_b + \vec{u})$ from $\Delta\rho$ at points in $\mathcal{D}_R(x, \vec{s}_b)$. Its value is

$$\begin{aligned}
\left\langle \mathrm{d}\Phi_L^{[\Delta\rho]} \right\rangle_{\mathcal{D}_{\mathfrak{R}_1}(x_b, \vec{s}_b)}(x, \vec{s}_b, R) &= \frac{\kappa\,\mathrm{d}x}{\pi\mathfrak{R}_1^2} \iint_{|\vec{u}|<\mathfrak{R}_1} \mathrm{d}^2u \iint_{\substack{|\vec{v}|>\mathfrak{R}_2 \\ |\vec{v}+\vec{u}|<R}} \mathrm{d}^2v\, \frac{\Delta\rho(\vec{u} + \vec{v}; x, \vec{s}_b, R)}{\sqrt{(x_b - x)^2 + v^2}} && \text{(F2)} \\
&= \frac{\kappa\,\mathrm{d}x}{\pi\mathfrak{R}_1^2} \iint_{|\vec{v}|>\mathfrak{R}_2} \frac{\mathrm{d}^2v}{\sqrt{(x_b - x)^2 + v^2}} \iint_{\substack{|\vec{u}|<\mathfrak{R}_1 \\ |\vec{u}+\vec{v}|<R}} \Delta\rho(\vec{u} + \vec{v}; x, \vec{s}_b, R)\,\mathrm{d}^2u && \text{(F3)} \\
&= \frac{\kappa\,\mathrm{d}x}{\pi\mathfrak{R}_1^2} \int_{v>\mathfrak{R}_2} \frac{\mathrm{d}v}{\sqrt{(x_b - x)^2 + v^2}} \int_0^{2\pi} \mathrm{d}\theta \iint_{\substack{|\vec{u}|<\mathfrak{R}_1 \\ |\vec{u}+\vec{v}|<R}} \Delta\rho(\vec{u} + \vec{v}(v, \theta); x, \vec{s}_b, R)\,\mathrm{d}^2u\,, && \text{(F4)}
\end{aligned}$$

where $\vec{v}(v, \theta)$ is $\vec{v}$ expressed in polar coordinates. The integral over $\vec{u}$ is an integral of $\Delta\rho$ on $\mathcal{D}_R(x, \vec{s}_b) \cap \mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v})$. When $\mathcal{D}_R(x, \vec{s}_b) \cap \mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v}) = \mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v})$ it vanishes in the large-$\mathfrak{R}_1$ limit. When $\mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v})$ is not a subset of $\mathcal{D}_R(x, \vec{s}_b)$, its integral on $\mathcal{D}_R(x, \vec{s}_b)$ is zero by definition and its average on $\mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v})$ vanishes in the large-$\mathfrak{R}_1$ limit. Therefore, in this limit, the average of $\Delta\rho$ on $\mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v}) \setminus \mathcal{D}_R(x, \vec{s}_b)$ is the negative of its average on $\mathcal{D}_R(x, \vec{s}_b) \setminus \mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v})$. The integral over θ is then a spatial average of $\Delta\rho$ on the annulus swept out by $\mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v}(v, \theta)) \setminus \mathcal{D}_R(x, \vec{s}_b)$ as θ varies between zero and 2π. It is also equal to the negative of the average of $\Delta\rho$ on the disc of radius $v - \mathfrak{R}_1$ centered at $(x, \vec{s}_b)$ consisting of all points in $\mathcal{D}_R(x, \vec{s}_b)$ that are not in $\mathcal{D}_{\mathfrak{R}_1}(x, \vec{s}_b + \vec{v}(v, \theta))$ for any value of θ. Therefore, when $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are large enough it vanishes.

REFERENCES

Aizu, K (1962), “Polarization, pyroelectricity, and ferroelectricity of ionic crystals,” Rev. Mod. Phys. **34**, 550–576.

Ashcroft, N W, and N. D. Mermin (1976), *Solid State Physics*, HRW international editions (Holt, Rinehart and Winston).

Bacskay, G B, J. R. Reimers, and S. Nordholm (1997), “The mechanism of covalent bonding,” J. Chem. Educ. **74** (12), 1494.

