Part III: Elementary Excitations in Solids

- Review of Lagrangian and Hamiltonian formulation of Classical Mechanics
- Canonical Quantization
- Lagrangian and Hamiltonian formulation for Fields
- Canonical Quantization of Fields (‘Second Quantization’)
- An example: Quantization of the Schrödinger Field
- Plasmons
- Phonons
- Photons, i.e., Quantization of the Electromagnetic Field
- Dielectric Properties of Semiconductors
Second quantization is necessary when we want to study many-particle effects. Historically, it was born out of the relativistic description of particles, since ordinary Quantum Mechanics (QM) could not handle it: The number of particles is fixed in conventional QM, but a relativistic description must allow for creation and annihilation of particles. It soon became apparent that this new description was extremely useful in describing solid-state phenomena and statistical mechanics, since second-quantization, dealing with fields, is intrinsically a description of systems with infinitely many (...well, very many indeed in solid-state systems, albeit not infinitely many...) degrees of freedom. It is thus very well suited to describe collective phenomena as elementary excitations above the ground state: Elementary excitations – such as phonons, plasmons, magnons, etc. – can be created (emissions) and destroyed (absorption), similarly to relativistic particles. In addition, for small perturbations away from equilibrium, the Hamiltonian can often be linearized and algebraically manipulated so that it takes the form of the Hamiltonian of a collection of harmonic oscillators (as we shall see for free electrons, plasmons, and phonons below). These manipulations often take the name of Bogoliubov transformations. So, the whole machinery of simple harmonic oscillators can be used to handle the much more complicated collective excitations.

In “first quantization” (i.e., conventional QM), we start by describing the system classically in the Lagrangian form. We then translate the description into the Hamiltonian form, so that the equations of motion are expressed in terms of Poisson brackets between the Hamiltonian, the dynamic variables (observables), and the canonical coordinates (generalized coordinates and conjugate momenta). Quantization is finally performed by promoting the dynamic variables to operators on a suitable Hilbert space and converting the Poisson brackets to commutators. In second quantization we follow a similar route: Think of a system as described by the generalized coordinates $\phi_i$, where $i$ runs over the degrees of freedom of the system. As the number of degrees of freedom approaches infinity, we replace the discrete index $i$ with a continuum of indices $r$ and our generalized coordinates become a field $\phi(r)$. Thus, we start from a Lagrangian description of the system in terms of Lagrangian fields (as done, classically, in elasticity theory), convert it to a Hamiltonian description, and convert the fields to operators on some (more complicated) Hilbert space promoting the Poisson brackets involving fields to commutators (or anticommutators, as discussed below). The use of the term “second quantization” is due to the fact that the first fields to be quantized were solutions of the Dirac equation. Thus, the procedure appeared to be a quantization of wavefunctions, already
quantized in conventional QM, and so it seemed that we had quantized the system twice. But this is not the right way of looking at it: What we have done is to have dropped altogether the concept of “particle” as primary concept and replaced it with field, from which particles (or “quasi-particles”, in solid state physics) will emerge as excitations of the field away from its ground state.

It is now clear that, in order to discuss second quantization even at an elementary level, we must deal with several topics: First we shall review the Lagrangian and Hamiltonian formulations of Classical Mechanics, as they are needed to perform the canonical quantization of a system with a finite number of degrees of freedom. Reviewing this procedure will be helpful, since the same procedure, as we just saw, will be followed in the more complicated case of infinitely many degrees of freedom. Having done that, we shall discuss the Lagrangian and Hamiltonian formulations of fields, and, as a final step, quantize them in the canonical way. The example of the simpler Schrödinger field will be the first case we shall discuss before moving to two example of great relevance to charge transport in solids: Plasmons and Phonons.

- Hamilton’s Principle.

There are several ways in which one may introduce the concepts of Lagrangian and derive laws of Classical Mechanics from them. The most ‘elegant’ is based on Hamilton’s principle. Let’s consider a system (see page 2) with \( N \) degrees of freedom. \( N \) may be thought as equal to \( 3n \), where \( n \) is the number of classical particles in three dimensions. Let \( q_i \) be their ‘generalized’ coordinates, with \( i = 1, N \). By ‘generalized’ we mean that these coordinates are related to the position of particle \( \alpha \) via relations of the form \( r_\alpha(q_1, q_2, \ldots, q_N) \). Thus, the coordinates \( q_i \) may be the particles positions in another coordinate system (polar, cylindrical, etc.), or may be the coordinates on the subspace on which the system is forced to be by constraints (such as particles on a surface), so that, in this case \( N < 3n \).

Let’s assume that forces induced by interactions among the particles or due to an external source (gravity, electric fields, etc.) can be described by a potential \( V(q) \) (where the \( q \) labels the \( N \)-dimensional vector
(\(q_1, q_2, \ldots, q_N\)) and let’s define the kinetic energy \(T\) of the system:

\[
T(\dot{q}) = \frac{1}{2} \sum_{\alpha} m_{\alpha} v_{\alpha}^2 = \frac{1}{2} \sum_{\alpha} m_{\alpha} \left( \sum_i \frac{\partial r_{\alpha}}{\partial q_i} \dot{q}_i + \frac{\partial r_{\alpha}}{\partial t} \right)^2 .
\]  

(231)

Now let’s define the Lagrangian:

\[
L(q, \dot{q}) = T(\dot{q}) - V(q) ,
\]

(232)

which should be consider as a ‘functional’ of all possible trajectories \(q(t), \dot{q}(t)\) of the system. Hamilton’s principle states that the motion of the system, \(q(t), \dot{q}(t)\), during the time interval \((t_1, t_2)\) is such that:

\[
\delta \int_{t_1}^{t_2} L(q, \dot{q}) dt = 0 ,
\]

(233)

that is: the integral takes an extremum value when we substitute the physical trajectory \(q(t), \dot{q}(t)\) in the integral above of the Lagrangian functional.

- **A primer on variational calculus:**

How do we solve the variational problem described by Eq. (233)? Let \(f(y, \dot{y}, x)\), with \(\dot{y} = dy/dx\) be a functional of \(y\) and \(\dot{y}\) and let’s look for the particular function \(y(x)\) for which the integral

\[
\mathcal{I} = \int_{x_1}^{x_2} f(y, \dot{y}, x) \, dx
\]

(234)

(where the ‘dot’ in \(\dot{y}\) now denotes differentiation with respect to \(x\)) takes an extremum value, with the boundary conditions \(y(x_1) = y_1\) and \(y(x_2) = y_2\). Let’s parametrize all possible ‘trajectories’ \(y(x)\) by a parameter \(\lambda\), so that the integral \(\mathcal{I}\) becomes a function of \(\lambda\),

\[
\mathcal{I}(\lambda) = \int_{x_1}^{x_2} f[y(x, \lambda), \dot{y}(x, \lambda), x] \, dx .
\]

(235)
Then, our problem asks us to solve
\[
\frac{\partial I(\lambda)}{\partial \lambda} = 0 .
\] (236)

Now:
\[
\frac{\partial I(\lambda)}{\partial \lambda} = \int_{x_1}^{x_2} \left\{ \frac{\partial f}{\partial y} \frac{\partial y}{\partial \lambda} + \frac{\partial f}{\partial \dot{y}} \frac{\partial \dot{y}}{\partial \lambda} \right\} \, dx .
\] (237)

The second integral can be integrated by parts:
\[
\int_{x_1}^{x_2} \frac{\partial f}{\partial \dot{y}} \frac{\partial^2 y}{\partial x \partial \lambda} \, dx = \left. \frac{\partial f}{\partial \dot{y}} \frac{\partial y}{\partial \lambda} \right|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{d}{dx} \left( \frac{\partial f}{\partial \dot{y}} \right) \frac{\partial y}{\partial \lambda} \, dx .
\] (238)

Our boundary conditions require all functions to take the same values \( y_1 \) and \( y_2 \) at \( x_1 \) and \( x_2 \), so the first term vanishes. Thus,
\[
\frac{\partial I(\lambda)}{\partial \lambda} = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}} \right) \frac{\partial y}{\partial \lambda} \, dx .
\] (239)

Now let’s consider the variation of \( y \) around an arbitrary value \( \lambda_0 \). Then:
\[
\left( \frac{\partial I}{\partial \lambda} \right)_0 \, d\lambda = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}} \right) \left( \frac{\partial y}{\partial \lambda} \right)_0 \, dx \, d\lambda .
\] (240)

Now set
\[
\delta I = \left( \frac{\partial I}{\partial \lambda} \right)_0 \, d\lambda ,
\] (241)
\[
\delta y = \left( \frac{\partial y}{\partial \lambda} \right)_0 \, d\lambda ,
\] (242)
and
\[
\delta \dot{y} = \left( \frac{\partial \dot{y}}{\partial \lambda} \right)_0 \, d\lambda .
\] (243)
Thus, Eq. (240) becomes:

$$\delta I = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}} \right) \delta y dx = 0 . \quad (244)$$

But now note that $\delta y$ is an arbitrary variation of the trajectory. The only way Eq. (244) can be satisfied is when the term within parentheses vanish, that is:

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial \dot{y}} = 0 . \quad (245)$$

This is known as the Euler equation (although there are so many ‘Euler equations' in so many areas of math and physics!).

- **Euler-Lagrange equations.**

Equation (233) can be solved in a very similar way, leading us to the Euler-Lagrange equations:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 . \quad (246)$$

Note immediately that this formulation of classical mechanics exhibits explicitly some conservation laws: If the Lagrangian $L$ does not depend explicitly on the variable $q_j$ for some $j$, then from Eq. (246) we see that

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \quad (247)$$

is a constant of motion (i.e., is a conserved quantity during the time evolution of the system). In this case $q_j$ is said to be a ‘cyclical coordinate’ while $p_i$ is called the ‘canonical generalized conjugate momentum’.
Hamiltonian and Hamilton equations.
This observation, as well as the definition of the conjugate momentum, gives the idea of transforming the Lagrangian \( L \) into a functional of \( q \) and \( p \), rather than of \( q \) and \( \dot{q} \). This can be done by the Legendre transform

\[
L(q, \dot{q}, t) \rightarrow H(q, p, t) = \sum_i p_i \dot{q}_i - L(q, \dot{q}, t).
\]  

(248)

This is the familiar ‘Hamiltonian’. In order to derive the equations of motion listed on page 3, note that

\[
dH = \sum_i \frac{\partial H}{\partial q_i} dq_i + \sum_i \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt.
\]  

(249)

On the other hand, from the definition (248) we have also:

\[
dH = \sum_i \dot{q}_i dp_i + \sum_i p_i d\dot{q}_i - \sum_i \frac{\partial L}{\partial q_i} dq_i - \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt.
\]  

(250)

Note that the second and third term in this equation cancel, thanks to the definition (247). Since \( \dot{p}_i = \partial L / \partial q_i \), Eq. (250) becomes:

\[
dH = \sum_i \dot{q}_i dp_i - \sum_i \dot{p}_i dq_i - \frac{\partial L}{\partial t} dt.
\]  

(251)

Comparing Eqns. (249) and (251) we arrive at the canonical Hamilton’s equations:

\[
\dot{q}_i = \frac{\partial H}{\partial p_i},
\]  

(252)

\[
\dot{p}_i = -\frac{\partial H}{\partial q_i},
\]  

(253)

together with

\[
\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.
\]  

(254)
• **Poisson Brackets.**

As we have seen before (see page 3), Hamilton’s equations can be reformulated in terms of the Poisson brackets:

\[
\{A, B\} = \sum_{i=1}^{n} \left( \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right).
\] (255)

Using this definition, it is trivial to see that Eqns. (252) and (253) can be written as:

\[
\frac{dq_i}{dt} = \{q_i, H\},
\] (256)

\[
\frac{dp_i}{dt} = \{p_i, H\},
\] (257)

while for a general dynamic variable \(A(q, p)\):

\[
\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}.
\] (258)

Note also the following relations which will be used when quantizing the system:

\[
\{q_i, p_j\} = \delta_{ij}, \quad \{q_i, q_j\} = 0, \quad \{p_i, p_j\} = 0.
\] (259)
Canonical Quantization

The axiomatization of Quantum Mechanics has brought us to the following procedure to describe quantum mechanically any system:

1. Define a Hilbert space $\mathcal{H}$ associated with the system. Vectors of this space are state-vectors of the system. Observables (or dynamic variables) are operators on this space.

2. Provide a Hamiltonian classical description of the system, identifying the canonical coordinates $q$ and momenta $p$.

3. Promote the canonical variables to operators acting on the Hilbert space $\mathcal{H}$,

\[
q_i \rightarrow \hat{q}_i, \quad p_i \rightarrow \hat{p}_i = -i\hbar \frac{\partial}{\partial q_i},
\]  

(260)

and convert Poisson brackets $\{\ldots, \ldots\}$ to commutators $[\ldots, \ldots]$ , so that:

\[
[q_i, p_j] = i\hbar \delta_{ij}, \quad [q_i, q_j] = 0, \quad [p_i, p_j] = 0.
\]

(261)

4. In the Heisenberg representation, the equations of motion for the dynamic variables are similar to the Hamilton equations (258):

\[
\frac{d\hat{A}}{dt} = \frac{i}{\hbar} [\hat{H}, \hat{A}] + \frac{\partial \hat{A}}{\partial t}.
\]

(262)
Lagrangian and Hamiltonian Formulation for Fields

- **Lagrangian Density.**
  We discuss here the transition from systems with a finite number of degrees of freedom to systems with infinitely many (actually, a continuum) of degrees of freedom. A typical example which might help to think of is the displacement of a violin string. At each location \( x \) along the string, the displacement of the string from its equilibrium position is \( \phi(x) \). We would like to find a Lagrangian and Hamiltonian description of the ‘plucked’ string. Clearly, we have way too many degrees of freedom, one for each location \( x \).

  The way to do it is to define a *Lagrangian density*:

  \[
  \mathcal{L} \left( \phi, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial t}, x, t \right),
  \]

  (263)

  so that the Lagrangian is obtained by integration over the length \( D \) of the string:

  \[
  L(t) = \int_0^D \mathcal{L} \left( \phi, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial t}, x, t \right) \, dx.
  \]

  (264)

  The equations of motion can be obtained from a generalized Hamilton principle requiring:

  \[
  \delta \mathcal{I} = \delta \int_{t_1}^{t_2} L(t) \, dt = \int_{t_1}^{t_2} \int_0^D \mathcal{L} \left( \phi, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial t}, x, t \right) \, dx \, dt = 0.
  \]

  (265)

  Following a generalized version of the procedure followed before (pages 121-124), this variational principle leads to Euler-Lagrange equations of the form

  \[
  \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} + \frac{d}{dx} \left( \frac{\partial \mathcal{L}}{\partial (\partial \phi/\partial x)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0.
  \]

  (266)
In three dimensions and in the case of several ‘fields’ $\phi_i$, the Euler-Lagrange equations become:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\phi}_i} + \sum_{k=1}^{3} \frac{d}{dx_k} \left( \frac{\partial L}{\partial (\partial \phi_i/\partial x_k)} \right) - \frac{\partial L}{\partial \phi_i} = 0 \quad \text{for } j = 1, 2, \ldots \quad (267)$$

Now let’s introduce the functional derivative of $L$ with respect to $\phi_i$ as:

$$\frac{\delta L}{\delta \phi_i} = \frac{\partial L}{\partial \phi_i} - \sum_{k=1}^{3} \frac{\partial}{\partial x_k} \left( \frac{\partial L}{\partial (\partial \phi_i/\partial x_k)} \right) \quad (268)$$

This notation has the advantage of allowing us to ignore the dependence of the Lagrangian density on the spatial derivatives of $\phi$. Similarly,

$$\frac{\delta L}{\delta \dot{\phi}_i} = \frac{\partial L}{\partial \dot{\phi}_i} \quad (269)$$

which is simpler looking, since $L$ does not depend on the spatial gradients of $\dot{\phi}_i$. Thus, Eq. (267) takes the simpler-looking form:

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{\phi}_i} - \frac{\delta L}{\delta \phi_i} = 0 , \quad (270)$$

which are very similar to the ‘usual’ Euler-Lagrange equations.

**Conjugate Momenta and Hamiltonian Density.**

We can push the analogy with the usual Hamiltonian theory even further: Let’s define the conjugate momenta (which are fields now!):

$$\pi_i = \frac{\delta L}{\delta \dot{\phi}_i} , \quad (271)$$

and the Hamiltonian:

$$H = \int \mathcal{H} \, dx , \quad (272)$$
where $\mathcal{H}$ is the Hamiltonian density:
\begin{equation}
\mathcal{H} = \sum_k \pi_k \dot{\phi}_k - \mathcal{L}.
\end{equation}

- **Hamilton Equations.**
  We can show that the Euler-Lagrange equations (270) become the Hamilton equations:
\begin{equation}
\frac{\delta H}{\delta \phi_k} = -\pi_k, \quad \frac{\delta H}{\delta \pi_k} = \dot{\phi}_k,
\end{equation}

Finally, given any function $A$ which can be represented as the integral of a density $\mathcal{A}$,
\begin{equation}
A = \int A \, dr,
\end{equation}

one can show that its time derivative is given by:
\begin{equation}
\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t},
\end{equation}

where the symbol $\{..., ...\}$ is the obvious generalization of the Poisson brackets:
\begin{equation}
\{A, B\} = \int \sum_k \left( \frac{\delta A}{\delta \phi_k} \frac{\delta B}{\delta \pi_k} - \frac{\delta A}{\delta \pi_k} \frac{\delta B}{\delta \phi_k} \right) \, dr = \int \{A((r)), B(r)\} \, dr.
\end{equation}

where the Poisson bracket at a given point has been defined as:
\begin{equation}
\{A(r), B(r')\} = \sum_k \left[ \frac{\delta A(r)}{\delta \phi_k} \frac{\delta B(r')}{\delta \pi_k} - \frac{\delta A(r)}{\delta \pi_k} \frac{\delta B(r')}{\delta \phi_k} \right].
\end{equation}

Note that the following relations hold for the Poisson brackets between canonical variables (see the analogous
relations Eq. (259) for systems with a finite number of degrees of freedom):

\[
\{\phi_i(r), \pi_j(r')\} = \delta_{ij} \delta(r - r'), \quad \{\phi_i(r), \phi_j(r')\} = 0, \quad \{\pi_i(r), \pi_j(r')\} = 0, \quad (279)
\]

- **Canonical Quantization of Fields.**

The procedure to be followed in order to quantize the field \( \phi \) is strictly parallel to the canonical quantization scheme outlined on page 127, although several issues are significantly more complicated:

1. We define/select/guess the Hilbert space associated with the system. As we shall see below, this task is not as straightforward as in the case of ‘conventional’ QM. Let’s defer to a later section a short discussion on this issue. For the moment let’s just say that we shall work on what’s called ‘Fock space’ \( \mathcal{F} \).

2. We re-express the field equations (Maxwell equations, the wave equation for the strings of our ‘quantum violin’, Dirac equation, or even Schrödinger equation, as we shall see below) in their Hamiltonian formulations.

3. We promote the fields \( \phi \) and \( \pi \) to operators acting on the Hilbert space \( \mathcal{F} \). These operators will satisfy the canonical commutation rules:

\[
[\phi_i(r), \pi_j(r')] = i\hbar \delta_{ij} \delta(r - r'), \quad [\phi_i(r), \phi_j(r')] = 0, \quad [\pi_i(r), \pi_j(r')] = 0. \quad (280)
\]

4. The equation of motion will follow from the ‘usual’ commutation rules with the Hamiltonian (in the Heisenberg representation).
The Schrödinger Field.