Bao, J L, P. Verma, and D. G. Truhlar (2018), “How well can density functional theory and pair-density functional theory predict the correct atomic charges for dissociation and accurate dissociation energetics of ionic bonds?” Phys. Chem. Chem. Phys. **20**, 23072–23078.

Baroni, S, S. de Gironcoli, A. Dal Corso, and P. Giannozzi (2001), “Phonons and related crystal properties from density-functional perturbation theory,” Rev. Mod. Phys. **73**, 515–562.

Bethe, H (1928), “Theorie der beugung von elektronen an kristallen,” Ann. Phys. (Berl.) **392**, 55–129.

Bloch, F (1929), “Über die quantenmechanik der elektronen in kristallgittern,” Z. Phys. **52**, 555–600.

Blount, E I (1962), “Formalisms of band theory,” Solid State Physics **13**, 305–373.

Blumenthal, L, J. M. Kahk, R. Sundararaman, P. Tangney, and J. Lischner (2017), “Energy level alignment at semiconductor–water interfaces from atomistic and continuum solvation models,” RSC Adv. **7**, 43660–43670.

Born, M, and K. Huang (1954), *Dynamical Theory of Crystal Lattices*, International Series of Monographs on Physics (Oxford University Press).

Born, M, and T. von Kármán (1912), “Über schwingungen im raumgittern,” Phys. Z. **13**, 297–309.

Boys, S F (1960), “Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another,” Rev. Mod. Phys. **32**, 296–299.

Bristowe, N C, P. Ghosez, P. B. Littlewood, and E. Artacho (2014), “The origin of two-dimensional electron gases at oxide interfaces: Insights from theory,” J. Phys.: Condens. Matter **26** (14), 143201.

Bristowe, N C, P. B. Littlewood, and E. Artacho (2011), “The net charge at interfaces between insulators,” J. Phys.: Condens. Matter **23** (8), 081001.

Buchwald, J, and R. Fox, Eds. (2013), *The Oxford Handbook of The History of Physics* (Oxford University Press, Oxford).

Cendagorta, J R, and T. Ichiye (2015), “The surface potential of the water–vapor interface from classical simulations,” J. Phys. Chem. B **119** (29), 9114–9122.

Cioslowski, J, and K. Strasburger (2021), “From Fredholm to Schrödinger via Eikonal: A new formalism for revealing unknown properties of natural orbitals,” J. Chem. Theory Comput. **17** (11), 6918–6933.

Cochran, W (1960), “Crystal stability and the theory of ferroelectricity,” Adv. Phys. **9** (36), 387–423.

Cochran, W, and R. A. Cowley (1962), “Dielectric constants and lattice vibrations,” J. Phys. Chem. Solids **23** (5), 447–450.

Cohen, R E (1992), “Origin of ferroelectricity in perovskite oxides,” Nature (6382), 136–138.

Cohen, R E (1993), “Ferroelectricity origins,” Nature **362**, 213.

Coiana, G, J. Lischner, and P. Tangney (2024), “Breakdown of phonon band theory in MgO,” Phys. Rev. B **109**, 014310.

Coleman, A J (1963), “Structure of fermion density matrices,” Rev. Mod. Phys. **35**, 668–686.

de Groot, S R, and J. Vlieger (1965), “Derivation of Maxwell’s equations: The statistical theory of the macroscopic equations,” Physica **31** (3), 254–268.

Edmiston, C, and K. Ruedenberg (1963), “Localized atomic and molecular orbitals,” Rev. Mod. Phys. **35**, 457–464.

Ferreira, L G, and N. J. Parada (1970), “Wannier functions and the phases of the Bloch functions,” Phys. Rev. B **2**, 1614–1618.

Finnis, M W (1998), “Accessing the excess: An atomistic approach to excesses at planar defects and dislocations in ordered compounds,” Phys. Status Solidi A **166** (1), 397–416.

Foster, J M, and S. F. Boys (1960a), “Canonical configurational interaction procedure,” Rev. Mod. Phys. **32**, 300–302.

Foster, J M, and S. F. Boys (1960b), “A quantum variational calculation for HCHO,” Rev. Mod. Phys. **32**, 303–304.