Let’s consider, as a first example, the quantization of the Schrödinger Field. This example – in addition to being relevant since it constitutes the basic framework needed to describe also electrons in a crystal – shows how the Hamiltonian of free (that is: non-interacting among themselves) particles can be written as the sum of Hamiltonians of harmonic oscillators, each oscillator corresponding to one particular single-electron state, each excited state of each oscillator representing the presence of one particle in that particular state. The main new ingredient we shall encounter is the anti-commutation rules (instead of ‘commutation’ rules) the electron operators must obey in order to satisfy Pauli’s exclusion principle.

• The Lagrangian and Hamiltonian Densities

As we said, the main idea is to consider the ‘usual’ Schrödinger equation

\[
i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(r, t) \psi,
\]

but consider the function \(\psi(r)\) as a ‘classical field’ (much like the electric field of Maxwell’s equation, or the displacement field of our violin string) rather than a ‘wavefunction’. ‘Classical’ here means ‘not yet quantized’. Having accepted that, let’s look for a Lagrangian density (functional of electronic wavefunction \(\psi\)) which gives us back the Schrödinger equation. It can be shown that:

\[
\mathcal{L}(\psi, \dot{\psi}, t) = i\hbar \psi^*(r, t) \frac{\partial \psi(r, t)}{\partial t} - \frac{\hbar^2}{2m} \nabla \psi^*(r, t) \cdot \nabla \psi(r, t) - V(r, t) \psi^*(r, t) \psi(r, t).
\]

then, using the usual variational procedure which yields the Euler equations, one finds that the field \(\psi(r, t)\) which corresponds to an extremum of the Lagrangian \(L(t) = \int d\mathbf{r} \mathcal{L}(\psi, \dot{\psi}, t)\) is a solution of the equation

\[
i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(r, t) \psi.
\]
i.e., Schrödinger equation, which is what we wanted. Actually, using Eq. (270) one would obtain the complex conjugate of Eq. (283). Equivalently, instead of varying $\psi$ to derive Eq. (270) from Eq. (265), one may vary the complex-conjugate of Eq. (282) over $\psi^*$ and obtain the Lagrange equation

$$\frac{d}{dt} \frac{\delta L^*}{\delta \psi^*_i} - \frac{\delta L^*}{\delta \psi^*_i} = 0,$$

from which Eq. (283) follows.

Now let’s define the canonical momentum density

$$\pi(r, t) = \frac{\delta L}{\delta (\partial \psi / \partial t)} = i \hbar \psi^*(r, t),$$

and the Hamiltonian density:

$$\mathcal{H}(\psi, \pi, t) = \pi(r, t) \frac{\partial \psi(r, t)}{\partial t} - \mathcal{L}(r, t) = -\frac{i\hbar}{2m} \nabla \pi \cdot \nabla \psi - \frac{i}{\hbar} V \pi \psi$$

We are now ready to quantize the field $\psi$.

*Canonical Quantization and Commutation Rules*

Second-quantization is now performed by promoting the functions $\psi$ and $\pi$ to operators satisfying the equal-time canonical commutation rules (CCR), which are a generalization of the canonical Poisson brackets between coordinates and momenta in the classical description of systems with a finite number of degrees of freedom:

$$[\psi(r, t), \pi(r', t)] = i\hbar \delta^{(3)}(r - r'), \quad [\psi(r, t), \psi(r', t)] = 0, \quad [\pi(r, t), \pi(r', t)] = 0,$$

which, in turn, imply:

$$[\psi(r, t), \psi^\dagger(r', t)] = \delta^{(3)}(r - r'), \quad [\psi(r, t), \psi(r', t)] = 0, \quad [\psi^\dagger(r, t), \psi^\dagger(r', t)] = 0.$$
Note that Hamilton’s equations become:

\[
\dot{\psi} = \frac{\delta H}{\delta \pi} = \frac{\partial H}{\partial \pi} - \sum_k \frac{\partial H}{\partial (\partial \pi/\partial x_k)} = -\frac{i}{\hbar} V \psi + \frac{i\hbar}{2m} \nabla^2 \psi , 
\]

which is the original Schrödinger equation, while the other equation,

\[
\dot{\pi} = - \frac{\delta H}{\delta \psi} = - \frac{\partial H}{\partial \psi} + \sum_k \frac{\partial H}{\partial (\partial \psi/\partial x_k)} = \frac{i}{\hbar} V \pi - \frac{i\hbar}{2m} \nabla^2 \pi , 
\]

is the complex conjugate equation, as can be seen using Eq. (285).

**The Hamiltonian as a collection of harmonic oscillators.**

Now let’s expand the field as:

\[
\psi(r, t) = \sum_{\mu} c_\mu(t) \, u_\mu(r) , 
\]

\[
\psi^\dagger(r, t) = \sum_{\mu} c^\dagger_\mu(t) \, u^*_\mu(r) , 
\]

where the (numeric) functions \(u_\mu\) are the stationary eigenfunctions of Eq. (283) with eigenvalues \(E_\mu\). The label \(\mu\) will actually be a set of three labels (+ spin, ignored for the time being), possibly belonging to a continuous spectrum, for electrons in three dimensions. It may be just the three components of the \(k\)-vector when \(V = 0\), it may be the set of integers \((n, l, m)\) for atomic orbitals, or any other set of ‘quantum numbers’. Now, the crucial step: Some algebra shows that the Hamiltonian (the spatial integral of the Hamiltonian density) can be written as:

\[
H(t) = \sum_{\mu} E_\mu \, c^\dagger_\mu(t) \, c_\mu(t) . 
\]

In order to derive this expression we need to derive the (anti)commutation rules for the operators \(c_\mu\) and \(c^\dagger_\mu\), which follow from the (anti)commutation rules for the field, Eqns. (287) and (288): As it will discussed below, the only non-vanishing (anti)commutators
all other (anti)commutators vanishing. We also need to remark that the expression given by Eq. (286) for the Hamiltonian density is ambiguous when interpreted in the operator sense, since the product $AB$ of two numbers can be written as $AB$ or $BA$. If $[A, B] \neq 0$, we must decide which of the two expressions we should use. Commonly the ‘symmetrized’ product $(AB + BA)/2$ is employed. Then we would reach the result:

$$H(t) = \sum_{\mu} \left[ E_{\mu} c_{\mu}^\dagger(t) c_{\mu}(t) + \frac{1}{2} \right] ,$$

instead of Eq. (293). The term $1/2$ will be discussed below, but we ignore it for the moment.

Note the similarity of this expression with the Hamiltonian of a harmonic oscillator. We are going to use this fact to draw an algebraic analogy with the ‘ladder operator’ formalism of the quantum harmonic oscillator to discover the eigenstates of the Hamiltonian. Let’s define the ‘position’ and ‘momentum’ operators: $\mu$:

$$q_{\mu} = \left( \frac{E_{\mu}}{2} \right)^{1/2} (c_{\mu} + c_{\mu}^\dagger) , \quad p_{\mu} = -i \left( \frac{E_{\mu}}{2} \right)^{1/2} (c_{\mu} - c_{\mu}^\dagger) .$$

Then, we see that the Hamiltonian can be rewritten as

$$H(t) = \sum_{\mu} \frac{1}{2} \left[ \left( p_{\mu}^2 + q_{\mu}^2 \right) - 1 \right] ,$$

which is indeed the Hamiltonian of a harmonic oscillator if we identify the operators $c_{\mu}$ and $c_{\mu}^\dagger$ with the ‘usual’ lowering (annihilator) and raising (creator) operators for each oscillator $\mu$. Note that the ”1” (a divergent term, whose origin we shall briefly discuss below) would disappear had we started from Eq. (295) instead of Eq. (293).

The Hamiltonian of a one-dimensional harmonic oscillator is:

$$H = \frac{p^2}{2M} + \frac{Kx^2}{2} .$$
Define the operators:

\[ a_{\pm} = \frac{1}{(2M)^{1/2}} p \pm i \left( \frac{K}{2} \right)^{1/2} x . \]  

Note that \( a^\dagger_+ = a_- \) and that, using the commutation rule \([p, x] = -i\hbar\):

\[ H = a_+ a_- + \frac{1}{2} \hbar \omega = a_- a_+ - \frac{1}{2} \hbar \omega , \]  

where \( \omega = (K/M)^{1/2} \). Note also that

\[ [a_+, a_-] = -\hbar \omega , \quad [H, a_{\pm}] = \pm \hbar \omega a_{\pm} . \]  

Using these algebraic relations, we can construct the entire spectrum of the Hamiltonian (298) using the ground state and the operators \( a_{\pm} \). Indeed, suppose that we know the ground state \(|0\rangle\), such that:

\[ H|0\rangle = E_0 |0\rangle . \]  

Using the first of Eqns. (301),

\[ Ha_+|0\rangle = (E_0 + \hbar \omega) a_+|0\rangle , \]  

which shows that starting from the ground-state \(|0\rangle\) we can ‘create’ excited states by applying repeatedly the ‘creation’ (or ‘raising’) operator \( a_+ \). What’s left to do in order to find the eigenvalues corresponding to these states is to find the ground-state energy \( E_0 \). Note that

\[ Ha_-|0\rangle = (E_0 - \hbar \omega)a_-|0\rangle . \]  

Since we have assumed that \( E_0 \) is the lowest possible energy, this equation implies that \( a_-|0\rangle \) must not exists, that is, \( a_-|0\rangle = 0 \). But from the first of the Eqns. (300),

\[ a_+ a_-|0\rangle = 0 \quad \rightarrow \quad \left( H - \frac{1}{2} \hbar \omega \right) |0\rangle = 0 . \]  

This has the form of the eigenvalue equation for the ground state, so that \( E_0 = \hbar \omega /2 \). Therefore, the eigenstates of \( H \) are the states \(|n\rangle = a_n^\dagger |0\rangle \) with eigenvalues \( E_n = (n + 1/2)\hbar \omega \).
The 'raising' operator $a_+$ corresponds to the creation operator $c_\mu^\dagger$, the 'lowering' operator $a_-$ to the annihilation operator $c_\mu$ defined above. They satisfy exactly the same algebraic and commutation relations.

The other observation we should make relates to the term ”-1” inside the square brackets in Eq. (297). This term corrects for the zero-point energy of the infinitely many oscillators $\mu$. It is clearly a troublesome diverging term. The reasons for its appearance are very deep and are briefly discussed below. It is usually referred to as the ‘energy of the vacuum’. It is just a constant (albeit an infinite one!) and if we decide to measure all energies from this value we can ignore it.

- **Anticommutation rules**
  We have performed canonical quantization by employing the commutation rules Eqns. (287) and (288) above. However, these do not guarantee that Pauli’s exclusion principle will be obeyed. Jordan and Wigner have shown that we can fix this problem by employing **anticommutation rules**

$$
[\psi(r, t), \pi(r', t)]_+ = i\hbar \delta^3(r-r'), \quad [\psi(r, t), \psi(r', t)]_+ = 0, \quad [\pi(r, t), \pi(r', t)]_+ = 0, \quad (306)
$$

(where $[A, B]_+ = AB + BA$) instead of Eq. (287) above. These imply also:

$$
[\psi(r, t), \psi^\dagger(r', t)]_+ = \delta^3(r-r'), \quad [\psi(r, t), \psi(r', t)]_+ = 0, \quad [\psi^\dagger(r, t), \psi^\dagger(r', t)]_+ = 0. \quad (307)
$$

The reason behind the selection of **anti**-commutation rules is very deep. Streater and Whightman, in a book entitled “PCT, Spin, Statistics, and all that” have shown that free particles of half-integer spin **must** be associated with a quantum field $\psi(r)$ such that the CAR, Eq. (307) are satisfied, or one would violate at least one of several “axioms” (relativistic invariance, causality, locality, etc.) which constitute the foundation of Quantum Field Theory (QFT). Similarly, for integer-spin particles, the field must be such that the “canonical
commutation rules” (CCR) are valid:

\[
[\psi(\mathbf{r}, t), \psi^\dagger(\mathbf{r}', t)]_\pm = \delta^{(3)}(\mathbf{r} - \mathbf{r}') ,
\]

(308)

all other commutators vanishing, as we saw before. One should add that the Schrödinger equation does not deal with spin, so we may, at this stage, chose commutators instead. Indeed, the same Schrödinger equation applies to all particles, fermions or not. But in chosing commutators, we would run into troubles, because we would not get Fermi-Dirac statistics. And since we will have to account for spin later, we might as well start with anticommutators.

- **Fock Space.**

In this framework, the states of a many-electron system belong to a complicated functional space built as follows. Let’s consider only free electrons, so that the index \( \mu \) is the 3D-wavevector \( \mathbf{k} \) and the functions \( u_\mu = u_\mathbf{k} \) are plane waves, and let’s start with the “vacuum” state \( |0\rangle \), that is, the ground-state of the system. This is the state in which each oscillator \( \mathbf{k} \) is in its ground state. Its energy diverges, because of the zero-point term \( 1/2 \) in Eq. (295) above. But we should think as subtracting this energy from the zero of our energy scale and carry on ignoring this problem. As we have just seen, the operator \( c_\mathbf{k} \) and its conjugate \( c_\mathbf{k}^\dagger \) can be viewed as lowering and raising operators, respectively. The latter can be thought as creating an electron in the eigenstate \( \mathbf{k} \) of the momentum operator when acting on the vacuum state. Thus, the wavefunction corresponding to this state will be simply:

\[
u_\mathbf{k}(\mathbf{r}) = \langle \mathbf{r}|c_\mathbf{k}^\dagger|0\rangle = \frac{1}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{r}}
\]

(309)

that is, just the corresponding one-particle plane-wave in the real-space representation. The operator \( c_\mathbf{k}^\dagger \) is called “creation” operator. Let’s call \( |\mathbf{k}\rangle = c_\mathbf{k}^\dagger|0\rangle \) the state in which there is only one electron in the system and this electron is in an eigenstate \( \mathbf{k} \). This is an eigenstate of the Hamiltonian with eigenvalue \( E_\mathbf{k} \). The conjugate operator \( c_\mathbf{k} \) destroys an electron in the state \( \mathbf{k} \). Thus \( c_\mathbf{k}|\mathbf{k}\rangle = |0\rangle \). The fact that we deal with fermions (and not bosons) translates into the property that the operators \( c_\mathbf{k} \) and \( c_{\mathbf{k}'}^\dagger \) satisfy the canonical anticommutation
rules:

\[ [c_k, c_{k'}^\dagger]^+ = \delta_{k,k'}, \quad [c_k, c_{k'}] = 0, \quad [c_{k'}^\dagger, c_k^\dagger] = 0. \]  

(310)

Note now that:

\[ c_k^\dagger c_k |0> = (1 - c_k c_k^\dagger) |0> = |0> - |0> = 0, \]

(311)

that is, the “zero” vector (not to be confused with the vacuum state!). Similarly:

\[ c_k^\dagger c_k |k> = c_k^\dagger |0> = |k>, \]  

(312)

so that the operator \( n_k = c_k^\dagger c_k \) is a sort of “number” operator, which counts the number of electrons in the state \( k \).

Note that if we try to add another \( k \)-electron to the state \( k \) we get:

\[ c_k^\dagger |k> = c_k^\dagger c_k^\dagger |0> = - c_k^\dagger c_k^\dagger |0>, \]

(313)

the last step relying on the CAR, so we must have \( c_k^\dagger |k> = 0 \) for this equation to be valid. Indeed, this is just Pauli’s principle: We can’t put two electrons in the same state. This follows directly from having used anti-commutation rules for the creation and annihilation operators.

Let’s also note the following:

\[ H |k> = \sum_{k'} E_{k'} c_{k'}^\dagger c_k^\dagger c_k |0> \]  

(314)

Let’s now use the anticommutation rule \( c_{k'} c_k^\dagger = \delta_{k,k'} - c_k^\dagger c_{k'} \):

\[ H |k> = \sum_{k'} E_{k'} c_k^\dagger (\delta_{k',k} - c_k c_{k'}) |0> \]  

(315)
Note that the second term inside the parentheses, \( c_{k'}^\dagger c_k^\dagger c_{k'} c_k |0\rangle \), vanishes, since the first operator on the right annihilates an electron \( k' \) in the vacuum state, which yields zero. Thus,

\[
H|k\rangle = E_k c_k^\dagger |0\rangle = E_k |k\rangle .
\]  

(316)

This shows that indeed the state \( |k\rangle = c_k^\dagger |0\rangle \) is an eigenstate of the Hamiltonian with eigenvalue \( E_k \).

From the discussion above it follows that an arbitrary state of our many-electron system can be written as:

\[
|\{N\lambda\}\rangle = |n_1 n_2 ... n_N\rangle = (c_1^\dagger)^{n_1} (c_2^\dagger)^{n_2} ... (c_N^\dagger)^{n_N} |0\rangle ,
\]

(317)

where the integers \( n_i \) are restricted to taking only the values 0 or 1 and the indices 1,2,...,\( N \) label wavevectors \( k_1, k_2, ... k_N \) (and also spin indices, if necessary). The notation \( |\{N\lambda\}\rangle \) identifies all possible combinations \( \lambda \) of \( N \)-electron states. Note that these states are already antisymmetric. For example, the wavefunction of a 2-electron state will be:

\[
< r', r | k', k > = < r', r | c_{k'}^\dagger c_k^\dagger |0\rangle = \frac{1}{\sqrt{2}} \frac{1}{(2\pi)^3} ( e^{ik' \cdot r'} e^{ik \cdot r} - e^{ik \cdot r'} e^{ik' \cdot r} ) .
\]

(318)

The general state \( |n_1 n_2 ... n_N\rangle \) is obviously an eigenstate of the Hamiltonian with eigenvalue \( n_1 E_1 + n_2 E_2 + ... \), as one can show following the derivation we have used above to reach Eq. (316).

So far we have considered a Schrödinger field, which does not account for spin. If we add it to the picture and label with \( \sigma = \pm 1/2 \) the spin variable, we have the new CAR:

\[
[c_k^\sigma, c_{k'}^{\sigma'}^\dagger]_+ = \delta_{\sigma,\sigma'} \delta_{k,k'} ,
\]

(319)

all other anticommutators vanishing. Everything carries as before, the only complication being related to the fact that for many-electron states the antisymmetrization of the wavefunctions must account for the quantum-mechanical rules of adding angular momenta. Thus, we will have 4 types of antisymmetrized 2-electron states:
The spin-singlet state, with a symmetric spatial component of the wavefunction, and three spin-triplet states, with an antisymmetric spatial component.

Starting from the “usual” single-particle Hilbert space $\mathcal{H}, L^2(\mathbb{R}^3) \otimes \mathcal{S}$, (where $\otimes$ means ‘tensor product of vector spaces and $\mathcal{S}$ is the 1/2-spin-space, which, using fancy group-theoretical terms, is a two-dimensional irreducible representation of the rotation group SU(2)), the total functional space associated with the many-electron system can be formally defined as:

$$\mathcal{H} \oplus \mathcal{H} \otimes \mathcal{H} \oplus \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} \ldots = \sum_{N=1}^{\infty} \oplus \prod_{j=1}^{N} \otimes \mathcal{H}, \quad (320)$$

which – after some proper topological closure and compactification, or some other detail like this, which only mathematicians could worry about – is called Fock space and I believe mathematicians are still struggling to understand how to define it “properly”.