Frias, W, and A. I. Smolyakov (2012), “Electromagnetic forces and internal stresses in dielectric media,” Phys. Rev. E **85**, 046606.

Gajdardziska-Josifovska, M, and A. H. Carim (1999), “Applications of electron holography,” in *Introduction to Electron Holography* (Springer US, Boston, MA) pp. 267–293.

Gajdardziska-Josifovska, M, M. R. McCartney, W. J. de Ruijter, D. J. Smith, J. K. Weiss, and J. M. Zuo (1993), “Accurate measurements of mean inner potential of crystal wedges using digital electron holograms,” Ultramicroscopy **50** (3), 285–299.

Giannozzi, P, S. de Gironcoli, P. Pavone, and S. Baroni (1991), “Ab initio calculation of phonon dispersions in semiconductors,” Phys. Rev. B **43**, 7231–7242.

Giesbertz, K J H, and R. van Leeuwen (2013), “Natural occupation numbers: When do they vanish?” J. Chem. Phys. **139** (10), 104109.

Gillespie, R J, and E. A. Robinson (2007), “Gilbert N. Lewis and the chemical bond: The electron pair and the octet rule from 1916 to the present day,” J. Comput. Chem. **28** (1), 87–97.

Goniakowski, J, F. Finocchi, and C. Noguera (2008), “Polarity of oxide surfaces and nanostructures,” Rep. Prog. Phys. **71** (1), 016501.

Goniakowski, J, and C. Noguera (2011), “Polarity at the nanoscale,” Phys. Rev. B **83**, 115413.

Goniakowski, J, and C. Noguera (2014), “Conditions for electronic reconstruction at stoichiometric polar/polar interfaces,” J. Phys.: Condens. Matter **26** (48), 485010.

Goniakowski, J, and C. Noguera (2016), “Insulating oxide surfaces and nanostructures,” C. R. Phys. **17** (3–4), 471–480.

Gonze, X, and C. Lee (1997), “Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory,” Phys. Rev. B **55**, 10355–10368.

Griffiths, D J (1999), *Introduction to Electrodynamics* (Prentice Hall).

de Groot, S R, and J. Vlieger (1964), “On the derivation of Maxwell’s equations,” Il Nuovo Cimento **33**, 1225–1227.

Grundmann, M (2016), *The Physics of Semiconductors - An Introduction Including Nanophysics and Applications*, 3rd ed. (Springer-Verlag Berlin Heidelberg).

Heaviside, O (1893), *Electromagnetic theory*, Vol. I (The Electrician Publishing).

Hörmann, N G, Z. Guo, F. Ambrosio, O. Andreussi, A. Pasquarello, and N. Marzari (2019), "Absolute band alignment at semiconductor-water interfaces using explicit and implicit descriptions for liquid water," NPJ Comput. Mater. **5**, 100.

Ibers, J A (1958), "Atomic scattering amplitudes for electrons," Acta Crystallogr. **11** (3), 178–183.

Jackson, J D (1962), *Classical electrodynamics*, 1st ed. (Wiley, New York, (NY)).

Jackson, J D (1975), *Classical Electrodynamics*, 2nd ed. (Wiley).

Jackson, J D (1998), *Classical Electrodynamics*, 3rd ed. (Wiley).

Jennison, D R, and A. B. Kunz (1976), "Electronic charge distribution and the degree of ionicity in crystalline NaF, NaCl, SrF_2, and $SrCl_2$ as found by the local-orbitals Hartree-Fock method," Phys. Rev. B **13**, 5597–5602.

Jones, W, and N. H. March (1973), *Theoretical solid state physics*, Vol. I (Dover Publications Inc., New York).

Kamenetskii, E O (1998), "Sampling theorem in macroscopic electrodynamics of crystal lattices," Phys. Rev. E **57**, 3556–3562.

Kathmann, S M (2021), "Electric fields and potentials in condensed phases," Phys. Chem. Chem. Phys. **23**, 23836–23849.

Kathmann, S M, I.-F. W. Kuo, C. J. Mundy, and G. K. Schenter (2011), "Understanding the surface potential of water," J. Phys. Chem. B **115** (15), 4369–4377.