Out of curiosity, it is interesting to note here a major difference between first and second quantization, difference which is the deep source of the infinite zero-point energy we saw above and of many other infinities which affect QFT: In ordinary QM (first quantization), there is only one Hilbert space – up to unitary transformations. That’s the reason Heisenberg’s matrix-mechanics and Schrödinger’s wave equation are actually the same thing, just viewed in two different, but equivalent, representations. Not so in many-body theory or QFT: For every Hamiltonian, there is only one representation which is “correct”. Unfortunately, Fock-space is the representation of free particles. As soon as we turn on any interactions, problems arise because we are in the wrong representation. Renormalization is a typical problem. This sad state of things is known as “Haag’s theorem” which is known to say that the interaction picture exists only if there is no interaction. “Constructive QFT” is the branch of theoretical physics which attempts to “construct” a proper representation for every interaction, thus bypassing renormalization problems.
As a second example, let’s consider the plasma oscillations which we have already encountered before (pages 116 and 117). This is a very interesting example which shows clearly how ‘elementary excitations’ emerge: We know how to solve the problem of electrons in a crystal when the system is in its ground state: The ions are ‘frozen’ in their equilibrium positions and the electrons are all in their lowest-energy state, filling all states up to the Fermi level. In tackling the problem of what happens when electrons acquire additional energy (via an external perturbation, typically) while interacting among themselves via long-range Coulomb interactions is quite a complicated task. So, we ‘linearize’ the problem by assuming that only first-order deviations from the ground-state occur. In retaining only the leading terms of the interaction Hamiltonian (in our case, the Coulomb potential energy) we retain only terms quadratic in the deviation from the ground state. By decompisong the system into normal modes (by taking Fourier transforms in the example below), the Hamiltonian will be the recast into the form of the sum of many Hamiltonians, each corresponding to a different wavelength, and each being the sum of a kinetic energy term (quadratic in the velocity or canonically conjugate momenta) and a potential energy term (quadratic in the displacent coordinates, which are the canonical coordinates). Thus, for each wavelength there will be a corresponding Hamiltonian of a single harmonic oscillator. We can now ‘quantize’ these oscillators using the Dirac algebraic method of raising (creation) and lowering (annihilation) operators, as we saw for the Schrödinger field. These are just the creation and annihilation operators for each wavelength. So, for each wavelength there may be no corresponding excitation at all (the oscillator is in its ground state), one excitation (the oscillator is in its first excited state), etc. Looking back at the Schrödinger field, there each excitation was interpreted as one electron ‘created ’ at a given wavelength. Here, in complete analogy, we shall call these excitations ‘quasi particles’. The suffix "on" attached to the name given to these excitations (plasmons, phonons, magnons, etc.) reflects this analogy.

So, let’s tackle the problem of the plasma exitations of a homogeneous, bulk electron gas. It will be instructive, at the end, to compare the resulting normalization of the plasmon field with the result of simpler semiclassical energetic arguments.
• Lagrangian Density.

The Lagrangian density of the electron gas can be written as:

$$\mathcal{L}(\dot{u}, \phi) = \frac{1}{2} n m \dot{u}(r) \cdot \dot{u}(r) - \frac{1}{2} \delta \rho(r) \phi(r) ,$$ (321)

where $en = \rho$ is the uniform charge density, $m$ the electron mass, $\phi(r)$ is the potential, and $u(r)$ is the displacement of the electron gas at position $r$. This displacement is related to the charge fluctuations $\delta \rho(r)$ via the continuity equation:

$$\frac{\delta \rho(r)}{\rho} = -\nabla \cdot u(r) .$$ (322)

Thus, Poisson equation can be written as:

$$\nabla^2 \phi(r) = \frac{\delta \rho(r)}{\epsilon_s} = -\frac{\rho}{\epsilon_s} \nabla \cdot u(r) ,$$ (323)

where $\epsilon_s$ is the dielectric constant of the solid.

• Going to Fourier transforms.

Let’s employ now Fourier decompositions of the fields:

$$\phi(r) = \sum_q \phi_q e^{i q \cdot r} ,$$ (324)

and

$$u(r) = \sum_q u_q e^{i q \cdot r} ,$$ (325)

Eq. (323) implies

$$-\sum_q q^2 \phi_q e^{i q \cdot r} = -i \frac{\rho}{\epsilon_s} \sum_q q \cdot u_q e^{i q \cdot r}$$ (326)
Since plasmons are longitudinal oscillations, \( \mathbf{q} \cdot \mathbf{u}_q = q u_q \), so that \( \phi_q = -i \rho u_q / (\epsilon_s q) \). Keeping in mind that the total energy must be real, we should replace the terms \( \mathbf{u}(\mathbf{r}) \cdot \mathbf{u}(\mathbf{r}) \) and \( \delta \rho(\mathbf{r}) \phi(\mathbf{r}) \) in the expression for the Lagrangian density Eq. (321) with \( \mathbf{u}(\mathbf{r})^* \cdot \mathbf{u}(\mathbf{r}) \) and \( \delta \rho(\mathbf{r})^* \phi(\mathbf{r}) \), respectively. Then, the kinetic energy can be written as:

\[
T = \frac{1}{2} \int d\mathbf{r} \; n \; m \; \mathbf{u}(\mathbf{r})^* \cdot \mathbf{u}(\mathbf{r}) = \frac{1}{2} n \; m \; \sum_q \mathbf{u}_q^* \cdot \mathbf{u}_q ,
\]

(327)

where the dots indicate derivatives with respect to time. The potential energy can be calculated ignoring second-order correction to the charge density, that is, by treating \( \rho(\mathbf{r}) = e n \) as a constant in Eq. (322), as we had done before when dealing with classical plasma oscillations. This is where the 'linearization' approximation comes into play. Then:

\[
V = \frac{1}{2} \int d\mathbf{r} \; \delta \rho(\mathbf{r})^* \phi(\mathbf{r}) = +\frac{1}{2} \frac{e^2 n^2}{\epsilon_s} \sum_q \mathbf{u}_q^* \cdot \mathbf{u}_q .
\]

(328)

Thus, from Eqns. (327) and (328), the Lagrangian for the system can be written as:

\[
L = T - V = \frac{1}{2} n \; m \; \sum_q \mathbf{u}_q^* \cdot \mathbf{u}_q - \frac{1}{2} \frac{e^2 n^2}{\epsilon_s} \sum_q \mathbf{u}_q^* \cdot \mathbf{u}_q .
\]

(329)

**The Hamiltonian.**

Let’s now define the Hamiltonian. First, we define the canonical conjugate momentum

\[
\pi_q = \frac{\partial L}{\partial \mathbf{u}_q} = \frac{\partial T}{\partial \mathbf{u}_q} = n \; m \; \mathbf{u}_q^* .
\]

(330)

Then, we apply the usual Legendre transform to get the Hamiltonian \( H = T + V \):

\[
H = \frac{1}{2} \sum_q \left\{ \frac{1}{nm} \pi_q^* \cdot \pi_q + \frac{e^2 n^2}{\epsilon_s} \mathbf{u}_q^* \cdot \mathbf{u}_q \right\} .
\]

(331)
- **Canonical Quantization.**
  
  Quantization can be now performed imposing the canonical commutation rules (CCR)
  
  \[
  [u_{i,q'}, \pi_{j,q}] = i\hbar \delta_{ij} \delta_{qq'} .
  \]
  \[ (332) \]
  
  Following a standard procedure, we rewrite the Hamiltonian as
  
  \[
  H = \frac{1}{2} \sum_q [ P_q^\dagger P_q + Q_q^\dagger Q_q ] ,
  \]
  \[ (333) \]
  
  having set:
  
  \[
  P_q = \frac{\pi_q}{(nm)^{1/2}} ,
  \]
  \[ (334) \]
  
  \[
  Q_q = \left( \frac{e^2n^2}{\epsilon_s} \right)^{1/2} u_q .
  \]
  \[ (335) \]
  
  Let’s now define, as usual, the creation and annihilation operators via the relations:
  
  \[
  Q_q = \left( \frac{\hbar \omega_P}{2} \right)^{1/2} (b_q + b_q^\dagger) ,
  \]
  \[ (336) \]
  
  \[
  P_q = i \left( \frac{\hbar \omega_P}{2} \right)^{1/2} (b_q - b_q^\dagger) ,
  \]
  \[ (337) \]
  
  where we have introduced the plasma frequency \( \omega_P = [(e^2n)/(\epsilon_s m)]^{1/2} \). Thus, finally, the Hamiltonian Eq. (331) becomes
  
  \[
  H = \frac{\hbar \omega_P}{2} \sum_q \left\{ b_q^\dagger b_q + b_q b_q^\dagger \right\} .
  \]
  \[ (338) \]
Using the definitions (334)-(337) and the CCR (332), the only nonvanishing commutator for the operators $b^\dagger_q$ and $b_q$ is

$$[b^\dagger_q, b_q] = -\frac{1}{\omega_P} \left( \frac{en}{\epsilon sm} \right)^{1/2} = -1.$$  

(339)

This shows that the operators $b^\dagger_q$ and $b_q$ behave like raising (creation) and lowering (annihilation) operators, so that the Hamiltonian Eq. (338) becomes a collection of harmonic oscillators

$$H = \hbar \omega_P \sum_q \left\{ b^\dagger_q b_q + \frac{1}{2} \right\} ,$$

(340)

while the displacement field and the potential can be expressed in terms of the creation and annihilation operators

$$u_q = \left( \frac{\hbar \omega_P \epsilon_s}{2e^2 n^2} \right)^{1/2} (b^\dagger_q + b_q) ,$$

(341)

$$\phi_q = -\frac{i}{q} \left( \frac{\hbar \omega_P}{2\epsilon_s} \right)^{1/2} (b^\dagger_q + b_q) .$$

(342)

As usual, the diverging contribution in Eq. (340) will be ignored and we shall write

$$H = \hbar \omega_P \sum_q b^\dagger_q b_q .$$

(343)

- **Fock space for bosons.**
  The Hilbert-like space associated with the system of bosons (as plasmons carry no spin) can be constructed essentially in the same way we saw above for the case of Schrödinger fermions. The obvious difference is that,
since we use commutation instead of anticommutation rules, we can put as many particles as we wish in each quantum state. Thus, the most general state of Fock space will have the form

\[ |\{N_\lambda\} > = |n_1n_2...n_N > = (b_1^\dagger)^{n_1}(b_2^\dagger)^{n_2}...(b_N^\dagger)^{n_N}|0> , \]  

where now states with \( n_i \neq 0,1 \) will also exist. Note that the squared norm of a state \(|n_q>\) will be \( n_q! \), so that states of the form (344) must be normalized diving by \( \sqrt{n_q!} \).

- **A calculation: The thermal expectation value of the plasmon field.**

Let’s assume that we are at thermal equilibrium. Since plasmons have zero spin, their population will obey the Bose-Einstein statistics and the number \( N_q \) of plasmons of momentum \( q \) present at temperature \( T \) will be given by:

\[ N_q = \left[ \exp \left( \frac{\hbar \omega_q}{k_BT} \right) - 1 \right]^{-1} . \]  

If we want to calculate the thermally-averaged rms amplitude of the field associated with the plasmons, we can first calculate the expectation value of the (squared) amplitude on the Fock state \(|n_q>=(n_q!)^{-1/2}(b_q^\dagger)^{n_q}|0>\) (where \(|0>\) is the vacuum) containing \( n_q \) plasmons of wave vector \( q \). This is:

\[ <n_q | \phi_q^\dagger \phi_q | n_q> = \frac{\hbar \omega_P}{2\epsilon_s} \frac{1}{q^2} <0|(b_q)^{n_q}(b_q+b_q^\dagger)(b_q^\dagger+b_q)(b_q^\dagger)^{n_q}|0> = \frac{\hbar \omega_P}{2\epsilon_s} \frac{1}{q^2} (1 + 2n_q) . \]  

We can then take the thermal average of this expression, so that:

\[ \langle \phi_q \rangle = \left( \frac{\hbar \omega_P}{2\epsilon_s} \frac{1}{q^2} \right)^{1/2} (1 + 2N_q)^{1/2} . \]  

Setting \( N_q = 0 \) above, it shows that plasmon zero-point motion at wave vector \( q \) is associated with a field of
rms amplitude

\[ |\phi_q| = \left(\frac{\hbar \omega_P}{2\epsilon_s}\right)^{1/2} \frac{1}{q}. \tag{348} \]

- **A semiclassical analogy.**

  Can we obtain this result without going through the procedure of canonical quantization? Stern and Ferrel have provided the following semiclassical argument, based on energetic considerations: Let us write the magnitude of the plasmon field at wave vector \( q \) as

  \[ E_{q,\omega}(r, t) = a_q \cos(q \cdot r - \omega pt), \tag{349} \]

  so that the electrostatic potential is

  \[ \phi_{q,\omega}(r, t) = -\frac{a_q}{q} \sin(q \cdot r - \omega pt). \tag{350} \]

  For harmonic oscillators such as plasmons, the time-averaged kinetic energy, \( <T> \), is equal to the time-averaged potential energy, \( <U> \). Thus, the total energy, \( W \), associated with the zero-point plasmon field over a volume \( \Omega \), *including* its own self-energy (which explains the absence of the factor 1/2 in front of the integral below), will be:

  \[ W = <T> + <U> = 2 <U> = 2 \left\langle \int_{\Omega} d\mathbf{r} \phi_{q,\omega}^*(\mathbf{r}) \rho_{q,\omega}(\mathbf{r}) \right\rangle = \epsilon_s |a_q|^2 \Omega, \tag{351} \]

  where \( \epsilon_s \) must be understood to be the high-frequency dielectric constant \( \epsilon^\infty \) in our case, since \( \rho_{q,\omega}(\mathbf{r}) \) is just a microscopic polarization charge, so that only the background (*i.e.*, valence) dielectric response should be considered. Quantum mechanically, this energy must represent the zero-point plasmon energy within the volume \( \Omega \), so that

  \[ W = \frac{1}{2} \hbar \omega_P \Omega, \]

  which implies

  \[ |\phi_q| = \frac{|a_q|}{q} = \left(\frac{\hbar \omega_P}{2\epsilon_s}\right)^{1/2} \frac{1}{q}. \tag{352} \]
in agreement with Eq. (348).
In the first part of the Notes we had a quick look at the ionic Hamiltonian (see pages 49 and 50), but we left it there, promising to come back and revisit this term. Here we shall do just that, as usual re-writing it in such a way that we will end up with a collection of harmonic oscillators.

- **Canonical Quantization**

Let $\mathbf{R}_{l\gamma}^{(0)}$ be the equilibrium position of the ion (labeled by the index $\gamma$ inside the unit cell $l$) and let $\delta\mathbf{R}_{l\gamma}$ its displacement from $\mathbf{R}_{l\gamma}^{(0)}$. Let also $\mathbf{R}_l$ be the coordinate-origin of the cell $l$. In terms of the momentum coordinate of ion $\gamma$ in cell $l$, $p_{l\gamma} = M_\gamma \delta \dot{\mathbf{R}}_{l\gamma}$, where $M_\gamma$ is the mass of the ion, the ionic Hamiltonian is:

$$H_{ion} = \frac{1}{2} \sum_{l\gamma} \left[ \frac{p_{l\gamma}^2}{M_\gamma} + \sum_{l' \neq l, \gamma' \gamma} \frac{\partial^2 V_{\gamma'}^{(lat)}(\mathbf{R}_{l\gamma}^{(0)} - \mathbf{R}_{l'\gamma'}^{(0)})}{\partial R_{l\gamma i} \partial R_{l'\gamma' j}} \delta R_{l\gamma i} \delta R_{l'\gamma' j} \right],$$  \hspace{1cm} (353)

where $V_{\gamma}^{(lat)}$ is the lattice potential of the ion $\gamma$. Note that here the indices $i$ and $j$ run over the spatial-coordinate indices $x, y, z$. Using Bloch theorem we can now expand $\delta\mathbf{R}_{l\gamma}$ in its spatial Fourier coefficients at wavevector $\mathbf{q}$ and assume a time dependence $e^{i\omega t}$:

$$\delta \mathbf{R}_{l\gamma} = \frac{1}{(NM)^{1/2}} \sum_{\mathbf{q}} \xi_{q\gamma} e^{i\mathbf{q} \cdot \mathbf{R}_l} e^{i\omega t},$$  \hspace{1cm} (354)

where $M$ is the mass of the unit cell and $N$ is the number of cells in the crystal. It should be noted that the periodicity of the lattice implies that the ‘polarization vector’ $\xi_{q\gamma}$ does not depend on the cell index $l$, the entire dependence on the cell being absorbed by the phase $e^{i\mathbf{q} \cdot \mathbf{R}_l}$. We also define the “lattice Fourier transform” of
the dynamic matrix as
\[ D_{ij\gamma\gamma'}(q) = \sum_{l''} \hat{D}_{ij\gamma\gamma'}(l'') e^{i q \cdot (R_l - R_{l''})}, \]  

where
\[ \hat{D}_{ij\gamma\gamma'}(l'') = -\frac{\partial^2 V^{(lat)}(R_{l\gamma}^{(0)} - R_{l'\gamma'}^{(0)})}{\partial R_{l\gamma i} \partial R_{l'\gamma' j}}, \]

where \( R_{l''} = R_l - R_{l''} \). Inserting the expansion (354) into the equations of motion determined by the Hamiltonian (353), for each \( q \) we obtain the homogeneous linear system
\[ \sum_{j\gamma\gamma'} \left[ \hat{D}_{ij\gamma\gamma'}(q) - \omega^2 \delta_{ij} \delta_{\gamma\gamma'} \right] \xi_{q\gamma'j} = 0, \]

where \( \hat{D}_{ij\gamma\gamma'}(q) = D_{ij\gamma\gamma'}(q)/(M_\gamma M_{\gamma'})^{1/2}. \) The associated secular equation determines the eigenfrequencies \( \omega_{q\eta} \) of oscillations labeled by a “branch index” \( \eta \) running over 3 (one LA and 2 TA) acoustic modes and over \( 3(N_{ions} - 1) \) optical modes, \( N_{ion} \) being the number of ions in the primitive cell. The figure on page 153 shows the dispersion for the 3 optical and 3 acoustic phonons in Si.

We can now write the general solution of the equations of motion as follows: Let’s normalize the polarization vectors \( \xi_{q\gamma} \) for each eigenmode \( \eta \) and call \( e_{q\gamma}^{(\eta)} \) the unit vectors so obtained. We can now express the general solution as a linear combination of these normalized eigenmodes, Eq. (354), with coefficients \( Q_{q\eta} \) which will play the role of canonical coordinates. Note that these coefficients are determined up to an overall multiplicative constant, which will be set by the CCR conditions below. We can view this as a simple rotation of the basis vectors. Thus, we can write for the “displacement field” \( \delta R_{l\gamma}^{(\eta)} \) for each mode \( \eta \):
\[ \delta R_{l\gamma}^{(\eta)} = \frac{1}{(NM)^{1/2}} \sum_{q\eta} e_{q\gamma}^{(\eta)} Q_{q\eta} e^{i q \cdot R_l}. \]
where $N$ is the number of cells in the normalization volume. We can now define a conjugate momentum

$$\mathbf{P}_{l\gamma}^{(\eta)} = \delta \mathbf{R}_{l\gamma}^{(\eta)} = \left( \frac{M}{N} \right)^{1/2} \sum_{q_\eta} e^{(\eta)}_{q_\gamma} P_{q_\eta} e^{i q \cdot \mathbf{R}_l} ,$$  \hspace{1cm} (359)

and quantize the system by promoting the canonical variables $P_{q_\eta}$ and $Q_{q_\eta}$ to operators on Fock space and imposing the CCR (as phonons do not carry spin)

$$[Q_{q_\eta}, P_{q'_{\eta'}}] = i\hbar \delta_{qq'} \delta_{\eta\eta'} , \quad [P_{q_\eta}, P_{q'_{\eta'}}] = 0 , \quad [Q_{q_\eta}, Q_{q'_{\eta'}}] = 0 .$$  \hspace{1cm} (360)

Then, using Eqns. (357)-(359), the ion Hamiltonian (353) becomes

$$H_{ion} = \frac{1}{2} \sum_{q_\eta} \left[ P_{q_\eta}^{\dagger} P_{q_\eta} + \omega_{q_\eta}^2 Q_{q_\eta}^{\dagger} Q_{q_\eta} \right] .$$  \hspace{1cm} (361)

Once more, this is the Hamiltonian of a collection of harmonic oscillators (phonons). Defining the annihilations and creation operators (going to the infinite-volume normalization):

$$Q_{q_\eta} = \left( \frac{\hbar}{2 \omega_{q_\eta}} \right)^{1/2} (b_{q_\eta} + b_{q_\eta}^{\dagger}) ,$$  \hspace{1cm} (362)

$$P_{q_\eta} = i \left( \frac{\hbar \omega_{q_\eta}}{2} \right)^{1/2} (b_{q_\eta} - b_{q_\eta}^{\dagger}) ,$$  \hspace{1cm} (363)

finally the phonon Hamiltonian can be rewritten in terms of the creation and annihilation operators as:

$$H_{ion} = \sum_{q_\eta} \hbar \omega_{q_\eta} b_{q_\eta}^{\dagger} b_{q_\eta}$$  \hspace{1cm} (364)

(ignoring as usual the diverging zero-point contribution) in analogy with Eqns. (293) and (343). Note that now, since phonons carry no spin, the creations and annihilation operators $b_{q_\eta}^{\dagger}$ and $b_{q_\eta}$ satisfy commutation rules like
Eq. (310) for the electron operators, but with commutators in place of anticommutators.
A simple example: The linear chain.