King-Smith, R D, and D. Vanderbilt (1993), "Theory of polarization of crystalline solids," Phys. Rev. B **47** (3), 1651–1654.

Kirkwood, J G (1936), "On the Theory of Dielectric Polarization," J. Chem. Phys. **4** (9), 592–601.

Kirkwood, J G (1940), "The local field in dielectrics," Ann. N. Y. Acad. Sci. **40** (5), 315–320.

Kittel, C (2004), *Introduction to Solid State Physics*, 8th ed. (Wiley).

Kohn, W (1973), "Construction of Wannier functions and applications to energy bands," Phys. Rev. B **7**, 4388–4398.

Kohn, W, and L. J. Sham (1965), "Self-consistent equations including exchange and correlation effects," Phys. Rev. **140**, A1133–A1138.

Kvashnin, A G, D. G. Kvashnin, and A. R. Oganov (2019), "Novel unexpected reconstructions of (100) and (111) surfaces of NaCl: Theoretical prediction," Sci. Rep. **9**, 14267.

Landauer, R (1960), "Pyroelectric effect in the cubic ZnS structure," J. Chem. Phys. **32** (6), 1784–1785.

Landauer, R (1981), "Pyroelectricity and piezoelectricity are not true volume effects," Solid State Commun. **40** (11), 971–974.

de Lange, O L, and R. E. Raab (2006), "Surprises in the multipole description of macroscopic electrodynamics," Am. J. Phys. **74** (4), 301–312.

de Lange, O L, R. E. Raab, and A. Welter (2012), "On the transition from microscopic to macroscopic electrodynamics," J. Math. Phys. **53** (1), 013513.

Larmor, J (1921), "On electro-crystalline properties as conditioned by atomic lattices," Proc. R. Soc. Lond. A **99**, 1–10.

van Leeuwen, R (2003), "Density functional approach to the many-body problem: Key concepts and exact functionals," (Academic Press) pp. 25–94.

Leung, K (2010), "Surface potential at the air-water interface computed using density functional theory," J. Phys. Chem. Lett. **1** (2), 496–499.

Lewis, G N (1916), "The atom and the molecule," J. Am. Chem. Soc. **38** (4), 762–785.

Li, B, A. Michaelides, and M. Scheffler (2007), "Density functional theory study of flat and stepped NaCl(001)," Phys. Rev. B **76**, 075401.

Littlewood, P B (1980), "On the calculation of the macroscopic polarisation induced by an optic phonon," J. Phys. C: Solid State Phys. **13** (26), 4893.

Littlewood, P B, and V. Heine (1979), "The infrared effective charge in IV-VI compounds. I. a simple one-dimensional model," J. Phys. C: Solid State Phys. **12** (21), 4431.

Lorentz, H A (1916), *The Theory of electrons and its applications to the phenomena of light and radiant heat*, 2nd ed. (Leipzig : B.G. Teubner ; New York : G.E. Stechert, New York).

Löwdin, P-O (1955), "Quantum theory of many-particle systems. I. physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction," Phys. Rev. **97**, 1474–1489.

Lyddane, R H, R. G. Sachs, and E. Teller (1941), "On the polar vibrations of alkali halides," Phys. Rev. **59**, 673–676.

Madsen, J, T. J. Pennycook, and T. Susi (2021), "*Ab initio* description of bonding for transmission electron microscopy," Ultramicroscopy , 113253.

Martin, R M (1974), "Comment on calculations of electric polarization in crystals," Phys. Rev. B **9**, 1998–1999.

Marzari, N, A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt (2012), "Maximally localized Wannier functions: Theory and applications," Rev. Mod. Phys. **84**, 1419–1475.

Marzari, N, and D. Vanderbilt (1997), "Maximally localized generalized Wannier functions for composite energy bands," Phys. Rev. B **56**, 12847–12865.

Maxwell, J C (1865), "A dynamical theory of the electromagnetic field," Phil. Trans. R. Soc. Lond. **155**, 459–512.

Maxwell, J C (1873), *A treatise on electricity and magnetism*, 3rd ed., Vol. I (Oxford, Clarendon Press).

Maxwell, J C (1892), *A treatise on electricity and magnetism*, 3rd ed., Vol. II (Oxford, Clarendon Press).

Mazur, P (1957), "On statistical mechanics and electromagnetic properties of matter," in *Adv. Chem. Phys.* (John Wiley & Sons, Ltd) pp. 309–360.

Mazur, P, and B. R. A. Nijboer (1953), "On the statistical mechanics of matter in an electromagnetic field. I: derivation of the Maxwell equations from electron theory," Physica **19** (1), 971–986.

McQuarrie, D A, P. A. Rock, and E. B. Gallogly (2011), *General Chemistry*, 4th ed. (Royal Society of Chemistry).

McWeeny, R (1960), "Some recent advances in density matrix theory," Rev. Mod. Phys. **32**, 335–369.

Miyake, S (1940), "On the mean inner potential of crystals," in *Proceedings of the Physico-Mathematical Society of Japan*, 3rd, Vol. 22, pp. 666–676.

Mukhopadhyay, S, J. Sun, A. Subedi, T. Siegrist, and D. J. Singh (2016), "Competing covalent and ionic bonding in Ge-Sb-Te phase change materials," Sci. Rep. **6**, 25981.

Noguera, C (2000), "Polar oxide surfaces," J. Phys.: Condens. Matter **12**, R367–R410.

Noguera, C, and J. Goniakowski (2013), "Polarity in oxide nano-objects," Chem. Rev. **113** (6), 4073–4105.

Ouellette, R J, and J. D. Rawn (2015), *Principles of Organic*

Chemistry (Elsevier Science).

Pauling, L (1960), *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry*, George Fisher Baker Non-Resident Lecture Series (Cornell University Press).

Peng, L-M (1999), "Electron atomic scattering factors and scattering potentials of crystals," Micron **30** (6), 625–648.

Pick, R M, M. H. Cohen, and R. M. Martin (1970), "Microscopic theory of force constants in the adiabatic approximation," Phys. Rev. B **1**, 910–920.

Pratt, L R (1992), "Contact potentials of solution interfaces: phase equilibrium and interfacial electric fields," J. Phys. Chem. **96** (1), 25–33.

Raab, R E, and O. L. de Lange (2005), "Transformed multipole theory of the response fields **D** and **H** to electric octopole–magnetic quadrupole order," Proc. R. Soc. A. **461**, 595–608.

Resta, R (1992), "Theory of the electric polarization in crystals," Ferroelectrics **136** (1), 51–55.

Resta, R (1993), "Macroscopic electric polarization as a geometric quantum phase," Europhys. lett. **22** (2), 133–138.

Resta, R (1994), "Macroscopic polarization in crystalline dielectrics: the geometric phase approach," Rev. Mod. Phys. **66**, 899–915.

Resta, R, and D. Vanderbilt (2007), "Theory of polarization: A modern approach," in *Physics of Ferroelectrics*, Topics in Applied Physics, Vol. 105 (Springer Berlin, Heidelberg) pp. 31–68.

Rez, D, P. Rez, and I. Grant (1994), "Dirac–Fock calculations of X-ray scattering factors and contributions to the mean inner potential for electron scattering," Acta Crystallogr. A **50** (4), 481–497.

Riley, K F, M. P. Hobson, and S. J. Bence (2006), *Mathematical Methods for Physics and Engineering: A Comprehensive Guide*, 3rd ed. (Cambridge University Press).

Robinson, F N H (1971), "The microscopic and macroscopic equations of the electromagnetic field," Physica **54** (3), 329–341.

Roche, J J (2000), "**B** and **H**, the intensity vectors of magnetism: A new approach to resolving a century-old controversy," Am. J. Phys. **68** (5), 438–449.

Russakoff, G (1970), "A derivation of the macroscopic Maxwell equations," Am. J. Phys. **38** (10), 1188–1195.

Saldin, D K, and J. C. H. Spence (1994), "On the mean inner potential in high- and low-energy electron diffraction," Ultramicroscopy **55** (4), 397–406.

Sanchez, A, and M. A. Ochando (1985), "Calculation of the mean inner potential," J. Phys. C: Solid State Phys. **18** (1), 33–41.