Let’s consider the simpler case of a 1D chain of atoms of mass \( M \). We can drop the indices \( i, j \) (since we are in 1D) as well as the indices \( \gamma \) (since we consider only one atom in each 1D cell). The ionic Hamiltonian becomes

\[
H_{ion} = \frac{1}{2} \sum_{l} \left[ \frac{P_{l}^{2}}{2M} + \sum_{l'} D_{l,l'} \delta R_{l} \delta R_{l'} \right].
\]

Note that:

\[
D_{l,l'} = D_{l-l',0}
\]

by translation symmetry and

\[
\sum_{l} D_{0,l} = 0 \quad \text{since for equal shifts } \delta R_{l} \text{ the crystal energy does not change}.
\]

Put it differently, since \( D_{0,l} \) is the force acting on the ion 0 due to ion \( l \), the equation above states that at equilibrium the total force due to all other ions acting on the ion 0 must vanish. The Hamilton equation for the ionic momentum will be:

\[
\dot{P}_{l} = -\frac{\partial H_{ion}}{\partial \delta R_{l}} = -\sum_{l'} D_{l,l'} \delta R_{l'},
\]

or

\[
M \frac{d^{2} \delta R_{l}}{dt^{2}} = -\sum_{l'} D_{l,l'} \delta R_{l'}.
\]

Let’s also make the simplifying assumption that each ion interacts only with its nearest neighbor (so that we can solve the problem ‘by hand’ and not numerically). Then:

\[
M \frac{d^{2} \delta R_{l}}{dt^{2}} = D_{l,l} \delta R_{l} + D_{l,l+1} \delta R_{l+1} + D_{l,l-1} \delta R_{l-1}.
\]

By translation symmetry \( D_{l,l} = D_{0,0} \). Without loss of generality we can consider the case \( l = 0 \). So we can
consider only the elements \( D_{0,0}, D_{0,\pm 1}, \) and \( D_{\pm 1,0} \). Using again the translation symmetry we have

\[
D_{0,1} = D_{-1,0},
\]

but since \( D_{ll'} = D_{l'l} \) (it’s only a switch in the order of the derivatives), we also have

\[
D_{0,1} = D_{1,0},
\]

so that \( D_{0,1} = D_{1,0} = D_{-1,0} = D_{0,-1} \).

By the property \( \sum_l D_{0,l} = 0 \) above we have

\[
D_{0,0} + D_{0,1} + D_{0,-1} = 0 \rightarrow D_{0,0} = -2 D_{1,0}.
\]

Let’s simply call \( D = D_{1,0} \) (the ‘spring’ constant of the inter-ionic force). Then the Hamilton equation of motion becomes simply:

\[
M \frac{d^2 \delta R_l}{dt^2} = D \left[ -2 \delta R_l + \delta R_{l+1} + \delta R_{l-1} \right].
\]

Now let’s set

\[
\delta R_l = e_q e^{iqal} e^{i\omega t},
\]

where \( a \) is the equilibrium distance between the ions in the chain (our ‘lattice constant’). Then:

\[
-M \omega^2 = D \left[ 2 \cos(qa) - 2 \right],
\]

becomes our trivial ‘secular equation’ from which we obtain the dispersion

\[
\omega = 2 \left| \sin \left( \frac{qa}{2} \right) \right| \left( \frac{D}{M} \right)^{1/2}.
\]

Note that the periodicity in \( 2\pi/a \) for \( q \) defines the BZ of the 1D chain.
• **Types of phonons.**

As noted, in a 3D cell having $N_{ions}$ ions, there will be 3 acoustic phonons (no relative motion between ions, the whole cell vibrates) and $3(N_{ions} - 3)$ optical phonons (characterized by relative displacement among ions in the same cell). The former are purely acoustic waves (like density waves) which effect electrons by modifying (linearly) the band structure with microscopic stress (they act like waves distorting the lattice constant as they travel). This effect is called ‘deformation potential’ interaction after Bardeen and Shockley. Optical phonons in polar materials (such as GaAs) carry an associated dipole field which scatters electrons in an obvious matter. Acoustic and optical phonons can be ‘transverse’ or ‘longitudinal’, depending on whether the displacement $\delta \mathbf{R}$ (also called ‘polarization’) is normal or parallel to the direction of travel $\mathbf{q}$. 

![Diagram of phonon types](image)
So far we have considered ‘excitations’ of the ions+electrons system. By this we mean that in dealing with phonons, for example, we have identified the ‘vacuum’ state $|0\rangle$ with the ground state of the ionic Hamiltonian (that is: ions at rest in their equilibrium positions). Excited states are realized by applying the creation operator to the ground state, thus obtaining states such as $|q\eta\rangle = b_{q\eta}^\dagger |0\rangle$ describing the excitation of a phonon of polarization $\eta$ and wavevector $q$. Similarly, plasmons are ‘collective’ excitations of the electronic Hamiltonian. And, in a way, even electronic states $|k\rangle = c_{k}^\dagger |0\rangle$ can be viewed as electrons excited from the valence band into the conduction band if we view the state $|0\rangle$ as the state in which all electrons populate the valence band (up to the Fermi level), which is the ground state of a semiconductor crystal at zero temperature. We could take as index $\mu$ the crystal momentum $k$ and a band index $n$, so that the functions $u_{kn}(r)$ on page 134 would be Bloch functions in band $n$, and $|kn\rangle = c_{kn}^\dagger |0\rangle$ would correspond to the excitation of an electron into the conduction band $n$.

In this section we temporarily abandon the environment of a crystal and consider instead the electromagnetic field in free space. The formalism will be the same, but now the ground state $|0\rangle$ will be electromagnetic vacuum. Considering the quantization of the electromagnetic field in vacuum will allow us to introduce ‘photons’ as particles (treated on the same footing as electrons before) and will allow us to later calculate the Einstein’s coefficients appearing in the matrix elements of radiative processes. Indeed, this is the first and foremost example of a field requiring ‘second quantization’ in order to reconcile the particle/wave nature of photons.

**Maxwell’s equations.**

Let’s re-write Maxwell’s equation in SI units:

\[
\nabla \cdot \mathbf{D} = \rho , \tag{365}
\]
\[
\nabla \cdot \mathbf{B} = 0 , \tag{366}
\]
\[ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (367) \]
\[ \nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \quad (368) \]

In vacuum and in the absence of external sources \((\mathbf{J} = 0, \rho = 0)\), recalling that \(\mathbf{D} = \varepsilon \mathbf{E}, \mathbf{B} = \mu \mathbf{H}\), and that \(\varepsilon \mu = 1/c^2\), we have:

\[ \nabla \cdot \mathbf{E} = 0, \quad (369) \]
\[ \nabla \cdot \mathbf{B} = 0, \quad (370) \]
\[ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (371) \]
\[ \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (372) \]

- **Lagrangian density.**

Recalling the expression for the energy of the electromagnetic field, we could immediately write down the Hamiltonian, express it in terms of Fourier components of the fields or potentials (that is, electromagnetic plane waves) and 'quantize' the result. However, let's follow the canonical procedure, starting from the definition of the Lagrangian density:

\[ \mathcal{L} = \frac{\varepsilon}{2} \mathbf{E}^2 - \frac{1}{2\mu} \mathbf{B}^2 = \frac{\varepsilon}{2} (\dot{\mathbf{A}} + \nabla \phi)^2 - \frac{1}{2\mu} (\nabla \times \mathbf{A})^2, \quad (373) \]

having introduced the usual scalar and vector potentials \(\phi\) and \(\mathbf{A}\) such that

\[ \mathbf{E} = -\dot{\mathbf{A}} - \nabla \phi, \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (374) \]

Note that Eqns. (370) and (371) are automatically satisfied thanks to the definitions (374). The remaining Maxwell's equations, Eqns. (369) and (372), must be derived from the Euler-Lagrange equations. The first one,

\[ \frac{\partial}{\partial t} \left( \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) + \sum_j \frac{\partial}{\partial x_j} \left( \frac{\partial \mathcal{L}}{\partial (\partial \phi/\partial x_j)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \quad (375) \]
\[ \nabla^2 \phi = 0, \quad (376) \]

which is Eq. (369). The second Lagrange equation, which we can write in a somewhat ‘sloppy’ vector notation as
\[
\frac{\partial}{\partial t} \left( \frac{\partial L}{\partial \dot{A}} \right) + \sum_j \frac{\partial}{\partial x_j} \left( \frac{\partial L}{\partial (A/\partial x_j)} \right) - \frac{\partial L}{\partial A} = 0, \quad (377)
\]
results in the equation:
\[
\epsilon \frac{\partial}{\partial t} (\dot{A} + \nabla \phi) - \frac{1}{\mu} \nabla \times \nabla \times A = 0, \quad (378)
\]
which is Eq. (372).

- **Hamiltonian density.**
  Let’s now define the canonical conjugate momentum density:
\[
P = \frac{\partial L}{\partial \dot{A}} = \epsilon (\dot{A} + \nabla \phi), \quad (379)
\]
so that the Hamiltonian density will be:
\[
\mathcal{H} = P \cdot \dot{A} - L = \frac{\epsilon}{2} (\dot{A} + \nabla \phi)^2 + \frac{1}{2\mu} (\nabla \times A)^2 - \epsilon (\dot{A} + \nabla \phi) \cdot \nabla \phi = \frac{1}{2} \left( \frac{P^2}{\epsilon} + \frac{(\nabla \times A)^2}{\mu} \right) - \epsilon (\dot{A} + \nabla \phi) \cdot \nabla \phi \quad (380)
\]
When integrated over the entire space – in order to get the Hamiltonian – the last term in the expression above vanishes. Indeed integrating it by parts:
\[
\epsilon \int (\dot{A} + \nabla \phi) \cdot \nabla \phi \, dr = \int P \cdot \nabla \phi \, dr = - \int \phi \nabla \cdot P \, dr = 0, \quad (381)
\]
by Green’s (or ‘divergence’ or ‘Gauss’) theorem, assuming that P vanishes at infinity or satisfies periodic
boundary conditions at the edge of the normalization volume. Therefore:

\[ H = \int \mathcal{H} \, dr = \frac{1}{2} \int \left[ \frac{P^2}{\epsilon} + \frac{(\nabla \times A)^2}{\mu} \right] \, dr . \]  

(382)

Let’s now represent the potentials as plane-waves of polarization \( \lambda \) and wavevector \( k \):

\[ A(r, t) = \frac{1}{V^{1/2}} \sum_{k\lambda} q_{k\lambda}(t) \, e_{k\lambda} \, e^{ik \cdot r} + \text{cc} , \]  

(383)

and

\[ P(r, t) = \frac{1}{V^{1/2}} \sum_{k\lambda} p_{k\lambda}(t) \, e_{k\lambda} \, e^{ik \cdot r} + \text{cc} , \]  

(384)

where \( V \) is the normalization volume and \( e_{k\lambda} \) is the polarization vector. Then, using the orthonormalization of the polarization vectors, the Hamiltonian can be written in terms of the ‘Fourier coefficients’ \( q_{k\lambda} \) and \( p_{k\lambda} \) as follows:

\[ H = \sum_{k\lambda} \left( \frac{1}{\epsilon} p_{k\lambda}^* p_{k\lambda} + \frac{k^2}{\mu} q_{k\lambda}^* q_{k\lambda} \right) . \]  

(385)

Equation (385) – as one could have seen also earlier in Eq. (382) – is taking the desired shape of the Hamiltonian of a collection of harmonic oscillators. Let’s now recast it in the usual form involving creation and annihilation operators.

- **Equations of motion.**
  In order to do that, let’s see what type of time-dependence the operators \( q_{k\lambda} \) and \( p_{k\lambda} \) will exhibit. The equations of motion for them can be derived by noticing that, after canonical quantization, they obey the CCR:

\[ [q_{k\lambda}, p_{k'\lambda'}^\dagger] = [q_{k\lambda}^\dagger, p_{k'\lambda'}] = i\hbar \, \delta_{k'k} \, \delta_{\lambda'\lambda} . \]  

(386)
Then:

\[
\dot{q}_{k\lambda} = \frac{1}{i\hbar} [q_{k\lambda}, H] = \frac{1}{i\hbar} \sum_{k'\lambda'} \frac{1}{\epsilon} [q_{k\lambda}, p_{k'\lambda'}^\dagger] \ p_{k'\lambda'} = \frac{1}{\epsilon} \sum_{k'\lambda'} \delta_{k'k} \delta_{\lambda'\lambda} \ p_{k'\lambda'} = \frac{p_{k\lambda}}{\epsilon} .
\] (387)

Similarly,

\[
\dot{p}_{k\lambda} = \frac{1}{i\hbar} [p_{k\lambda}, H] = -\frac{k^2}{\mu} q_{k\lambda} .
\] (388)

Taking the time derivative of Eq. (387) and using Eq. (388):

\[
\ddot{q}_{k\lambda} = \frac{\dot{p}_{k\lambda}}{\epsilon} = -\frac{k^2}{\epsilon\mu} q_{k\lambda} ,
\] (389)

which has the general solution:

\[
q_{k\lambda}(t) = A_{k\lambda} e^{ikc t} + A_{k\lambda}^\dagger e^{-ikc t} .
\] (390)

Inserting this result into Eq. (387) we also get:

\[
p_{k\lambda}(t) = ic\epsilon k A_{k\lambda} e^{ikc t} - ic\epsilon k A_{k\lambda}^\dagger e^{-ikc t} .
\] (391)

The ‘coefficients’ \( A_{k\lambda} \) and \( A_{k\lambda}^\dagger \) have all of the properties of annihilation and creation operators except for an overall normalization. Indeed, from the CCRs (386) and from the definitions (390) and (391) we have:

\[
[A_{k\lambda}, A_{k'\lambda'}^\dagger] = -\frac{\hbar}{2\epsilon c\epsilon} \delta_{k'k} \delta_{\lambda'\lambda'} .
\] (392)

So, let’s define instead:

\[
q_{k\lambda}(t) = \left( \frac{\hbar}{2k\epsilon c} \right)^{1/2} [a_{k\lambda} e^{ikc t} + a_{k\lambda}^\dagger e^{-ikc t}] ,
\] (393)
\[ p_{k\lambda}(t) = ik \left( \frac{\epsilon}{\mu} \right)^{1/2} \left( \frac{\hbar}{2kce} \right)^{1/2} \left[ a_{k\lambda} e^{ikct} - a_{k\lambda}^\dagger e^{-ikct} \right], \]  

(394)

or, inverting them:

\[ a_{k\lambda} = e^{-ikct} \left( \frac{kce}{2\hbar} \right)^{1/2} \left[ q_{k\lambda} - \frac{i}{k} \left( \frac{\mu}{\epsilon} \right)^{1/2} p_{k\lambda} \right], \]  

(395)

and

\[ a_{k\lambda}^\dagger = e^{ikct} \left( \frac{kce}{2\hbar} \right)^{1/2} \left[ q_{k\lambda} + \frac{i}{k} \left( \frac{\mu}{\epsilon} \right)^{1/2} p_{k\lambda} \right], \]  

(396)

Then, the CCRs become:

\[ [a_{k\lambda}, a_{k'\lambda'}^\dagger] = \delta_{k'k} \delta_{\lambda\lambda'}, \]  

(397)

and the Hamiltonian (385) becomes:

\[ H = \sum_{k\lambda} \hbar ck \left[ a_{k\lambda}^\dagger a_{k\lambda} + \frac{1}{2} \right] \rightarrow \sum_{k\lambda} \hbar \omega_k a_{k\lambda}^\dagger a_{k\lambda}, \]  

(398)

having ignored the 'usual' diverging zero-point term and having used the fact that \( kc = \omega_k \) is the energy of a photon of frequency \( \nu = \omega_k/(2\pi) \).

- **The quantum field and the importance of second quantization.**

From Eq. (393) we see how the 'position' operator \( q_{k\lambda} \) can be expressed in terms of the single-photon operators \( a_{k\lambda} \) and \( a_{k\lambda}^\dagger \). Similarly, Eq. (394) relates the 'momentum' operator \( p_{k\lambda} \) in terms of the same quantities. In turn, Eqns. (383) and (384) tell us how to express the vector and scalar (via the momentum operator \( \Phi_{k\lambda} \)) potentials in terms of the operators \( q_{k\lambda} \) and \( p_{k\lambda} \). The net result is that we can express the vector and scalar potentials in terms of \( a_{k\lambda} \) and \( a_{k\lambda}^\dagger \). For example:

\[ A(r, t) = \frac{1}{V} \sum_{k\lambda} \left( \frac{\hbar}{2kce} \right)^{1/2} \left[ a_{k\lambda} e^{ikct} + a_{k\lambda}^\dagger \right] e_{k\lambda} e^{i(k \cdot r - kct)}. \]  

(399)
From this – and the corresponding equation for $\dot{A} + \nabla \phi$ – we can derive the amplitude of the electric and magnetic fields:

$$B(r,t) = \nabla \times A(r,t) = \frac{i}{V} \sum_{k\lambda} \left( \frac{\hbar}{2kc\epsilon} \right)^{1/2} \left[ a_{k\lambda} e^{ikct} + a^\dagger_{k\lambda} e^{-ikct} \right] (k \times e_{k\lambda}) e^{ik\cdot r}, \quad (400)$$

and

$$E(r,t) = \frac{P(r,t)}{\epsilon} = \frac{i}{V} \sum_{k\lambda} \frac{k}{\epsilon} \left( \frac{\hbar}{2kc\mu} \right)^{1/2} \left[ a_{k\lambda} e^{ikct} - a^\dagger_{k\lambda} e^{-ikct} \right] e_{k\lambda} e^{ik\cdot r} . \quad (401)$$

**This is the basic result of second quantization:** It tells us the amplitude of the field associated to a single quantum of excitation. Without second quantization, how could we know the strength of the electric field associated to a photon? Similar relations apply to electrons: Eq. (342) gives us the amplitude of the (scalar) potential associated to a single quantum of plasma excitations, Eq. (362) together with Eq. (358) allows us to calculate the displacement of each ion in terms of the phonon creation and annihilation operators:

$$\delta R_{l\gamma}^{(\eta)} = \frac{1}{(NM_\gamma)^{1/2}} \sum_{q\eta} e_{q\gamma}^{(\eta)} \left( \frac{\hbar}{2\omega_{q\eta}} \right)^{1/2} (b_{q\eta} + b_{q\eta}^\dagger) e^{iq\cdot R_l} . \quad (402)$$

The algebraic rules allowing us to evaluate easily expectation values of the creation and annihilation operators on the ground state $|0>$ will ultimately allow us to calculate the strength of interactions between electrons and these excitations (phonons, plasmons, photons), result which we could not have accomplished without the information, given us by second quantization, regarding the strength of the field carried by a single excitation.
In this section we discuss the response of the semiconductor to a perturbing potential. The practical reason which forces us to discuss this subject is, of course, our need to handle the problem of carriers which scatter with the external potential. Unfortunately, any external perturbation causes not only the ‘single-particle’ scattering process we are interested in, but also a ‘many-body’ rearrangement of all electrons, valence and conduction, as well as of the ions (especially if the crystal is ionic). Thus, we must consider both the electronic and the ionic dielectric properties. This rearrangement modifies the scattering potential, so that the scattering event depends on these many-body effects as well.

In introductory textbooks of electromagnetic theory, the effect is ‘lumped’ into a dielectric constant of the medium: If \( \mathbf{E} \) is the ‘external’ electric field applied to a dielectric medium, this field will induce a redistribution of the charges in the medium which will give rise to an additional polarization field \( \mathbf{P} \). Assuming a linear relation between the perturbation \( \mathbf{E} \) and the ‘response’ \( \mathbf{P} \) (a typical example of what is indeed known as ‘linear response theory’), we write:

\[
\mathbf{P} = \chi \mathbf{E},
\]

where \( \chi \) (the Greek letter ‘chi’) is called ‘dielectric susceptibility’. Therefore, the total electric displacement field \( \mathbf{D} \) resulting from both the applied field \( \mathbf{E} \) and the response \( \mathbf{P} \) of the medium will be:

\[
\mathbf{D} = \epsilon_0 (\mathbf{E} + \mathbf{P}) = \epsilon_0 (\mathbf{E} + \chi \mathbf{E}) = \epsilon_0 (1 + \chi) \mathbf{E} = \epsilon \mathbf{E},
\]

where \( \epsilon \), the dielectric constant, absorbs the effect of the polarization charges.

Such a ‘macroscopic averaging’ may be too much of a simplification in our case. In general, the response of the system depends on both the wavelength and the frequency of the external perturbation. In addition, semiconductors are not isotropic media. Thus a single dielectric ‘constant’ is actually a frequency and wavevector dependent matrix. We shall derive it in a few cases assuming for the spatial and temporal dependence of the perturbation a simple harmonic form of the type \( e^{i\mathbf{q} \cdot \mathbf{r}} e^{i\omega t} \). Assuming a ‘linear response’, we can always Fourier-decompose any external perturbation function of space and time, analyze each Fourier component, one at the time, derive the response of the system for a particular value of \( \mathbf{q} \) and \( \omega \), and add the response at all wavelengths and frequencies.
Let’s consider first the response of the ions, which we shall treat in rather simple terms.

- **Ionic Response.**
  Let’s consider first the response of the ions, which we shall treat in rather simple terms.

- **Optical phonons in a chain.**
  Let’s go back to our example of a linear chain of ions (page 154) and modify the problem slightly, so that we now have two ions per ‘cell’ as illustrated in the figure below. The ions are connected by springs with elastic constant $D$, as before. The ions have alternating masses, $M$ and $m$, as shown, with $M > m$. The equilibrium distance between two adjacent atoms is $a$.

  The equations of motion can be derived by writing two equations, one for the displacement $u_1$ of the atoms of mass $M$ and another one for the displacement $u_2$ of the atoms of mass $m$. So, denoting by $u_{1,j}$ the displacement of the ion of mass $M$ at location $ja$ and by $u_{2,j}$ the displacement of the ion of mass $m$ at...
location \((j + 1)a\), then the equations of motion are:

\[
M \frac{d^2 u_{1,j}}{dt^2} = D(u_{2,j} - u_{1,j}) + D(u_{2,j-1} - u_{1,j}) .
\]  

(405)

\[
m \frac{d^2 u_{2,j}}{dt^2} = D(u_{1,j+1} - u_{2,j}) + D(u_{1,j} - u_{2,j}) .
\]  

(406)

Let's define:

\[
u_{1,j} = v_1 e^{i(2ja - \omega t)} ,
\]

(407)

\[
u_{2,j} = v_2 e^{i(2ja - \omega t)} .
\]

(408)

Then the equations of motion become:

\[-M \omega^2 v_1 = Dv_2 - Dv_1 + Dv_2 e^{-2ia} = Dv_1 ,
\]

(409)

\[-m \omega^2 v_2 = Dv_1 e^{2ia} - Dv_2 + Dv_1 - Dv_2 ,
\]

(410)

or:

\[(M \omega^2 - 2D)v_1 + D(1 + e^{-2ia})v_2 = 0 ,
\]

(411)

\[D(1 + e^{2ia})v_1 + (m \omega^2 - 2D)v_2 = 0 .
\]

(412)

This \(2 \times 2\) linear homogeneous system (which is exactly the system given by Eq. (357) in the more general case considered above) admits nontrivial solutions only if the determinant of the coefficients vanish, that is:

\[(M \omega^2 - 2D)(m \omega^2 - 2D) - D^2(1 + e^{-2ia})(1 + e^{2ia}) = 0 .
\]

(413)

The two solutions are:

\[
\omega^2_{\pm} = \frac{D}{mM} \left\{ (m + M) \pm \left[ (m + M)^2 - 4mM \sin^2(qa) \right]^{1/2} \right\} .
\]

(414)
The choice of the ‘plus’ sign leads to an ‘optical’ branch:

$$\omega^2_+ \rightarrow \begin{cases} 
2D \left( \frac{1}{m} + \frac{1}{M} \right) & \text{for } q \rightarrow 0 \\
\frac{2D}{m} & \text{for } q \rightarrow \frac{\pi}{2a} .
\end{cases} \quad (415)$$

The choice of the ‘minus’ sign yields an acoustic mode:

$$\omega^2_- \rightarrow \begin{cases} 
\frac{D}{M+m} a^2 q^2 & \text{for } q \rightarrow 0 \\
\frac{2D}{M} & \text{for } q \rightarrow \frac{\pi}{2a} .
\end{cases} \quad (416)$$

The optical solution corresponds to a mode in which the two different ions oscillate out of phase, the acoustic solution to a mode in which the different ions oscillate in phase.

**Response in the infrared.**

Let’s go back to Eqns. (409) and (410). Let’s consider the long-wavelength (infrared) limit $qa \rightarrow 0$ and add the perturbation of an electric field $E e^{-i\omega t}$: Assuming a positive charge $e$ for the ion with mass $M$ and negative charge $e$ for the ion with mass $m$, we get:

$$-M \omega^2 v_1 = 2D(v_2 - v_1) + eE , \quad (417)$$

$$-m \omega^2 v_2 = 2D(v_1 - v_2) - eE . \quad (418)$$

Let’s divide the first equation by $M$, the second one by $m$, and subtract the second from the first. We then get:

$$v_2 - v_1 = \frac{eE/\mu}{(2D/\mu) - \omega^2} , \quad (419)$$

where $\mu$ is the ‘reduced mass’ $1/(m^{-1} + M^{-1})$. From Eq. (415) we recognize that $2D/\mu = \omega^2_{TO}$, the squared frequency of the optical mode in the bi-atomic linear chain considered above.

Now, the application of the external field $E$ has caused a polarization of the pair, the negative ion moving in
the opposite direction of the positive ion. Thus, this polarization results in an additional electric field

$$P_{ion} = \frac{1}{\epsilon(\infty)} eN(v_2 - v_1) = \frac{1}{\epsilon(\infty)} \frac{e^2 N/\mu}{\omega_{TO}^2 - \omega^2} E,$$

(420)

where \(N\) is the number of ion pairs per unit volume and \(\epsilon(\infty)\) is the dielectric constant in absence of the ionic response (so, at a frequency large enough so that the ions cannot respond). Note the resonant behavior at the frequency \(\omega_{TO}\). In general, this resonant behavior of Eq. (420) translates into a frequency dependence of the form:

$$\epsilon_{ion}(\omega) = \epsilon(\infty) + [\epsilon(0) - \epsilon(\infty)] \frac{\omega_{TO}^2}{\omega_{TO}^2 - \omega^2},$$

(421)

where \(\omega_{TO}\) in the case of a 3D crystal (more general than our simple 1D chain) is the frequency of the transverse optical modes at long wavelengths. What Eq. (421) tells us is that the dielectric ‘constant’ approaches a smaller value \(\epsilon(\infty)\) at frequencies so large that the ions do not respond to the external perturbation. Thus \(\epsilon(\infty) = \epsilon_s\), the ‘static’ electronic dielectric constant of the semiconductor (see below, Eq. (441)). As the frequency decreases, the ions begin to respond and, after having gone through the resonance at \(\omega_{TO}\), their response results in a larger dielectric ‘constant’ \(\epsilon(0)\). The factor \(\epsilon(0) - \epsilon(\infty) = e^2 N/(\mu \omega_{TO}^2)\) represents the contribution of ions to the dielectric response and it is also known as the ‘oscillator strength’. Note that Eq. (421) can also be written as

$$\epsilon_{ion}(\omega) = \epsilon(\infty) \frac{\omega_{LO}^2 - \omega^2}{\omega_{TO}^2 - \omega^2},$$

(422)

where \(\omega_{LO}\) is such that \(\epsilon_{ion}(\omega_{LO}) = 0\). This shows that the dielectric constant is negative for \(\omega_{TO} < \omega < \omega_{LO}\), which means that electromagnetic waves do not propagate through the crystal (they will be damped, as the polarization of the crystal is in the opposite direction of the external field, thus cancelling it). To see this mathematically, recall that in a dielectric medium the relation between wavevector
and frequency of an electromagnetic wave is:

\[ k^2 = \frac{\epsilon(\omega)}{\epsilon_0 c^2} \omega^2. \]  (423)

Whenever \( \epsilon < 0 \), \( k \) will acquire an imaginary part which tells us that the wave becomes damped.

We have yet to understand the physical meaning of \( \omega_{LO} \). Since

\[ \nabla \cdot \mathbf{D} = \epsilon(\omega) \nabla \cdot \mathbf{E} = 0, \]  (424)

assuming \( \mathbf{E} = \mathbf{E}_0 e^{i k \cdot \mathbf{r}} \), for transverse modes \( k \cdot \mathbf{E}_0 = 0 \), so that Eq. (424) is trivially satisfied when \( \epsilon(\omega) \neq 0 \). However, when we deal with longitudinal modes, we satisfy Eq. (424) with \( \epsilon(\omega) = 0 \). We see that longitudinal modes correspond to the zeros of the dielectric function, transverse modes correspond to its poles. We also see from Eq. (421) and the property \( \epsilon(\omega_{LO}) = 0 \) that

\[ \omega_{LO} = \left[ \frac{\epsilon(0)}{\epsilon(\infty)} \right]^{1/2} \omega_{TO}. \]  (425)

This expression is known as the Lyddane-Sachs-Teller relation.

Note, finally, that since \( \mathbf{D}(\omega) = \epsilon(\omega) \mathbf{E}(\omega) \), and since in the absence of external charges \( \mathbf{D}(\omega) = 0 \), transverse modes must have a vanishing microscopic electric field \( \mathbf{E}(\omega) \). Instead, longitudinal modes can have \( \mathbf{E}(\omega) \neq 0 \) and still have \( \mathbf{D}(\omega) = 0 \), since \( \epsilon(\omega_{LO}) = 0 \). Thus, electrons will interact via polar interactions only with longitudinal modes.

- **Electronic Response.**

  Let’s now consider the response of the electrons under an external perturbing potential. We shall not discuss the problem in its full generality: Even for the case of a homogeneous and isotropic electron gas the problem has a long and complicated history. Instead, we shall rely on two main simplifications.
1. First, we shall restrict our discussion to the so-called longitudinal response. We shall assume that our medium is isotropic (in some case thanks to a suitable transformation, such as the Herring-Vogt transformation of k-space coordinates usually employed to ‘isotropize’ the Si equienergy ellipsoids), or we shall only consider the longitudinal polarization field, that is, the microscopic field (generated by the redistribution of the charges under the external field) along the direction of the external field itself. Transverse components, such as those which may arise from transverse vector potentials in dynamic situations, will be ignored. A notable example of a situation where ‘local transverse fields’ may be important is discussed when dealing with a microscopic theory of the electron-phonon interaction.

2. The second approximation we shall embrace consists in assuming a ‘quasi-equilibrium’ carrier distribution. This approximation, dictated exclusively by practical computing difficulties, will be discussed below.

- **Linear screening in the Random Phase Approximation.**
  Let’s consider a ‘free’ semiconductor, i.e. a semiconductor described by the Schrödinger equation (141). Let’s denote by \( H_0 \) the free Hamiltonian at the left-hand side of Eq. (88), by \( \psi_{\mathbf{k},\mu}(\mathbf{r}) = \langle \mathbf{r} | \mathbf{k}, \mu \rangle \) the (Bloch) plane waves which are the eigenfunctions of the lattice Hamiltonian \( H_0 \) with eigenvalues \( E_\mu(\mathbf{k}) \) for each band \( \mu \).

  If we apply an external perturbation described by the potential \( \varphi^{(ex)}(\mathbf{r}, t) \), the charges in the semiconductor will rearrange themselves, causing an additional ‘polarization’ (or ‘screening’) potential \( \varphi^{(pol)}(\mathbf{r}, t) \) which modifies the original external potential. Thus, the ‘actual’ potential in the material will be the sum of the external perturbation and of the polarization potential,

\[
\varphi(\mathbf{r}, t) = \varphi^{(ex)}(\mathbf{r}, t) + \varphi^{(pol)}(\mathbf{r}, t).
\]

We are interested in describing the total potential \( \varphi \), since this is the effective perturbation acting on the material. We may think of \( \varphi^{(ex)} \) as the ‘bare’ potential of a positively charged donor impurity in n-type Si, described by the \( 1/r \) Coulomb potential, while \( \varphi \) will be the screened potential felt by the conduction electrons, resulting from single positive charge of the impurity ion and from the negative cloud of free conduction electrons attracted by the impurity. Two approximations are very commonly made to
reduce the complexity of the problem: First, the Fourier components of the various potentials are considered independently. Cross-terms mixing different ‘wavelengths’ are ignored, on the grounds that their phases will vary wildly, resulting in their cancellation. This is the so-called Random Phase Approximation (RPA). An equivalent way of stating this approximation is to say that each electron will respond to the ‘average’ potential caused by the other electrons. We thus transform this many-body problem to a single-particle problem, thanks to this ‘mean-field’ approximation. Secondly, we assume that both the external and the polarization potentials are weak perturbations of the ‘free’ Hamiltonian $H_0$. This is the so-called ‘linear screening approximation’ which allows us to express the polarization charge in simple first order perturbation theory.

Thanks to the RPA, we can state the problem more simply: Let’s expand the external perturbation in plane waves:

$$\varphi^{(ex)}(r, t) = \sum_q \varphi_q^{(ex)} e^{-i\mathbf{q}\cdot\mathbf{r} - i\omega t}, \quad (427)$$

having assumed a simple harmonic time dependence. Then, we are interested in finding the ‘dielectric’ function $\epsilon^{(r)}(\mathbf{q}, \omega) = \epsilon(\mathbf{q}, \omega)/\epsilon_0$ where $\epsilon_0$ is the vacuum permittivity, such that

$$\varphi_q^{(ex)} = \varphi_q - \varphi_q^{(pol)} = \epsilon^{(r)}(\mathbf{q}, \omega) \varphi_q, \quad (428)$$

or $\varphi_q = \varphi_q^{(ex)}/\epsilon(\mathbf{q}, \omega)$ and so $\epsilon(\mathbf{q}, \omega)/\epsilon_0 = 1 - \varphi_q^{(pol)}/\varphi_q$. The usefulness of this expression lies in the fact that, once we know the dielectric function $\epsilon(\mathbf{q}, \omega)$, the knowledge of the simple external potential is sufficient to determine the net response of the system, without having to re-evaluate the self-consistent redistribution of the charges internally to the system. We shall proceed as follows: First, we shall employ the approximation of linear response to determine the ‘new’ wavefunctions of the system under the perturbation of the net potential $\varphi$, assuming that it is known. From the perturbed wavefunctions we shall derive the ‘polarization’ charge,

$$\rho^{(pol)}(r, t) = \sum_q \delta \rho_q e^{-i\mathbf{q}\cdot\mathbf{r} - i\omega t}, \quad (429)$$

and from this the polarization potential. Summing the external and polarization potentials, we shall solve for the net potential self-consistently.
Let’s now express the wavefunctions, \( \phi_{k,\mu}(r, t) \), of the perturbed system (lattice Hamiltonian plus perturbing potential) in terms of the unperturbed wavefunctions \( \psi_{k,\mu}(r, t) = \langle r, t | k, \mu \rangle \), using first order perturbation theory (see pages 17, 18):

\[
\phi_{k,\mu}(r, t) = \psi_{k,\mu}(r, t) + \delta \psi_{k,\mu}(r, t) = \psi_{k,\mu}(r, t) + \sum_{k',\mu'} \frac{e^{\langle k', \mu' | \varphi | k, \mu \rangle}}{E_{\mu}(k) - E_{\mu'}(k') + \hbar \omega + i \hbar s} \psi_{k',\mu'}(r, t).
\]

(430)

and

\[
\phi^*_{k,\mu}(r, t) = \psi^*_{k,\mu}(r, t) + \delta \psi^*_{k,\mu}(r, t) = \psi^*_{k,\mu}(r, t) + \sum_{k',\mu'} \frac{e^{\langle k, \mu | \varphi | k', \mu' \rangle}}{E_{\mu}(k) - E_{\mu'}(k') - \hbar \omega - i \hbar s} \psi^*_{k',\mu'}(r, t).
\]

(431)

These expressions are likely to present a surprise, as they contains a ‘strange’ imaginary term \( \pm i \hbar s \) in the denominator. Its origin lies in the fact that when we have a continuous spectrum, we must ‘regularize’ the behavior of the coefficient appearing in the equation above when \( k' \rightarrow k \). The standard ‘trick’ is to add a small imaginary part to the denominator, do the calculations, and at the end take the limit \( s \rightarrow 0 \). This is also known as taking the ‘principal part’ of the integral. We shall discuss it at length below, understanding what it means physically. For the time being, we may ignore it.