Saunders, V R, C. Freyria-Fava, R. Dovesi, L. Salasco, and C. Roetti (1992), "On the electrostatic potential in crystalline systems where the charge density is expanded in gaussian functions," Mol. Phys. **77** (4), 629–665.

Schram, K (1960), "Quantum statistical derivation of the macroscopic Maxwell equations," Physica **26** (12), 1080–1090.

Sokhan, V P, and D. J. Tildesley (1997a), "The free surface of water: Molecular orientation, surface potential, and nonlinear susceptibility," Mol. Phys. **92**, 625.

Sokhan, V P, and D. J. Tildesley (1997b), "The free surface of water: molecular orientation, surface potential and nonlinear susceptibility," Mol. Phys. **92** (4), 625–640.

Souza, I, N. Marzari, and D. Vanderbilt (2001), "Maximally localized Wannier functions for entangled energy bands," Phys. Rev. B **65**, 035109.

Spence, J C H (1993), "On the accurate measurement of structure-factor amplitudes and phases by electron diffraction," Acta Crystallogr., Sect. A: Found. Crystallogr. **49** (2), 231–260.

Spence, J C H (1999), "The future of atomic resolution electron microscopy for materials science," Mater. Sci. Eng., R **26** (1), 1 – 49.

Stengel, M (2011), "Electrostatic stability of insulating surfaces: Theory and applications," Phys. Rev. B **84**, 205432.

Stengel, M, and D. Vanderbilt (2009), "Berry-phase theory of polar discontinuities at oxide-oxide interfaces," Phys. Rev. B **80**, 241103.

Stillinger, F H, Jr, and A. Ben-Naim (1967), "Liquid—vapor interface potential for water," J. Chem. Phys. **47** (11), 4431–4437.

Tagantsev, A K (1991), "Electric polarization in crystals and its response to thermal and elastic perturbations," Phase Transitions **35** (3-4), 119–203.

Tasker, P W (1979), "The stability of ionic crystal surfaces," J. Phys. C: Solid State Phys. **12** (22), 4977.

Toumpanaki, E, D. U. Shah, and S. J. Eichhorn (2021), "Beyond what meets the eye: Imaging and imagining wood mechanical–structural properties," Adv. Mat. **33** (28), 2001613.

Vanderbilt, D (2018), *Berry Phases in Electronic Structure Theory: Electric Polarization, Orbital Magnetization and Topological Insulators* (Cambridge University Press).

Vanderbilt, D, and R. D. King-Smith (1993), "Electric polarization as a bulk quantity and its relation to surface charge," Phys. Rev. B **48**, 4442–4455.

Vinogradov, A P, and A. V. Aivazyan (1999), "Scaling theory for homogenization of the Maxwell equations," Phys. Rev. E **60**, 987–993.

van Vleck, J H (1937), "On the role of dipole-dipole coupling in dielectric media," J. Chem. Phys. **5** (7), 556–568.

Vogl, P (1978), "Dynamical effective charges in semiconductors: A pseudopotential approach," J. Phys. C: Solid State Phys. **11** (2), 251.

Wannier, G H (1937), "The structure of electronic excitation levels in insulating crystals," Phys. Rev. **52**, 191–197.

Wilson, M A, A. Pohorille, and L. R. Pratt (1987), "Molecular dynamics of the water liquid-vapor interface," J. Phys. Chem. **91** (19), 4873–4878.

Wilson, M A, A. Pohorille, and L. R. Pratt (1988), "Surface potential of the water liquid-vapor interface," J. Chem. Phys. **88** (5), 3281–3285.

Wilson, M A, A. Pohorille, and L. R. Pratt (1989), "Comment on "Study on the liquid-vapor interface of water. 1. Simulation results of the thermodyanmic properties and orientational structure," J. Chem. Phys. **90**, 5211.

Woo, J W F (1971), "Piezoelectricity under hydrostatic pressure," Phys. Rev. B **4**, 1218–1220.

Yesibolati, M N, S. Laganà, H. Sun, M. Beleggia, S. M. Kathmann, T. Kasama, and K. Mølhave (2020), "Mean inner potential of liquid water," Phys. Rev. Lett. **124**, 065502.

Zürcher, U (2018), *Electrostatics at the Molecular Level* (IOP Concise Physics and Morgan & Claypool Publishers)