From Eq. (427) and some manipulations with Bloch waves, we have:

\[
\langle k', \mu' | \varphi | k, \mu \rangle = \varphi_{k-k'} \sum_G \langle k' + G, \mu' | e^{i(k-k') \cdot r} | k, \mu \rangle.
\]

(432)

The ‘new’ charge density is obtained by summing the individual contributions \( e^{|\phi_{k,\mu}|^2} \) over all occupied states \( \phi_{k,\mu} \). Denoting by \( p(k, \mu) \) the occupation of each state – not necessarily the equilibrium occupation
\[
\rho(r, t) = e \sum_{k, \mu} p(k, \mu) \left| \phi_{k, \mu}(r, t) \right|^2 = e \sum_{k, \mu} p(k, \mu) \left| \psi_{k, \mu}(r, t) + \delta \psi_{k, \mu}(r, t) \right|^2 \approx \rho_0(r, t) + e^2 \sum_{k, \mu} p(k, \mu) \psi^*_{k, \mu}(r, t) \sum_{G, k', \mu'} \frac{\varphi_{k-k'} \left\langle k' + G, \mu' \right| e^{i(k-k') \cdot r} \left| k, \mu \right\rangle}{E_{\mu}(k) - E_{\mu'}(k') + \hbar \omega + i \hbar s} \psi_{k', \mu'}(r, t) + cc,
\]

where \( cc \) means ‘complex conjugate’. Now notice that under the sums we may swap the dummy variables \( k \) with \( k' \) and \( \mu \) with \( \mu' \) in the \( cc \) term. Finally, expressing the product \( \psi^*_{k, \mu}(r, t) \psi_{k', \mu'}(r, t) \) in terms of its Fourier transform,

\[
\psi^*_{k, \mu}(r, t) \psi_{k', \mu'}(r, t) = \sum \left\langle k, \mu \right| e^{i\mathbf{q} \cdot \mathbf{r}} \left| k', \mu' \right\rangle e^{-i\mathbf{q} \cdot \mathbf{r}} = \sum G' \left\langle k, \mu \right| e^{i(k' - k) \cdot \mathbf{r}} \left| k' + G', \mu' \right\rangle e^{-i(k' - k) \cdot \mathbf{r}}
\]

and setting \( \mathbf{q} = k' - k \), we get:

\[
\delta \rho(r, t) = e^2 \sum_{k, \mu} \sum_{G, \mathbf{q}, \mu'} \frac{[p(k, \mu) - p(k + \mathbf{q}, \mu')]|\left\langle k + \mathbf{q} + G, \mu' \right| e^{i\mathbf{q} \cdot \mathbf{r}} \left| k, \mu \right\rangle|^2}{E_{\mu}(k) - E_{\mu'}(k + \mathbf{q}) + \hbar \omega + i \hbar s} \varphi_{\mathbf{q}} e^{-i\mathbf{q} \cdot \mathbf{r}}. \tag{434}
\]

Comparing Eq. (429) with Eq. (433), we have:

\[
\delta \rho_{\mathbf{q}} = e^2 \sum_{G, k, \mu, \mu'} \frac{[p(k, \mu) - p(k + \mathbf{q}, \mu')]|\left\langle k + \mathbf{q} + G, \mu' \right| e^{i\mathbf{q} \cdot \mathbf{r}} \left| k, \mu \right\rangle|^2}{E_{\mu}(k) - E_{\mu'}(k + \mathbf{q}) + \hbar \omega + i \hbar s} \varphi_{\mathbf{q}}. \tag{435}
\]

By Poisson equation, \( \nabla^2 \varphi^{(pol)} = -\delta \rho / \epsilon_0 \), we can easily express the Fourier components of the polarization
potential in terms of the Fourier components of the charge variation $\delta \rho_q$:

$$
\varphi^{(pol)}_q = \frac{\delta \rho_q}{q^2 \epsilon_0}.
$$

(436)

Combining Eqns. (428), (435), and (436), we can finally identify the dielectric function with:

$$
\epsilon^{(r)}(q, \omega) = 1 - \frac{e^2}{q^2 \epsilon_0} \sum_{G,k,\mu,\mu'} \frac{[p(k, \mu) - p(k + q, \mu')]|\langle k + q + G, \mu' | e^{iqr} | k, \mu \rangle|^2}{E_{\mu}(k) - E_{\mu'}(k + q) + \hbar \omega + i\hbar \delta}.
$$

(437)

− **Static Dielectric Constant.**

We are interested in situations in which the occupation of the states, $p(k, \mu)$, is the occupation of the perfect crystal at zero temperature, plus a small perturbation representing the free carriers. Thus,

$$
p(k, \mu) = p_0(k, \mu) + f(k, \mu),
$$

(438)

where $p_0$ is unity for all states in the valence bands, zero otherwise, and $f$ is the distribution of free electrons (holes) in the conduction (valence) bands. It is convenient to separate the contributions of the two different populations, since the population $p_0$ is largely unaffected by the free carriers, and the free carriers themselves matter almost exclusively in our context.

Considering the dielectric response of the valence electrons, we note that almost all of the perturbing potentials $\varphi$ we have to deal with are characterized by frequencies much smaller than the gap of the semiconductor. The Coulomb potential responsible for interband impact ionization represents a notable exception. Since contributions to the sum in Eq. (437) can be nonzero only when $p(k, \mu) = 1$ and $p(k + q, \mu') = 0$, or vice versa, the largest contributions will come from occupied (empty) states $|k, \mu\rangle$ and empty (occupied) states $|k + q, \mu'\rangle$ whose energy difference is the smallest. In other words, only states across the valence-conduction gap ought to be considered. Thus we can perform the sum over the topmost valence band and the lowest-lying
conduction band, \( \mu = v, \mu' = c \). We can use the ‘sum rule’ (which we shall not prove):

\[
\sum_f (E_f - E_i) |\langle f|e^{iq\cdot r}|i\rangle|^2 = \frac{\hbar^2 q^2}{2m},
\]

which is an extension of the sum rule for oscillator strengths. Assuming \( E_c(k + q) - E_v(k) \approx E_{gap} \), which holds approximately for wavelengths \( 1/q \) long enough to result in almost ‘vertical’ transitions between \( k \) and \( k + q \), and assuming also, as mentioned above, that \( \hbar \omega \ll E_{gap} \), we have:

\[
\sum_G |\langle k + q + G, c|e^{iq\cdot r}|k, v\rangle|^2 = \frac{\hbar^2 q^2}{2mE_{gap}}.
\]

Here \( E_{gap} \) is some average separation between the valence and conduction band. In practice, since the largest contributions to the sum in Eq. (437) come from regions of high DOS, the direct gap at the symmetry point \( X \) represents a good approximation to the quantity \( E_{gap} \) considered here. Thus, from Eqns. (437) and (440), accounting for a factor of 2 resulting from the cases \( p(k, c) = 1, p(k + q, v) = 0, \) and \( p(k, v) = 0, p(k + q, c) = 1 \), we have:

\[
\epsilon^{(r)}(q, 0) \approx 1 + \left( \frac{\hbar \omega_P}{E_{gap}} \right)^2 \equiv \epsilon_s^{(r)},
\]

where the bulk, valence band plasma frequency is given by:

\[
\omega_P = \left( \frac{e^2 n_v}{\epsilon_0 m} \right)^{1/2}.
\]

This is the static dielectric constant of the semiconductor, which approaches a constant at the long wavelength we have considered.
Response of the Free Carriers.

Considering now the response of the free carriers, we restrict ourselves to carriers in a single band, since the corresponding denominator is often (but not always!) the smallest. We shall also consider here only perturbations such that \( q \ll G \), so that the matrix element of \( e^{i\mathbf{q}\cdot\mathbf{r}} \) between states in the same band becomes unity, that is,

\[
\sum_{\mathbf{G}} \left| \langle \mathbf{k} + \mathbf{q} + \mathbf{G}, \mu | e^{i\mathbf{q}\cdot\mathbf{r}} | \mathbf{k}, \mu \rangle \right|^2 \simeq \sum_{\mathbf{G}} \left| \langle \mathbf{k} + \mathbf{q} + \mathbf{G}, \mu | \mathbf{k}, \mu \rangle \right|^2 \simeq 1 . \tag{443}
\]

A notable case for which this approximation fails is the case of zone-edge phonons, whose screening must be treated retaining the matrix element above. Then, dropping for now the band indices, we can now rewrite Eq. (437) as follows:

\[
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{\epsilon^2}{q^2 \epsilon_s} \sum_{\mathbf{k}} \frac{f(k) - f(k + q)}{E(k) - E(k + q) + \hbar \omega + i\hbar s} . \tag{444}
\]

Except for the static dielectric constant \( \epsilon_s \) in place of the vacuum value, \( \epsilon_0 \), this is exactly the so-called Lindhard’s expression valid for a free electron gas. As a matter of convenience, we introduce a ‘screening wavevector’ \( \beta(q, \omega) \) and recast Eq. (444) in the form:

\[
\epsilon(q, \omega) = \epsilon_s \left[ 1 + \frac{\beta^2(q, \omega)}{q^2} \right] . \tag{445}
\]

Thus, the final answer to our initial problem (“What is the ‘net’ potential \( \varphi(r, t) \), once the external perturbation \( \varphi^{(ex)}(r, t) \) is known?”) can be expressed as:

\[
\varphi(r, t) = \sum_{\mathbf{q}} \frac{\varphi_q^{(ex)}}{\epsilon(r)(q, \omega)} \epsilon^{i\mathbf{q}\cdot\mathbf{r}} e^{i\omega t} . \tag{446}
\]
The evaluation of the electronic component of the dielectric function requires the knowledge of the electron distribution function \( p(k) \). In transport problems this may actually be the main unknown, so that self-consistent schemes are required. But even when \( p \) is known (as in thermodynamic equilibrium), the evaluation of the RPA expression Eq. (437) or even (444) is, in general, quite complicated. Therefore, we shall now consider various cases in which this expression, ultimately needed to screen the scattering potentials, can be evaluated in closed form. This can be done by employing a simple effective mass, parabolic band approximations for the conduction band and assuming that the distribution function is well approximated by its thermal equilibrium value. Indeed, screening is particular important in high density regions of the devices. In these regions the large density results in small electric fields and, consequently, in moderate carrier heating. Thus, we set \( f(k) \approx f_0(k) \).

We shall now derive well-known expressions for the dielectric constant in the static- and long-wavelength limits, obtaining the Debye-Hückel/Thomas-Fermi expression and the semiclassical plasma frequency, respectively. Afterwards, we shall derive more general expressions — valid at all wavelengths and frequencies — in the two limits of a nondegenerate, high temperature electron gas, and of a degenerate gas at zero temperature.

1. **Static Debye-Hückel/Thomas-Fermi Screening.**

   For static perturbations (e.g. impurity scattering), we set \( \hbar \omega = 0 \) and employ the following approximations valid at long wavelengths:

   \[
   E(k + q) - E(k) \approx q \cdot \nabla_k E(k),
   \]

   \[
   f_0(k + q) - f_0(k) \approx q \cdot \frac{\partial f_0}{\partial E} \nabla_k E(k),
   \]

   so that

   \[
   \epsilon(q, 0) \approx \epsilon_s - \frac{e^2}{q^2} \int dE \mathcal{D}(E) \left( -\frac{\partial f_0}{\partial E} \right) = \epsilon_s + \frac{e^2}{q^2} \int dE \mathcal{D}(E) \left( -\frac{\partial f_0}{\partial E} \right),
   \]

   (449)
where $\mathcal{D}(E)$ is the density of states at energy $E$. Thus, we obtain the static screening parameter at long wavelengths:

$$
\beta_s^2 = \frac{e^2}{\epsilon_s} \int dE \mathcal{D}(E) \left( \frac{\partial f_0}{\partial E_F} \right) = \frac{e^2}{\epsilon_s} \frac{\partial n}{\partial E_F}.
$$

(450)

In the nondegenerate case (i.e., at very high temperatures), $\partial n/\partial E_F \simeq n/(k_BT)$ and we have the Debye-Hückel expression:

$$
\beta_{DH}^2 \simeq \frac{e^2 n}{\epsilon_s k_BT}.
$$

(451)

In the opposite case of a degenerate gas (low temperature), $\partial f/\partial E_F \simeq \delta(E - E_F)$ and we have the Thomas-Fermi expression:

$$
\beta_{TF}^2 \simeq \frac{e^2}{\epsilon_s} \mathcal{D}(E_F) = \frac{e^2}{\epsilon_s} \frac{3n}{2E_F}.
$$

(452)

2. Plasma Oscillations

We can consider now the opposite limit of dynamic effects at long wavelengths, i.e., the limit $q \to 0$. Let’s recast Eq. (444) as follows (ignoring the imaginary part):

$$
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2}{\epsilon_s q^2} \sum_k \left\{ \frac{f(k)}{E(k) - E(k + q) + \hbar \omega} - \frac{f(k + q)}{E(k) - E(k + q) + \hbar \omega} \right\}.
$$

(453)

Inside the second sum let’s set $k' = -k - q$:

$$
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2}{\epsilon_s q^2} \left\{ \sum_k \frac{f(k)}{E(k) - E(k + q) + \hbar \omega} - \sum_{k'} \frac{f(-k')}{E(-k' - q) - E(-k') + \hbar \omega} \right\}.
$$

(454)
Since \( E(-k') = E(k') \) and \( f(-k') = f(k') \):

\[
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2}{\epsilon_s q^2} \left\{ \sum_k \frac{f(k)}{E(k) - E(k + q) + \hbar \omega} - \sum_{k'} \frac{f(k')}{E(k' + q) - E(k') + \hbar \omega} \right\}.
\] (455)

Relabeling \( k \) the dummy variable \( k' \) inside the second term:

\[
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2}{\epsilon_s q^2} \sum_k f(k) \left\{ \frac{1}{E(k) - E(k + q) + \hbar \omega} - \frac{1}{E(k + q) - E(k) + \hbar \omega} \right\} =
\]

\[
= 1 - \frac{e^2}{\epsilon_s q^2} \sum_k f(k) \left\{ \frac{1}{E(k) - E(k + q) + \hbar \omega} + \frac{1}{E(k) - E(k + q) - \hbar \omega} \right\},
\] (456)

which we can write as:

\[
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 + \frac{e^2}{\epsilon_s q^2} \sum_k \frac{2f(k)[E(k) - E(k + q)]}{(\hbar \omega)^2 - [E(k) - E(k + q)]^2}.
\] (457)

Now let's expand for small deviation \( q \) from \( k \) as we had done above in Eq. (448), but retaining also terms of order \( q^2 \) (for reasons we shall understand soon enough):

\[
E(k) - E(k + q) \approx -\nabla_k E(k) \cdot q - \frac{1}{2} q \cdot \nabla^2 E(k) \cdot q.
\] (458)

Inserting this into Eq. (457) we see that the term linear in \( q \) integrates out to zero, since it is odd in \( q \). Thus the first non-vanishing term will be \((1/2)q \cdot \nabla^2 E(k) \cdot q = (q^2/2)\partial^2 E/\partial q^2 = \hbar^2 q^2/(2m^*)\). Thus:

\[
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2}{\epsilon_s q^2} \sum_k \frac{f(k)\hbar^2 q^2}{m^*(\hbar \omega)^2}.
\] (459)
Since we have ignored spin,
\[ \sum_k f(k) = n, \]  
the electron density. Thus, Eq. (459) becomes:
\[ \frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2 m}{\epsilon_s \omega^2} = 1 - \frac{\omega_p^2}{\omega^2}, \]  
where \( \omega_p \) is the plasma frequency associated to the free carriers. It is given by the expression after Eq. (337), with the density of electrons in the valence bands replaced by the (much smaller) density of free carriers. The singularity of the dielectric function identifies the frequency of the plasma modes. By expanding the numerator to higher orders in \( q \) and considering also the imaginary part of \( \epsilon \), one could draw very close analogy between this (quantum) treatment of the eigenmodes of the electron gas and the classical picture we drew before when dealing with the rms amplitude of the field associated with plasmons.

3. Nondegenerate Screening at High Temperature

In the nondegenerate limit (i.e. \( f_0(k) \simeq \exp[-E(k)/(k_B T)] \)), we can do better, by finding an expression which describes a general dynamical situation for arbitrary wavelengths, and which reduces to the Debye-Hückel expression in the static and long-wavelengths limits. Following Fetter and Walecka, we can convert the sum in Eq. (444) into an integral over \( k \). We obtain (the derivation is outlined on page 184):
\[ \Re[\beta^2(q, \omega)] \simeq \beta_{DH}^2 \left( \frac{mk_B T}{2\hbar^2 q^2} \right)^{1/2} \times \]
\[ \left\{ \Phi \left[ \left( \frac{m}{2k_B T} \right)^{1/2} \left( \frac{\omega}{q} + \frac{\hbar q}{2m} \right) \right] - \Phi \left[ \left( \frac{m}{2k_B T} \right)^{1/2} \left( \frac{\omega}{q} - \frac{\hbar q}{2m} \right) \right] \right\} \]
\[ \text{Im} [\beta^2(q, \omega)] \approx \beta_{DH}^2 \left( \frac{\pi m \omega^2}{2k_B T q^2} \right)^{1/2} \exp \left( - \frac{m \omega^2}{2k_B T q^2} - \frac{\hbar^2 q^2}{8mk_B T} \right) \frac{\sinh[\hbar \omega/(2k_BT)]}{\hbar \omega/(2k_BT)} , \] (463)

where

\[ \Phi(x) = \frac{1}{\pi^{1/2}} \mathcal{P} \int_{-\infty}^{+\infty} dy \frac{e^{-y^2}}{x-y} \] (464)

is the real part of the Plasma dispersion function and \( m \) is replaced by the effective effective mass of the lowest-lying valley or conduction band (or valence, for holes). Useful asymptotic and series expansions for the function \( \Phi \) are:

\[ \Phi(x) \approx \begin{cases} \frac{1}{x} \left( 1 + \frac{1}{2} x^{-2} + \ldots \right) & x \gg 1 \\ 2x \left( 1 - \frac{2}{3} x^2 + \ldots \right) & x \ll 1 \end{cases} , \]

so that

\[ g_1(x) \approx \begin{cases} \frac{8\pi}{x^2} \left( 1 + \frac{8\pi}{x^2} + \ldots \right) & x \gg 1 \\ 1 - x^2 \frac{2}{24\pi} + \ldots & x \ll 1 \end{cases} , \]

In the static limit, \( \omega \to 0 \), we have:

\[ \beta(q, \omega \to 0)^2 \to \beta_{DH}^2 g_1(q\ell) , \] (465)

where

\[ \ell = \left( \frac{2\pi \hbar^2}{mk_B T} \right)^{1/2} \]

is the thermal wavelength and

\[ g_1(x) = \frac{2\pi^{1/2}}{x} \Phi \left( \frac{x}{4\pi^{1/2}} \right) , \]

At long wavelengths, we recover the static Debye-Hückel expression, \( \beta(q \to 0, 0) \to \beta_{DH} \), while at short distances screening effects disappear as \( 1/q^2 \). The imaginary part of the dielectric function is related to the lifetime of the electron response: As a function of \( \omega \) it peaks at the plasma frequency, while the full-width
at half-maximum corresponds to the lifetime plasmons have before decaying into single-particle excitations (electrons ‘surfing’ the waves), process called ‘Landau damping’. The figures below show some of these features as well as the validity of the high-$T$, non-degenerate approximation compared to an ‘exact’ numerical integration of the RPA expression Eq. (437).
Si 300K $n_g = 10^{18}$ cm$^{-3}$

$\epsilon_g(q,\omega=\omega_P)$ ($\epsilon_0$)

Re($\epsilon$), Im($\epsilon$), high-T limit, numerical RPA

$\omega_P (10^{14}$ s$^{-1}$)

$\epsilon_g(q=\beta_{s,3D},\omega)$ ($\epsilon_0$)

Re($\epsilon$), Im($\epsilon$)

Si 300K $n_g = 10^{20}$ cm$^{-3}$

$\epsilon_g(q=\beta_{s,3D},\omega)$ ($\epsilon_0$)

Re($\epsilon$), Im($\epsilon$), high-T limit, numerical RPA

$\omega_P (10^{15}$ s$^{-1}$)
Derivation of Eqns. (462) and (463).

First of all, we must prove the following identity:

\[
\frac{1}{x - x_0 \pm i\eta} = \mathcal{P} \mp i\pi \delta(x - x_0),
\] (466)

where \( \mathcal{P} \) means that we should take the ‘Cauchy principal value’ (or 'part') when integrating.

In order to prove this, we must recall Cauchy’s theorem and the ‘residue’ theorem: If \( f(z) \) is a function of the complex variable \( z \) with poles at \( z_j \) (that is: singularities of the ‘benign’ form \( f(z) \sim a_j/(z - z_j) \) where \( a_j = \text{Res}[f(z_j)] \) is called the ‘residue’ of \( f \) at \( z_j \)), then:

\[
\oint f(z) \, dz = 2\pi i \sum_j \text{Res}[f(z_j)],
\] (467)

where the integration contour is a closed loop, the integration path runs counter-clockwise around the contour, and the sum extends over all residues enclosed by the contour. This is the ‘residue’ theorem. In particular, if \( f \) is analytic (i.e., does not have singularities) inside the contour, then the integral along the closed loop will vanish (Cauchy’s theorem).

Now let’s consider the integral:

\[
\oint_C \frac{f(z)}{z - x_0} \, dz,
\] (468)

where \( f(z) \) is function which is analytic everywhere (which means \( f \) has infinitely many derivatives on the entire complex plane) and the closed contour \( C \) is illustrated below:
Since we do not enclose the singularity within the contour,

\[
\oint_C \frac{f(z)}{z-x_0} \, dz = 0 .
\]  

(469)

Note that the residue of the integrand at \( z = x_0 \) is \( f(x_0) \) and, so, integrating along the full little circle around the pole \( x_0 \) we get \( 2\pi i f(x_0) \). If we integrate clockwise around the upper half of the circle – as indicated in the plot above – we get \( -i\pi f(x_0) \). Therefore, breaking the integration contour into its components:

\[
\oint_C \frac{f(z)}{z-x_0} \, dz = \lim_{\eta \to 0} \int_{-\infty}^{x_0-\eta} \frac{f(x)}{x-x_0} \, dx + \int_{C_0} \frac{f(z)}{z-x_0} \, dz + \lim_{\eta \to 0} \int_{x_0+\eta}^{\infty} \frac{f(x)}{x-x_0} \, dx + \int_{C_\infty} \frac{f(z)}{z-x_0} \, dz =
\]

\[
= \mathcal{P} \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} \, dx - i\pi f(x_0) + \int_{C_\infty} \frac{f(z)}{z-x_0} \, dz = 0 ,
\]  

(470)

having defined the Cauchy principal value:

\[
\mathcal{P} \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} \, dx = \lim_{\eta \to 0} \left( \int_{-\infty}^{x_0-\eta} \frac{f(x)}{x-x_0} \, dx + \int_{x_0+\eta}^{\infty} \frac{f(x)}{x-x_0} \, dx \right) .
\]  

(471)

But the integration along the entire contour can be written as the integration over the upper semi-circle at infinity and the integration along the real axis:

\[
\oint_C \frac{f(z)}{z-x_0} \, dz = \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} \, dx + \int_{C_\infty} \frac{f(z)}{z-x_0} \, dz ,
\]  

(472)

and so, from Eqns. (470) and (471):

\[
\int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} \, dx = \mathcal{P} \int_{-\infty}^{\infty} \frac{f(x)}{z-x_0} \, dz - i\pi f(x_0) .
\]  

(473)

This may seem like a roundabout way of proving our statement, since we could have simply re-written the integral over the real axis in the desired form without having to deal with the integration over the closed contour. However, having integrated over the entire contour \( C \), we can now note that if we had chosen to integrate counterclockwise around the lower-half little circle \( C_0 \) around \( x_0 \) we would have obtained a contribution of \( +i\pi f(x_0) \), but we would have enclosed the pole inside the big contour, so that we would have had on the left hand side of Eq. (470) the term \( 2i\pi f(x_0) \) instead of 0. Bringing the term to the left-hand
side, it would add up with \( +i\pi f(x_0) \) giving \(-i\pi f(x_0) \), as before. This shows that we can go around the singularity either clockwise or counter-clockwise and still obtain the same result.

Yet, one arbitrariness still remains. So far we have chosen to close the contour by integrating over the path \( C_\infty \) in the upper half of the complex plane. This can be rephrased mathematically by moving the singularity at \( z = x_0 \) slightly below the real axis, thus setting \( x_0 \rightarrow x_0 - i\eta \), so that closing the loop along \( c_\infty \) would not give us any residue. Thus, it is often convenient to rewrite Eq. (473) as:

\[
\int_{-\infty}^{\infty} \frac{f(x)}{x-x_0+i\eta} \, dz = \mathcal{P} \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} \, dz - i\pi f(x_0),
\]

which is equivalent to Eq. (466) after integration.

On the contrary, we might as well have chosen to close the loop by integrating over the lower half of the complex plane by setting \( x_0 \rightarrow x_0 + i\eta \). We would have obtained the same result, but the sign of the \( \delta \)-function term would have been opposite:

\[
\int_{-\infty}^{\infty} \frac{f(x)}{x-x_0-i\eta} \, dz = \mathcal{P} \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} \, dz + i\pi f(x_0).
\]

What does the choice of the sign mean physically? To understand this, we must go back to time-dependent perturbation theory: From page 19 of the notes, assuming that the perturbation was present at \( t \rightarrow -\infty \), the fifth equation may be re-written as

\[
a_k(t) = -\frac{1}{\hbar} \int_{-\infty}^{t} H_{kj}(t') \, e^{-i(E_k/\hbar - E_j/\hbar + \omega) t'} \, dt'.
\]

Replacing \( k \rightarrow k, j \rightarrow k + q, H_{kj} \rightarrow \varphi_q \) in our continuum case, we see that the denominator \( E(k - E(k + q) + \hbar \omega \) in Eq. (444) originates from the integration above, having assumed that the perturbation \( \varphi^{(ex)}(r, t) \) has the harmonic time dependence \( e^{-i\omega t} \) shown in Eq. (427). Therefore, adding a term \( \pm i\hbar s \) to \( E(k - E(k + q) + \hbar \omega \) is equivalent to assuming that the time-dependence of the perturbation is of the form \( e^{\pm st} e^{-i\omega t} \). Physically this is called an ‘adiabatic turn-on’ of the perturbation: If we select \( E(k - E(k + q) + \hbar \omega + i\hbar s \), the perturbation acquires an exponential damping backwards in time, which means that the perturbation was turned on slowly in the past (at \( t \rightarrow \infty \)) and the perturbed wavefunction responds after the perturbation has been turned on. This corresponds to the choice of a \textit{retarded} wavefunction. Conversely, the opposite choice for the sign would correspond to an \textit{advanced} wavefunction. In general, it is the particular boundary or initial condition of our problem which fixes the sign-choice for the ‘regularizing’ term \( e^{\pm st} \). For example: If we were dealing with emission and absorption of phonons, we could consider the retarded wavefunction in the case of absorption (the phonon present at \( t \rightarrow -\infty \) results in a perturbed wavefunction at later times), the advanced wavefunction in the case of emission (the wavefunction causes the presence of the perturbing phonon at later times). In both cases, the turn-on and turn-off is called ‘adiabatic’ because the limit \( s \rightarrow 0 \) is always implied and this corresponds to a turn-on (or turn-off) which is infinitely slow. In our case, the choice of the
retarded wavefunction leads us to the sign-choice we have made so far.

Let's now consider Eq. (444) and convert the sum over $k$ into an integral:

$$\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{e^2}{q^2 \epsilon_s} \int \frac{dk}{(2\pi)^3} \frac{f(k) - f(k + q)}{E(k) - E(k + q) + \hbar \omega + i\hbar s}.$$  \hfill (477)

Let's consider the integral

$$\mathcal{I} = \int \frac{dk}{(2\pi)^3} \frac{f(k) - f(k + q)}{E(k) - E(k + q) + \hbar \omega + i\hbar s}. \hfill (478)$$

Let's define

$$k' = k + \frac{1}{2}q,$$  \hfill (479)

so that

$$k = k' - \frac{1}{2}q, \quad k + q = k' + \frac{1}{2}q, \quad dk = dk'. \hfill (480)$$

Then:

$$\mathcal{I} = \int \frac{dk'}{(2\pi)^3} \frac{f(k - q/2) - f(k' + q/2)}{E(k' - q/2) - E(k' + q/2) + \hbar \omega + i\hbar s}. \hfill (481)$$

Since we integrate over the entire $k$-space, we can let $k' \rightarrow -k'$ inside the first integral. Since $f(k) = f(-k)$, we can rewrite this integral as:

$$\mathcal{I} = \int \frac{dk'}{(2\pi)^3} f(k + q/2) \left\{ \frac{1}{E(k' + q/2) - E(k' - q/2) + \hbar \omega + i\hbar s} - \frac{1}{E(k' - q/2) - E(k' + q/2) + \hbar \omega + i\hbar s} \right\}. \hfill (482)$$

Now:

$$E(k' + q/2) = \frac{\hbar^2}{2m} \left( k^2 + \frac{1}{4}q^2 + k \cdot q \right),$$  \hfill (483)

$$E(k' - q/2) = \frac{\hbar^2}{2m} \left( k^2 + \frac{1}{4}q^2 - k \cdot q \right),$$  \hfill (484)

so:

$$\mathcal{I} = \int \frac{dk'}{(2\pi)^3} f(k + q/2) \left\{ \frac{1}{\hbar \omega + i\hbar s + (\hbar^2/m)k \cdot q} - \frac{1}{\hbar \omega + i\hbar s - (\hbar^2/m)k \cdot q} \right\}. \hfill (485)$$
Let's now use Boltzmann statistics with Fermi level $E_F$ (indeed we are within the high-temperature, non-degenerate approximation) and introduce cylindrical coordinates with the $z$ axis along the direction of $\mathbf{q}$:

$$k = k_\perp + \hat{q}k_z ,$$  \hspace{1cm} (486)

where $\hat{q}$ is the unit vector along the direction of $\mathbf{q}$. Then:

$$I = e^{\beta E_F} \int \frac{dk_\perp}{(2\pi)^2} \exp \left(-\beta \frac{\hbar^2 k_\perp^2}{2m} \right) \int \frac{dk_z}{2\pi} \exp \left(-\beta \frac{\hbar^2 (k_z + q/2)^2}{2m} \right) \times$$

$$\times \left\{ \frac{1}{\hbar\omega + i\hbar s + (\hbar^2/m)k_z q} - \frac{1}{\hbar\omega + i\hbar s - (\hbar^2/m)k_z q} \right\} ,$$  \hspace{1cm} (487)

where, as usual, $\beta = 1/(k_B T)$. The integration over $\mathbf{k}_\perp$ can be done easily: Going to polar coordinates and setting $E_\perp = \hbar^2 k_\perp^2/(2m)$, so that $k_\perp dk_\perp = (m/\hbar^2)dE_\perp$, we have:

$$\int \frac{dk_\perp}{(2\pi)^2} \exp \left(-\beta \frac{\hbar^2 k_\perp^2}{2m} \right) = \int_0^{2\pi} d\phi \int_0^\infty \frac{dk_\perp}{(2\pi)^2} k_\perp e^{-\beta \hbar^2 k_\perp^2/(2m)}$$

$$= \frac{1}{2\pi} \frac{m}{\hbar^2} \int_0^\infty dE_\perp e^{-\beta E_\perp} = \frac{m}{2\pi\hbar^2\beta} .$$  \hspace{1cm} (488)

In order to handle the remaining integral over $k_z$, let's set the new integration variable (writing now $k$ for $k_z$):

$$x^2 = \frac{\beta \hbar^2}{2m} \left( k + \frac{q}{2} \right)^2 , \hspace{1cm} k + \frac{q}{2} = \left( \frac{2m}{\beta \hbar^2} \right)^{1/2} x , \hspace{1cm} dk = \left( \frac{2m}{\beta \hbar^2} \right)^{1/2} dx , \hspace{1cm} \frac{\hbar^2 q}{m} = \left( \frac{2\hbar^2}{m\beta} \right)^{1/2} qx - E_q ,$$  \hspace{1cm} (489)
where \( E_q = \hbar^2 q^2 / (2m) \). Thus we have:

\[
\mathcal{I} = \frac{m e^{\beta E_F}}{2\pi \beta \hbar^2} \left( \frac{2m}{\beta \hbar^2} \right)^{1/2} \int_{-\infty}^{\infty} dx \, e^{-x^2} \left\{ \frac{1}{\hbar \omega + i\hbar \omega + [2\hbar^2 / \beta \hbar^2]^{1/2} q x - E_q + [2\hbar^2 / (m\beta)]^{1/2} q x} \right\},
\]

or, setting

\[
x_\pm = \left( \frac{m\beta}{2\hbar^2} \right)^{1/2} \frac{1}{q} \left[ \hbar \omega \pm \frac{\hbar^2 q^2}{2m} + i\hbar \right],
\]

we now have:

\[
\mathcal{I} = \frac{m e^{\beta E_F}}{2\pi \beta \hbar^2} \frac{m}{2\pi \hbar^2 q} \int_{-\infty}^{\infty} dx \left\{ \frac{e^{-x^2}}{x_\pm + x} - \frac{e^{-x^2}}{x_+ + x} \right\}.
\]

Now we make use of Eq. (466). Considering first the real part, defining the 'plasma dispersion function'

\[
\Phi(y) = \frac{1}{\pi^{1/2}} \mathcal{P} \int_{-\infty}^{\infty} dx \, e^{-x^2 / (x - y)}.
\]

after a little algebra we get

\[
\text{Re}(\mathcal{I}) = \frac{n}{\hbar q} \left( \frac{1}{2} m \beta \right)^{1/2} \left[ \Phi(y_-) - \Phi(y_+) \right],
\]

where

\[
y_\pm = \left( \frac{\beta m}{2} \right)^{1/2} \left( \frac{\omega \pm \hbar q}{2m} \right).
\]

This leads to Eq. (462). The imaginary part, Eq. (463), should follow quite easily now from the imaginary part of Eq. (492) using (466) as well.

4. **Degenerate Screening at Low Temperature.**

In the opposite limit of a degenerate electron gas at low temperature, we can approximate the Fermi function
\( f_0(k) \) with a step function at the Fermi level. Then, the screening wavevector becomes:

\[
\beta(q, \omega)^2 \approx \beta_{TF}^2 \left\{ \frac{1}{2} + \frac{1}{8\eta^3} \left[ (\gamma + \eta^2)^2 - \eta^2 \right] \ln \left| \frac{\eta + \eta^2 + \gamma}{\eta - \eta^2 - \gamma} \right| + \\
+ \left[ (\gamma - \eta^2)^2 - \eta^2 \right] \ln \left| \frac{\eta + \eta^2 - \gamma}{\eta - \eta^2 + \gamma} \right| \right\},
\]

(496)

where \( \eta = q/(2k_F) \) and \( \gamma = (\hbar \omega + is)/(4E_F) \). In the limit \( q \to 0 \) we find:

\[
\frac{\epsilon(q, \omega)}{\epsilon_s} = 1 - \frac{\omega_p^2}{\omega^2} \left( 1 + \frac{3v_Fq^2}{5\omega^2} \right),
\]

where \( v_F = \hbar k_F/m \) is the Fermi velocity, which shows how higher order corrections to Eq. (459) introduce a (weak) dispersion for the plasma modes. In the static limit, Eq. (496) becomes:

\[
\beta(q, 0)^2 \approx \beta_{TF}^2 \left\{ \frac{1}{2} + \frac{4k_F^2 - q^2}{8k_F q} \ln \left| \frac{2k_F + q}{2k_F - q} \right| \right\}.
\]

(497)

As \( q \to 0 \) this expression tends to the Thomas-Fermi expression Eq. (452), while screening effects become weaker at short distances, similarly to what implied by the high temperature limit, Eq. (465).
Screened Impurity and Friedel’s Sum Rule

A particularly important application is given by the static potential of an ionized impurity, such as dopant impurities in semiconductors. Let’s assume that the impurity has a net charge $eZ$, and let’s keep the discussion as simple as possible. From Eq. (444), the full expression Eq. (462), the static limit Eq. (465), and from the approximate expressions for the plasma dispersion function, we have:

$$
\varphi(\mathbf{r}) = \frac{eZ}{\epsilon_s} \sum_q \frac{e^{iq \cdot r}}{q^2 + \beta(q, 0)^2} = \frac{eZ}{\epsilon_s r 4\pi^2} \int_{-\infty}^{\infty} dq \frac{e^{iqr}}{q^2 + \beta(q, 0)^2}, \quad (498)
$$

having assumed that the screening wavevector depends only on the magnitude of $q$ since the impurity potential does not depend on time. At large distances away from the impurity ($r \gg 1/\beta$), we can approximate $\beta$ with its long-wavelength limit, Eq. (450), obtaining the well-known Yukawa-type potential:

$$
\varphi(\mathbf{r}) = \frac{eZ}{4\pi \epsilon_s r e^{-\beta_s r}}. \quad (499)
$$

As we have seen, the particular form of the screening parameter $\beta(q, 0)$ may be quite complicated, depending on the approximation embraced. Thus, expression (499) above will fail at short distances, when the $q$-dependence of $\beta(q, 0)$, as given in Eq. (465), is important. Therefore, the scattering cross section obtained from Eq. (499) above may be inaccurate. However, we can follow a completely different path to derive a particular condition on $\beta$, valid generally for arbitrary potentials. We shall consider the simpler case of zero temperature, but the general formulation at arbitrary temperature can be obtained by modifying all sums up to the Fermi level as sums over all states weighted by their thermal occupation.

In the absence of the impurity, the electron gas consists of states occupied to the Fermi level. Instead of considering plane waves in a box, as usually done, we can consider a spherical volume and express the plane waves in spherical harmonics.
We write the Laplace operator in polar coordinates and write Schrödinger equation as:

\[-\frac{\hbar^2}{2m^*} \frac{1}{r^2 \sin \theta} \left[ \sin \theta \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial^2 \psi}{\partial \phi^2} \right] + V(r) \psi = E \psi.\] (500)

Separating variables, we find that the azimuthal component \( \Phi(\phi) \) must satisfy the equation

\[
\frac{1}{\Phi(\phi)} \frac{\partial^2 \Phi}{\partial \phi^2} = \text{constant}.
\] (501)

Since we require single-valued solutions, the constant must be equal to \(-m^2\), where \(m\) is an integer, so that \(\Phi(\phi) = e^{\pm im \phi}\).

The polar component \(\Theta(\theta)\) must satisfy the equation:

\[
\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \Theta}{\partial \theta} \right) - \frac{m^2}{\sin^2 \theta} \Theta = -l(l+1) \Theta,
\] (502)

where the ‘separation constant’ has been set to the total angular momentum quantum number \(l(l+1)\). Finally, the radial component of the wavefunction must satisfy the equation:

\[-\frac{\hbar^2}{2m^*} \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \left[ \frac{\hbar^2 l(l+1)}{2m^*r^2} + V(r) \right] R(r) = E R(r).\] (503)

At large distances \(r\) the potential \(V\) becomes negligible and the solution of Eq. (503) is the spherical Bessel function \(j_{kl}(r)\) (where as usual \(k = (2m^*E)^{1/2}/\hbar\)). Its asymptotic behavior at large \(r\) is \(\sim sin(kr - \pi l/2)/r\).

Each state is then labeled by the magnitude of the \(k\)-vector and by its angular momentum quantum number \(l\):

\[
\psi_{k,l}(r, \theta) \sim \frac{1}{r} \sin \left( kr - \frac{1}{2}l\pi \right) P_l(\cos \theta),
\] (504)

(where \(P_l(x)\) is the Legendre polynomial of order \(l\)) at large distances. We assume that the wavefunctions vanish on some boundary \(r = R\), where \(R\) is sufficiently large so that the (screened) potential vanish.
Therefore, the allowed values of $k$ are labeled by an integer $n$ such that $knR = n\pi + \pi l/2$. The number of electrons, $N_R$, in the volume can be obtained by counting all states up to the Fermi level, that is:

$$N_R \sim 2 \sum_{n<n_F, l \leq n} (2l + 1),$$

(505)

where a factor 2 comes from spin, the factor $(2l + 1)$ represents the number of all possible angular momentum states for each $l$, and the maximum 'allowed' radial quantum number $n_F$ is set by the Fermi level, i.e., $n_F = k_FR/\pi - l/2$. If we now introduce the impurity potential, the radial component of the new wavefunction obeys the equation:

$$\left[ -\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) + \frac{l(l + 1)}{r^2} + \frac{2me}{\hbar^2} \varphi(r) - k^2 \right] R_{l,k}(r) = 0.$$  

(506)

Now, thanks to the spherical symmetry of the scattering potential, we know that at large distances the angular dependence of the wavefunction must remain unaltered. All the external potential can do is to shift the phases of the partial radial waves. This is known as the partial wave method. Therefore, at large distances the wavefunctions can be approximated by:

$$\phi_{k,l}(r, \theta) \sim \frac{1}{r} \sin \left[ kr - \frac{1}{2} l\pi + \eta_l(k) \right] P_l(\cos \theta),$$

(507)

where $\eta_l(k)$ is the phase shift of the partial wave, given by

$$\sin[\eta_l(k)] = -\frac{2me}{\hbar^2} \int_0^{\infty} dr \ r^2 j_l(kr) R_{l,k}(r) \varphi(r).$$

(508)

Note that in the Born approximation, we assume that all shifts are small and we substitute $R_{l,k}(r)$ with its
'free particle' expression $R_{l,k}(r) \simeq j_l(kr)$, so that

$$\eta_l^{\text{Born}}(k) \simeq -\frac{2mek}{\hbar^2} \int_0^\infty dr \, r^2 j_l^2(kr) \varphi(r).$$

(509)

It is easy to see that the phase-shift results in additional electrons being placed in the volume: For every $l$-state, the maximum index $n_F$ is increased by $\eta_l(k_F)/\pi$. Therefore, the change in the number of electrons in the volume is given by:

$$\delta N_R = \frac{2}{\pi} \sum_l (2l + 1) \eta_l(k_F).$$

(510)

This extra charge must accounted for by the electrons required to screen the perturbing potential. Therefore, since precisely $Z$ electrons are needed to neutralize the charge of the impurity at large distances, we must have:

$$Z = \frac{2}{\pi} \sum_l (2l + 1) \eta_l(k_F).$$

(511)

This is Friedel's sum rule. This result depends only on the assumption that the Fermi level is unchanged far away from the impurity and that the perturbing potential must be neutralized by excess carriers at some finite distance.

In the presence of a general potential, we can compute the scattering cross section as follows. We start with a screened potential of the Yukawa form (499), where $\beta$ is now treated as a parameter to be determined. We compute the phase shifts and fix $\beta$ so that Friedel’s sum rule, Eq. (511), is satisfied. Finally, we obtain the cross section from the well-known formula

$$\sigma(k, \theta) = \frac{1}{k^2} \left| \sum_l (2l + 1) \sin[\eta_l(k)] e^{i\eta_l(k)} P_l(\cos \theta) \right|^2.$$

(512)

This result is usually more accurate than the use of the matrix element $\varphi_q$, as in the Born approximation. In particular, the difference between attractive and repulsive Coulomb scattering (which vanish in the Born
approximation), the charge oscillations associated with either the wavefunctions (507) or the screening parameter (465), and the general dependence of $\beta$ on $q$ will be accounted for.

### Two-Dimensional Screening

The response of a two-dimensional electron (or hole) gas to an external perturbation can be analyzed in the same way. There are, however, slight complications caused mainly by the fact that carriers can occupy many subbands. Therefore, it is useful to derive here a general expression for screening caused by carriers in inversion layers or quantum wells. A complete discussion can be found in the Ando-Fowler-Stern Rev. Mod. Phys. article as well as in Ridley’s text. Here we shall simplify somewhat the picture. In particular, we shall consider here only longitudinal screening. Intersubband excitations, considered by Dahl and Sham in a famous paper of theirs, can play a nonnegligible role, particularly in (resonant) dynamic situations and give rise to a transverse component of the polarizability, since intersubband transitions cause a redistribution of the charge density in the direction normal to plane of quantization.

Assuming that the external perturbation is isotropic in the plane of confinement, it is convenient to express it in terms of its Fourier-Bessel expansion:

$$\varphi^{(ex)}(\mathbf{R}, z, t) = 2\pi \int dQ Q J_0(QR) \varphi^{(ex)}_Q(z) e^{i\omega t},$$

(513)

where $J_0$ is the spherical Bessel function of order 0 and the variables $\mathbf{R}$ and $\mathbf{Q}$ denote the coordinate and wavevector in the quantization plane and $z$ is the coordinate normal to the plane. The electron wavefunctions in each subband (we shall label subbands with Greek letters $\lambda$, $\mu$, $\nu$, since we shall not consider bands here and no confusion should arise) can be written as:

$$\psi_{K,\mu}(\mathbf{R}, z) = e^{i\mathbf{K} \cdot \mathbf{R}} \zeta_\mu(z) \sum_G e^{i\mathbf{G} \cdot \mathbf{r}} u_G(k, 0) \equiv |K, \mu\rangle.$$

(514)

Under the net perturbation $\varphi = \varphi^{(ex)} + \varphi^{(pol)}$, the wavefunctions are modified, similarly to Eq. (430), as
follows:

$$\psi_{K,\mu}(\mathbf{R}, z) \to \phi_{K,\mu}(\mathbf{R}, z) = \psi_{K,\mu}(\mathbf{R}, z) + \sum_{K',\mu'} \frac{e\langle K', \mu' | \varphi | K, \mu \rangle}{E_{\mu}(K) - E_{\mu'}(K') + \hbar \omega + i\hbar s} \psi_{K',\mu'}(\mathbf{R}, z).$$

(515)

We have noticed before that the matrix element $\langle K', \mu' | \varphi | K, \mu \rangle$ is, in principle, a quite complicated object, since it involves the Bloch wavefunctions themselves. However, assuming as we did before that the external perturbation is of wavelength much longer than any nonzero $\mathbf{G}$-vector, the overlap term over Bloch factors yields unity and

$$\langle K', \mu' | \varphi | K, \mu \rangle \simeq \varphi_{K-K',\mu\mu'} \equiv \int dz \, \zeta_{\mu}(z) \varphi_{K-K'}(z) \zeta_{\mu'}(z).$$

(516)

In complete analogy with the derivation of Eq. (435), we have for the Fourier-Bessel components of the induced charge density:

$$\delta \rho_Q(z) = 2e^2 \sum_{K,\mu,\mu'} \left[ g_{\mu} f^{(2D)}(K, \mu) - g_{\mu'} f^{(2D)}(K + Q, \mu') \right] \frac{\varphi_{Q,\mu\mu'}(z) \zeta_{\mu}(z) \zeta_{\mu'}(z)}{E_{\mu}(K) - E_{\mu'}(K + Q) + \hbar \omega + i\hbar s},$$

(517)

where we have indicated by $g_{\mu}$ the degeneracy of the $\mu$-th subband. The form of the Poisson equation replacing Eq. (436) in a two-dimensional context is less transparent. Indeed, both in the case of inversion layers as well as of quantum wells, the discontinuity of the permittivity $\epsilon_s$ across interfaces, such as the Si-SiO$_2$ or GaAs-Al$_x$Ga$_{1-x}$As interfaces, should be handled by the appropriate boundary conditions, reflecting the presence of image charges. In order to keep the discussion at a general level, we need to consider the Green’s function $G_Q(z, z')$ such that

$$\left( \frac{d^2}{dz^2} - Q^2 \right) G_Q(z, z') = \delta(z - z'),$$

(518)
subject to the boundary conditions appropriate to the geometry at hand. A notable example is the geometry of two semi-infinite media (namely, Si and SiO$_2$ characterized by the dielectric constants $\epsilon_s$ and $\epsilon_{ox}$), for which we have (see Ando-Fowler-Stern):

$$G_Q(z, z') = \frac{1}{2Q} \left[ e^{-Q|z-z'|} + \frac{\epsilon_s - \epsilon_{ox}}{\epsilon_s + \epsilon_{ox}} e^{-Q(z+z')} \right]. \quad (519)$$

Expressions for more complicated geometries have been given by Dahl and Sham. Then, from Eq. (518) we can obtain the self-consistent equation for the net potential:

$$\varphi_Q(z) = \varphi_Q^{(ex)}(z) + \int dz' G_Q(z, z') \frac{\delta \rho_Q(z')}{\epsilon_s}. \quad (520)$$

Only the matrix elements over the envelope wavefunctions $\zeta_\mu(z)$ are required in order to evaluate the scattering rates caused by the perturbation $\varphi^{(ex)}$. From Eqns. (516), (517), and (520) we have:

$$\varphi_{Q,\mu\mu'} = \varphi_{Q,\mu\mu'}^{(ex)} - \sum_{\lambda,\lambda'} \frac{\beta_{\lambda\lambda'}^{(2D)}(Q, \omega)}{Q} G_{Q,\mu\mu',\lambda\lambda'} \varphi_{Q,\lambda\lambda'}, \quad (521)$$

where the ‘form factor’ $G$ is given by:

$$G_{Q,\mu\mu',\lambda\lambda'} = 2 \int dz dz' Q G_Q(z, z') \zeta_\mu(z) \zeta_{\mu'}(z) \zeta_\lambda(z') \zeta_{\lambda'}(z'), \quad (522)$$

and the two-dimensional screening wavevector is given by:

$$\beta_{\mu\mu'}^{(2D)}(Q, \omega) = \frac{e^2}{2\epsilon_s} \sum_K \frac{g_{\mu} f^{(2D)}(K, \mu) - g_{\mu'} f^{(2D)}(K + Q, \mu')}{E_{\mu'}(K + Q) - E_{\mu}(K) - \hbar \omega - i\hbar s}. \quad (523)$$
The self-consistent Eq. (521) can be recast in terms of a dielectric matrix:

\[
\varphi_{Q,\mu\mu'} = \sum_{\lambda\lambda'} [\epsilon_{\mu\mu',\lambda\lambda'}(Q, \omega)]^{-1} \varphi_{Q,\lambda\lambda'}^{(ex)},
\]

where

\[
\epsilon_{\mu\mu',\lambda\lambda'}(Q, \omega) = \delta_{\mu\mu'} \delta_{\lambda\lambda'} - \frac{\beta^{(2D)}(Q, \omega)}{Q} G_{Q,\mu\mu',\lambda\lambda'}. \tag{525}
\]

The problem of finding the net potential from the external perturbation is now translated to the problem of inverting this dielectric matrix.

As described by Eq. (524), the problem is a complicated one. A relative simplification can be obtained by noting that terms of \( G_{Q,\mu,\mu',\lambda\lambda'} \) with \( \lambda' \) different from \( \lambda \) do not contribute to a net change of the charge in the layer, but simply to a polarization of the charge in the direction normal to the plane of quantization. This is seen by integrating Eq. (517) along \( z \): for \( \lambda \neq \lambda' \) we have \( \delta \rho_Q = 0 \). Thus we can consider only the ‘longitudinal’ dielectric function, by retaining only ‘diagonal’ terms \( G_{Q,\mu,\mu',\lambda\lambda} \) and set \( \beta^{(2D)}_{\lambda\lambda} \equiv \beta^{(2D)}_\lambda \).

Thus, the change of the charge density becomes proportional to the expectation value of the potential in each subband, which the customary assumption. The problem given by Eq. (524) can now be solved in two steps.

First, the ‘diagonal’ problem \( \mu = \mu' \) is solved by inverting the matrix \( \epsilon_{\mu\mu,\lambda\lambda}(Q, \omega) \). Once the diagonal matrix elements \( \varphi_{Q,\mu\mu,\lambda\lambda} \) are known, the off-diagonal terms can be evaluated directly from Eq. (524).

Equation (523) is very similar to the Lindhart expression, Eq. (444), and we can obtain closed-form expression for it in the same limits considered above. We shall consider only the diagonal (in the subband indices) screening wavevector \( \beta^{(2D)}_{Q,\mu\mu'} \).

1. **Static Screening, Long-Wavelength Limit**

For static perturbations, we proceed along the path followed to derive Eq. (447) obtaining:

\[
\beta^{(2D)}_{s,\mu} = \frac{e^2}{2\epsilon_s} \int dE \mathcal{D}_\mu(E) \left[ \frac{\partial f^{(2D)}(k, \mu)}{\partial E_F} \right] = \frac{e^2}{2\epsilon_s} \frac{\partial n_\mu}{\partial E_F}, \tag{526}
\]
where $D_\mu$ and $n_\mu$ refer to the DOS and density in the $\mu$-th subband, respectively. In the nondegenerate case (i.e., at high temperatures), $\partial n_\mu / \partial E_F \simeq n_\mu / (k_B T)$ and we obtain the two-dimensional equivalent of the Debye-Hückel limit:

$$\beta_{DH,\mu}^{(2D)} \simeq \frac{e^2 n_\mu}{2\epsilon_s k_B T}.$$  \hspace{1cm} (527)

In the opposite case of a degenerate gas (low temperature), $\partial f / \partial E_F \simeq \delta(E - E_F)$ and we have equivalent of the Thomas-Fermi expression:

$$\beta_{TF,\mu}^{(2D)} \simeq \frac{e^2 g_\mu m_{d,\mu}}{2\epsilon_s 2\epsilon_s \pi \hbar^2},$$  \hspace{1cm} (528)

for occupied subbands, $m_{d,\mu}$ being the DOS-effective mass in the $\mu$-th subband. Note that this expression implies, somewhat counter-intuitively, that the screening wavevector is independent of the density of free carriers. This depends on our perhaps oversimplified assumption of ignoring higher order terms in $Q$ in order to derive Eq. (526). We shall see below (Eq. (537)) that a more accurate derivation overcomes this difficulty, at least partially. However, it is generally true that screening in the two-dimensional quantum limit depends weakly on the density of free carriers, as compared to the bulk case.

2. Two-Dimensional Plasma Oscillations

Proceeding in analogy with Eqns. (459) and (461), we obtain in the limit of long wavelengths and high frequencies, $\hbar \omega \gg E_\mu(K + Q) - E_\mu(K)$ for all states of interest:

$$\beta_\mu^{(2D)}(Q, \omega) \simeq \frac{e^2 g_\mu n_\mu}{2\epsilon_s m_\mu \omega^2 Q^2}.$$  \hspace{1cm} (529)

This expression, when substituted into the dielectric matrix, Eq. (525), is still quite complicated. In the simple case of only one subband ($\mu = 0$) occupied, and in the isotropic case, we have:

$$\frac{\epsilon_0(Q, \omega)}{\epsilon_s} \simeq 1 - \frac{\omega_p^{(2D)}(Q)^2}{\omega^2},$$  \hspace{1cm} (530)
where the ‘plasma frequency’ now exhibits a strong dispersion, unlike the almost dispersionless behavior of the bulk plasmons:

\[ \omega_p(Q)^2 = \frac{e^2 n_0}{2\epsilon_s m_0} Q \mathcal{G}_{Q,0,0,0} \approx \frac{e^2 n_0}{2\epsilon m_0} Q, \]  

(531)

where in deriving the last expression we have ignored the thickness of the two-dimensional layer \((\text{i.e. } \zeta_0(z) \approx \delta(z))\) to evaluate explicitly the form factor \(\mathcal{G}\), we have considered the geometry of two semi-infinite media \((\text{i.e., the Green function is given by Eq. (519)})\), and we have set \(\bar{\epsilon} = (\epsilon_s + \epsilon_{ox})/2\). The presence of boundaries will alter the dispersion of the plasma modes, because of the \(Q\)-dependence of the form factor \(\mathcal{G}\). A notable case if the presence of a gate insulator of finite thickness: In the limit of a very thin insulator, the dependence of the plasma frequency on \(Q\) will actually become linear (see Ando-Fowler-Stern).

3. **Nondegenerate Screening at High Temperature**

In the high temperature limit, we can follow once more Fetter and Walecka as we have done before in the case of three dimensions, the only difference being an overall factor \(2/\ell_\mu\). Thus, in analogy with Eqns. (462) and (463), we have:

\[
\text{Re}[\beta^{(2D)}_\mu(Q, \omega)] \approx \beta^{(2D)}_{DH,\mu} \frac{\pi^{1/2}}{Q\ell_\mu} \times \left\{ \Phi \left[ \left( \frac{m_\mu}{2k_B T} \right)^{1/2} \left( \frac{\omega}{Q} + \frac{\hbar Q}{2m_\mu} \right) \right] - \Phi \left[ \left( \frac{m_\mu}{2k_B T} \right)^{1/2} \left( \frac{\omega}{Q} - \frac{\hbar Q}{2m_\mu} \right) \right] \right\} \]

(532)

\[
\text{Im}[\beta^{(2D)}_\mu(Q, \omega)] \approx \beta^{(2D)}_{DH,\mu} \frac{\pi \hbar \omega}{Q\ell_\mu k_B T} \exp \left( - \frac{m_\mu \omega^2}{2k_B T Q^2} - \frac{\hbar^2 Q^2}{8m_\mu k_B T} \right) \frac{\sinh[\hbar \omega/(2k_B T)]}{\hbar \omega/(2k_B T)}, \]

(533)

In the long-wavelength limit we have:

\[ \beta^{(2D)}_\mu(Q, \omega \to 0) \approx \beta^{(2D)}_{DH,\mu} g_1(Q\ell_\mu). \]

(534)

4. **Screening in the Quantum Limit**

In the so-called ‘extreme quantum limit’ of zero temperature, the two-dimensional electron gas is fully degenerate and we assume that only the ground-state subband is occupied. Under these conditions Stern
has obtained an analytic expression for the dielectric function:

\[ \beta^{(2D)}(Q, \omega) = \frac{e^2 m_0 n_0}{\hbar^2 k_F Q} \left\{ \left[ C_+(A_+^2 - 1)^{1/2} + C_-(A_-^2 - 1)^{1/2} - \frac{Q}{2k_F} \right] + i \left[ D_-(1 - A_+^2)^{1/2} + D_-(1 - A_-^2)^{1/2} \right] \right\}, \]

(535)

where:

\[ A_\pm = \frac{Q}{2k_F} \pm \frac{\omega}{v_F Q}, \]

and

\[ \left\{ \begin{array}{c} C_\pm = \text{sgn}(A_\pm) \\
D_\pm = 0 \end{array} \right\} |A_\pm| > 1, \]

\[ \left\{ \begin{array}{c} C_\pm = 0 \\
D_\pm = 1 \end{array} \right\} |A_\pm| < 1. \]

In the long-wavelength limit, \( Q \to 0 \), the expression above provides higher order (in \( Q \)) corrections to the dispersion of the plasma modes, exactly as we found in three dimensions:

\[ \frac{\epsilon^{(2D)}(Q, \omega)}{\epsilon_s} \to 1 - \left[ \omega_p^{(2D)}(Q)^2 + \frac{3\epsilon_s}{4\epsilon} Q^2 v_F^2 \right] \frac{1}{\omega^2}, \]

(536)

while in the static limit, \( \omega \to 0 \), we find:

\[ \beta^{(2D)}(Q, \omega) \to \left\{ \begin{array}{c} \beta^{(2D)}_{TF} \\
\beta^{(2D)}_{TF} \left\{ 1 - [1 - (2k_F/Q)^2]^{1/2} \right\} \end{array} \right\} Q < 2k_F, \]

\[ Q > 2k_F, \]

(537)

thus overcoming the counter-intuitive results of Eq. (528) that the screening length does not depend on density in the degenerate limit.

Finally, we consider the two-dimensional version of the screened impurity potential and Friedel’s sum rule: By arguments similar to those used to derive Eq. (511) we have, in the limit of zero temperature and only one
subband occupied:

\[ Z = \frac{2}{\pi} g_0 \sum_{l=-\infty}^{\infty} \eta_l(k_F), \]  

(538)

where \( g_0 \) is the degeneracy of the ground-state subband, while the cross section derived from the partial waves analysis is:

\[ \sigma(k, \theta) = \frac{2}{\pi k} \sum_{l=-\infty}^{\infty} \left| e^{i[l\theta + \eta_l(k)]} \sin[\eta_l(k)] \right|^2. \]  

(539)